SELINA
SCIENCE FOR EVIDENCE-BASED AND SUSTAINABLE DECISIONS ABOUT NATURAL CAPITAL

D5.2 Enhancing the spatial and temporal resolution of ecosystem accounts using satellite data

30/04/2025

Lead Author/s:
Andrea Peters, Michelle Watson, Lori Giagnacovo, Bruno Smets, Marcela Quinoñones, Diego Barbulo, Boris Kooij, Artur Gil, Filipe Bernardo, Panayotis Dimopoulos and Ioannis P. Kokkoris

Funded by
the European Union

SELINA receives funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101060415. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.

Prepared under contract from the European Commission
Grant agreement No. 101060415
EU Horizon 2020 Research and Innovation Actions

Project acronym:	**SELINA**
Project full title:	**Science for Evidence-based and sustainabLe decIsions about NAtural capital**
Project duration:	01.07.2022 – 30.06.2027 (60 months)
Project coordinator:	Prof. Dr. Benjamin Burkhard, Gottfried Wilhelm Leibniz University Hannover
Call:	HORIZON-CL6-2021-BIODIV-01
Deliverable title:	Enhancing the spatial and temporal resolution of ecosystem accounts using satellite data
Deliverable n°:	D5.2
WP responsible:	WP5
Nature of the deliverable:	Report
Dissemination level:	Public
Lead beneficiary:	WU
Citation:	Peters, A., Watson, M., Giagnacovo, L., Smets, B., Quiñones, M., Barbulo, D., Kooij, B., Gil, A., Bernardo, F., Dimopoulos, P. & Kokkoris, I.P. (2025). SELINA D5.2 Enhancing the spatial and temporal resolution of ecosystem accounts using satellite data.
Due date of deliverable:	34
Actual submission date:	34

Deliverable status:

Version	Status	Date	Authors	Reviewers
1.0	First draft for review	1st April 2025	T5.2 taskforce	Nicolas Grondard (WU)
1.1	Second Draft for review	April 15th 2025	T5.2 taskforce	Benjamin Burhard (LeibnizUniversity Hannover) Fernado Santos- Martin (Rey Juan Carlos University)
1.2	Final	April 30th 2025	T5.2 taskforce	

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Table of Contents

1. Preface 21

2. Executive Summary 22

3. List of abbreviations 25

4. Introduction 26

4.1. Ecosystem Accounting 26

4.2. Aims & Objectives 28

4.3. Expected outcomes 29

5. Test sites 30

5.1. Peloponnese, Greece 30

5.2. São Miguel, Azores (Portugal) 32

6. Data 35

6.1. Existing data 35

6.1.1. National data 36

6.1.1.1. Peloponnese TS 36

6.1.1.2. São Miguel TS 36

6.1.2. CLMS data 38

6.2. *In situ* data collection 39

6.2.1. Peloponnese TS 39

6.2.2. São Miguel TS 42

7. Methodology 44

7.1. Ecosystem extent mapping 44

7.1.1. National-centric approach 44

7.1.1.1. São Miguel TS 45

7.1.1.2. Peloponnese TS 51

7.1.2. Vegetation centric approach 55

7.1.2.1. Extent mapping workflow 55

7.1.3. Mapping EUNIS habitats 56

7.1.4. Mapping EU extent 58

7.1.5. EUNIS habitat mapping for Peloponnese 64

7.1.5.1. Training data preparation 64
7.1.5.2. Post-processing steps 66
7.1.5.3. Level 1 EUNIS Habitat map 67
7.1.5.4. Level 2 EUNIS Habitat Map 70
7.1.5.5. Level 3 EUNIS Habitat Map 72
7.1.5.6. Validation 75
7.1.6. EUNIS Habitat Mapping for São Miguel 79
7.1.6.1. Training data preparation 79
7.1.6.2. Post-processing steps 82
7.1.6.3. Level 1 EUNIS Habitat map 84
7.1.6.4. Level 2 EUNIS Habitat Map 85
7.1.6.5. Level 3 EUNIS Habitat Map 88
7.1.6.6. Validation 91
7.2. Ecosystem condition mapping 95
7.2.1. PEOPLE-EA approach for forest condition index 95
7.2.2. Peloponnese – ARIES for PEOPLE-EA 102
7.2.3. São Miguel – Forest condition assessment 105
7.3. Forest Carbon Accounting 112
7.3.1. Carbon stock maps using remote sensing derived products 113
7.3.2. Carbon Flux mapping using GPP 119

8. Results 125

8.1. Ecosystem extent mapping 125
8.1.1. National-Centric Approach 125
8.1.1.1. Peloponnese 125
8.1.1.2. São Miguel 131
8.1.2. Vegetation-centric approach 134
8.1.2.1. Peloponnese extent account 134
8.1.2.2. São Miguel Extent account 137
8.2. Ecosystem condition mapping 140
8.2.1. Peloponnese Forest Condition Index 140
8.2.2. São Miguel Forest Condition Index 144
8.3. Forest Carbon Accounting 147
8.3.1. Carbon stock maps using remote sensing-derived products 147
8.3.1.1. Results and validation of Forest structural maps 147
8.3.1.2. SarSentry Deforestation and degradation time series 151

8.3.1.3. Biomass Maps 155
8.3.1.4. Carbon Maps 159
8.3.1.5. Carbon emissions by carbon stock changes 160
8.3.2. Carbon Flux mapping using GPP 163
9. Discussion 165
9.1. Ecosystem extent – National-centric approach...165
9.2. Ecosystem extent – Complementing the vegetation-centric approach with forest structure mapping 166
9.3. Ecosystem condition mapping – Forest Condition Index 167
9.4. Carbon flux mapping..172
9.5. Carbon stock mapping using remote sensing-derived products 176
9.6. Links with other SELINA work packages 178
9.6.1. SELINA WP3..178
9.6.2. SELINA WP4..179
9.7. Uptake by SELINA Demonstration Projects and Compendium of Guidance (CoG) 179
9.8. Contribution to policy making aligning with the SELINA Framework...................180
10. Recommendations 183
10.1. Ecosystem extent accounts 184
10.2. Forest Condition Accounts...186
10.3. Carbon accounts 186
10.4. Recommendation on the use of new remote sensing technology...................187
10.5. Recommendation Regarding the use of ETA typology 188
10.6. Recommendations from Peloponnese test site (UPATRAS) 188
10.7. Recommendation from São Miguel test site (FGF) 189
11. Acknowledgements... 191
12. References 192
13. Annex.. 201
13.1. Peloponnese Crosswalk 202
13.1.1. Peloponnese CLMS to ETA crosswalk 202
13.1.1.1. CLMS Crosswalk 204
13.1.1.2. CLMS Typologies 234
13.1.1.3. European Ecosystem Typology 244
13.1.2. Peloponnese National to ETA crosswalk...251

13.1.2.1. National Crosswalk....251
13.2. São Miguel Crosswalk....261
13.2.1. São Miguel CLMS to ETA Crosswalk....261
13.2.1.1. CLMS Crosswalk....262
13.2.1.2. CLMS Typologies....276
13.2.1.3. European Ecosystem Typology....280
13.2.2. São Miguel National to ETA Crosswalk....287
13.2.2.1. Land Use to ETA....287
13.2.2.2. Land Use Crosswalk....288
13.2.2.3. Land Use Translation....289
13.2.2.4. European Ecosystem Typology....290
13.2.3. São Miguel Forest Inventory (2024) to ETA....308
13.2.3.1. Forest Inventory to ETA Crosswalk....309
13.2.3.2. Forest Inventory Translation....309
13.2.3.3. Forest Inventory ID Crosswalked....309
13.3. Feature list for habitat mapping....310
13.4. EUNIS habitat classes used in this report....316
13.5. Overview of training data for habitat mapping....320
13.5.1. Overview of training data for habitat mapping of Peloponnese TS....320
13.5.2. Overview of training data for habitat mapping of São Miguel TS....322
13.6. Raw Forest Condition variable accounts for São Miguel 2017-2023....324

Tables

Table 1: Key characteristics of the Greek Test Site in Peloponnese. 31

Table 2: Key characteristics of São Miguel Island Test Site, Azores. 34

Table 3: Type of bias created by data availability disparity, affecting mapping and modelling processes...35

Table 4: National datasets used for mapping Peloponnese ecosystems. 36

Table 5: National datasets used for mapping São Miguel ecosystems. 37

Table 6: CLMS data used for ecosystem extent mapping. 39

Table 7. Overview of provided input data for training point selection. 39

Table 8: Example of the crosswalk between the Forest Inventory 2024 towards the European Ecosystem Typology (table should be read from top to bottom, column-wise). Each column corresponds to elements for defining forest structure and composition in FI 2024. 47

Table 9: Crosswalk between São Miguel Land Use 2018 and the European Ecosystem Typology classes (including the most detailed typology level where the match could be established). 48

Table 10: Land Use 2018 data, crosswalked to the European Ecosystem Typology and refined with CLMS data used for more detailed mapping. 50

Table 11: Examples of a 2-step crosswalk from national N2K via EUNIS to the European Ecosystem Typology. 53

Table 12: Example of the crosswalk hierarchy between CLMS data and the ETA. 53

Table 13: Classes crosswalked to ETA that were disaggregated to a lower level using CLMS data. 55

Table 14: Examples of disaggregating national data to a lower European Ecosystem Typology using CLMS data. 55

Table 15: EU extent mapping crosswalk Vegetation Centric approach – level 1. 60

Table 16: EU extent mapping crosswalk Vegetation Centric approach – Level 2. 61

Table 17: EU extent mapping crosswalk Vegetation Centric approach – level 3 Forest and Woodland and Coastal ecosystems. 62

Table 18: Specified desired sample size for each EUNIS level 1 habitat class, derived by an area-wise proportional analysis how this class occupies the MAES LIFE-IP map in Peloponnese. 65
Table 19: Translation of EUNIS habitats at level 1 to classes in CLC+ Backbone map. 66
Table 20: Model accuracies of winning Catboost models generated to map EUNIS habitat level 2 classes within overarching EUNIS habitat level 1 class 70
Table 21: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 1 classes. The list displays the features from top to bottom in order of importance with highest importance at the top. 70
Table 22: Model accuracies of winning Catboost models generated to map EUNIS habitat level 3 classes within overarching EUNIS habitat level 2 class 72
Table 23: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 2 classes. The list displays the features from top to bottom in order of importance with highest importance at the top. 73
Table 24: Area distribution [ha] of EUNIS level 3 habitat classes mapped within Peloponnese. 75
Table 25: Area distribution of EUNIS level 1 habitat classes mapped in the CLMS Coastal Zone map. Fraction per class was calculated by dividing the area per class by the total area on the island. The desired sample size was generated by multiplying the fraction by 50.000 (i.e., the desired total number of training points). 79
Table 26: Model accuracies of winning Catboost models generated to map EUNIS habitat level 2 classes within overarching EUNIS habitat level 1 class. 85
Table 27: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 1 classes. The list displays the features from top to bottom in order of importance with highest importance at the top. 86
Table 28: Model accuracies of winning Catboost models generated to map EUNIS habitat L3 classes within overarching EUNIS habitat L2 class. 88
Table 29: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 2 classes. Features ordered from top to bottom in order of importance. 88
Table 30: Relevant misclassifications identified in the habitat mapping of São Miguel. 93
Table 31: Overview of forest condition variables. 96

Table 32: Forest condition indicators ranked on 4 criteria (spatial resolution, temporal resolution, temporal frequency, and dataset quality). When data from more than one indicator is considered equal in one criterion, each indicator is attributed to the average of the positions they would represent (e.g. for spatial resolution NDWI, AGB and FCP in the table below). 102
Table 33: Number of pixels within the provided forest reference sites after the filtering process illustrated in Figure 43. 108
Table 34: Net primary productivity (NPP) condition indicator account per forest type for 2018-2023. 109
Table 35: Aboveground biomass (AGB) condition indicator account per forest type for 2018-2023. 110
Table 36: Normalized difference water index (NDWI) condition indicator account per forest type for 2018-2023. 110
Table 37: Forest connectivity (FC) condition indicator account per forest type for 2018-2023. 110
Table 38: Threatened Forest Bird Species Diversity (TFBSD) condition indicator account per forest type for 2018-2023. 111
Table 39: Soil organic carbon (SOC) condition indicator account per forest type for 2018-2023. 111
Table 40: Summary of products, remote sensing data, map resolutions and reference dates. Information applies for both study sites. 114
Table 41: Overview of datasets used to map the corresponding ecosystem type class. 120
Table 42: Calibrated Light Use Efficiency coefficient per ecosystem type class. 122
Table 43: Root-to-shoot ratio values used to derive aboveground carbon accumulation (ANPP) from net primary productivity (NPP) per ecosystem type class. 124
Table 44: Results of the ecosystem extent delineation at level 1 of the European Ecosystem Typology using three mapping approaches for the Peloponnese TS. 126
Table 45: Results of ecosystem extent delineation to all levels of the European Ecosystem Typology using three mapping approaches for the Peloponnese TS. 127
Table 46: Shows the results of the ecosystem extent delineation at level 1 of the European Ecosystem Typology using three mapping approaches for the São Miguel TS. 131

Table 47: Results of ecosystem extent delineation to all levels of the European Ecosystem Typology using three mapping approaches for the Sao Miguel TS. 132
Table 48: The Ecosystem Extent Account for Peloponnese for the year 2020. 135
Table 49: The Ecosystem Extent Account for São Miguel for the year 2020. 138
Table 50: Forest condition account (mean forest condition indices per class) for forest types in Peloponnese for 2018-2022..140
Table 51: Forest condition account (mean forest condition indices per class) for forest types (according to mapped EUNIS habitat level 2 classes on the habitat map) in Sao Miguel for period 2018-2023. T1: Broadleaved deciduous forest, T2: Broadleaved evergreen forest, T3: Coniferous Forest & S4: Temperate shrub heathland. 144
Table 52: Overview of the mean total aboveground and belowground carbon sequestration (ANPP and BNPP respectively) for 2022 and 2023 summed, per ecosystem type class in São Miguel and Peloponnese. 164
Table 53: Comparison between the national-centric and vegetation-centric approaches for ecosystem extent accounts. 184
Table 54: Comparison between annual carbon stock maps and annual carbon sequestration (uptake maps). 186

Figures

Figure 1: General chart of SELINA WP5. Task 5.2 applies technologies and methodologies to create innovative mapping products that can support the SEEA EA.....................................29
Figure 2: Key expected outputs proposed in the frame of Task 5.2......................................30
Figure 3: Orthophoto map of Peloponnese peninsula (EPSG: 3035). 32
Figure 4: Orthophoto map of São Miguel Island (EPSG: 3035)...34
Figure 5: Land Use 2018 map of São Miguel (DRAAC 2018)...37
Figure 6: Forest Inventory 2024 draft (partial) data used to map São Miguel. 38
Figure 7: The processing workflow for developing wall-to-wall ecosystem extent maps for São Miguel to the European Ecosystem Typology...46
Figure 8: The processing workflow for developing wall-to-wall ecosystem extent maps for Peloponnese to the European Ecosystem Typology. 52
Figure 9: Overview workflow to generate ecosystem extent maps by the vegetation centric approach. 56
Figure 10: Illustrative overview of the different modules in VITO's habitat mapping process: 1) Feature Extraction Module, 2) Feature Selection and Training Module, 3) Inference Module and 4) Post-processing Module...57
Figure 11: Overview workflow of the transformation to the EU Extent Typology.................59
Figure 12: MAES LIFE-IP map classified to EUNIS level 1 habitat types (Kokkoris et al. 2020). 65
Figure 13: Normalized confusion matrix of winning Catboost model to map habitats at EUNIS level 1, for Peloponnese. The model accuracy contains 83.27%..68
Figure 14: Predicted EUNIS level 1 habitat map of Peloponnese. See Annex 9.4 to explain the habitat codes in the legend. 69
Figure 15: Area distribution [ha] of EUNIS level 1 habitat classes mapped within Peloponnese. See Annex 9.4 to explain the habitat codes in the legend. 69
Figure 16: Normalized confusion matrix for mapping EUNIS level 2 habitat classes within EUNIS level 1 habitat T. The model accuracy contains 90,60%...70
Figure 17: Predicted EUNIS level 2 habitat map of Peloponnese. See Annex 13.4 for the legends of the habitat codes. 71

Figure 18: Area distribution [ha] of EUNIS level 2 habitat classes mapped within Peloponnese. 72
Figure 19: Normalized confusion matrix for mapping EUNIS level habitat classes within EUNIS level 2 habitat N1. The model accuracy contains 97.63% 72
Figure 20: Predicted EUNIS level 3 habitat map of Peloponnese. See Annex 13.4 for the legends of the habitat codes. 74
Figure 21: Zoom-in on area in coastal areas of Peloponnese to check visually the correspondence of the predicted EUNIS level 3 habitat map with a Google Satellite image. See Figure 20 to link the color in the habitat map to a EUNIS habitat class. 74
Figure 22: Photointerpretation validation areas in Peloponnese (red rectangles). 77
Figure 23: The most extensive area of the priority habitat type “Species-rich Nardus grasslands, on silicious substrates in mountain areas (and submountain areas in Continental Europe)”, at the northern part of the TS (Mt. Panachaiko). This area has been included in the “Grassland (non-classified)” class (see green polygons in the ecosystem type map on the right). 78
Figure 24: Inland marshes and wetland (purple), sparsely vegetated ecosystems (pink), beaches and sandy dunes (cyan) and coastal dunes (yellow) in the southernmost part of Peloponnese, well captured and mapped by the EO model (see map on the right). 78
Figure 25: Ecosystems in the National Park of Chelmos-Vouraikos (inside the river gorge). Sclerophyllous vegetation, riparian forests and coniferous forests are well captured, however rocky cliffs are underestimated. 78
Figure 26: Normalized confusion matrix of winning Catboost model to map habitats at EUNIS L1 for São Miguel. The model accuracy contains 87.01%. 84
Figure 27: Predicted EUNIS level 1 habitat map of São Miguel. See Annex 13.4 for the legends of the habitat codes. 84
Figure 28: Area distribution [ha] of EUNIS level 1 habitat classes mapped within São Miguel. See Annex 9.4 to explain the habitat codes in the legend. 85
Figure 29: Normalized confusion matrix for mapping EUNIS level 2 habitat classes within EUNIS level 1 habitat N. The model accuracy contains 83.93%. 85
Figure 30: Predicted EUNIS level 2 habitat map of São Miguel. See Annex 9.4 to explain the habitat codes in the legend. 87

Figure 31: Area distribution [ha] of the natural vegetation EUNIS L2 habitat classes in São Miguel. 87
Figure 32: Normalized confusion matrix for mapping EUNIS level habitat classes within EUNIS level 2 habitat T2. The model accuracy contains 97.63%. 88
Figure 33: Predicted EUNIS level 3 habitat map of São Miguel. See Annex 9.4 to explain the habitat codes in the legend. 89
Figure 34: Zoom-in on area in East of São Miguel to visually check the correspondence of the predicted EUNIS level 3 habitat map with Google Aerial imagery. See Figure 33 to link the colours in the habitat map to a EUNIS habitat class. 90
Figure 35: Area distribution [ha] of the natural vegetation EUNIS L3 habitat classes in São Miguel. 90
Figure 36: Normalized confusion matrix for [left] - Original classification of external validation points vs. the predicted classification in the habitat map at level 1. Class U was not mapped on the modelled habitat map. The map accuracy based on these validation points is 78,44%. [Right] Original classification of external validation points vs. the classification on the Coastal Zone map of CLMS (2018) at level 1. Class Q is not mapped on the CLMS Coastal Zone map. The map accuracy is 66,67%. 91
Figure 37: Normalized confusion matrix for the original classification of external validation points vs. the predicted classification in the habitat map at level 2. 92
Figure 38: Normalized confusion matrix for [left] - original classification of external validation points for Broadleaved forest (BF, i.e., T1 and T2 merged), Coniferous forest (CF) and S4 points vs. the predicted classification of the forest classes in the level 2 habitat map, and [right] original classification of external validation points for Broadleaved forest (BF, i.e., T1 and T2 merged), CF and S4 points (map accuracy considering only BF and CF: 76,58%, map accuracy including all classes: 70,58%) vs. the classification on the CLMS FTY layer (2018) (map accuracy: 69,48%). 93
Figure 39: SEEA EA Ecosystem Condition Typology per group (abiotic, biotic and landscape level) and per associated classes explained (Vallecillo Rodriguez et al. 2022). 96
Figure 40: Visualization of extraction of 98th (i.e., Upper Reference) and 2nd percentile (i.e., Lower Reference) per condition variable for rescaling variables to indicators (Bruelheide et al. 2024). 101

Figure 41: European biogeographical regions classified by the European Environmental Agency (EEA 2024). 103
Figure 42: Examples of output ARIES for PEOPLE-EA Explorer - Net Primary Productivity variable condition account [above] and Net Primary Productivity Indicator account [below]. 104
Figure 43: Example of output ARIES for PEOPLE-EA Explorer, showing the spatial raster map of forest condition index for 2018 [above] and the tabular forest condition account for 2018-2022 per forest type [below]. 105
Figure 44: Locations of the forest reference sites per forest type and heathland in São Miguel. 107
Figure 45: Distributions of tree cover density (TCD) within the provided forest reference sites per forest types: T1 (Broadleaved deciduous forest) [top-left], T2 (Broadleaved evergreen forest) [top-right], T3 (Coniferous Forest) [bottom-left] and S4 (Temperate shrub heathland) [bottom-right]. 107
Figure 46: Illustration of the filtering process: filtering out areas from the provided forest reference sites where tree cover density is lower than 50% for T1, T2 and T3, and where tree cover density is lower than 20% for S4. 108
Figure 47: Illustration of all accounting areas per forest type: all area on São Miguel for which a forest condition index will be generated, comprising the areas used as forest reference sites and all remaining forest. 109
Figure 48: SarVision methodology for the unsupervised carbon stock and emission mapping. 114
Figure 49: Biomass stratification system proposed by the High Carbon Stock (HCS) classification system. (https://highcarbonstock.org/what-is-the-high-carbon-stock-approach/) 115
Figure 50: Flowchart of the forest structure mapping methodology. 116
Figure 51: Main steps involved in the processing of Sentinel-1 radar images during the creation of thematical deforestation and forest degradation data using the SarSentry system. 117
Figure 52: Steps for biomass mapping...118

Figure 53: Ecosystem types for which the carbon sequestration between end of 2021 and end of 2023 are to be generated in São Miguel [above] and Peloponnese [below] (sources in Table 41). 120
Figure 54: Annual Total GPP for 2022 and 2023 in São Miguel for classes: wetland, broadleaf forest, coniferous forest, low woody and evergreen forest. 122
Figure 55: Annual Total GPP for 2022 and 2023 in Peloponnese for classes: wetland, broadleaf forest, coniferous forest, low woody and evergreen forest. 123
Figure 56: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using National data only (EPSG:3035). 128
Figure 57: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using National data enhanced with CLMS data (EPSG:3035). 129
Figure 58: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using CLMS data only (EPSG:3035). 130
Figure 59: The ETA ecosystem extent map for São Miguel TS using National data only (EPSG:3035). 133
Figure 60: The ETA ecosystem extent map for São Miguel TS enhanced with CLMS data (EPSG:3035). 133
Figure 61: The ETA ecosystem extent map for São Miguel TS using CLMS data only (EPSG:3035). 134
Figure 62: Ecosystem extent map 2020 for Peloponnese using the Sentinel / Vegetation centric approach (EPSG:3035). 135
Figure 63: The European Ecosystem Typology ecosystem extent map for São Miguel TS using Sentinel data / vegetation centric approach (EPSG:3035). 137
Figure 64: The European Ecosystem Typology ecosystem extent Quality Flag for São Miguel TS using Sentinel data / vegetation centric approach (EPSG:3035). 138
Figure 65: Evolution of mean forest condition index through time (2018-2022) per forest type in Peloponnese for Transitional woodland shrub, broadleaf forest, coniferous forest and mixed forest. 141
Figure 66: Spatial map of the forest condition index in 2022 for Peloponnese. 141
Figure 67: Net change map of forest condition index between 2018 and 2022 for Peloponnese. The redder the pixels are colored, the more a decline in forest condition index

was observed. The greener the pixels are colored, the more an increase in forest condition. 142
Figure 68: Change maps of forest condition index in pairs of consecutive years between 2018 and 2022 for Peloponnese. 143
Figure 69: Evolution of mean forest condition index through time (2018-2023) per forest type in São Miguel: T1 Broadleaved deciduous forest, T2 Broadleaved evergreen forest, T3 Coniferous Forest and S4 Temperate shrub heathland. 144
Figure 70: Net change map of forest condition index between 2018 and 2023 for Sao Miguel. The redder the pixels are colored, the more a decline in forest condition index was observed. The greener the pixels are colored, the more an increase in forest condition index was observed. 145
Figure 71: Change maps of forest condition index in consecutive years between 2018 and 2023 for São Miguel. 146
Figure 72: Forest structural maps of São Miguel for 2017, 2021 and 2023, the whole island (on the left), a detail over the east volcanic area where the Peatland bogs were mapped (on the right). 147
Figure 73: Confusion matrix presenting the validation results for the FSM 2023 in São Miguel. Map presents the location of randomly selected points used for map validation using high resolution Planet data as a reference. Overall accuracy was estimated in 87.5% with a calculated kappa coefficient of 0.8. 148
Figure 74: Forest Structural Maps of Peloponnesus 2021-2023. Structural differences can be seen specially the transformation of forest (green) into shrubland classes (orange). This forest loss and forest degradation were mainly attributed to fires, the whole area (at the top) and detail over an area of change between the two years (at the bottom). 150
Figure 75: Confusion matrix presenting the validation results for the FSM 2023 in Peloponnese. Map presents the location of randomly selected points used for map validation using high resolution Planet data as a reference. Overall accuracy was estimated in 82% with a kappa coefficient of 0.8. 151
Figure 76: SarSentry results for São Miguel. Observation periods correspond to 2017-2021 and 2017-2023. SarSentry thematic products show the yearly forest change detections (deforestation, degradation and regrowth); the whole island (at the top) and detail over the east volcanic area where changes in the forest cover were detected (at the bottom). 153

Figure 77: SarSentry thematic results for Peloponnesus for the 2017-2021 and the 2017-2023 periods; SarSentry thematic products show the yearly forest change detections (deforestation, degradation and regrowth); the whole peninsula (at the top) and detail over the central area where changes in the forest cover were detected (at the bottom). 154
Figure 78: Location of GEDI footprints acquire over the area of São Miguel during years 2020, 2021 and 2022. In total, 57200 points with canopy height data (h100) were used for the SarCarbon analysis, covering most of the Island. Histograms showing the distribution of canopy height data for each of the forest structural classes of the structural maps are shown. 155
Figure 79: Final biomass maps for São Miguel. Data per pixel is expressed in tons /ha. The legend is aggregating the biomass values of a float map into classes. Values range from 5 ton/ha to a maximum of 270 ton/ha; the whole island (at the top) and a detail from the mountainous eastern sections are shown (at the bottom). 156
Figure 80: Location of GEDI footprints acquired over Peloponnese in 2019, 2020 and 2021. In total, 908.000 points with canopy height data (h100) were used for the SarCarbon analysis, covering most of the peninsula. Histograms showing the distribution of canopy height data for each of the forest structural classes of the structural maps are shown. 157
Figure 81: Final biomass maps estimated for Peloponnese. Data per pixel is expressed in tons/ha. The legend is aggregating the biomass values of a float map into ranges. Values range from 5 ton/ha to 270 ton/ha; the whole area (at the top) and detail of areas were the biomass changes between the years(at the bottom). 158
Figure 82: Final carbon maps for São Miguel. Data per pixel is expressed in tons C/ha. The legend is aggregating the carbons values of a float map into classes. Values range from 5 ton C/ha to a maximum of 100 ton C/ha; the whole island (at the top) and detail over a mountainous region of the island (at the bottom). 159
Figure 83: Final carbon maps for Peloponnese. Data per pixel is expressed in tons C/ha. The legend is aggregating the carbons values of a float map into classes. Values range from 5 ton C/ha to 100 ton C/ha; the whole area (at the top) and details over the central region where changes were observed (at the bottom). 160
Figure 84: Carbon difference (flux) map for São Miguel expressing data in tons of carbon per hectare (ton C/ha). The legend aggregates the carbon values from a floating-point raster map into distinct carbon classes; the whole island (at the top) and detail at the east end

showing red areas where carbon emissions are due to deforestation (at the bottom). This calculation represents the change detected in two years. 161
Figure 85: Carbon difference (flux) map for Peloponesse expressing data in tons of carbon per hectare (ton C/ha). The legend aggregates the carbon values from a floating-point raster map into distinct carbon classes; the whole peninsula (at the top) and detail at the central region showing red areas where carbon emissions are due to forest change (at the bottom). This calculation represents the change detected in two years. 162
Figure 86: Total aboveground and belowground carbon sequestration for São Miguel by summing the aboveground carbon accumulation (ANPP) in 2022-2023 and belowground carbon accumulation (BNPP) in 2022-2023, respectively. 163
Figure 87: Total aboveground and belowground carbon sequestration for Peloponnese by summing the aboveground carbon accumulation (ANPP) in 2022-2023 and belowground carbon accumulation (BNPP) in 2022-2023, respectively. 164
Figure 88: Side comparison of the forest structure map [left] and the EUNIS habitat map [right] in the same mountainous and forested area of São Miguel. 167
Figure 89: Evolution of the mean Net Primary Productivity (NPP) condition indicator plotted with the evolution of the mean forest condition index through time per forest type in Sáo Miguel for T1 Broadleaved deciduous forest, T2 Broadleaved evergreen forest, T3 Coniferous forest and S4 Temperate shrub heathland. 168
Figure 90: Average accumulated precipitation in summer months from 2016 to 2024 in São Miguel Island. Data gathered from the Hydrometeorological Network of the Azores (https://redehidro.ambiente.azores.gov.pt/). 169
Figure 91: Average monthly temperatures in summer months from 2016 to 2024 in São Miguel Island. Data gathered from the Hydrometeorological Network of the Azores (https://redehidro.ambiente.azores.gov.pt/). 170
Figure 92: [Left] Forest condition index change map for 2020-2021 (bottom left) overlaid on orthographic relief image (ESRI, bottom top)) for São Miguel. [Right] Zoom-in on area in the East of São Miguel that demonstrates that most of the decline in forest condition index from 2020 to 2021 happened on higher elevations, following the mountain ridges. 170
Figure 93: Carbon sequestration difference map between VITO and SarVision approaches. 173

Figure 94: Zoom-in to an area in São Miguel where the difference map indicated high divergence for carbon sequestration between SarVision and VITO approaches. 174
Figure 95: Example of divergence in carbon sequestration results between SarVision and VITO. [Top] VITO ANPP results show aboveground increment in 2022 and 2023. [Middle] SarVision carbon stock maps showing an abrupt stock decrease. [Bottom] Forest degradation maps attribute this abrupt decline in carbon stock is due to (old) deforestations events. 175
Figure 96: Details of the series of maps using remote sensing derive products created for São Miguel [above] and for Peloponnesus [below]., necessary for the carbon stock and carbon flux accounting for a certain area. 177

1. Preface

The importance of biodiversity, natural capital and healthy ecosystems and the services they supply has increasingly been acknowledged in diverse policy initiatives (e.g., the EU nature restoration and amending Regulation from 2024, EU Biodiversity Strategies 2020 and 2030, Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), UN’s Natural Capital and Ecosystem Services Accounting (SEEA EA), Intergovernmental Panel on Climate Change (IPCC) and Convention on Biological Diversity (CBD)).

The EU Horizon Research and Innovation Action “Science for Evidence-based and sustainabLe decIsions about NAtural capital” (SELINA) aims to provide robust information and guidance that can be harnessed by different stakeholder groups to support transformative change in the EU, to halt biodiversity decline, to support ecosystem restoration and to secure the sustainable supply and use of essential Ecosystem Services (ES) in the EU by 2030.

SELINA builds upon the Mapping and Assessment of Ecosystems and their Services (MAES) initiative that has provided the conceptual, methodological, data and knowledge base for comprehensive assessments on different spatial scales, including the EU-wide assessment (Maes et al. 2020) and assessments in EU member states. Knowledge and data for different ecosystem types are increasingly available.

Within these overall project aims, Work Package (WP) 5 “Ecosystem Accounting” integrates insights from ecosystem conditions and ecosystem services into the UN System of Environmental Economic Accounting - Ecosystem Accounting (SEEA EA) framework. It addresses key challenges such as incorporating negative externalities, improving the accounts’ spatial and temporal resolution, and exploring how different valuation methods influence ES and asset values.

Task 5.2 focused on enhancing the spatial and temporal resolution of ecosystem accounts through the integration of satellite data, remote sensing techniques, and next generation modelling methodologies. Significant data required for implementing the SEEA EA framework can be efficiently obtained from these three sources in a cost-effective and timely manner. Outcomes from Deliverable D5.2 “Enhancing the spatial and temporal resolution of ecosystem accounts using satellite data” include state-of-the-art products designed to support ecosystem extent, condition, and carbon accounting. To demonstrate the applicability and effectiveness of these products, initially, three study sites were selected: the island of La Réunion department of France, the island of São Miguel in the Azores, and the Peloponnese Peninsula in Greece. However, the partners from the University of La Réunion expressed their preference to contribute to other tasks, as they already possess remote sensing data for their area. Therefore, only two Test Sites (TS) were ultimately selected:

- The island of São Miguel in the Azores, Portugal.
- The Peloponnese region in Greece.

2. Executive Summary

Context and Objectives

The increasing global emphasis on sustainable development and climate action, led by frameworks such as the United Nations Framework Convention on Climate Change (UNFCCC) and the System of Environmental-Economic Accounting - Ecosystem Accounting (SEEA EA), calls for reliable and harmonized data on ecosystems. Ecosystem accounting, particularly under the SEEA EA framework, supports this effort by tracking the extent, condition, and services of ecosystems in a structured and policy-relevant manner.

The SELINA (Science for Evidence-based and sustainabLe decIsions about NAtural capital) project aims to generate robust, policy-relevant data and tools that empower decision-makers to tackle biodiversity loss, promote ecosystem restoration, and ensure the long-term supply and sustainable use of vital Ecosystem Services (ES) across the EU by 2030. Building on the achievements of the Mapping and Assessment of Ecosystems and their Services (MAES) initiative, SELINA extends the conceptual, methodological, and data infrastructure required for ecosystem assessments at multiple spatial levels—ranging from EU-wide overviews to national, regional and local-scale evaluations.

Within this broader framework, Task 5.2 focuses on enhancing the spatial and temporal resolution of ecosystem accounts by integrating available European maps and classification systems aligned to the European ecosystem typology (ETA), cutting-edge Remote Sensing (RS) techniques, and advanced spatial modelling approaches. These innovations support more dynamic and responsive ecosystem monitoring, aligning with the SEEA EA's goals of harmonized, scalable, and cost-effective ecosystem accounting.

This report outlines innovations and applications in high-resolution ecosystem accounting using RS and spatial modelling in 2 SELINA test sites: Peloponnesus (Greece) and São Miguel (Azores, Portugal).

Methodology Overview

Under SELINA Task T5.2, ecosystem accounting was developed for three key components: extent, condition, and carbon fluxes/stocks. The approaches integrated national datasets, Copernicus Land Monitoring Service (CLMS) data, EUNIS habitat classification and RS products (e.g., Sentinel 1-2, GEDI LiDAR) to generate wall-to-wall maps aligning with the ETA.

1. Ecosystem Extent Mapping employed two approaches:
    a. National-Centric: Uses national data as a baseline, enhanced with CLMS data to fill gaps and standardize outputs.
    b. Vegetation-Centric: Uses Artificial Intelligence (AI) to classify EUNIS habitats based on RS-derived features, supporting consistent, scalable mapping. These were trained and tested on field-validated data to produce wall-to-wall 10 m

resolution habitat maps, after which a crosswalk to EU typology created ecosystem extent maps.

2. Ecosystem Condition Accounts were developed using the PEOPLE-EA forest condition index (FCI), combining indicators such as soil organic carbon, tree cover density, biodiversity, and productivity. Indicators were rescaled and weighted to produce a composite index representing forest health relative to an optimal reference state. Data harmonization was a key aspect, including the use of legacy datasets like LUCAS, SoilGrids, and MODIS to maintain consistency across spatial and temporal scales.
3. Carbon Accounting included:
    a. *Remote sensing-based method*: Uses structural vegetation stratification, GEDI canopy height data, and RS imagery to map Above-Ground Biomass (AGB) and infer carbon flux. The SarSentry algorithm was used to detect deforestation and degradation using Sentinel-1 radar time series.
    b. *GPP-based Method*: Estimates Gross Primary Productivity using the Light Use Efficiency (LUE) model, with Sentinel-2 and meteorological data. GPP was converted to Net Primary Productivity (NPP), and further to Aboveground NPP (ANPP) using biome-specific root-to-shoot ratios derived from IPCC guidelines.

Results and Analysis

- Peloponnesus: Due to limited national data, extensive use of CLMS was required. The National & CLMS approach improved mapping resolution and coverage, enabling 85% crosswalk to the ETA. Forest condition was generally stable, with indices around 0.55–0.60 and minor declines attributed to known data artefacts (sensor transitions). Carbon flux assessments revealed emissions primarily linked to fire and degradation. Deforestation and degradation detection using SarSentry confirmed these trends, with most change hotspots located near forest edges and fire-prone zones.
- São Miguel: Benefited from more comprehensive national datasets, allowing accurate ecosystem extent and forest condition mapping. Forest condition index varied across forest types, with natural vegetation generally in better condition than managed or exotic plantations. Some degradation was identified in the central axis of the island, coinciding with abiotic disturbances such as landslides. Carbon maps, validated via canopy height estimates and IPCC data, showed biomass ranges up to 270 tons/ha. Wetlands and highland bogs were successfully mapped as distinct carbon sinks, thanks to high-resolution RS products and detailed land cover classifications.

Innovations and Tools

- Development of automated RS-based workflows for habitat and Ecosystem Types classification, integrating over 150 predictors and machine learning models (e.g., CatBoost).
- Integration of AI (CatBoost models, CNNs) to classify habitats and ecosystem types at 10 m resolution.

- Deployment of the ARIES for PEOPLE-EA tool for FCI computation, enabling year-by-year condition mapping across biogeographical zones.
- Use of SarSentry algorithm for detecting forest change dynamics via radar time series, supporting temporal change tracking in dense vegetation areas.
- Application of allometric models to convert canopy height into biomass and carbon estimates, improving accuracy in stock and flux assessments.
- Validation accuracies of 82–87.5% were achieved in classification outputs.

Limitations include incomplete national data, challenges in validating RS-based forest degradation, and generalizability constraints in remote islands like São Miguel. Where field data was lacking, IPCC biomass guidelines were applied. Validation was conducted using confusion matrices from high-resolution satellite imagery, with overall kappa coefficients above 0.8 for both sites. Errors were most frequent in shrubland classes, whereas forest classifications showed strong consistency.

Policy Implications and Recommendations

- The demonstrated methods are scalable and cost-effective, supporting e.g. the EU Biodiversity Strategies, Green Deal goals, and national SEEA EA reporting.
- Continued investment in remote sensing infrastructure and data harmonization is crucial, especially for so far underrepresented regions.
- Adoption of AI-driven tools will enhance the temporal and spatial resolution of ecosystem monitoring, facilitating annual reporting.
- Strengthening linkages with local expertise is essential to validate and contextualize automated outputs.
- Develop training and knowledge-sharing platforms for ecosystem accounting teams at national and regional levels.
- Encourage inclusion of ecosystem degradation indicators into economic and land-use planning systems.

Conclusion

This report illustrates a viable path toward operational, harmonized ecosystem accounting across Europe. By combining national and EU datasets with cutting-edge RS and modelling techniques, it is possible to develop high-resolution, policy-relevant ecosystem accounts that inform sustainable environmental and economic decision-making. These methods offer a foundation for expanding ecosystem monitoring at different scales and temporal resolutions and integrating natural capital into planning and climate policy frameworks.

Future directions include refining AI models, expanding reference site databases, integrating local citizen science for validation, and scaling methodologies to continental or global levels under initiatives like the ESA World Ecosystem Extent Dynamics project. With these tools, ecosystem accounting can evolve into a central pillar of environmental management.

3. List of abbreviations

AGB	Above Ground Biomass
BGB	Below Ground Biomass
CBD	Convention on Biological Diversity
CLMS	Copernicus Land Monitoring Service
DMP	Dry Matter Productivity
DOM	Dead Organic Matter
DP	Demonstration Project (of SELINA)
EEA	European Environmental Agency
EO	Earth Observation
ES	Ecosystem Services
ESA	European Space Agency
ETA	European Ecosystem Typology for Accounting
EU	European Union
EUNIS	European nature information system
GDMP	Gross Dry Matter Productivity
IPBES	Intergovernmental Platform on Biodiversity and Ecosystem Services
IPCC	Intergovernmental Panel on Climate Change
JAXA	Japan Aerospace Exploration Agency
LULUCF	Land Use, Land-Use Change and Forestry
MAES	Mapping and assessment of Ecosystem and their Services
NASA	National Aeronautics and Space Administration
SEEA EA	System of Environmental Economic Accounting - Ecosystem Accounting
SOM	Soil organic matter
TS	Test site (of SELINA)
UNFCCC	United Nations Framework Convention on Climate Change
DP	Demonstration projects
CoG	Compendium of Guidance

4. Introduction

4.1. Ecosystem Accounting

The United Nations Framework Convention on Climate Change (UNFCCC) provides a global platform to combat climate change by stabilizing Green-House Gas (GHG) emissions. It emphasizes the vital role ecosystems — such as forests, wetlands, and peatlands — play as carbon sinks that absorb and store atmospheric CO_2.

To effectively integrate ecosystems into climate and environmental policy, the System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA) offers a standardized framework to measure and monitor ecosystem extent, condition, services, and monetary values. These accounts support evidence-based policymaking and help align economic planning with environmental sustainability. The key account types include:

- Ecosystem Extent Accounts: Map and track the size and spatial distribution of ecosystems over time, helping detect land cover change and habitat loss.
- Ecosystem Condition Accounts: Assess ecosystem health through indicators like vegetation cover, soil quality, water availability, and biodiversity status.
- Ecosystem Services Accounts: Quantify the benefits ecosystems provide — ranging from provisioning (e.g., food, timber) to regulating (e.g., carbon storage, flood control) and cultural (e.g., recreation, heritage) services.
- Monetary Accounts: Estimate the economic value of ecosystem services and natural capital for inclusion in national accounts.
- Thematic account: Thematic accounts focus on a specific environmental or natural resource topic and systematically organize data to monitor changes, understand impacts, and support decision-making.

There is a growing synergy between ecosystem accounting and RS technologies, which offer repeatable, consistent, and scalable data critical for spatially explicit accounting. Over the last decades, agencies like ESA, NASA, and JAXA have launched numerous satellites equipped with optical, radar, and LiDAR sensors, enabling high-resolution monitoring of vegetation cover, land use, biomass, and ecosystem changes. In addition, Several European datasets and classification systems support the operationalization of SEEA EA. These products can be of different types according to the definition and the use.

Habitat Maps: Habitat is a terrestrial or aquatic area distinguished by geographic, abiotic, and biotic features, whether entirely natural or semi-natural. In the framework of Natura 2000, habitats are defined as natural or semi-natural environments that are important for the conservation of biodiversity, particularly for species and habitat types of European interest.

Available products include:

- Natura 2000: Protected habitat areas under EU regulation from European environmental agency (EEA).
- EUNIS Habitat Classification: A detailed and harmonized habitat typology developed by the EEA, available across Europe and linking to systems like MAES, CORINE, and IUCN.

Ecosystem Maps: Dynamic complex of plant, animal, and micro-organism communities and their non-living environment interacting as a functional unit. This definition follows the one provided by the Convention on Biological Diversity (CBD) and is used throughout the EU Biodiversity Strategy and the MAES initiative. Available products like:

- European Ecosystem Maps (MAES) EU wide with basic ecosystem types by EEA.

Land Use/Land Cover (LULC) maps: Geospatial representations that classify and display the physical and functional characteristics of the Earth's surface. They combine two related but distinct concepts: Land Cover refers to the physical material on the surface of the Earth, such as forests, grasslands, water bodies, croplands, or built-up areas and Land Use that refers to the human purpose or function associated with a particular area, such as agriculture, residential development, industrial activity, or recreation. Available products are:

- Corine Land Cover: (Coordination of Information on the Environment – Europe) High resolution land use land cover (LULC) data from Copernicus EEA source.
- Copernicus High resolution layers: Detail LULC components EU wide from Copernicus source.
- Copernicus Land Monitoring Service (CLMS): Provides harmonized, regularly updated geospatial data for land cover, land use, and changes, widely used in ecosystem extent and condition analysis.

At the heart of European ecosystem accounting is the European Ecosystem Typology (ETA), developed by the EEA. ETA provides a hierarchical classification system — from broad ecosystem types (e.g., forest, grassland) at Level 1–2 to more detailed subtypes (e.g., broadleaved forest, managed grasslands) at Level 3+. It is compatible with land cover products (e.g., CORINE) and designed to align with remote sensing data for standardizing and comparing accounts across scales.

To use various datasets within the ETA framework, a crosswalking process is required. This involves aligning national or thematic classification systems with the ETA typology to ensure consistency and comparability of spatial ecosystem data. This process was explored in detail in Deliverable 3.1 of this project.

The application of RS and advanced spatial modelling enables the creation of accurate, policy-relevant ecosystem accounts:

- Extent Accounts: Use land cover/use data (e.g., CORINE, Sentinel, CLMS) to delineate and monitor ecosystem types.
- Condition Accounts: Derive abiotic, biotic and landscape indicators using vegetation indices (e.g., NDVI), canopy height (LiDAR), or soil/water parameters (radar/optical).

- Ecosystem Services Accounts: Combine RS data with climate, soil, and socioeconomic models to quantify services such as crop yields, flood mitigation, or carbon sequestration.

Integrating remote sensing innovations, and other available datasets, and modelling techniques improves the spatial and temporal accuracy of ecosystem accounts. These advances reduce costs, increase scalability, and enhance the relevance of SEEA EA in climate, biodiversity, and sustainable development policies. Ultimately, this integrated approach ensures a more comprehensive understanding of ecosystem dynamics, supports regular updates to national accounts, and strengthens the scientific and policy foundation for natural capital management and climate change response.

4.2. Aims & Objectives

Task 5.2 (Figure 1) aims to enhance the temporal and spatial resolution of ecosystem accounts using satellite data to improve their policy relevance and ensure their alignment with other reporting and monitoring systems, such as those under the UNFCCC. Additionally, this task seeks to provide recommendations for improving SEEA EA by integrating high-resolution, high-quality accounting data, thereby strengthening its applicability for environmental and economic decision-making.

Task 5.2 focuses on three key ecosystem accounting components: (1) ecosystem extent, (2) ecosystem condition, and (3) carbon stocks and fluxes, providing original maps and information based on advanced, state of the art spatial modelling and remote sensing techniques. Two SELINA Test Sites (TS) were selected for the application of these innovative methods: São Miguel Island, Azores, Portugal, and the peninsula of Peloponnesus, Greece.

The demonstration of state-of-the-art remote sensing–derived and advanced spatial modelling products to the SEEA EA scientific community supports the implementation of frameworks where data can be efficiently integrated in a cost-effective and timely manner.

The high costs associated with traditional land survey methods and field validation campaigns stand in contrast to the comparatively low cost and scalability of remote sensing–based data production. Continuous improvements in spatial and temporal resolution, RS and spatial modelling techniques will enable the development of long-term monitoring strategies for ecosystem extent, condition, and service accounts offering UpToDate information with improved temporal and spatial resolutions. Complementarity of available methods not only enhances data accessibility but also strengthens the scientific and policy relevance of ecosystem accounting by ensuring regular, consistent updates across spatial scales.

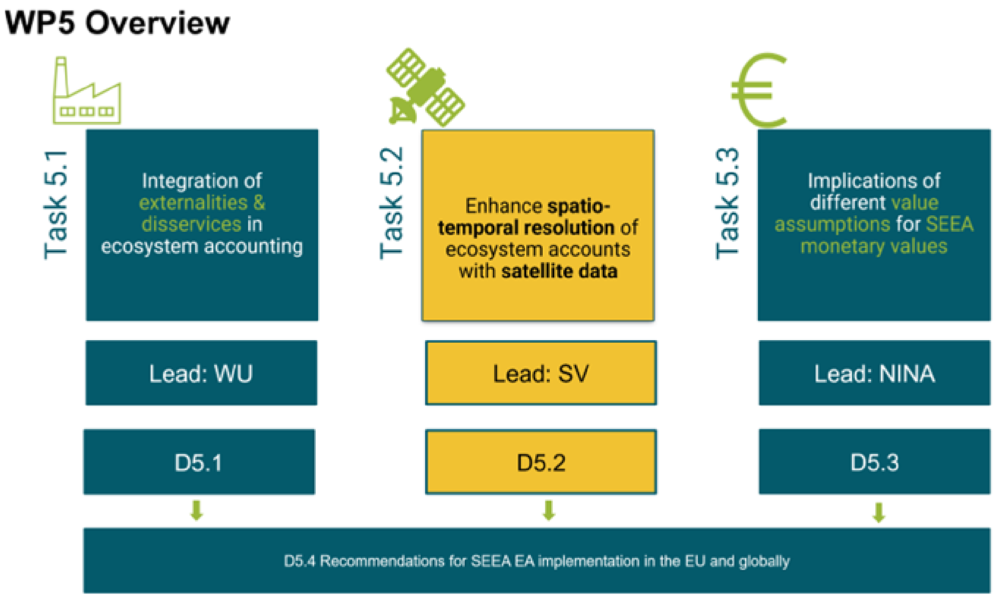


Figure 1: General chart of SELINA WP5. Task 5.2 applies technologies and methodologies to create innovative mapping products that can support the SEEA EA.

4.3. Expected outcomes

Developing high-resolution and high-accuracy spatial data serves multiple purposes. Firstly, it helps establish a baseline to assess the complexity and condition of ecosystems, providing a benchmark for future comparisons. Additionally, it enables the estimation of biophysical variables like carbon stocks and carbon fluxes, facilitating the creation of monitoring schemes and the development of policy scenarios for improved environmental decision-making. The main outputs delivered by Task 5.2. (Figure 2) are the following:

- Developing High-Resolution Ecosystem Type Maps – Utilize multi-source remote sensing (RS) data, including optical (Sentinel-2), SAR (Sentinel-1), multi/hyperspectral, and LiDAR (GEDI), to produce ecosystem extent maps.
- Mapping and assessing Forest and Habitat Conditions – Generate ecosystem condition and habitat maps to track changes and degradation over time.
- Analyzing Deforestation and Forest Degradation – Develop historical deforestation and degradation maps to understand land-use changes, assessing forest condition and their impact on carbon stokes and fluxes.
- Quantifying Biomass and Carbon Stocks – Create biomass maps (2021-2023) and carbon stock maps (2021, 2022, 2023) to assess carbon sequestration and emissions.Estimate Carbon Fluxes – Present methodologies for carbon flux estimations. Production of tables to track carbon gains and losses in the ecosystem.
- Linking Aboveground and Belowground Biomass (AGB-BGB) – Establish relationships between aboveground biomass (AGB) and belowground biomass (BGB) for improved carbon accounting models.

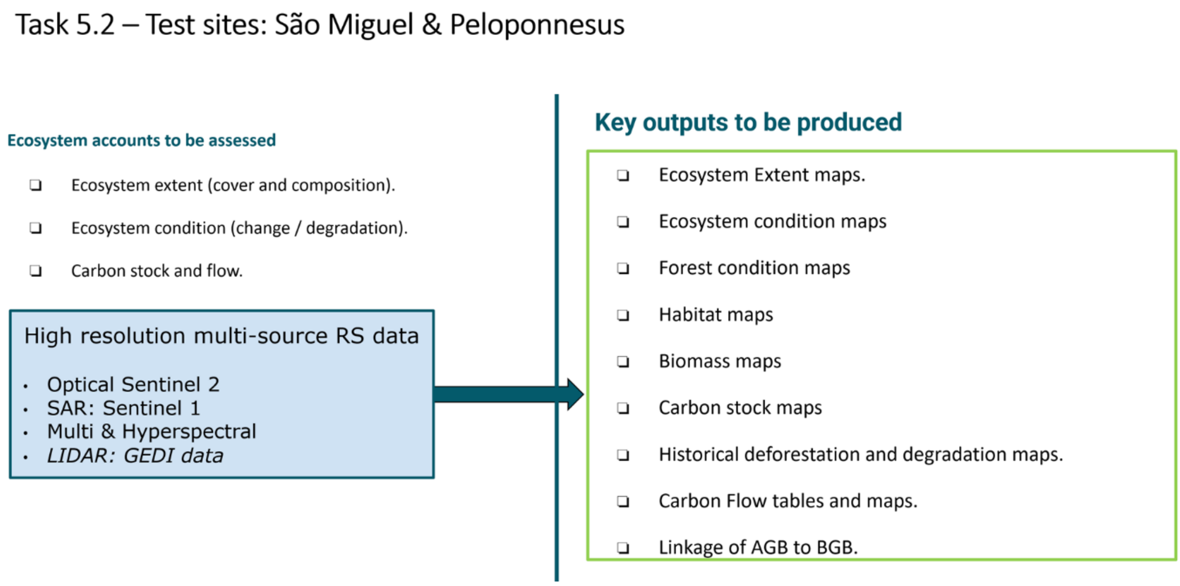


Figure 2: Key expected outputs proposed in the frame of Task 5.2.

5. Test sites

5.1. Peloponnese, Greece

The Peloponnese (Figure 3) is the largest peninsula of Greece (21,549 km²) and the southernmost part of the Greek mainland, with a population of around 1.2 million, including the third-largest Greek city, Patras. It is one of the 13 floristic regions of Greece, hosting more than 4,024 plant taxa, of which 522 are Greek endemics (approximately 32% of the total Greek endemics), and 579 have a restricted distribution. The region also includes extensive areas of the endemic Greek forest of *Abies cephalonica* (Greek fir), as well as the only occurrence in the EU of *Juniperus drupacea* forests, which are found exclusively in southeastern Peloponnese, restricted to Mount Parnon. This gives additional value to this priority forest habitat for its conservation and management, as well as to Greece's responsibility to protect its only occurrence in the EU.

It hosts a variety of ecosystems, ranging from coastal dunes and thermophilous Aleppo pine forests and phrygana to oro-Mediterranean ecosystems. Important and protected coastal, inland, and mountainous wetlands are also present, forming a complex landscape that has been under human modification since ancient times. Wilderness is mainly found in the Natura 2000 areas and National Parks of the region, including all wetland ecosystems and the highest mountains (most with peaks above 2,000 m).

Table 1: Key characteristics of the Greek Test Site in Peloponnese.

Key characteristics of Peloponnese Test Site
• High biodiversity value (e.g. hosts ca. 32% of Greek endemic plant taxa and unique habitat types at the EU level). • Complex landscape shaped by human activity since the antiquity. • Severe land use modification from natural to agricultural over the last 70 years • Affected by mega fires and numerous severe wildfires. • Tourism and nature recreation hotspot.

The lowland landscape and plains have been transformed over the last 70 years from semi-natural areas (grazing land and traditional cultivations) into intensively cultivated areas, including wetland drainage, river management for irrigation, and the impact of numerous wildfires — including mega-fires, with the most extensive and severe occurring in 2007 — that altered the area's physiognomy. Moreover, in the last 3 years, severe storms and heavy rainfall impacted the region (especially the western part), destroyed cultivations and infrastructure, leading to flooding and soil erosion, mainly in areas recently hit by forest fires.

In addition, the ever-growing tourism sector in Greece includes the Peloponnese among the must-visit destinations, offering not only sandy beach holidays but also hosting significant historical sites of global importance, such as Ancient Olympia, Sparta, and Mycenae. As a result, the demand for tourism infrastructure is growing, especially in coastal areas. Furthermore, wind farms and photovoltaics are being intensively developed in the region, transforming land use and degrading wilderness — particularly due to the extensive road networks required for large-scale wind park development on mountain tops.

However, the region's strategy includes the promotion of sustainable tourism and agricultural development (e.g. eco- and agro-tourism, skiing, etc.) to help distribute visitors throughout the different seasons of the year.

The Peloponnese was selected as a case study because a vast amount of *in situ* vegetation and ecosystem type data is available, which can be used to train Earth Observation (EO) algorithms and validate results. Additionally, data on ecosystem condition are also available via the Habitat Directive Article 17 reporting database (conservation degree at plot) (Hellenic Ministry of Environment and Energy 2016 and ongoing assessment), as well as from the MAES national project (LIFE IP 4 NATURA) and the relevant MAES_GR online database (ecosystem condition at plot) (Kokkoris et al. 2021 and ongoing database updates).

For these reasons (Table 1), Peloponnese has been considered a valuable case study for implementing ecosystem extent and condition accounts, to document, capture, and record changes for policymaking and management. The outcomes should highlight areas (hotspots) where changes in extent and/or condition require attention, and where measures and actions should be taken — especially in protected areas, endemic forests, and priority habitats.

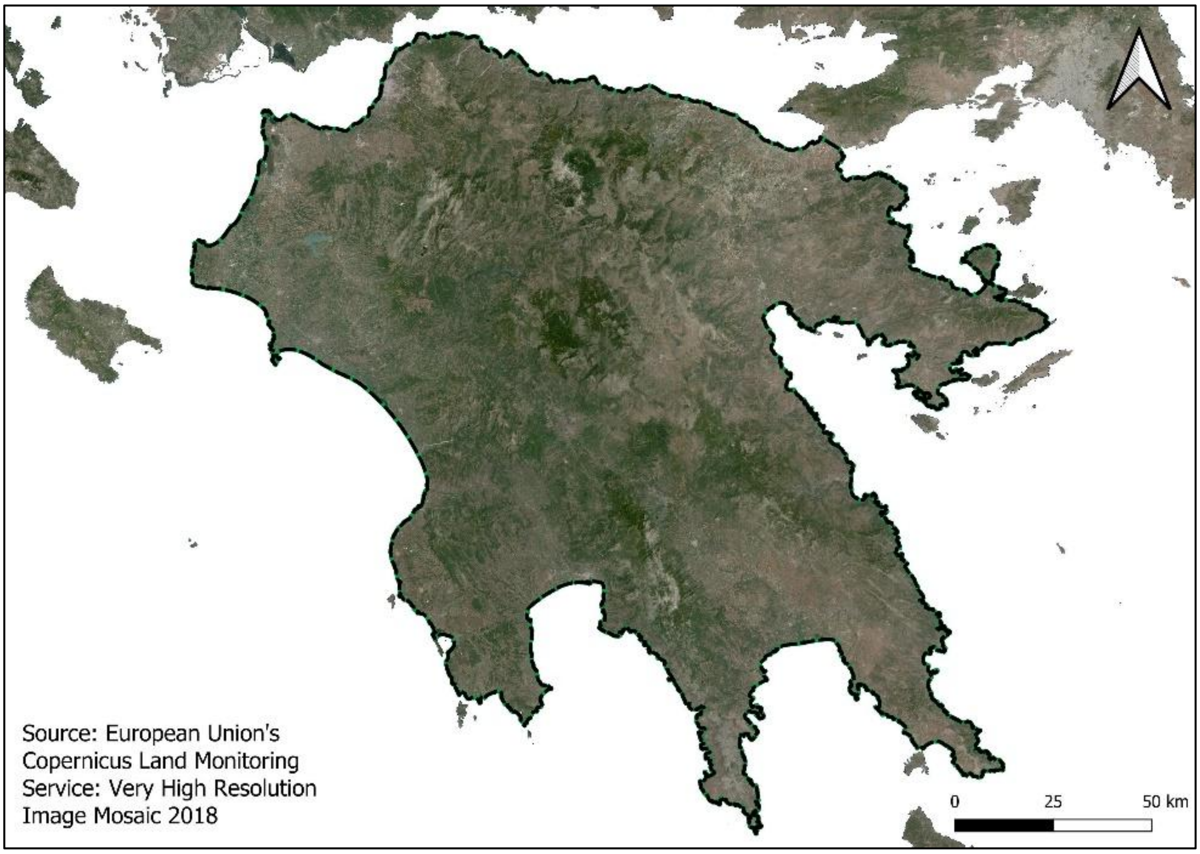


Figure 3: Orthophoto map of Peloponnese peninsula (EPSG: 3035).

5.2. São Miguel, Azores (Portugal)

The island of São Miguel (Figure 4) is the largest (area: ~745 km^2) and most populated (~140.000 inhabitants) of the Azores archipelago, one of the outermost EU regions located in the north-Atlantic Ocean and belonging to Portuguese territory. The nine Azorean islands are a group of peaks of volcanic origin rising from the Mid-Atlantic Ridge, along the Azores Triple Junction, where the North American, Eurasian, and African tectonic plates meet. São Miguel was formed from east to west, beginning with the uplift of the Nordeste (northeast) volcanic complex estimated around 4 million years ago, followed by the activity of the younger central and western volcanoes, whose explosive eruptions and subsequent caldera-forming collapses have shaped its volcanic architecture and ongoing geothermal activity. São Miguel's geological history of both destructive and constructive (effusive lava flows) volcanic processes established a complex landscape shaped by ongoing tectonic and volcanic activities.

Since their human colonization in the XV century, the Azores islands suffered intensive deforestation and conversion of native habitats. The native vegetation of São Miguel Island was typical of the subtropical temperate Macaronesian biogeographic region - predominantly laurel forest, known as "Laurisilva", complemented with arborized *Sphagnum* bogs, grasslands, and coastal scrub. Paleoecological studies (involving pollen analysis from caldera lake sediments) further confirm that São Miguel was originally covered by a dense, multi-layered evergreen forest, dominated by endemic and native species such as broadleaved *Laurus azorica*, *Erica azorica*, *Ilex perado* and coniferous *Juniperus brevifolia*. High-altitude areas supported cool and humid stunted cloud forests due to persistent cover by low-lying clouds and fog at canopy level, coupled with punishing winds, forming dense mats of bryophytes (mosses and liverworts), lichens and ferns with rich associated biodiversity. Human colonization led to rapid and widespread deforestation, primarily for agriculture, timber, and pastureland, drastically reducing the extent of native laurel forest, natural ungrazed grasslands and wetlands in all but the most steep or inaccessible areas.

The widespread land-use conversion was closely followed by the introduction of exotic invasive alien species (IAS), well-suited by the island's fertile volcanic soils as well as temperate oceanic, mild and humid climate, akin to their own native locations (Australia, East Asia or tropical/sub-tropical Americas), enjoying year-round growth under the nutrient-rich volcanic substrate and lack of extreme seasonal stresses (frost or drought). Consequently, these species were easily established and became dominant in many landscapes, further pressuring the degradation of native habitats. During the XX century, a more recent logging wave for timber production led to reforestation by large-scale and fast-growing monoculture plantations with the Japanese cedar (*Cryptomeria japonica*), and on a smaller scale, with Australian *Acacia spp*. and *Eucalyptus spp*. plantations. Other introduced exotic species like *Pittosporum undulatum*, *Hedychum gardnarium*, *Hydrangea macrophylla* or *Gunnera tinctoria* became highly invasive and spontaneously widespread, displacing and outcompeting the native flora due to their aggressive plant traits: faster growth and reproductive output; efficient dispersal; canopy light-altering properties; allelopathic alterations of soil chemistry; thriving on shaded understory layers, quicker colonization of bare or highly artificialized lands.

São Miguel island is included in the Natura 2000 ecological network of the EU territory through the implementation of protected sites preconized in Directives 79/409/EEC (Birds Directive) and 92/43/EEC (Habitats Directive). São Miguel has three land and one coastal protected site for the conservation of the remnants of laurel forest and its endemic Azorean bullfinch bird (*Pyrrhula murina*), endemic heaths, wetland bogs, crater lakes and surrounding slopes, lava flows, and coastal cliffs. Contemporary restoration initiatives are mainly targeted at these protected sites with both native reforestation efforts and controlling measures against invading exotic species. Sustainable and scientifically informed forest management practices are needed for achieving these goals while providing sufficient timber production and optimizing the supply of other key forest-related ES such as carbon sequestration.

The combination of natural hazards (earthquakes, volcanic eruptions, landslides, rockfalls, subsidence, floods, coastal erosion) and human-driven pressures affecting ecosystem extent, condition and services supply on the island, require improving the spatial and temporal resolution of data feeding into ecosystem accounts for capturing its complex and fast-paced ecological dynamics. Higher spatial resolution allows more detailed mapping of fine-scale habitat features, such as small remnants of native laurel forest, peat bogs, and riparian zones, while higher temporal resolution, with frequent revisiting times, allows tracking the rapid spread of invasive species, detecting land-use change and degradation of native habitats, monitoring vegetation recovery in restoration areas or measuring the impacts of sudden events like storms and landslides.

For these reasons (Table 2), São Miguel is a suitable and distinct TS for SELINA's Task 5.2 goals, representative of the unique challenges of EU outermost regions, where, unlike continental regions, ecological pressures are exacerbated by limited land availability, isolation and high endemism, requiring highly precise and updated assessments for increased policy relevance and responsive action. The expected outcomes of this task for this TS, related to developing highly detailed carbon accounts in forests and peatlands, should provide an exemplary demonstration of improved integrated ecosystem assessments by:

- combining optical and radar RS data for enhanced spatial and temporal resolution.
- leveraging field validated data for increased robustness.
- harmonizing habitat classifications towards the ETA for increased consistency.

Table 2: Key characteristics of São Miguel Island Test Site, Azores.

Key characteristics of São Miguel Test Site
• Volcanic origin with a complex landscape shaped by explosive eruptions, lava flows, and ongoing tectonic activity. • Originally covered by dense laurel forests, *Sphagnum* bogs, grasslands, and coastal scrub typical of the Macaronesian biogeographic region. • Severe deforestation and habitat destruction over time due to agriculture, timber extraction, and pasture expansion, driving biodiversity and endemism loss. • Heavily impacted by invasive alien species displacing native flora. • XX century reforestation dominated by exotic monocultures (*Cryptomeria*). • Includes Natura 2000 sites protecting remnants of laurel forest, crater lakes, endemic species, and coastal habitats. • Ecosystem condition impacted by natural hazards and human pressures.

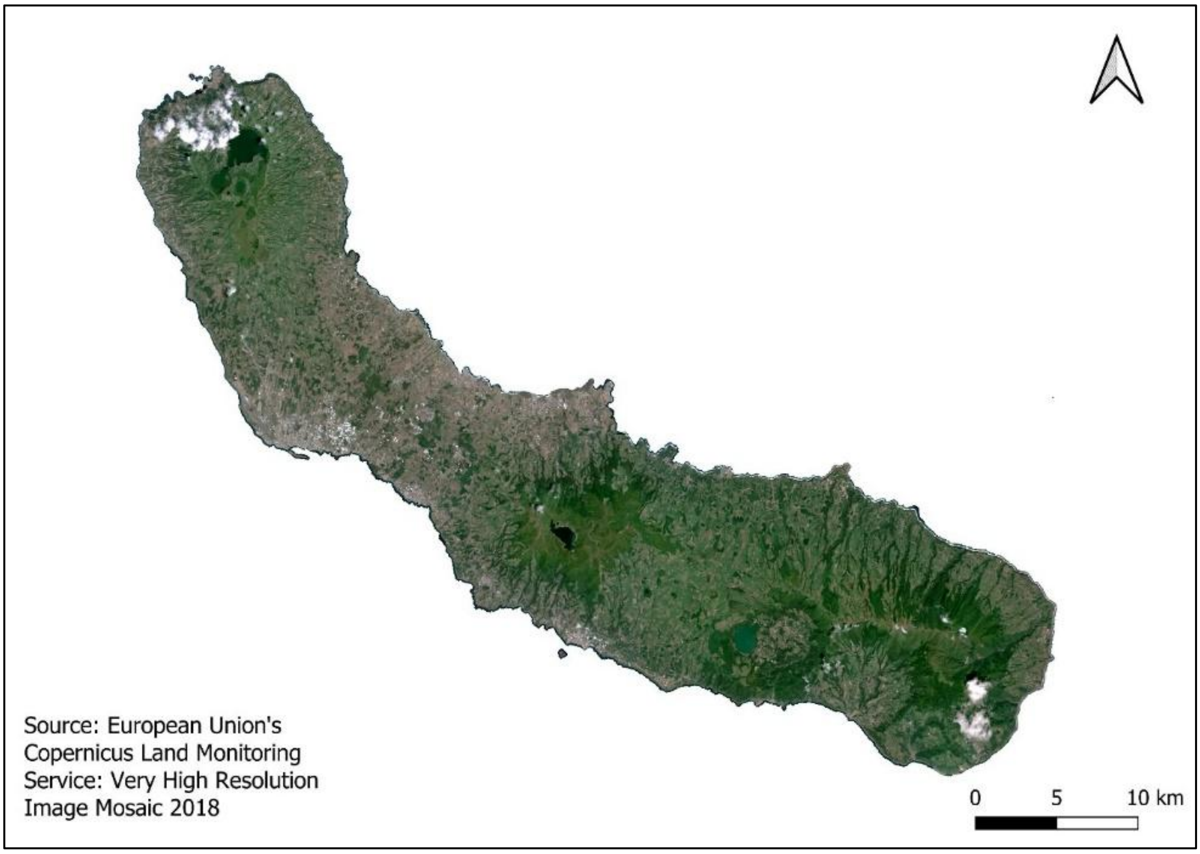


Figure 4: Orthophoto map of São Miguel Island (EPSG: 3035).

6. Data

This Section outlines the data used for the different analysis accomplished and is divided into:

- The input data products until June 2024.
- A description of how these were used.

6.1. Existing data

Even though most European countries have national datasets and share data produced by the EEA or Copernicus, there are differences in both the quality and extent of the data. Some areas are oversampled, while others may be underrepresented due to low coverage or the rarity of certain features. This disparity in data availability between countries and regions can introduce various types of bias, which may affect mapping and modeling processes. Table 3 shows some of these biases and its consequences.

Table 3: Type of bias created by data availability disparity, affecting mapping and modelling processes.

Type of Bias	Explanation	Consequences
Sampling Bias	Sites with more comprehensive data may be overrepresented in analysis.	Skewed results toward well-documented regions; underperformance in poorly mapped areas may go unnoticed.
Model Overfitting	Machine learning models may overfit data-rich regions.	Poor generalization to data-poor or structurally different regions.
Spatial Bias	Geographic or ecological regions with limited data might not be accurately captured.	Misrepresentation of land cover or habitat distribution across Europe.
Validation Bias	Model validation is stronger in regions with better data, leading to misleading performance metrics.	Overestimated accuracy of classification or predictions.
Temporal Bias	If data are uneven not only in quantity but also in time (e.g., recent data in some areas, old in others).	Inconsistent comparisons across sites; false trends or stability detection.
Ecological Bias	Certain ecosystems (e.g., remote wetlands, alpine regions) may lack high resolution or recent data.	These ecosystems may be undervalued or misclassified in conservation planning.

For Task 5.2., national datasets were complemented with other European datasets with the idea of improving the quality and accuracy of the spatial information.

6.1.1. National data

6.1.1.1. Peloponnese TS

For the Peloponnese test site, datasets at national level were made available by the University of Patras (Table 4). A vector boundary file (boundary) was provided to delineate the area to be mapped (the delineation can also be seen in Figure 3). A habitat dataset (N2K) of the Habitat's Directive Annex I habitats inside Natura 2000 sites formed the full complement of national data available for Peloponnese. The N2K dataset only partially covered the extent of Peloponnese, reiterating the need to investigate other appropriate datasets to fill gaps and complete wall-to-wall mapping.

Table 4: National datasets used for mapping Peloponnese ecosystems.

Abbreviation	Data Name	Reference year	Data format	Resolution
Boundary	Peloponnese_coastline	2025	vector	-
N2K	Natura2000_Habitat_types _(Scale_1_5000)	2016	vector	-

Section 7.1.1 explains in more detail how these datasets were used to crosslink and map ecosystem extent following the European Ecosystem Typology.

6.1.1.2. São Miguel TS

Datasets used for São Miguel at national level are listed in Table 5. A wall-to-wall dataset at national level for São Miguel is the LU 2018 map (Figure 5), a Land Cover/ Land Use map of the Autonomous Region of the Azores (DRA, 2018). It was produced based on SPOT6 and SPOT7 satellite imagery from 2015-2016 and quality controlled using very high-resolution satellite images and additional vector data derived from official cartography.

Another layer used in the analysis is the Forest Inventory Map of 2024 (Figure 6), focusing on specifically mapping of forest species. As of the time of this report, the map is still under validation by the forest services, using drone and *in situ* data. Thus, only a partial validated dataset was available for the analysis, which includes information on scrubs, forests, inland waters, wetlands, and urban areas, along with detailed data on dominant and subordinate species, as well as whether the vegetation is native or exotic, spontaneous or cultivated.

Table 5: National datasets used for mapping São Miguel ecosystems.

Abbreviation	Name	Reference year	Data format	Resolution
LU	Land Use Map 2018 Azores	2018	vector	MMU 0.95 ha MMW 20 m
FI	Forest Inventory Map 2024	2024	vector	MMU 0.05 ha MMW 2-20 m

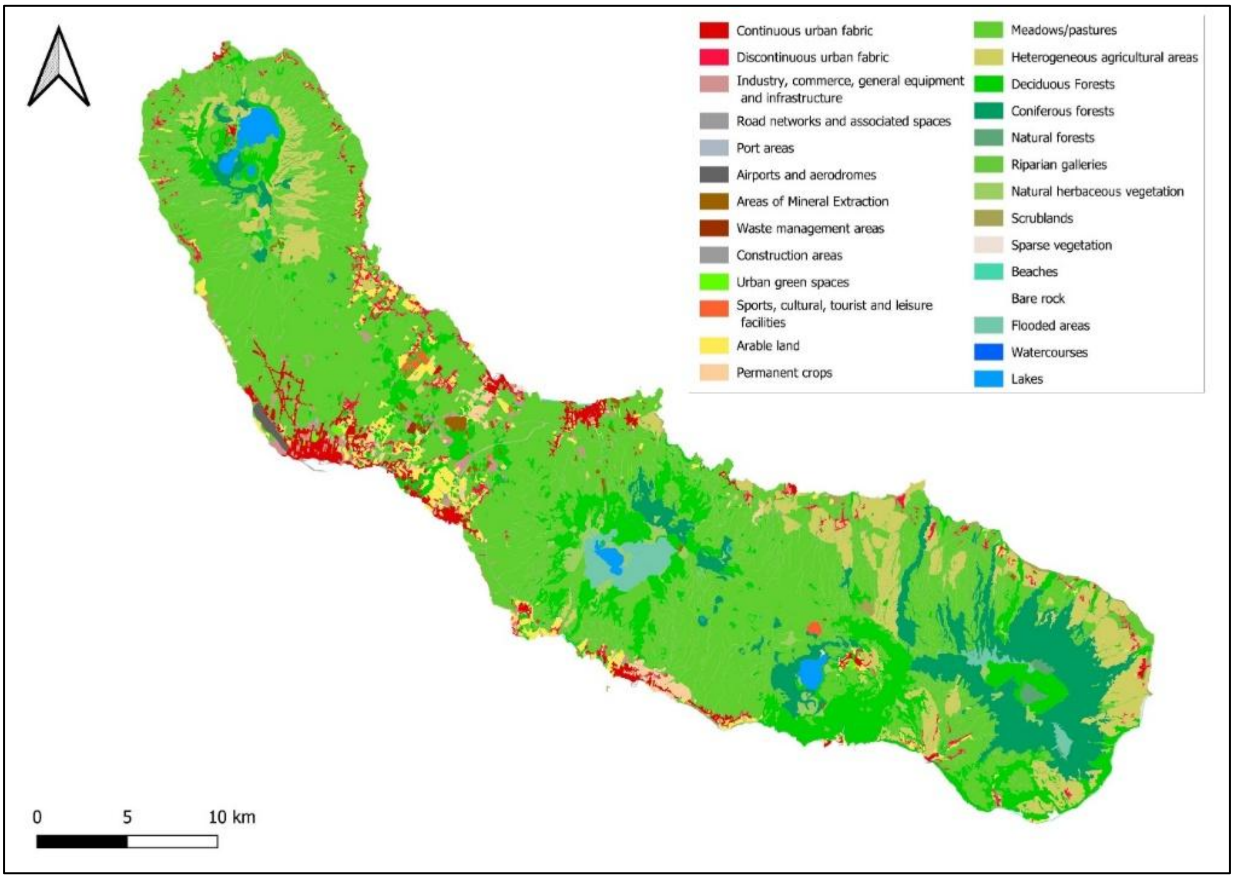


Figure 5: Land Use 2018 map of São Miguel (DRAAC 2018).

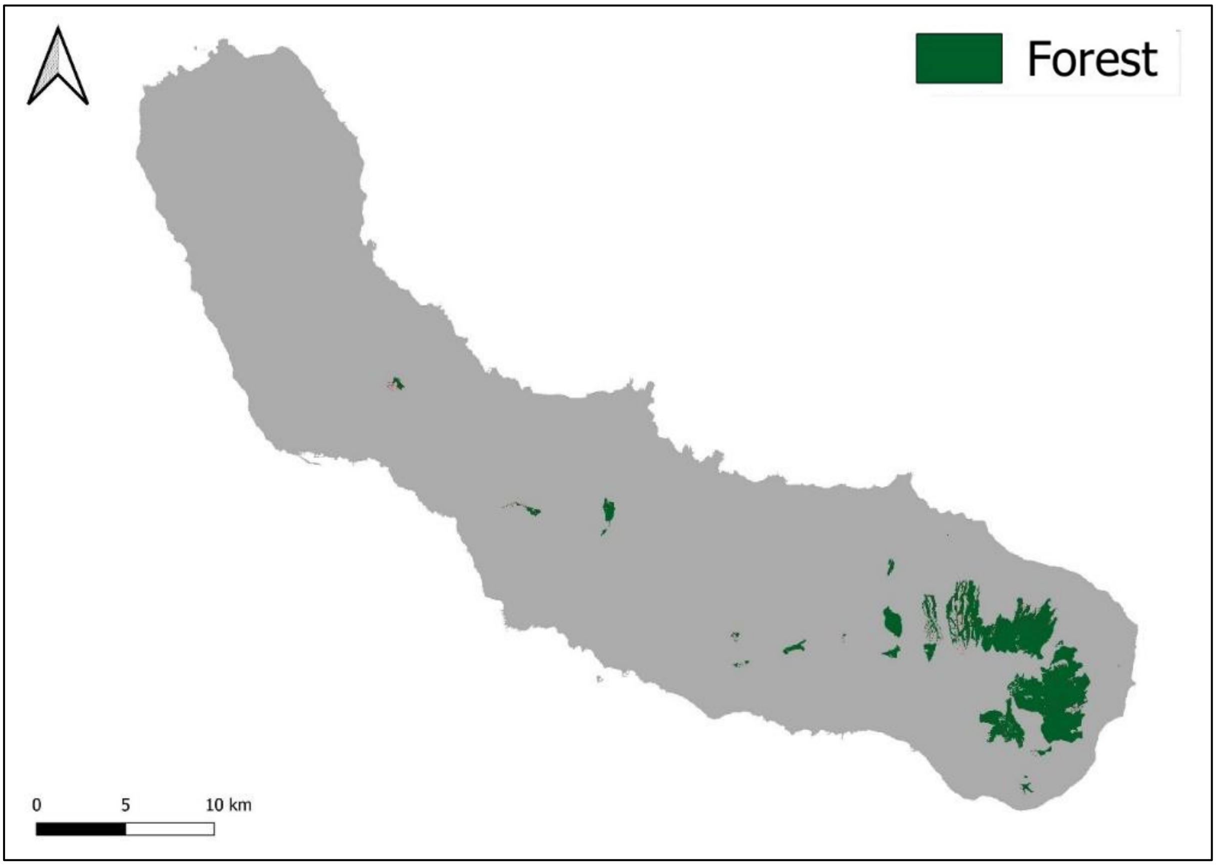


Figure 6: Forest Inventory 2024 draft (partial) data used to map São Miguel.

6.1.2. CLMS data

The Earth Observation (EO) datasets that were found to be most appropriate for this exercise were Copernicus Land Monitoring Service (CLMS) data products. CLMS provides comprehensive, high-resolution geospatial data on land cover, land use, and environmental conditions across Europe. It delivers essential information derived from satellite imagery, *in situ* data and modelling. These products are useful due to their regular update cycles, high consistency, spatial detail and free accessibility.

The CLMS products that were used are listed under Table 6. They were organized according to their relevant components, as found under the CLMS portfolio. The Priority Area Mapping (PAM) products provide high-resolution, detailed land cover and land use data for specific regions of interest. Available datasets include: Coastal Zones (cz), Riparian Zones (rz) and Urban Atlas (ua). The second set of CLMS products are grouped under the heading "Land Cover and Land Use Mapping": High-Resolution Layers such as the Imperviousness layer (IMD), the Forest type layer (FTY), CLCplus Backbone (CLCplus BB) and CORINE Land Cover (CLC). Finally, the Reference and Validation Data include the EU-Hydro layer.

Table 6: CLMS data used for ecosystem extent mapping.

Components	Abbreviation	Data Name	Reference year	Data format	Resolution
Priority area mapping	cz	Coastal Zones	2018	vector	0.25 ha
	rz	Riparian Zones	2018	vector	0.25 ha
	ua	Urban Atlas	2018	vector	0.25 / 1 ha
Land Cover & Land Use	IMD (LC)	High Resolution Layer Imperviousness	2018	raster	10m
	FTY (LU)	High Resolution Layer Forest Type Layer	2018	raster	10/100m
	CLCplus BB (LC)	CLCplus Backbone	2018	raster	10m
	CLC (LC+LU)	CORINE Land Cover	2018	raster	100m
Reference and Validation Data	EU-Hydro (LU)	EU-Hydro	2018	vector	-

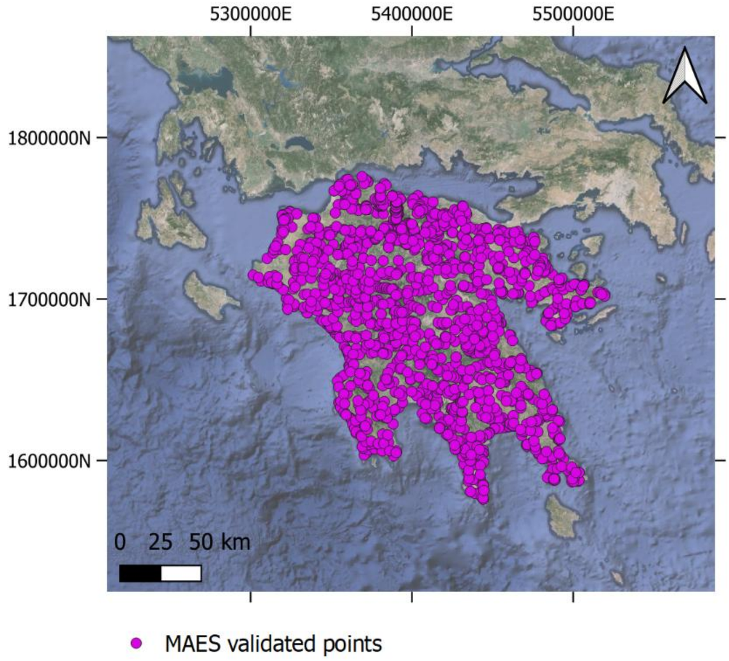
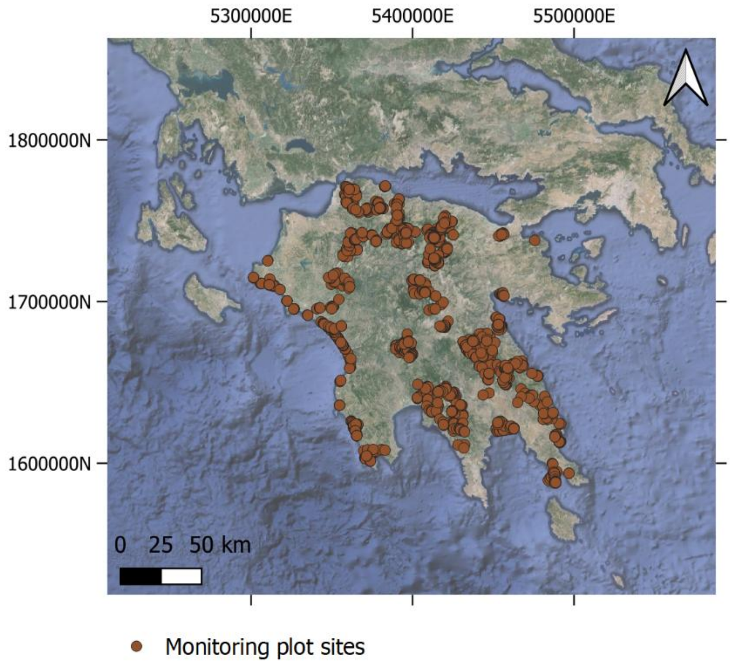
6.2. *In situ* data collection

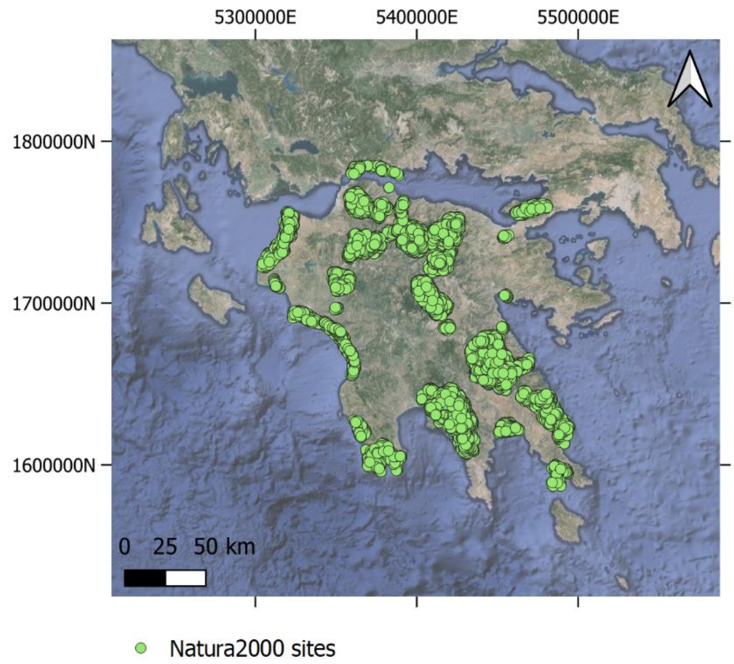
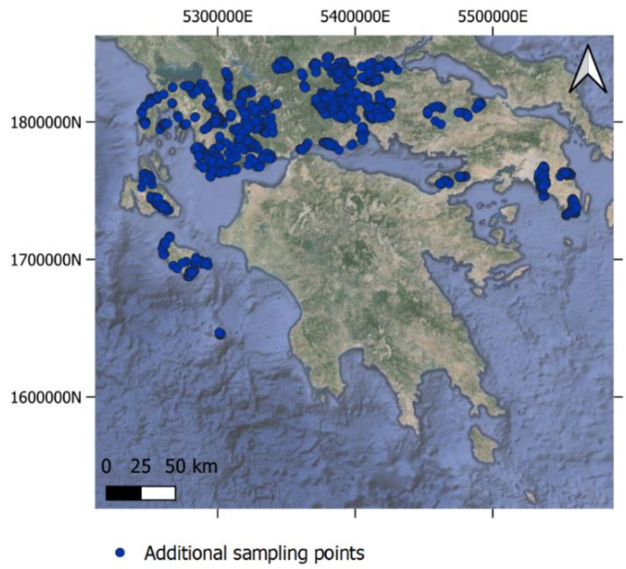
6.2.1. Peloponnese TS

The input data for training point selection for the Peloponnese TS consisted of 4 datasets (Table 7). These point locations were classified into habitat classes according to the Annex I typology. The classes were translated into EUNIS2021 habitat classes by expert knowledge from Peloponnese TS local partner.

Table 7. Overview of provided input data for training point selection.

Dataset (point vectors)	Description
Validated points - MAES map (Verde et al. 2020)	Validated in the field during LIFE-IP 4 NATURA (LIFE Integrated Project). Original classification in Annex I format. Source: MAES_GR online platform export (Developed for the LIFE IP 4 NATURA Project) (Kokkoris et al. 2021 and ongoing database update).

	
Monitoring plot sites	Situated mainly within and near Natura2000 sites Original classification in Annex I format. Source: Natura 2000 Monitoring plots database (2016) 
Natura2000 sites	Selected and validated points of Natura2000 polygon map Original classification in Annex I format. Source: Natura 2000 SAC Habitat type mapping (2016) (points extracted as centroids from each habitat type mapping polygon).

	
Additional sampling points	Additional validated point vector dataset, covering surrounding Greek domain around Peloponnese, to add training input for coastal zones, grassland and shrub type habitats Original classification in Annex I format. Source: Natura 2000 habitat monitoring plots database (2016) 

6.2.2. São Miguel TS

The training data for mapping ecosystem types in the São Miguel TS was compiled using a methodical selection of the latest versions of three local reference datasets, produced by Azorean regional authorities (1. and 2. are the same previously listed in Table 5):

1. Land Use Map 2018, developed by the Regional Directorate of the Environment and Climatic Action (DRAAC 2018).
2. Forest Inventory Map 2024 (ongoing), developed by the Regional Directorate for Forest Resources and Spatial Planning (DRRFOT 2024).
3. Field Cartography of Natura 2000 Habitats and Species 2022, produced within the scope of the LIFE IP AZORES NATURA project of the Secretary General for the Environment and Climatic Action of the Azorean Government (Dias and Pereira 2022).

Habitat classes in these datasets were translated into EUNIS 2021 or EUNIS 2012 Level 3 using expert interpretation. The process included all classes that could be confidently matched to a EUNIS L3 equivalent, regardless of their spatial representation or abundance, to ensure comprehensive ecological coverage.

Although training points were generally selected using a stratified sampling method, no class was excluded a priori. Underrepresented or rare classes were sampled manually, applying a 50-meter minimum spacing rule between points. Annex 13.5.2 lists all source classes, their dataset of origin, and the number of sampling points per class for São Miguel TS.

The Natura 2000 habitat cartography was a prioritized source for several reasons:

- It is the only fully field-validated and finalized dataset.
- It uses Annex I habitat codes from the EU Habitats Directive (92/43/EEC), with direct crosswalk to EUNIS classes.
- Its mapping methodology was specifically tailored to detect and delimit habitats in their reference, undisturbed state, aligned with Annex I definitions.
- It covers all nine islands of the Azores archipelago, unlike the forest inventory, which is still incomplete and currently limited to São Miguel.

Representativity of Annex I habitats underrepresented in São Miguel, such as Azorean laurel forests and endemic Macaronesian heaths, was improved by incorporating Natura 2000 habitat cartography data from Pico and Terceira islands, given their ecological similarity:

- These islands belong to the same Macaronesian biogeographical region.
- They are volcanic in origin, with similar soil conditions, humid Atlantic climate, and altitudinal ecological zonation.
- They share a common pool of native and endemic species and show similar patterns of habitat degradation and historical disturbance.

This ecological alignment legitimizes the inclusion of sampling points from other islands to supplement rare or fragmented habitats in São Miguel.

A fourth dataset was introduced to address specific gaps in grassland and agricultural class coverage: the TerceiraLandUse map. It provided additional validated points for:

- Cultivated broadleaved orchards.
- Secondary semi-natural mesic grasslands.
- Artificial grasslands under rotational use, commonly found throughout the Azores.

These rotational grasslands are typically reseeded and rotated with maize for cattle feed, a practice supporting soil fertility and weed control, but they are challenging to map due to:

- Frequent cloud cover, which disrupts satellite-based multi-temporal analysis.
- Spectral confusion in remote sensing, where, recently grazed or harvested plots may resemble fallow land and maize fields may appear similar to permanent crops.

Given these limitations, the forest inventory classified all rotational grasslands under a single general grasslands class, provided the land's primary use was grazing. The EUNIS classification currently lacks a specific subclass to represent this dynamic land use. After consultation with the EEA EUNIS team, it was deemed acceptable to map these areas to EUNIS class V31 (agriculturally improved grasslands), under the condition that maize cultivation does not exceed 3 months every year.

Annex I habitats were crosswalked directly from the Natura 2000 dataset to EUNIS Level 3. When overlapping with other datasets, Natura 2000 always took precedence. For non-Annex Forest and heathland classes, the forest inventory was preferred due to known inaccuracies in the land use map—for example, overestimation of deciduous forest — and the availability of a fully validated subset in the eastern part of São Miguel known as the “forest perimeter”. This area includes public forest and heathland lands, some of which are certified by the Forest Stewardship Council (FSC) for responsible forestry.

For agricultural and artificial classes not captured by Natura 2000, the forest inventory and land use map were used interchangeably, depending on polygon precision, best EUNIS class match (by code, description, and diagnostic species), or expert judgment. For instance, the road network was better vectorized in the forest inventory, while the land use map uniquely distinguished between continuous and discontinuous urban fabric. All training point locations were visually validated and corrected using Google Earth imagery where vector polygons from the source datasets were noticeably imprecise.

7. Methodology

With the aim of developing methodologies to enable the semi-automated production of SEEA EA accounts, Section 4 describes the methods developed for the two TSs to:

- Delineate ecosystem extent accounts.
- Develop an ecosystem condition accounting process.
- Develop a process for forest carbon accounting.

The target typology is the ETA for reporting on European Environmental Economic Accounts (Regulation (EU) No 691/2011).

7.1. Ecosystem extent mapping

Ecosystem extent accounts are the basis for ecosystem accounting (EUROSTAT 2024). There are multiple approaches for delineating ecosystem extents. The quality of the output is dependent on several factors, such as availability of data, quality of data, availability of expertise (where crosslinking ecosystem typologies is needed), access to spatial analysis expertise, to name a few. It is generally accepted that data available at national level, e.g. habitat surveys and monitoring programs, are the best type of information as a basis for extent delineation. However, it must be acknowledged that this type of data is not always available or is only partially available. Deferring to additional EU-wide data products to enhance national data, it is oftentimes necessary to ensure a consistent and reliable EU-wide ecosystem accounting. Sections 7.1.1 and 7.1.2 explore two methodologies:

- A **national-centric approach** that prioritizes the availability of nationally available datasets to be enhanced with freely available CLMS data.
- A **vegetation-centric approach** modelling habitat maps to derive ecosystem extents.

7.1.1. National-centric approach

The aim of the national centric approach is to determine the feasible mapping level for an ecosystem extent map based on the ETA using 1) national data only 2) national combined with CLMS data and 3) CLMS data only. Where national data is unavailable or the crosswalked ETA level is limited to Level 1 or 2, CLMS data can enhance the map by filling data gaps and enabling mapping at a more detailed level than possible with national data alone.

Datasets compiled at national level are generally prioritized as the best information available, often being compiled from *in situ* data, partial or complete field surveys or programs and generally verified by experts on local habitats. However, this information may be limited in terms of not being regularly updated or only specific ecosystem types focused on. As continuous surveying is costly, the use of openly available data to complement national data is an option to complete wall-to-wall mapping. The CLMS data portfolio provides a dual set of

European products: a set of wall-to-wall high-resolution land cover raster layers and a second set of vector products focused on certain areas of priority interest, such as coastal zones, riparian zones and urban areas. These products are free to use and regularly updated, making them an ideal solution to both facilitate crosswalking between ecosystem class information and the provision of spatial data for extent delineation.

This Section describes the methodology followed for developing complete ecosystem extent maps to the ETA for the two TS. As the methodologies differed slightly for both, due to the amount and quality of national data available as a starting point, they are described separately.

For both TS, the methodology resulted in three extent mapping outcomes:

- Using national data only.
- Using CLMS data only (”Blank canvas approach”).
- Using national data enhanced with CLMS.

The mapping process for both TS was undertaken in QGIS (QGIS Development Team 2024) and RStudio version 2024.04.2 Build 764 (Posit team 2024). For a detailed overview of how the datasets were combined, the crosswalks, scripts and processing steps have been shared and documented for each TS in sections 13.1 and 13.2 of the Annex of this report. The following methodologies focus on the crosswalking process for each TS while also providing examples of how the datasets were combined. The results and statistics extracted are described further in Section 8.1.1.

7.1.1.1. São Miguel TS

The workflow below on Figure 7 outlines the process of mapping ecosystem extents for São Miguel by integrating multiple data inputs. It begins with the collection and preprocessing of national and CLMS datasets. During the pre-processing stage, all datasets were rasterized to 5 m and reprojected to EPSG:3035. These inputs were also crosswalked to the ETA and combined using local knowledge from São Miguel TS local partner with S4E expertise to generate the three ecosystem extent maps.

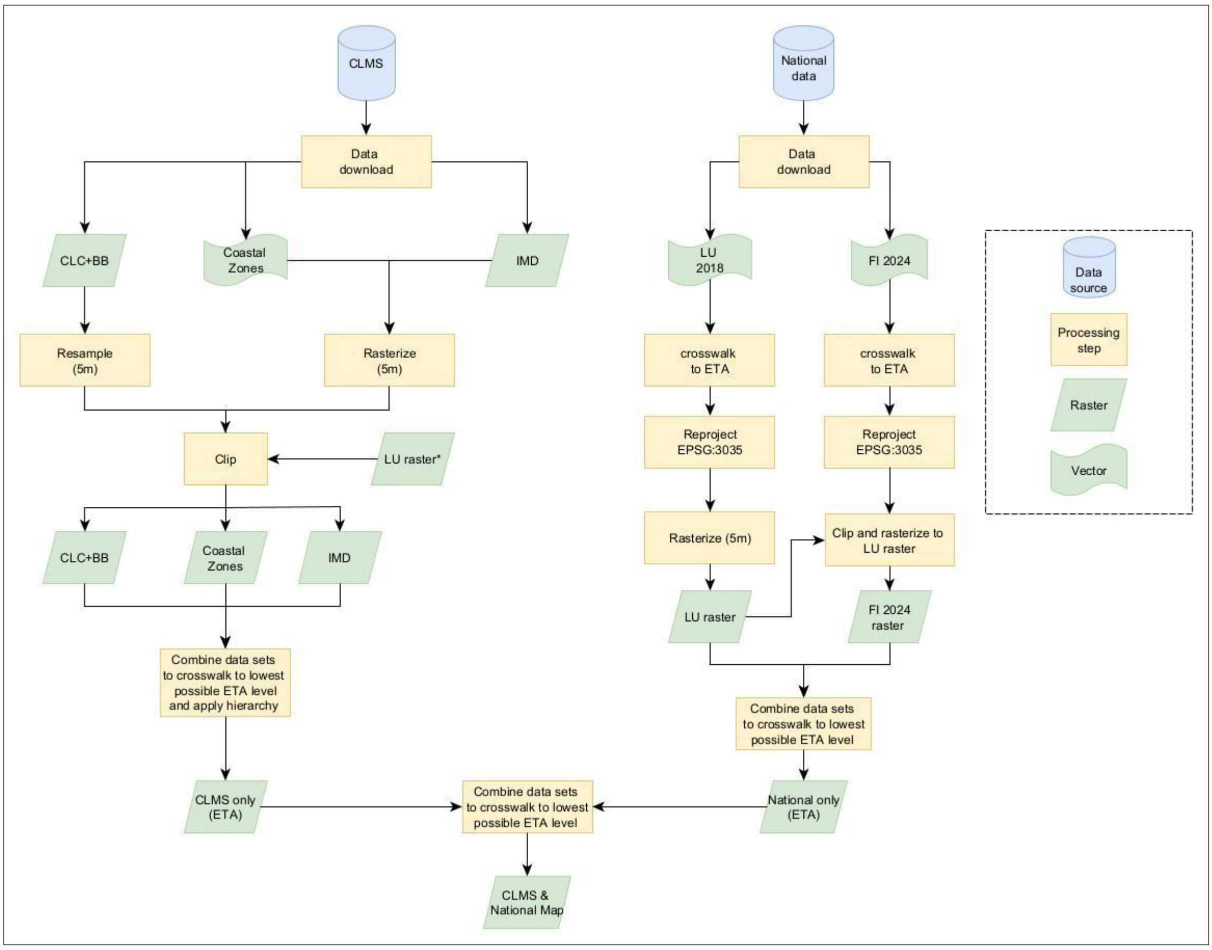


Figure 7: The processing workflow for developing wall-to-wall ecosystem extent maps for São Miguel to the European Ecosystem Typology.

The broad steps of the data selection and cross-walking process for mapping São Miguel are described below.

a) Identification, cross-walking and mapping of national datasets to the ETA

An initial review of national datasets was undertaken to establish a baseline (see Section 6.1.1.2). In this step, the two national data sets, Forest Inventory 2024 (FI 2024) and Land Use 2018 (LU 2018) were crosswalked to the ETA and subsequently combined to create the 'National only map' for São Miguel.

The Forest Inventory 2024 compiles information on land use and land cover of São Miguel. It contains information on scrubs, forests, inland waters, wetlands and urban areas. It also includes detailed information on dominating and dominated species, if vegetation is native or exotic, spontaneous or cultivated. This is the most detailed forest data set available at national level. However, due to ongoing data validation at the time of mapping, only a partial dataset could be used. The validated portions used included forest and road classes, both of which were suitable for mapping at Level 3 ETA. To undertake the crosswalk, all unique column combinations were extracted from the data set and cross-walked to the ETA. This was a substantial effort, given that there are 512 unique combinations of forest classes. Table 8 shows examples of the crosswalk between the unique combinations of the FI and the ETA.

Table 8: Example of the crosswalk between the Forest Inventory 2024 towards the European Ecosystem Typology (table should be read from top to bottom, column-wise). Each column corresponds to elements for defining forest structure and composition in FI 2024.

Forest inventory 2018 - Database attributes	Combination 1	Combination 2	Combination 3	Combination 4
Level 1 - land use (IFUsolo_ID)	Forest	Forest	Forest	Forest
Level 2 - land use (IFOcSo_ID)	Exotic vegetation	Exotic vegetation	Native vegetation	Native vegetation
Level 3 - vegetation structure (IFEvert_ID)	multi-stratified	multi-stratified	multi-stratified	Uni-stratified
Level 4 - nature of vegetation cover (IFNcob_ID)	Spontaneous	Spontaneous	Cultivated	Cultivated
Level 5 - fragmentation (IF_frag_ID)				
Level 6 - composition (IFcomp_ID)	mixed	pure	mixed	pure
Most dominant species (Espécie dominante)	*Pittosporum undulatum*	*Pittosporum undulatum*	*Prunus azorica*	*Juniperus brevifolia*
Most dominant species - 2nd priority (Espécie dominada)	*Laurus azorica*	*Pittosporum undulatum*	*Frangula azorica*	*Juniperus brevifolia*
Auxiliary species (Eaux3)	*Clethra arborea*	*Laurus azorica*	*Cryptomeria japonica*	Not applicable
Auxiliary species (Eaux4)	*Cryptomeria japonica*	*Ilex azorica*	*Quercus robur*	Not applicable
Auxiliary species (Eaux5)	*Acacia melanoxylon*	*Acacia melanoxylon*	*Ilex azorica*	Not applicable
Overgrown dominant species	*Hedychium garderianum*	Not applicable	*Viburnum treleasei*	Not applicable
Overgrown dominated species	*Hedychium garderianum*	Not applicable	*Erica azorica*	Not applicable
Resulting ETC level 3 class	4.4.2 Mixed Forest dominated by broadleaved species	4.3.6 Other broadleaved evergreen forest	4.6.2 Mixed plantations	4.6.1 Monoculture plantations

The Land Use 2018 (LU 2018) data set (DRA, 2018) was also deemed valuable to map the ecosystems of São Miguel. It provides detailed classifications of land use categories, such as urban areas, pasture, agriculture, forests, water bodies and other classes. This data set was also used to delineate the boundary of São Miguel. For the LU 2018 data set, an initial visual crosswalk was undertaken (Table 9) based on class labels and definitions. All classes could be crosswalked 1:1, but some classes were not that straightforward and required a more detailed analysis of the class definition (inclusions, exclusions etc.) and occasionally an aggregation to a less detailed level in the ETA. An example of this is with the level 3 Land Use class 2018 ‘1.2.1 Industry, Commerce, General Equipment, and Infrastructures’ which could only be mapped at level 1 namely ‘1 Settlements and other artificial areas’ in the ETA. The reason for the aggregation to level 1 was that the ETA requires the class to be split up into continuous and discontinuous areas to be mapped at a lower level.

Table 9: Crosswalk between São Miguel Land Use 2018 and the European Ecosystem Typology classes (including the most detailed typology level where the match could be established).

Azores Land Use 2018	European Ecosystem Typology (ETA)		
	Level 3	Level 2	Level 1
1.1.1. Continuous Urban Fabric	1.1.1 Continuous residential area		
1.1.2. Discontinuous Urban Fabric	1.2.1 Discontinuous residential area		
1.2.1. Industry, Commerce, General Equipment, and Infrastructures			1. Settlements and other artificial areas
1.2.2. Road Networks and Associated Spaces	1.3.1 Road and rail networks and associated land		
1.2.3. Port Areas	1.3.2 Port areas		
1.2.4. Airports and Aerodromes	1.3.3 Airports		
1.3.1. Areas of Mineral Extraction	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)		
1.3.2. Waste Management Areas	1.3.6 Dump areas		
1.3.3. Construction Areas	1.3.7 Construction sites		
1.4.1. Urban Green Spaces	1.4.1 Parks (including Zoos and botanical gardens)		
1.4.2. Sports, Cultural, Tourist, and Leisure Facilities	1.4.2 Sports and recreation sites		
2.1.1. Arable Land		2.1 Annual cropland	
2.1.2. Permanent Crops		2.3 Permanent crops	
2.1.3. Meadows/Pastures		3.1 Sown pastures and fields (modified grassland)	
2.1.4. Heterogeneous Agricultural Areas		2.5 Mixed farmland	
3.1.1. Deciduous Forests		4.1 Broadleaved deciduous forest	
3.1.2. Coniferous Forests		4.2 Coniferous forests	
3.1.3. Natural Forests	4.3.3 Macaronesian laurophyllous forest		
3.1.4. Riparian Galleries	4.1.1 Riparian forest and woodland		

3.1.5. Natural Herbaceous Vegetation		3.2 Natural and semi-natural grassland	
3.1.6. Scrublands	5.2.3 Temperate and Mediterranean lowland shrub and heathland		
3.2.1. Sparse Vegetation			6. Sparsely vegetated ecosystems
3.2.2. Beaches		11.2 Coastal dunes, beaches and sandy and muddy shores	
3.2.4. Bare Rock		6.1 Bare rocks	
4.1.1. Flooded Areas			7. Inland wetlands
5.1.1. Watercourses			8. Rivers and canals
5.1.2. Lakes			9. Lakes and reservoirs

After completing the crosswalks for both datasets, they were combined into a common layer, giving FI 2024 priority over LU 2018 where both input data sets provided information for the same location, resulting in the creation of the 'National only' map of São Miguel. At this stage, based on class definitions, some national-level classes could only be crosswalked to a higher level in the ETA. In the next step, the availability of CLMS data allowed for the disaggregation of these higher-level classes, enabling mapping at a more detailed level for the National and CLMS approach.

b) Identification, crosswalking and mapping CLMS data to the ETA

To establish the degree to which ecosystem extents can be mapped based purely on CLMS data, a "blank canvas approach" was tested. The goal was to map at the lowest level of the ETA as possible (level 3).

The datasets used for this mapping included Coastal Zones (CZ), Imperviousness (IMD) and CLCplus Backbone (CLCplus BB). Due to the size of the island of São Miguel, the Coastal Zones (CZ) layer covered the entire island as it extends as a buffer up to 10 km inland from the coast, making it an ideal dataset to use with detailed class information. The CZ layer served as the primary data source and geometric framework for mapping the ecosystem extent of São Miguel. When other CLMS products were more suitable, they were integrated within the existing CZ geometries. For instance, since evergreen forests were not mapped under the CZ layer but are required under the ETA, the CLCplus BB dataset was employed to map this forest class within the existing CZ forest boundaries. Under the ETA, settlement and commercial and industrial zones are categorized into continuous and discontinuous areas at level 3, thus selected urban classes from the CZ layer were reclassified utilizing the IMD data set. All relevant classes existing within São Miguel were crosswalked to the ETA and then used for mapping within the boundary layer created from the LU 2018 data set.

The collection of these datasets that were crosswalked to the ETA were useful for the next step, which was to complement the 'National only map' with CLMS data.

c) Improving national data with CLMS, creating the CLMS and National map.

The main goal of this mapping exercise was to determine which CLMS data could be used to improve or map the existing classes from the national data at a higher level. Based on a visual comparison of the LU 2018 dataset with aerial images, the mapping consistency of arable land, meadows/pastures and heterogeneous agricultural areas was found inconsistent. Consequently, these classes required improved discrimination. The CLMS Costal Zone (CZ) and CLCplus BB layers were suitable to improve the mapping of these classes. The High Resolution Layer "Imperviousness" can distinguish between continuous and discontinuous industry and commerce surfaces, thus enabling the CZ class '1.2.1. Industry, Commerce, General Equipment, and Infrastructures' to be mapped to the ETA level 3 classes '1.1.2 Continuous commercial and industrial area' and '1.2.2 Discontinuous commercial and industrial area'.

After identification of CLMS data sources, an exercise was undertaken to establish which was the most appropriate to use to map each typology class. Table 10 shows which LU 2018 classes required additional data to crosswalk and map to the higher-level ETA. The FI 2024 dataset was crosswalked to level 3 of the ETA which is why it did not require additional CLMS data.

Table 10: Land Use 2018 data, crosswalked to the European Ecosystem Typology and refined with CLMS data used for more detailed mapping.

Source Nomenclature	**Supporting CLMS datasets for improved match between source and target nomenclature**		**Target Nomenclature**
Azores LU 2018 (already crosswalked to an initial ETA class)	Coastal Zones	CLCplus BB	European Ecosystem Typology (ETA)
210 Annual cropland		6 Permanent Herbaceous	310 Sown Pastures and Fields (modified grassland)
		7 Periodically Herbaceous	210 Annual Cropland
310 Sown pastures and fields (modified grassland)		6 Permanent Herbaceous	310 Sown Pastures and Fields (modified grassland)
		7 Periodically Herbaceous	210 Annual Cropland
210 Annual cropland	23200 Complex cultivation patterns		251 Mosaic farmland (comprising cropland, grassland and (semi-) natural components)
	23300 Land principally occupied by agriculture with significant areas of natural vegetation		
	23200 Complex cultivation patterns		

310 Sown pastures and fields (modified grassland)	23300 Land principally occupied by agriculture with significant areas of natural vegetation		
230 Permanent crops	21200 Greenhouses		151 Permanent greenhouses
410 Broadleaved deciduous forest	34000 Transitional woodland and scrub		451 Transitional woodland/forest land
411 Riparian forest and woodland			
420 Coniferous forest			
433 Macaronesian Laurophyllous forest			
900 Lakes and reservoirs	82200 Reservoirs		921 Artificial reservoirs
1120 Coastal dunes, beaches and sandy and muddy shores	62111 Sandy beach		1122 Beaches and sandy shores

Below are some examples of how CLMS data sets were used to refine the LU 2018 national mapping to assigning detailed level ETA classes.

Example 1: The ETA level 3 class '451 Transitional woodland/forest land' is not defined as a separate class under the national data. However, the CZ layer delineates and maps this at level 3 as '34000 Transitional woodland and scrub'. Thus, all LU 2018 forest classes that were mapped under the CZ data set as '34000 Transitional woodland and scrub' were crosswalked and mapped as '451 Transitional woodland/forest land'.

Example 2: The national data classifies 'greenhouses' under the ETA level 2 class 'permanent crops'. Under the ETA this class is mapped under a different level 1 class "settlements and other artificial areas' and at a lower level (Level 3) as '151 Permanent greenhouses. The CZ layer was able to isolate and map greenhouses using the '21200 Greenhouses' class.

Example 3: The national data do not consistently differentiate between level 2 cropland class '210 Annual cropland' and level 2 grasslands '350 Sown pastures and fields'. Level 3 class '251 Mosaic farmland' is frequently mapped under either of the latter two Level 2 classes. To address this, CLCplus BB is used to distinguish between 210 and 350, while CZ was used to map 251 Mosaic farmland (comprising cropland, grassland and (semi-) natural components). Once the classes that could be improved or mapped to a lower level were identified, the crosswalked CLMS data was used to substitute the areas of the 'National only map'.

7.1.1.2. Peloponnese TS

The workflow below (Figure 8) outlines the process of mapping ecosystem extents for Peloponnese by integrating multiple data inputs that will be described below. It begins with the collection and preprocessing of national and CLMS datasets. During the preprocessing process all datasets were rasterized to 10 m and reprojected to EPSG:3035. These inputs were then crosswalked and combined by local knowledge from Peloponnese TS local partner with S4E expertise to generate three ecosystem extent maps. Due to the limitation of data at national level, there was a need for the identification of numerous additional CLMS data products, greater than those employed for São Miguel.

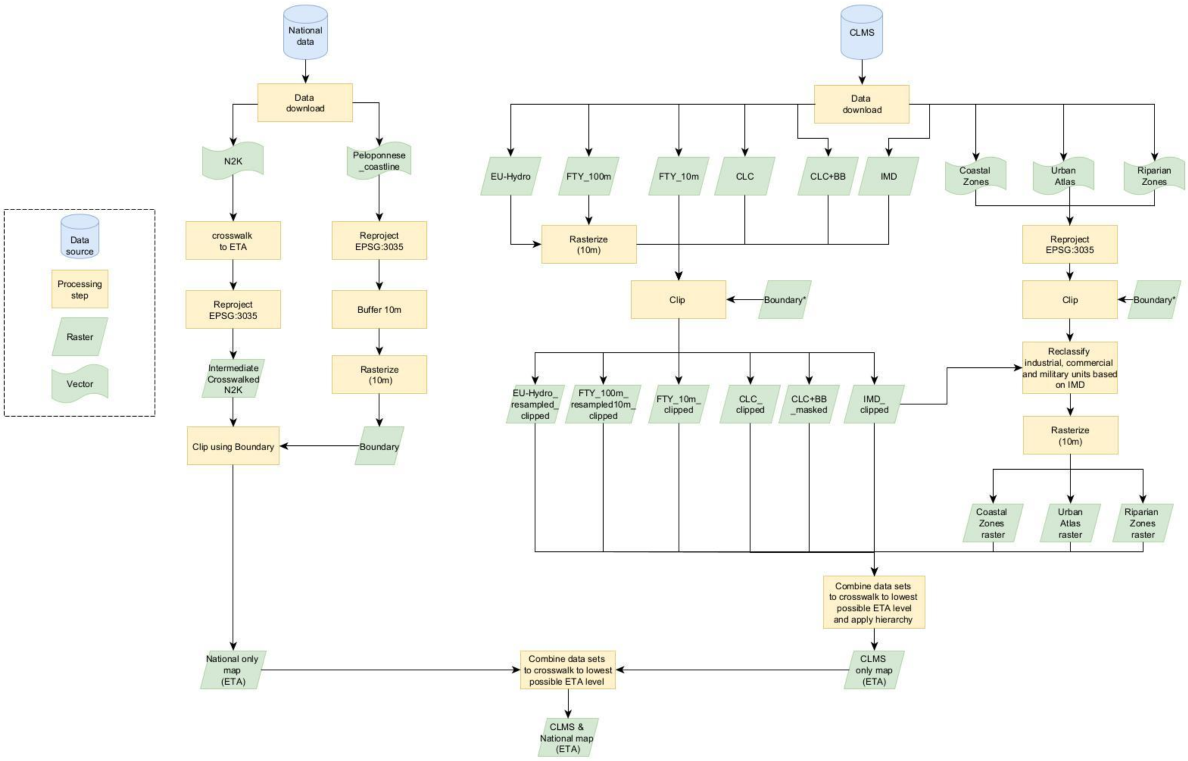


Figure 8: The processing workflow for developing wall-to-wall ecosystem extent maps for Peloponnese to the European Ecosystem Typology.

To produce all the maps, the Peloponnese coastline was used as the boundary file for Peloponnese. This vector data set was buffered by 10 m and rasterized to 10 m.

The broad steps of the data selection and crosswalking process for mapping Peloponnese ecosystem extent are described below.

a) Identification, crosswalking and mapping of national datasets to the ETA

An initial review of national datasets was done to establish a baseline (see Section 6.1.1).

The Greek Natura 2000 Habitat Types (N2K) was the only dataset at national level providing information on ecosystem types. It contains information on the distribution of Annex I habitats (i.e. reporting under Article 17 of the EU Habitats Directive). However, the dataset was limited to Natura 2000 sites, leaving the remaining areas of Peloponnese unmapped. Crosswalking the national N2K data set to the ETA is not a straightforward task. The Natura 2000 habitat types are based on species composition which contains a selected list of habitat types for protection under the EU Habitats Directive. Therefore, an intermediary step was needed whereby an existing crosswalk to the EUNIS habitat classification was used, and this was crosswalked to the ETA (Table 11). This step allowed deciding which N2K classes in the N2K datasets could be crosslinked and at which level of the ETA it could be crosswalked to.

Table 11: Examples of a 2-step crosswalk from national N2K via EUNIS to the European Ecosystem Typology.

Greek Natura 2000 Habitat Types	EUNIS code and habitat description	European Ecosystem Typology
1050 Cereals excluding rice, excluding maize	V13 Arable land with unmixed crops grown by low-intensity agricultural methods	2.1 Annual cropland
1065 Highly modified coniferous forests, in particular plantations	T3N Coniferous plantation of site-native trees	4.2.9 Highly modified coniferous forests, in particular plantations
1310 Inland salt steppes	R61 Mediterranean inland salt steppe	3.2.5 Inland salt steppes

b) Identification, crosswalking and mapping CLMS data to the ETA

Due to the larger size of Peloponnese compared to São Miguel, more data was required to map the whole peninsula solely using CLMS data (using a "blank canvas approach"). Thus, all the CLMS datasets existing within the boundary of the peninsula, listed in Table 3, were used for mapping. The classes of each data set were crosswalked into the relevant ETA class, sometimes combining several CLMS data sets to enable crosswalking.

A hierarchical framework for classifying specific land use classes was created to systematically populate a blank map of the Peloponnese with CLMS datasets. This enabled complete wall-to-wall mapping. The order of the hierarchy was determined based on expert knowledge and the likelihood of a given data source class overwriting another. For example, greenhouses were mapped separately within the riparian zones (RZ) and coastal zone (CZ) datasets and are also represented individually within the ETA dataset, so they are placed at the top of the hierarchy. In contrast, the urban atlas (UA) categorizes greenhouses as part of industry and commerce, which places them lower in the hierarchy to prevent the greenhouses from being incorrectly overwritten into the wrong ETA class. An extract of the mapping hierarchy, where CLMS data sets are combined and crosswalked to the ETA, is shown in Table 12.

Table 12: Example of the crosswalk hierarchy between CLMS data and the ETA.

Hierarchy	Dataset	UA	CZ	RZ	IMD	Crosswalked European Ecosystem Typology
1	rpz			2120		1.5.1 Permanent Greenhouses
2	cz		21200			1.5.1 Permanent Greenhouses
3	ua	11300			80-100	1.1.2 Continuous commercial and industrial area

4	ua	11300	1-79	1.2.2 Discontinuous commercial and industrial area
5	ua	11100		1.1.1 Continuous residential area
6	ua	11210		1.2.1 Discontinuous residential area
7	ua	11220		1.2.1 Discontinuous residential area
8	ua	11230		1.2.1 Discontinuous residential area

In certain cases, a crosswalk from the CLMS data to the ETA was not possible. The following are the exceptions:

Priority Area Mapping (PAM) – Water

As the ETA categorizes water into level 1 classes 8. Rivers and Canals and 9. Lakes and Reservoirs, areas mapped as 'Water' under the Priority Area Mapping (PAM) were left mapped as such because they could not be further differentiated into the ETA Level 1 classes.

Imperviousness Layer

The Imperviousness Layer (IMD) provides information on sealing density, expressed as a percentage ranging from 0% to 100%. In areas where PAM (Prescribed Aggregation Methodology) data for settlement areas was unavailable, classification used the IMD dataset. These areas were delineated as either Continuous Settlement (80–100% imperviousness) or Discontinuous Settlement (0–79% imperviousness), based on the sealing density values

CLCplus Backbone

In certain regions, the CLCplus Backbone was the only available dataset for mapping purposes, due to the absence of other key sources such as PAM or FTY. As a result, it was used as a gap filler and retained its original nomenclature, since it could not be crosswalked to the Ecosystem Type Account (ETA). This limitation stems from the fact that CLCplus is a land cover dataset, whereas the ETA framework is based on ecosystem typologies, making direct alignment between the two classifications unfeasible.

c) Improving national data with CLMS creating the CLMS & National map

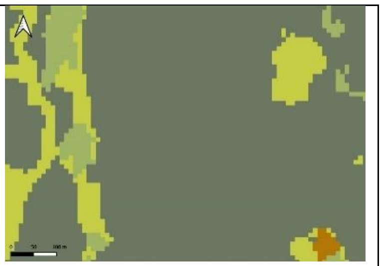
The mapping approach for Peloponnese differed from that of São Miguel. As previously mentioned, national mapping data (N2K) exists only for part of the peninsula, leaving gaps or unmapped areas that could be filled using CLMS that was crosswalked in the step above. In addition, where N2K data existed, these classes were improved (mapped to a lower level e.g., level 2 or 3) using CLMS data. The N2K classes that were crosswalked to the ETA and disaggregated to a lower ETA level are shown in Table 13.

Table 13: Classes crosswalked to ETA that were disaggregated to a lower level using CLMS data.

National crosswalked class
1. Settlements and other artificial areas
1.3 Infrastructure
2.3 Permanent crops
3.2 Natural and semi-natural grassland
5. Heathlands and shrub
6. Sparsely vegetated ecosystems
6.1 Bare rocks
7. Inland wetlands
9. Lakes and reservoirs
11.4 Coastal saltmarshes and salines

An example of how CLMS data was used to map certain ETA classes at a lower level is shown in Table 14. Here national data was complemented by CLMS data to map olive groves at level 3 from the previous aggregated level 2 permanent crops (National data). Olive groves were mapped here using the Riparian zones, Coastal Zones and CORINE datasets.

Table 14: Examples of disaggregating national data to a lower European Ecosystem Typology using CLMS data.

 Aerial image (VHR 2018) of olive groves	 Lime green area represents area mapped as 2.3 Permanent crops (National)	 Area in grey has been mapped from 2.3 Permanent crops (National) to 2.3.1 Olives (National data complimented with CLMS data)

7.1.2. Vegetation centric approach

7.1.2.1. Extent mapping workflow

In the Vegetation centric approach, a habitat map in EUNIS typology was created to derive an ecosystem extent map from this by using a developed crosswalk table between the habitat and ecosystem extent classes following the ETA, in combination with supplementary data sources (e.g. land use maps). The developed extent map for one specific year (e.g. 2021) serves as a ‘base reference’ map and can later be ‘updated’ by considering changing hotspots. The base map is composed of three important components:

- land cover physical characteristics, typically derived from land cover maps.
- ecosystem or plant functional type characteristics, according to the EUNIS typology.
- land use characteristics.

The information of these three components was combined to create an ecosystem extent map at Level 2 or Level 3 of the ETA typology (Figure 9). Once the change hotspots are detected, a similar mapping as done for the base map is to apply to find the new ecosystem extent class.

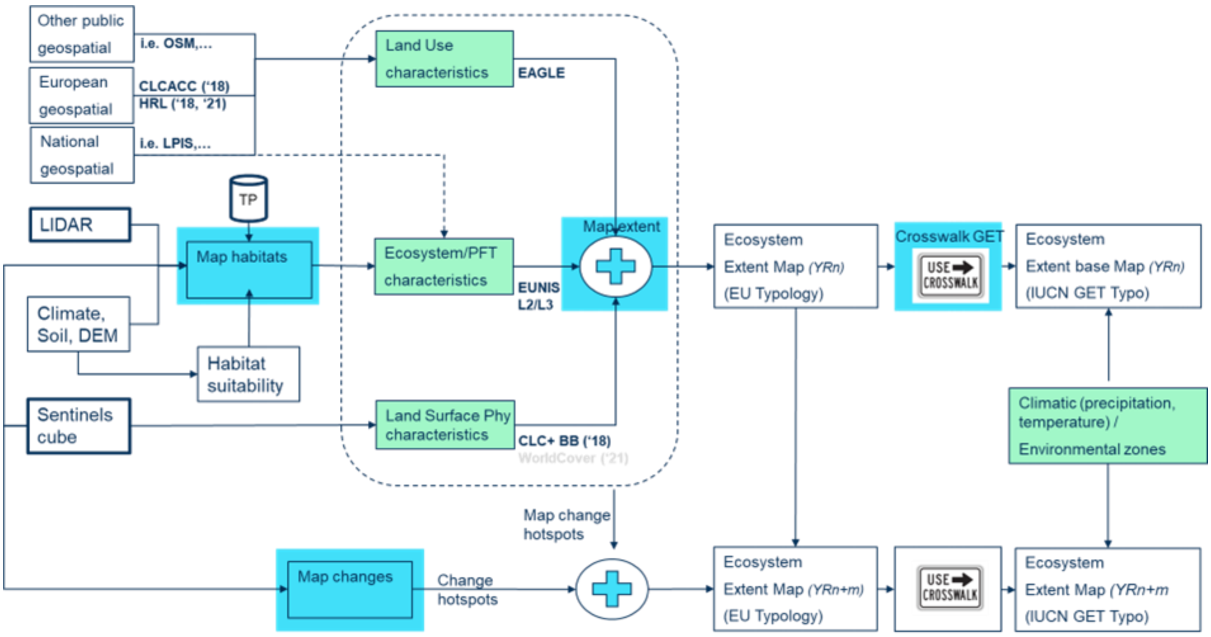


Figure 9: Overview workflow to generate ecosystem extent maps by the vegetation centric approach.

7.1.3. Mapping EUNIS habitats

An automated workflow was developed to generate the habitat map, consisting of four modules, as shown in Figure 10:

1. Feature Extraction Module, preparing a set of predictors derived from remote sensing and other datasets.
2. Feature Selection and Training Module, using the field validated and classified habitat input data as training points and generate models using the extracted features.
3. Inference Module, generating the full wall-2-wall habitat maps at different hierarchical levels applying the models from the previous step.
4. Post-Processing Module, selecting the best fit pixel from the hierarchical habitat maps in combination with external data (e.g. habitat suitability maps).

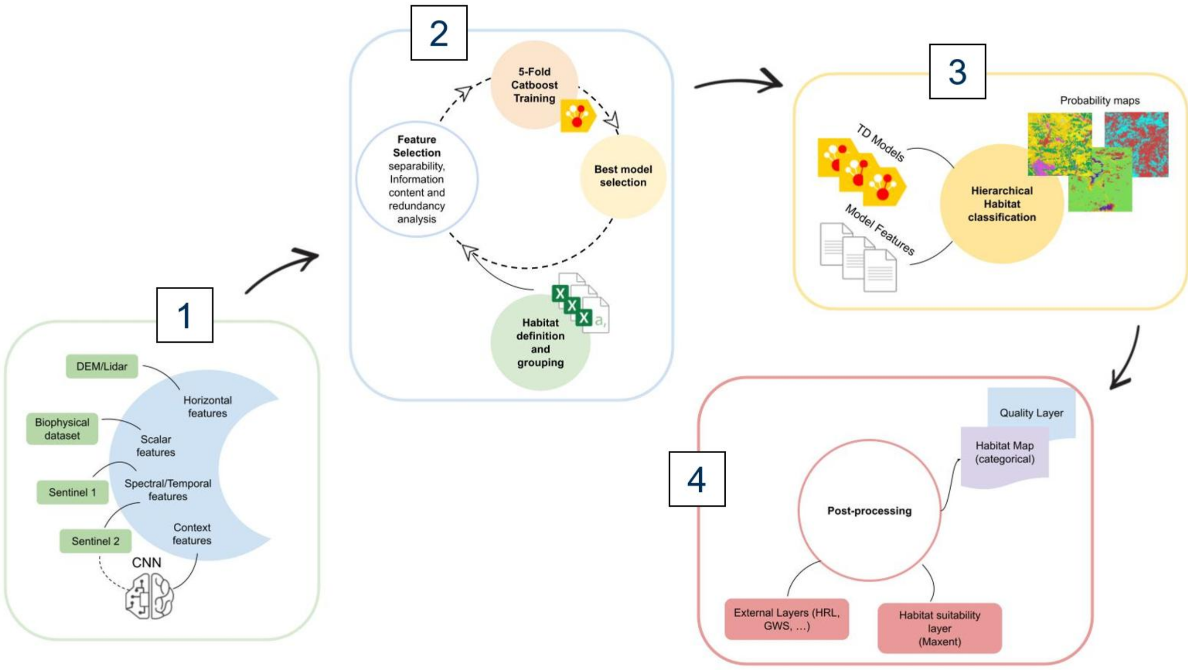


Figure 10: Illustrative overview of the different modules in VITO's habitat mapping process: 1) Feature Extraction Module, 2) Feature Selection and Training Module, 3) Inference Module and 4) Post-processing Module.

The **Feature Extraction Module** is responsible for preparing geospatial data (raster data) that provides information on the characteristics of habitats. The characteristics were categorized into four different feature sets providing a total of 150 features (aka model predictors). A full feature list is provided in Section 13.3 of the Annex.

- 3 Horizontal features, derived from Digital Elevation Model (DEM) data (altitude and slope) and LIDAR data (vegetation height);
- 25 Scalar features derived from climate data (snow covered days, mean temperature during growing season, number of growing days above 5°C, precipitation in growing season, annual precipitation), soil data (bulk density, cation exchange capacity, coarse fragments, organic carbon concentration, sand texture, clay texture, pH) and other data sources (distance to inland water, inundation occurrence, population density, vegetation phenological parameters, leaf area indices);
- 106 Spatial-temporal features, capturing both spatial and temporal information from optical (Sentinel-2) and radar (Sentinel-1) earth observation data, as is the 10th, 50th and 90th percentile and inter-quantile range of spectral bands and a number of derived indices using a temporal series of 1 year of input data, plus a harmonics time-series for the Normalized Difference Vegetation Index (NDVI);
- 6 Context features, capturing the spatial landscape information from the habitat derived from optical (Sentinel-2) earth observation data through an Artificial Intelligence (Conventional Neural Network or CNN) model.

The **Feature Selection and Training Module** uses reference or training data (field-validated point locations with habitat classification) to train models per EUNIS level on the extracted features (see previous section) associated with the geolocations of the training data. For each hierarchical level, the best features (to distinguish the habitats) were selected through a Machine Learning approach using a five-folded Catboost model. This model is an improved version of the well-known Random Forest model and provides a gradient boosting framework (Dorogush et al. 2017) to solve categorical features using a permutation driven alternative compared to classical algorithms. Every hierarchical level is represented by a specific trained model and the best selected features (Figure 10).

The **Inference Module** performs the hierarchical habitat classification by applying the models and will generate stacked raster files containing probability maps (pixel values between 0 and 100%) per habitat class distinguished by the model. The module then selects per pixel the habitat class with the highest probability to create wall-2-wall maps with discrete pixel-wise classes at 10 m spatial resolution. s

Finally, the **Post-processing Module** revises the discrete pixel classification in the Inference Module by applying several decision rules considering the probability maps and other datasets (such as existing land cover maps, CLMS High Resolution Layer (HRL) maps and habitat suitability maps). The output is again a discrete wall-2-wall habitat map with possible pixel reclassifications and an accompanying quality layer that specifies the final decision rule that was applied and how much confidence is attributed to that decision rule.

More information on the workflow can be found in EuropaBON Deliverable 5.2 (Past-to-present EBV modelled datasets and status indicator for selected terrestrial habitats in the Habitats Directive), available through https://riojournal.com/topical_collection/145/ (Bruelheide et al. 2024). A full list of habitat classes used for the TS in this task and mentioned in this report can be found in section 13.4 of the Annex. As explained above, the workflow is based on Artificial Intelligence, hence the training of the models is an, if not utmost, important aspect to generate high accurate Habitat maps. Therefore, special attention was given to the sampling of the training reference data, as described further ahead.

7.1.4. Mapping EU extent

The habitat maps, as explained in the previous Chapter, were thereafter merged with other data sources either to include more non-natural habitats or to confirm the quality of the habitat maps. During this process (Figure 11), the transformation to the EU Extent Typology (ETA) was done. Priority was given to ancillary layers for non-natural classes (settlements and other artificial areas, cropland, etc.), while priority was given to EUNIS habitat maps for natural classes (grassland, forest and woodland, etc.).

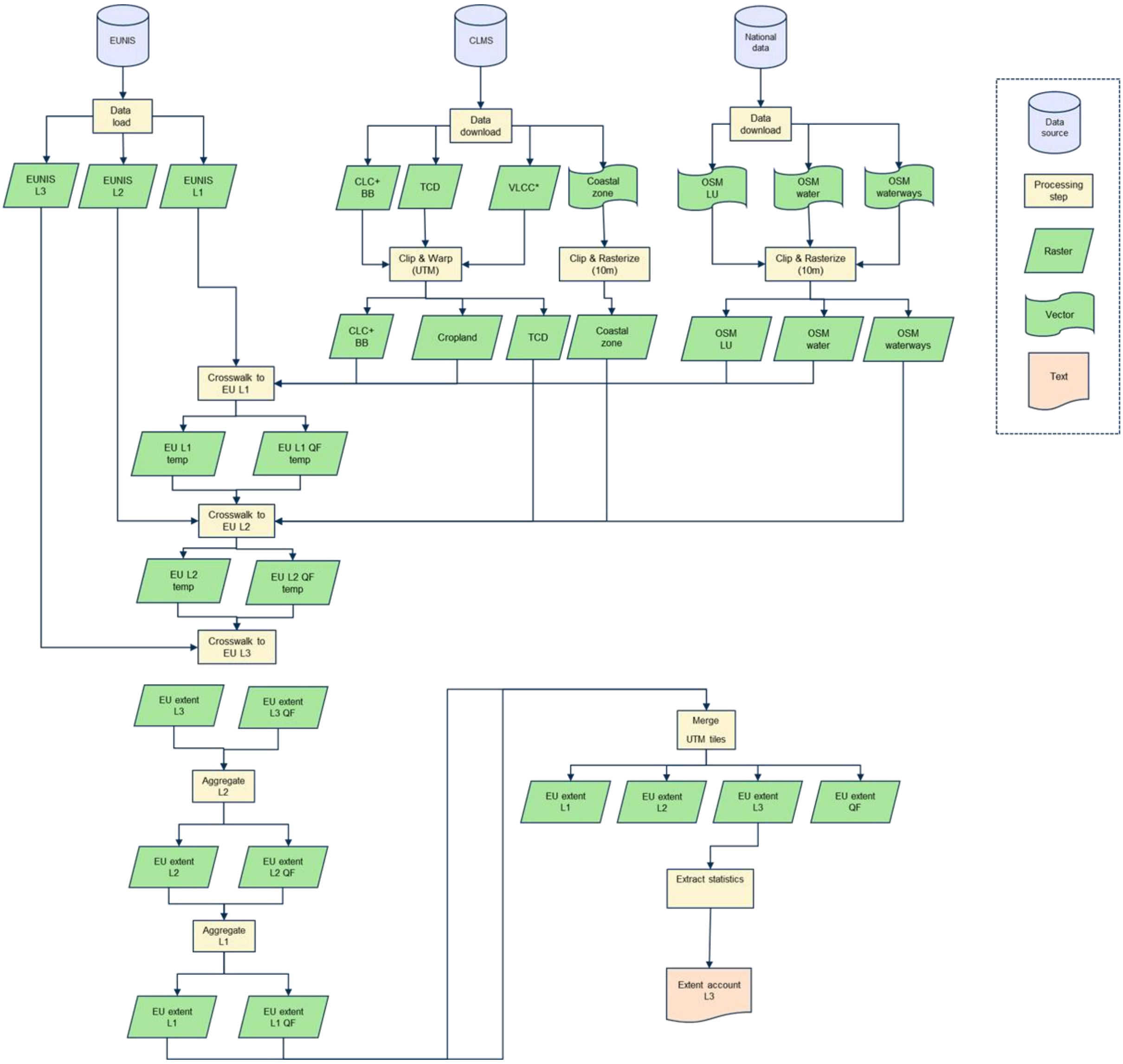


Figure 11: Overview workflow of the transformation to the EU Extent Typology.

The crosswalking rules are described hereafter. The first step was mapping at ETA level 1, which is closely related to land surface characteristics. Therefore, rules were based on using primarily land surface characteristics for non-vegetated classes and habitat maps for vegetated classes.

More details on the rules for level 1 mapping are shown in Table 15, featuring the ecosystem extent class, the three input sources, the extent rule and the final extent map raster value. It is important to notice that the extent rules are executed in a specific order and this order defines also the mutual exclusive class of the extent map. The order of the rules is shown with the number in brackets before the actual extent rule. A potential secondary data source, which is not included in the rule itself, was used for the Quality flag. If both the primary source and the secondary source have the same information, then the level of the quality is raised.

Table 15: EU extent mapping crosswalk Vegetation Centric approach – level 1.

EU ecosystem typology level 1	EU/Global data sources for augmentation	Extent RULE	Raster value
1. Settlements and other artificial areas	CLC+ backbone (sealed=1)	(1) CLC+ backbone is sealed or (13) EUNIS = J	1000
2. Cropland	ESA worldcover (crop=), future HRL_crop	(4) ESA worldcover is cropland	2000
3. Grassland (pastures, semi-natural and natural grasslands)		(7) EUNIS is class R or CLC+ is period grassland or (14) EUNIS=X and CLC+ is grassland	3000
4. Forest and woodlands	CLC+ backbone (woody trees if X)	(5) EUNIS is class T or (16) EUNIS=X and CLC+ is forest	4000
5. Heathland and shrub	CLC+ backbone (woody trees if X)	(8) EUNIS is class S or (15) EUNIS=X and CLC+ is woody plants	5000
6. Sparsely vegetated ecosystems	CLC+ backbone is non or sparse vegetated (9) or lichen and mosses (10)	(11) CLC+ backbone is sparse or lichen/mosses	6000
7. Inland wetlands		(6) EUNIS is class Q	7000
8. Rivers and canals	CLC+ backbone (water = 10)	(3) CLC+ backbone is water or OpenStreetMap_waterways is canal, river or stream	8000
9. Lakes and reservoirs	CLC+ backbone (water = 10)	(2) CLC+ backbone is water	9000
10. Marine inlets and transitional waters		skipped	10000
11. Coastal beaches, dunes and wetlands	Continental coastal zone mask (1km)	(9) EUNIS is class N and within coastal zone or (10) ET=settlement and other artificial areas within coastal zone mask	11000
12. Marine ecosystems		skipped	12000

Some limitations require particular attention:

- CLC+ backbone layer, used as input for land surface characteristics, has no cropland class, therefore the global WorldCover map (ESA) was used. The former layer is from 2018, while the latter is for 2020, which could cause some errors if land is converted from/to cropland within this period.
- CLC+ backbone layer has a single class for water, so the OpenStreetMap layer was used to distinguish rivers & canals (ecosystem type 8) to lakes and reservoirs (ecosystem type 9). The OpenStreetMap layer is currently used to generalize the mapping across full Europe. However, it can be replaced by national layers.
- The EUNIS 2012 typology includes a complex class (X), which combines different vegetation types outside the R, S, and T classes. If we would not crosswalk them, some pixels would remain unclassified at Level 1. Therefore, specific rules were added to map these complex classes using the land cover characteristics (CLC+ backbone layer).
- The EUNIS 2021 typology includes a human man-made class (V), which is typically mapped to cropland. At level 1 this EUNIS input is not required.

After applying the rules at Level 1, a set of rules was defined to decompose every (except artificial) ecosystem extent pixel to Level 2. These rules are shown in Table 16.

Table 16: EU extent mapping crosswalk Vegetation Centric approach – Level 2.

EU ecosystem typology level 2	EU/Global data sources for augmentation	Extent RULE	Raster value
2.1 Annual cropland	CLMS_HRL_TCD (<30%)	CLMS_HRL_TCD <30%	2100
2.2 Rice fields			2200
2.3 Permanent crops	CLMS_HRL_TCD (>70%)	CLMS_HRL_TCD (>70%)	2300
2.4 Agro-forestry areas	??	EUNIS is class V5	2400
2.5 Mixed farmland	CLMS_HRL_TCD (30%<70%)	CLMS_HRL_TCD (30%<70%)	2500
2.6 Other farmland		EUNIS is class VV1	2600
2.7 Hedgerows + cropland tree rows		next version, test future HRL_SWF	2700
3.1 Sown pastures and fields (modified grasslands)		EUNIS is class R2 or (R1 if OSM is meadows)	3100
3.2 Natural and semi-natural grasslands		EUNIS is R3 to R7 or R1 if OSM is not meadows	3200
3.3 Hedgerows + grassland tree rows		next version, test future HRL_SWF	3300
4.1 Broadleaved deciduous forest		EUNIS is class T1 minus (T1H, T1K) minus mixed (FTY)	4100
4.2 Coniferous forests		EUNIS is class T3 minus (T3M,T3N) minus mixed (FTY)	4200
4.3 Broadleaved evergreen forest		EUNIS is class T2 minus (T29, T2A)	4300
4.4 Mixed forests	HRL_FTY	FTY at 1OOm = 3 (mixed)	4400
4.5 Transitional Forest and woodland shrub		EUNIS is V6	4500
4.6 Plantations		EUNIS is one of class (T1K, T29, T2A, T3M, T3N)	4600
5.1 Tundra		EUNIS is S1	5100
5.2 Heathland + (sub) alpine shrubs		EUNIS is S2, S3 or S4	5200
5.3 Sclerophyllous vegetation		EUNIS is S5, S6, S7 or S8	5300
6.1 Bare rocks		EUNIS is H2 or H3	6100
6.2 Semi-desert, desert and other sparsely vegetated areas		EUNIS is H5	6200
6.2 Semi-desert, desert and other sparsely vegetated areas		EUNIS is H5	6200
6.3 Ice sheets, glaciers and perennial snowfields			6300
7.1 Inland marshes on mineral soil		EUNIS is D5 or D6	7100
7.2 Mires, bogs and fens		EUNIS is D1 or D2 or D3 or D4	7200
8.1 Rivers		OSM_waterways is river or stream	8100
8.2 Canals, ditches and drains		OSM_waterways is canal	8200
9.1 Lakes		OSM_water is not reservoir	9100
9.2 Artificial reservoirs		OSM_water is reservoir	9200
9.3 Geothermal pools and wetlands (Iceland)			9300
11.1 Artificial shorelines	CLC+ Backbone (sealed) , Coastal mask	CLC+ backbone is sealed, and ET coastal mask is true	11100
11.2 Coastal dunes, beaches and sandy and muddy shores	Coastal mask	EUNIS is N1 and ET coastal mask is true	11200
11.3 Rocky shores	Coastal mask	EUNIS iS N3 and ET coastal mask is True	11300
11.4 Coastal saltmarshes and salines			11400

Some limitations require particular attention:

- Detecting L2 classes for settlements and other artificial areas was not further explored. To distinguish settlement areas and infrastructure areas, the HRL impervious layer can be used following the protocol as defined in Corine. The urban green space can be detected with the combination of the Copernicus VPP layers.
- Rice fields were not further investigated and mapped, as they were not occurring in our test areas.
- The distinction between agroforestry and other farmland areas currently relies on the EUNIS classification system. However, as agroforestry falls under a man-made vegetation class, it is more difficult to accurately distinguish using Earth Observation (EO) data alone. Therefore, it is recommended to further explore the integration of land use information to improve classification accuracy.
- The detection of small linear features such as hedgerows and tree rows in cropland or grassland was not further tested, however Copernicus provides a small woody feature layer that could be explored to be integrated.
- Transitional forest and woodland shrub were currently derived from EUNIS V6. As with agroforestry, this is a man-made vegetation class and further investigation is required to potentially use other EUNIS classes for clear-felling and/or transitional woodland.
- Plantations were derived from some specific EUNIS T classes but require further validation to check its consistency. Complementing with land-use information could improve its accuracy.
- To distinct rivers from canals/ditches and lakes from artificial reservoirs, the Open Street Map layer was used. This layer can be used for the entire continent, but its accuracy needs to be evaluated. Furthermore, the vector layer was currently rasterized at the 10 m pixel level, while some further selection is required especially for rivers and canals (some are larger while others are smaller). Adding complementary land use (or cadaster) information can improve the results.
- Ice sheets, glaciers and perennial snowfields were not further explored, as they were not occurring in our test areas.
- Marine ecosystems and Marine inlets and transitional waters were not explored.

After applying the rules at Level 2, a set of rules was defined to decompose the forest and woodlands, and coastal beaches dunes and wetlands ecosystem extent to Level 3. These rules are shown in Table 17. Further work is required to define the rules to decompose also other ecosystem types to Level 3.

Table 17: EU extent mapping crosswalk Vegetation Centric approach – level 3 Forest and Woodland and Coastal ecosystems.

EU Ecosystem typology level 3	**Extent RULE**	**Raster value**
Forest and Woodland		
4.1.1 Riparian Forest and woodland	EUNIS is T11 to T14	4101
4.1.2 Broadleaved swamp woodland on non-acid and acid peat	EUNIS is T15 or T16	4102
4.1.3 *Fagus*-dominated forest	EUNIS is T17 or T18	4103
4.1.4 Submediterranean and Mediterranean thermophilous deciduous forest	EUNIS is T19 or T1A	4104

4.1.5 Acidophilous [*Quercus*]- dominated woodland	EUNIS is T1B	4105
4.1.6 Temperate and boreal and Southern European *Betula* and *Populus tremula* forest on mineral soils	EUNIS is T1C or T1C	4106
4.1.7 Other broadleaved deciduous forest, excluding highly modified plantations	EUNIS is T1E to T1K	4107
4.1.8 Highly modified broadleaved deciduous forests including stands of non-native trees species that have long been established in European ecosystems stands		4108
4.2.1 Boreal and temperate fir and spruce forest	if EUNIS old, add envZ to split 4.2.1 and 4.2.2	4201
4.2.2 Mediterranean mountain fir and spruce forest	EUNIS is T32 or T33	4202
4.2.3 Temperate subalpine *Larix*, *Pinus cembra* and *Pinus uncinata* forest	EUNIS is T34	4203
4.2.4 Pine Forest, excluding mires, non-thermophilous	EUNIS is T35, T36, T37, T38 or T39	4204
4.2.5 Mediterranean thermophilous lowland pine forest	EUNIS is T3A	4205
4.2.6 Spruce, pine and larch mire forest	EUNIS is T3K or T3J	4206
4.2.7 Taiga forests	EUNIS is T3F, T3G or T3H	4207
4.2.8 Other coniferous forests, excluding plantations	EUNIS is T38, T3B, T3C or T3D or T3L	4208
4.2.9 Highly modified coniferous forests including stands of non-native trees species that have long been established in European ecosystems stands	EUNIS is T3M or T3N	4209
4.3.1 Mediterranean evergreen *Quercus* Forest	EUNIS is T21	4301
4.3.2 Mainland laurophyllous forest	EUNIS is T22	4302
4.3.3 Macaronesian laurophyllous forest	EUNIS is T23	4303
4.3.4 *Olea europaea-Ceratonia* siliqua forest	EUNIS is T24	4304
4.3.5 Palm groves	EUNIS is T25 or T26	4305
4.3.6 Other broadleaved evergreen forests	EUNIS is T27 or T28	4306
4.3.7 Highly modified broadleaved evergreen forests including stands of non-native trees species that have long been established in European ecosystems stands	EUNIS is T2A	4307
4.4.1 Mixed forests dominated by coniferous species	ET level-2 is coniferous AND FTY is mixed	4401
4.4.2 Mixed forests dominated by broadleaved species	ET level-2 is deciduous AND FTY is mixed	4402
4.4.3 Other mixed forests including stands of non-native trees species that have long been established in European ecosystems stands	ET level-2 is evergreen AND FTY is mixed	4403
4.5.1 Transitional woodland/forest land including recently felled or clear-cut, burnt, replanted or newly afforested		4501
4.6.1 Monoculture plantations of non-native tree species		4601
4.6.2. Mixed plantations of a few species of non-European coniferous and broadleaved trees with underdeveloped undergrowth. Forest stands of single or mixed species consisting of native and/or non-native trees species that have long been established in European ecosystems and have diverse undergrowth typical for forest ecosystems should be classified as part of types 4.1 to 4.4)		4602
Coastal		
11.1.1 Artificial shorelines	same as Level 2	0
11.2.1 Coastal dunes	EUNIS is N14 or N16 or N1B or N1G	11201
11.2.2 Beaches and sandy shores	EUNIS is N12	11202
11.2.3 Muddy shores	EUNIS is N1J	11203

11.3.1 Coastal shingle		11301
11.3.2 Rock cliffs, ledges and shores		11302
11.4.1 Coastal saltmarshes		11401
11.4.2 Salines		11402

Some limitations require particular attention:

- EUNIS (level 3) maps provide a very high thematic detail and can be used easily to be crosswalked to the EU typology. This was expected since the EU typology is derived from EUNIS at L3. Furthermore, several EUNIS classes need to be grouped for this mapping, hence the EUNIS maps have an additional value to provide more details compared to the EU extent maps and hence can be positioned as complementary. Countries could use this to further split some L3 classes according to their needs, which is still in line with the EU typology.
- The highly modified deciduous forest class was not detected, as no training data was available for this class. Note the highly modified coniferous and broadleaved evergreen were classified, but further investigation is required in their accuracy.
- Mixed forests class is currently based on the Copernicus HRL Forest Type (FTY) dataset. To identify mixed forest types, it is suggested to add a 'moving box' filer approach in the workflow instead of using the 100m Copernicus HRL Forest Type (FTY) dataset.
- As explained in L2 mapping, the mapping of plantations was unreliable, hence we did not further distinct plantations into L3.
- The distinction of coastal L3 classes is highly dependent (like other ecosystem types) on the availability of EUNIS training data. It was only possible to distinguish coastal dunes, beaches and sandy and muddy shores. Special care should be taken to gather also training data for rocky shores and coastal salt marshes and salines.

7.1.5. EUNIS habitat mapping for Peloponnese

7.1.5.1. Training data preparation

The MAES map was used to retrieve the area distribution of EUNIS2021 Level 1 habitat classes over Peloponnese. The MAES map (Kokkoris et al. 2020) has its own LIFE-IP classification typology (Figure 12), which was translated into EUNIS2021 Level 1 classes. Since the current EUNIS2021 habitat typology does not cover urban or industrial areas and inland surface water, classes ‘J’ and ‘C’ were added from the EUNIS2012 habitat typology to provide a wall-to-wall habitat mapping. A proportional area-wise desired sample size was calculated per EUNIS2021 L1 habitat class, aimed at a total amount of 20.000 sample points (Table 18).

Table 18: Specified desired sample size for each EUNIS level 1 habitat class, derived by an area-wise proportional analysis how this class occupies the MAES LIFE-IP map in Peloponnese.

EUNIS2021 Level 1 habitat class	MAES - LIFE-IP classes contained in EUNIS class	Desired sample size
J	1, 2, 23, 24	551
C	22	111
M	21	/
N	16	50
Q	18, 19, 20	50
R	12	479
S	13, 14	8803
T	5, 6, 7, 8, 9, 10, 11, 31	3275
U	15, 17	230
V	3, 4, 30	6538

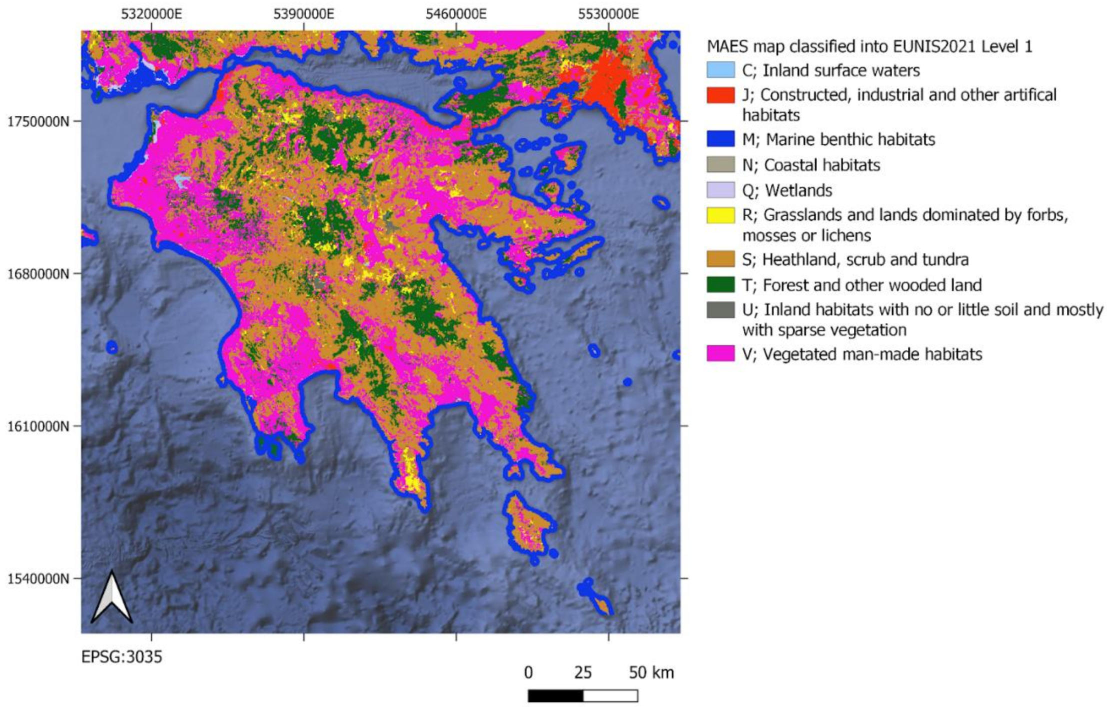


Figure 12: MAES LIFE-IP map classified to EUNIS level 1 habitat types (Kokkoris et al. 2020).

The four datasets, described in Table 7, were combined and counts per unique EUNIS2021 Level 3 habitat class were extracted. Then all Level 3 classes for the same Level 1 class were summed up and the fraction of each Level 3 habitat class in its associated Level 1 class was calculated. Then, this fraction was multiplied with the desired sample size per EUNIS2021 Level 1 habitat class. As a result, the proportional area-wise amount of sample points for each EUNIS2021 Level 3 habitat class was found. This amount was raised to a minimum amount of 20 points if the resulting amount for the Level 3 class was under 20. Next, a random sample point selection was run to extract for each Level 3 habitat class the desired number of points, considering a minimum distance of 100 meters between all selected points (not only between points within one habitat class but also between points of different habitat classes). For Level 1 classes R, S and V not enough training points were selected, due to the criterion of minimum distance between the points. For classes R and S, this was not a substantial amount, so it was decided to neglect the lack of training points. However, for class V still many points were missing to create a representative area sample for the region. Also, the four datasets did not contain data on classes J and C. Therefore, the remaining necessary points for classes J, C and V were extracted from the MAES map, considering a minimum distance of 100 meters between all selected points (also considering the previously selected training points).

The combined set of all selected training points by this described method resulted in a total set of 17.400 training points. The originally set desired total number of training points (i.e., 20.000 points) was not reached, due to the minimum distance criteria. Annex 13.5.1 shows how many points per habitat class were selected, as also if the training points of a particular class were used in the hierarchical habitat mapping at Level 1, 2 or 3. For habitat mapping at Level 1, all training points are evidently used. At Level 2, training points for classes J, C and V are left out since they were extracted from the MAES map and were mapped up to Level 1. Furthermore, we did not have the intention to map classes J and C further than Level 1. For all V-classes at Level 3, the training points were removed from habitat mapping at Level 3 since there was no intention to map V-classes any further than Level 2.

The sets of training data at level 1, level 2 and level 3 were used to train the models to classify the study area of Peloponnese for each of the EUNIS categories in Annex 13.5.1. The inference module applies to the models on the full study area and creates the probability maps per class trained in the model, for each pixel. In the first attempt to create a wall-to-wall habitat map, the class with the highest probability was chosen for each pixel. How a more advanced post-processing was executed for the EUNIS level 1 habitat map is described in Section 7.1.5.2. Afterwards, the level 2 classified maps were merged in the level 1 classified map, and the level 3 classified maps were merged in the resulting level 2 map. Note that the output of the inference (i.e., the pixel-wise probability layers for each class that was trained by the model) allow many other decision rules to be applied in potential post-processing steps (for also level 2 and level 3), in case detailed reference/comparison maps are available.

7.1.5.2. Post-processing steps

The post-processing steps for Peloponnese were kept limited to the EUNIS level 1 habitat map. We used the CLMS CLC+ Backbone map (2018) to compare pixel-wise the predicted EUNIS habitat classification at level 1 with the classification by CLMS (EEA, 2023). Table 19 demonstrates the translation of EUNIS habitats at level 1 we made to the classes by CLC+.

Table 19: Translation of EUNIS habitats at level 1 to classes in CLC+ Backbone map.

EUNIS level 1 habitat	Linked CLC+ Backbone classes
C	• 10: Water
J	• 1: Sealed

N	• 9: Non- and sparsely vegetated
Q	• 5: Low-growing woody plants (bushes, shrubs) • 6: Permanent herbaceous • 7: Periodically herbaceous • 8: Lichens and mosses • 9: Non and sparsely vegetated • 10: Water
R	• 6: Permanent herbaceous • 7: Periodically herbaceous • 8: Lichens and mosses
S	• 5: Low-growing woody plants (bushes, shrubs)
T	• 2: Woody - needle leaved trees • 3: Woody - broadleaved deciduous trees • 4: Woody - broadleaved evergreen trees
U	• 9: Non- and sparsely vegetated
V	• 5: Low-growing woody plants (bushes, shrubs) • 6: Permanent herbaceous • 7: Periodically herbaceous • 9: Non- and sparsely vegetated

The post-processing followed a set-up decision rule. First, it was checked if the original pixel classification (the habitat class with the highest probability) matches one of the linked categories of CLCplus. If there is a match, this original winning habitat classification was maintained. If not, we checked the habitat classification of the second winning class (habitat with second highest probability). If this second winning class has a match with one of the linked CLCplus classes, an additional check was carried out. Only if this second winning class has over 40% probability and a difference from the first winning class smaller than 10%, we concluded that the model had significant confusion in mapping this pixel and the probability is high that the second winning class might be also very applicable. If the criteria were met, we assigned the second winning class as the new habitat classification for that pixel. If not, we kept the original classification. In the case that both the first and second winning classes did not match one of the linked CLCplus classes, we retained the original classification, since we lack enough information to ensure a reclassification with high confidence.

7.1.5.3. Level 1 EUNIS Habitat map

For model training at level 1, 55 features were selected. The top five most important features were: 1) DEM-slo-20m, 2) contextft_3 (Sentinel-2 data), 3) gdd5, 4) DEM-alt-20m 5) bio12 (see feature list in section 13.3 of the Annex, which explains the abbreviations that are used here and further onwards). The normalized confusion matrix for the Level 1 habitat modeler shows that the accuracy of mapping for all classes was very good, though the accuracy of mapping peatlands and sparse vegetation is lower (Figure 13). Classes Q and U were represented in small amounts in the training data; however, peatlands did not occur largely in Peloponnese and adequate training data on class U was limited. Overall, the modeler's accuracy of EUNIS2021 Level 1 habitat mapping was 83,27 %.

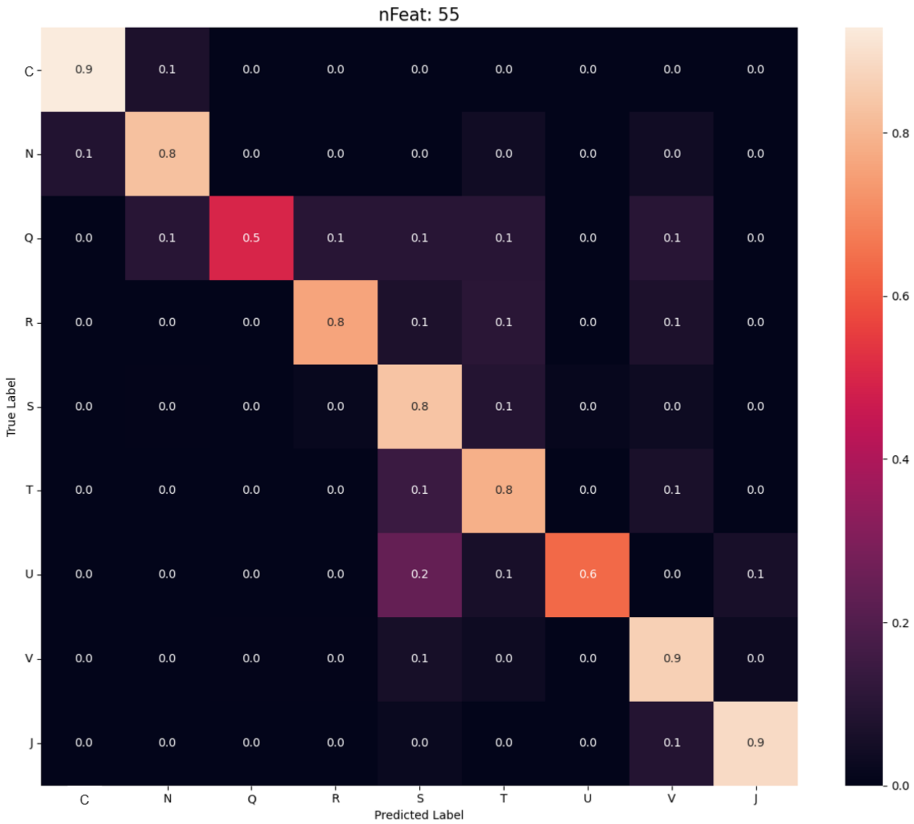


Figure 13: Normalized confusion matrix of winning Catboost model to map habitats at EUNIS level 1, for Peloponnese. The model accuracy contains 83.27%.

Figure 14 presents the wall-to-wall habitat mapping for the Peloponnese region, classified at Level 1 EUNIS 2021. According to the habitat map, the Peloponnese — with a total mapped area of 2,234.561 hectares — is predominantly covered by agricultural land (Figure 15). Embedded within this agricultural matrix are large patches of forest and shrubland, contributing to the region's ecological diversity.

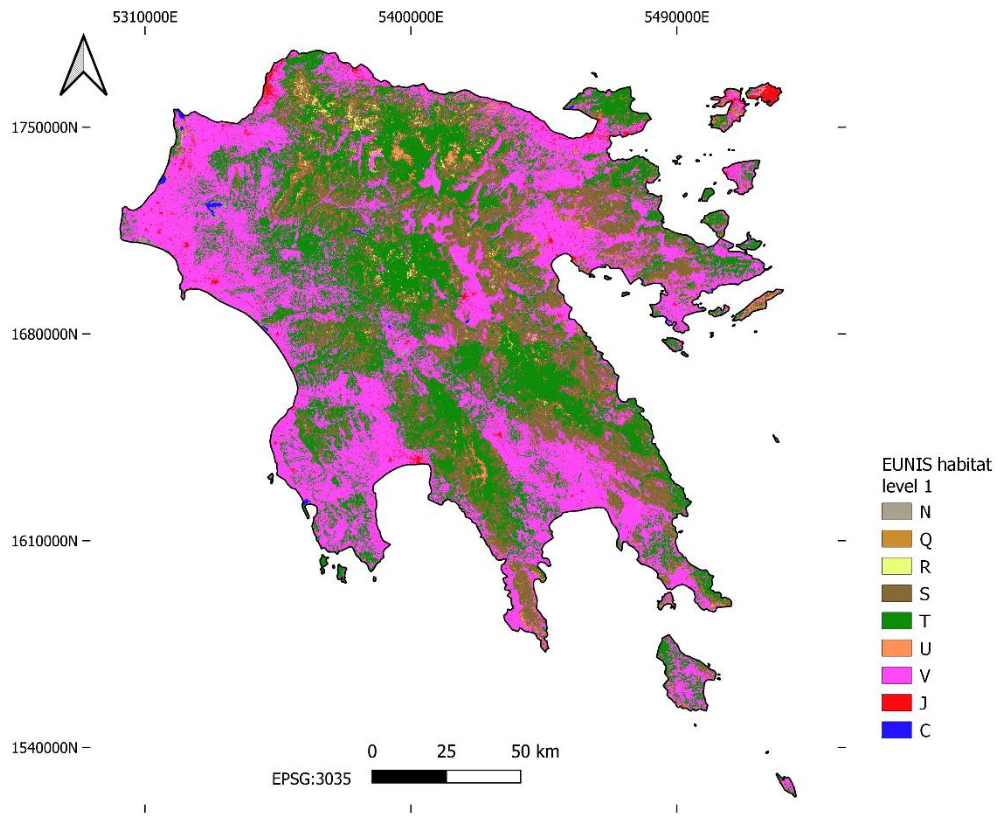


Figure 14: Predicted EUNIS level 1 habitat map of Peloponnese. See Annex 9.4 to explain the habitat codes in the legend.

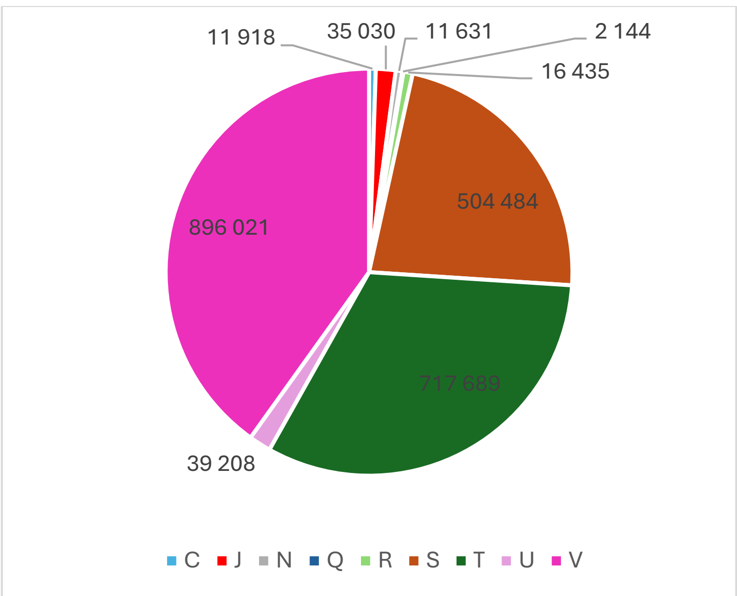


Figure 15: Area distribution [ha] of EUNIS level 1 habitat classes mapped within Peloponnese. See Annex 9.4 to explain the habitat codes in the legend.

7.1.5.4. Level 2 EUNIS Habitat Map

At level 2, the model accuracies exhibited very good results (Table 20). We zoom in to the normalized confusion matrix of the model for class T that shows that class T1 (Broadleaved deciduous forest) had highest chances to be mistaken for other level 2 forest types T2 (Broadleaved evergreen forest) and T3 (Coniferous Forest) by the model (Figure 16).

Table 20: Model accuracies of winning Catboost models generated to map EUNIS habitat level 2 classes within overarching EUNIS habitat level 1 class

Overarching EUNIS habitat level 1 class	Model accuracy [%] to map the associated L2 classes
N	95,83
Q	90,91
R	100,00
S	94,48
T	90,60
U	93,94
V	94,74

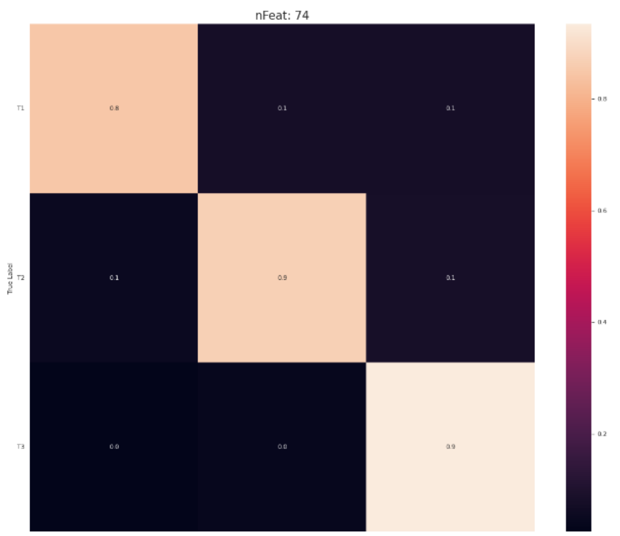


Figure 16: Normalized confusion matrix for mapping EUNIS level 2 habitat classes within EUNIS level 1 habitat T. The model accuracy contains 90,60%.

Table 21 contains an overview of the top 10 most important features with predictive power in the model to differentiate the level 2 classes within the associated level 1 class. The topographic features 'slope' and 'altitude' were important for each model. Besides, at level 2 the Remote Sensing derived features (e.g. B06-p90-20m, anir-iqr-20m, …) played a bigger role compared to the level 1 model. Also, the contextual features are identified to take up relevance in classifying coastal habitats.

Table 21: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 1 classes. The list displays the features from top to bottom in order of importance with highest importance at the top.

Class-R	Class-N	Class-T
gdd5	contextft_8	DEM-alt-20m
DEM-alt-20m	DEM-slo-20m	gsp
dist	B06-p90-20m	veg-height
bio12	contextft_14	gdd5
DEM-slo-20m	contextft_10	anir-iqr-20m
gst	gst	dist
pop18	nbr2-p10-20m	vppgup
lai01	contextft_11	DEM-slo-20m
vh_vv-iqr-20m	VV-p50-20m	VV-iqr-20m
ndre3-p10-20m	contextft_6	B11-p10-20m

Figure 17 illustrates the output of the habitat mapping of EUNIS habitat classes at level 2. Looking at the natural vegetation classes, classes S6 (Garrigue), S7 (Spiny Mediterranean heaths (phrygana, hedgehog-heaths and related coastal cliff vegetation), T1 (Broadleaved deciduous forest), T2 (Broadleaved evergreen forest) and T3 (Coniferous Forest) occupied most of the area in the peninsula (Figure 18). The Q-classes (for wetlands and peatland) were absent and also the grassland habitats were rather rare.

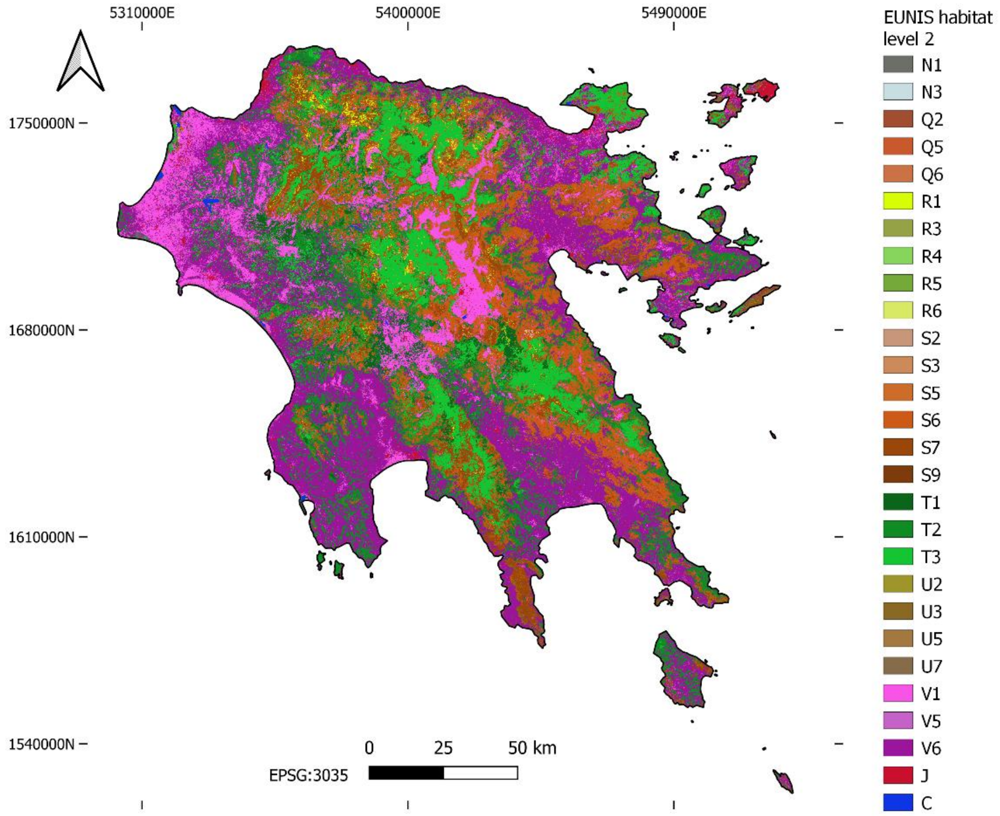


Figure 17: Predicted EUNIS level 2 habitat map of Peloponnese. See Annex 13.4 for the legends of the habitat codes.

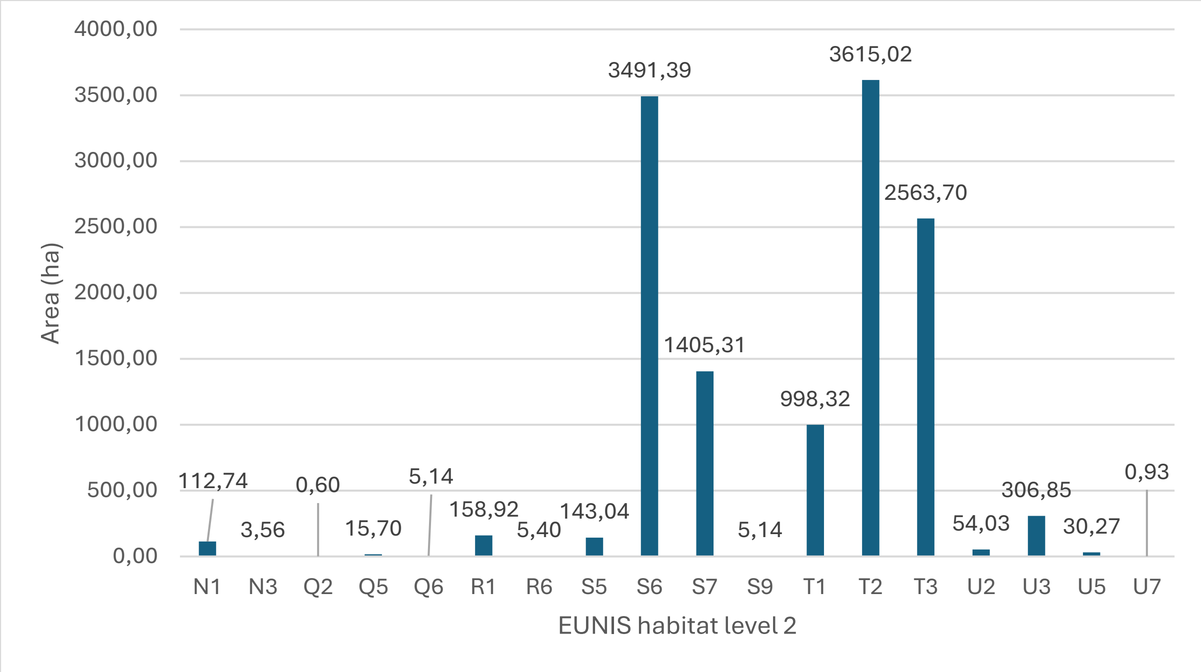


Figure 18: Area distribution [ha] of EUNIS level 2 habitat classes mapped within Peloponnese.

7.1.5.5. Level 3 EUNIS Habitat Map

At level 3, the model accuracies had great results (Table 22). We zoom in to the normalized confusion matrix of the model for class N1 showing that classes N16 (Mediterranean and Macaronesian coastal dune grassland (grey dune), N1G (Mediterranean coniferous coastal dune forest) and N1J (Mediterranean and Black Sea moist and wet dune slack) can be misclassified to N1B (Mediterranean and Black Sea coastal dune scrub) (Figure 19).

Table 22: Model accuracies of winning Catboost models generated to map EUNIS habitat level 3 classes within overarching EUNIS habitat level 2 class

Overarching EUNIS habitat level 2 class	Model accuracy [%] to map the associated level 3 classes
N1	81,82
R1	100,00
S5	98,88
S7	98,64
T1	96,92
T2	91,23
T3	86,55

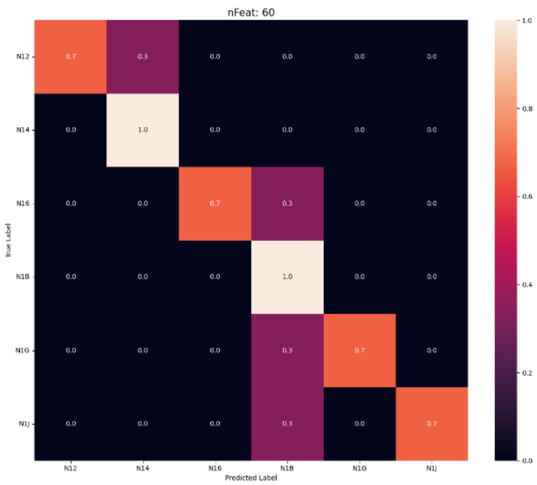


Figure 19: Normalized confusion matrix for mapping EUNIS level habitat classes within EUNIS level 2 habitat N1. The model accuracy contains 97.63%

Table 23 contains an overview of the top 10 most important features with predictive power in the model to differentiate the level 3 classes within the associated level 2 class. Compared to the more predictive features in the level 2 models, more Remote Sensing derived products come up in the top 10 most important features for each level 3 model. For forest classes at level 3 within T2 (Broadleaved evergreen forest), the biophysical parameters maintained their important role. For coastal habitats within N1 (Coastal dunes and sandy shores), Sentinel-1 backscatter parameters (e.g. VV-p10-20m) exhibit importance.

Table 23: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 2 classes. The list displays the features from top to bottom in order of importance with highest importance at the top.

Class-N1	Class-S5	Class-T2
DEM-alt-20m	DEM-alt-20m	dist
VV-iqr-20m	soc	gdd5
B02-p90-10m	gdd5	gst
gst	phh20	rep-iqr-20m
rvi-p10-20m	gst	gsp
VH-p50-20m	contextft_8	pop18
ndwi-p50-10m	evi-p10-10m	lai01
VV-p10-20m	ndvi-iqr-20m	bio12
ndre2-p50-20m	nbr-iqr-20m	nirv-iqr-10m
contextft_2	contextft_3	contextft_16

Figure 20 exhibits the final habitat map of Peloponnese at EUNIS level 3, with further zoom-ins showcasing the level of detail mapped in coastal areas (Figure 21). A visual comparison with a Google Satellite image of the same location was performed that showed good agreement between the modeled habitat map and the satellite image. At level 3, the EUNIS habitat classes S62 (Western acidophilous garrigue), T21 (Mediterranean evergreen *Quercus* Forest), T24 (*Olea europaea-Ceratonia siliqua* forest), T2B (Mediterranean evergreen Quercus Forest in Greece) and T3P (Mediterranean mountain *Abies* forest in Greece) are the most dominant classes (table 22). Where possible, the EUNIS habitats were matched to the IUCN Red list of habitats status and discovered that Peloponnese contains a set of habitat types with 'Endangered' and 'Vulnerable' status (such as N14, N16, N1B, R61, S54 and T14) (Table 24). Linking habitat types to IUCN Red List status is a significant outcome, providing a bridge between mapping outputs and biodiversity policy/action. The proportional coverage was derived by dividing the area of the natural habitat types at level 3 by 22345,57 km^2 (i.e., the total area of the Peloponnese peninsula for which habitat mapping was carried out).

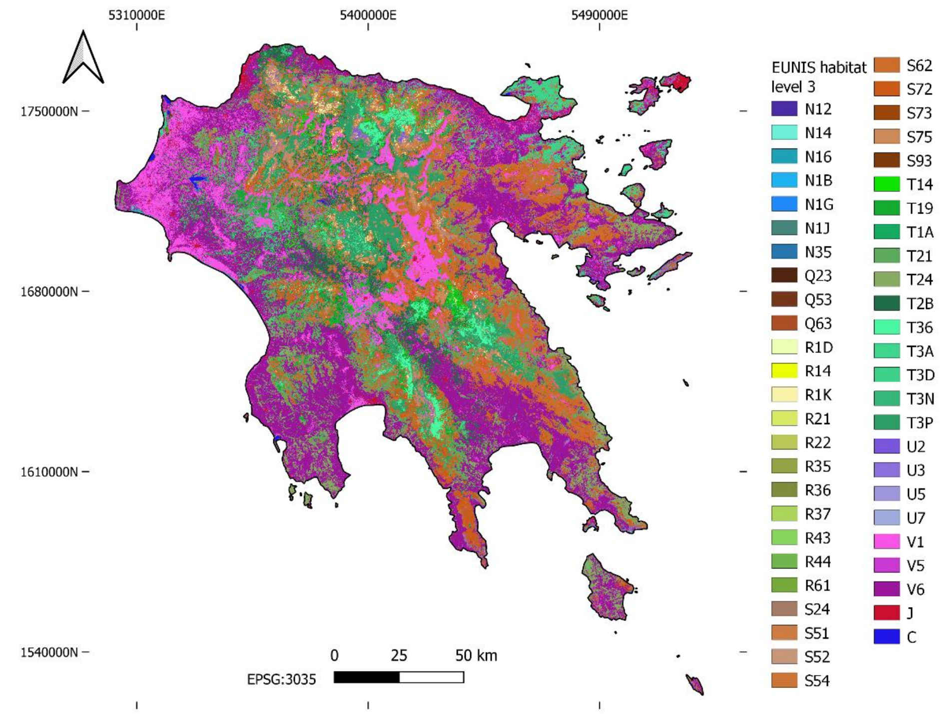


Figure 20: Predicted EUNIS level 3 habitat map of Peloponnese. See Annex 13.4 for the legends of the habitat codes.

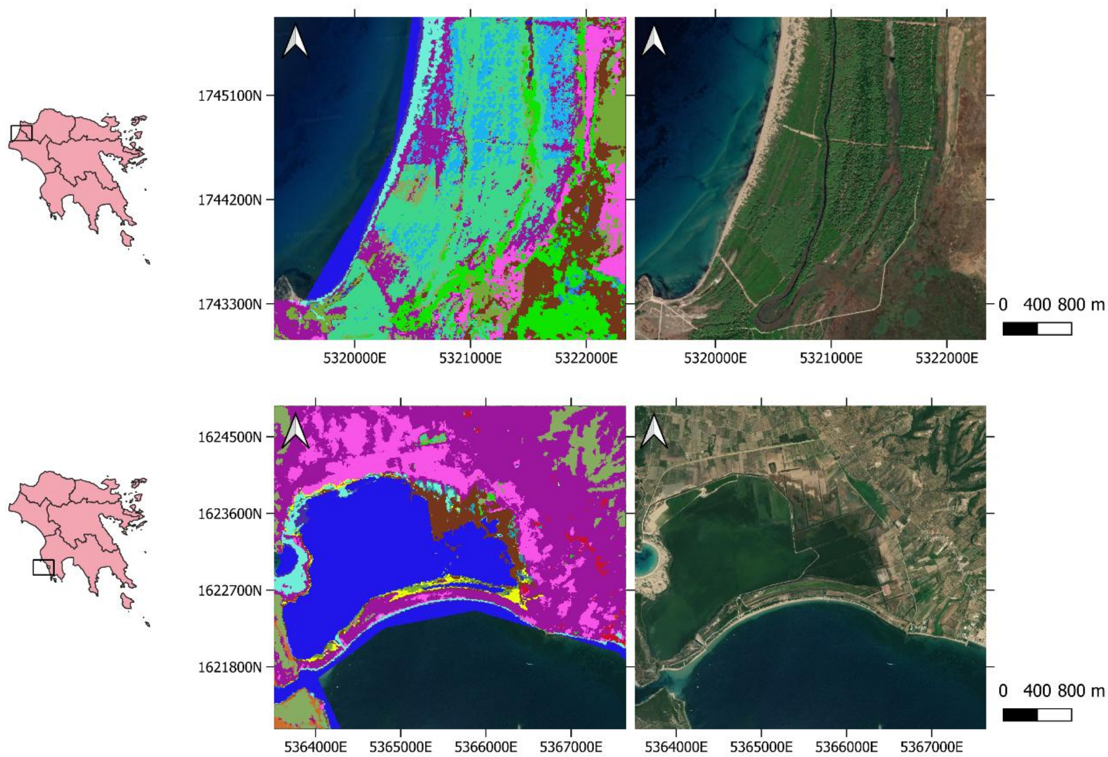


Figure 21: Zoom-in on area in coastal areas of Peloponnese to check visually the correspondence of the predicted EUNIS level 3 habitat map with a Google Satellite image. See Figure 20 to link the color in the habitat map to a EUNIS habitat class.

Table 24: Area distribution [ha] of EUNIS level 3 habitat classes mapped within Peloponnese.

EUNIS L3 habitat	Area [km²] / proportion	IUCN Red list of Habitats status	EUNIS L3 habitat	Area [km²] / proportion	IUCN Red list of Habitats status
N12	19,87 / 0,09%	Near threatened	S54	3,56 / 0,02%	Vulnerable
N14	38,08 / 0,17%	Vulnerable	S62	3491,73 / 15,63%	Least concern
N16	19,98 / 0,09%	Endangered	S72	858,15 / 3,84%	Least concern
N1B	4,81 / 0,02%	Vulnerable	S75	547,27 / 2,45%	Least concern
N1G	9,96 / 0,04%	Least concern	S93	5,12 / 0,02%	Least concern
N1J	20,08 / 0,09%	Least concern	T14	429,17 / 1,92%	Vulnerable
N35	3,56 / 0,02 %	Data deficient	T19	561,66 / 2,51%	Least concern
Q23	0,60 / 0,00%	/	T1A	7,12 / 0,03%	Least concern
Q53	15,69 / 0,07%	/	T21	1229,84 / 5,50%	Least concern
Q63	5,11 / 0,02%	/	T24	1368,38 / 6,12%	Least concern
R1D	128,05 / 0,57%	Least concern	T2B	1016,76 / 4,55%	/
R1E	4,25 / 0,02%	Least concern	T36	428,42 / 1,92%	Least concern
R1K	26,65 / 0,12%	Least concern	T3A	671,13 / 3,00%	Least concern
R61	5,40 / 0,02%	Vulnerable	T3D	16,92 / 0,08%	Least concern
S51	138,29 / 0,62%	Least concern	T3N	27,97 / 0,13%	/
S52	1,31 / 0,01%	Least concern	T3P	1419,34 / 6,35%	/

7.1.5.6. Validation

The validation method followed the classic photointerpretation scheme focusing on well-known and studied natural, semi-natural, and anthropogenic ecosystems (cultivations, settlements, infrastructure, etc.).

The selected areas (Figure 22) for evaluating outcomes included sites where:

- recent ecosystem type identification was made via UPATRAS students' PhD theses, field work and analysis.
- valid field data from well-studied areas via recent and ongoing botanical explorations in the region of Peloponnese are available.

Field surveys dedicated to the project task were also conducted, focusing on less studied areas, based on the transect method, where ecosystem documentation was made via *in situ* observation. The selection of the evaluation sites covers all types of ecosystems present in the region, from coastal to high-altitude habitats.

The information from the field identification of habitat classes was compared with the results from the SELINA habitat map and extent map for Peloponnese. This evaluation relied on a qualitative approach that explored how well the EO habitat models and crosswalk to EU extent typology captured real-life conditions (in this case, how well extent mapping was performed). The goal was not only to have an overview of the mapping performance, but also to focus on areas where known issues may be present (e.g. screes, rocky cliffs, temporary ponds and wetland habitats whose phenological characteristics depend on hydrological conditions, plus grassland types that were mapped in the "grassland – unclassified" category).

Main findings showed that the classes inland marshes and wetlands, sparsely vegetated ecosystems, beaches and sandy dunes, and coastal dunes are well captured and mapped by the EO model. Coniferous forests (fir and pine forests) are also well captured and mapped, as well as sclerophyllous vegetation and riparian forests.

However, rocky cliffs are considered underestimated, and different habitat types of grasslands are mapped in the general “Grassland (non-classified)” class (see e.g. Figure 23, Figure 24, and Figure 25). In general, this evaluation confirmed the results of the accuracy assessment (see Sections 7.1.5.3, 7.1.5.4, 7.1.5.5), although done in a qualitative way.

Furthermore, the model accuracies used to map habitat classes at levels 1, 2, and 3 probably do not represent reliable map accuracies. Underestimation of rocky cliffs and lack of detail in grassland mapping are still limitations with origins in both the habitat training dataset and the derived habitat models. Therefore, the comparison between the habitat model accuracies and the well-known reference areas reveals some overlooked biases.

Additional analyses with an external validation dataset could have provided more reliable information about the actual map accuracies, which could have been useful to link to the validation based on the photointerpretation scheme. However, the training data selection for habitat mapping left only a small number of points per habitat class that could have been used in an external validation. It was thus concluded that this set of points was too limited to generate reliable map accuracy, as it could be biased towards the locations of the remaining points.

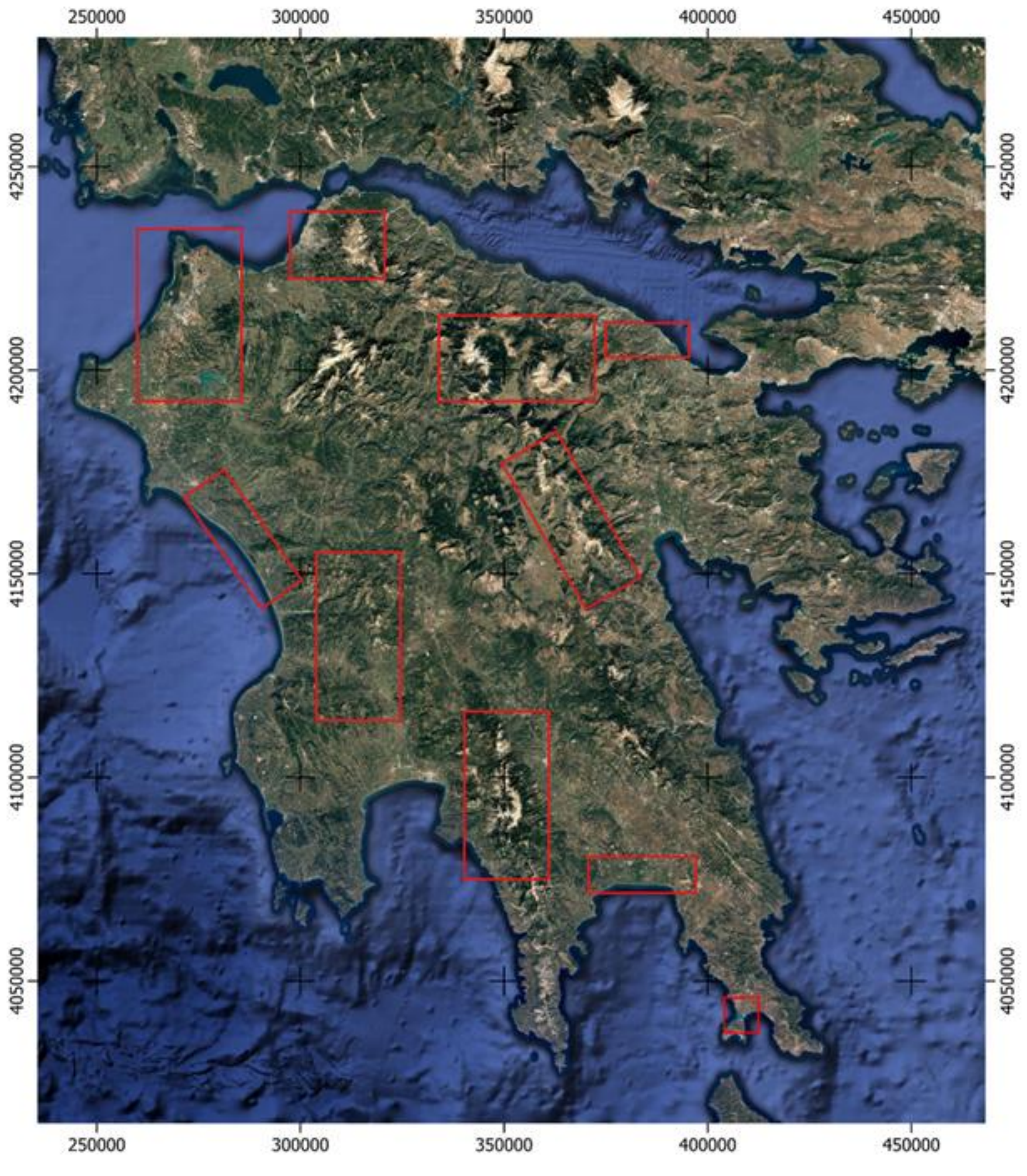


Figure 22: Photointerpretation validation areas in Peloponnese (red rectangles).

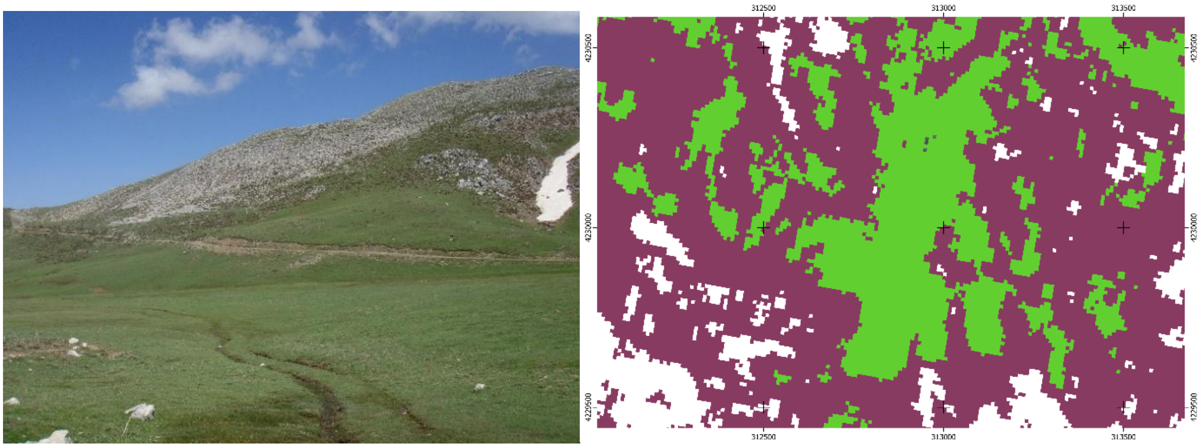


Figure 23: The most extensive area of the priority habitat type "Species-rich Nardus grasslands, on silicious substrates in mountain areas (and submountain areas in Continental Europe)", at the northern part of the TS (Mt. Panachaiko). This area has been included in the "Grassland (non-classified)" class (see green polygons in the ecosystem type map on the right).

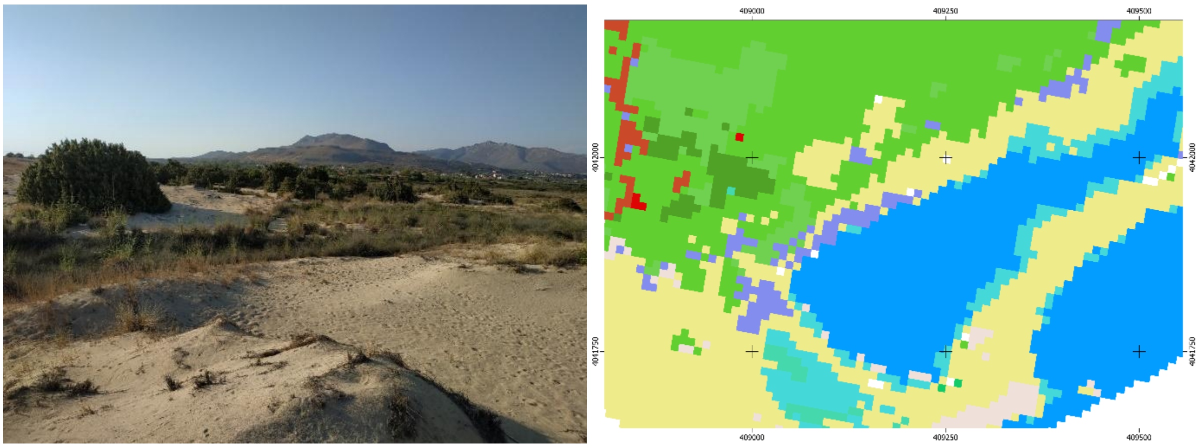


Figure 24: Inland marshes and wetland (purple), sparsely vegetated ecosystems (pink), beaches and sandy dunes (cyan) and coastal dunes (yellow) in the southernmost part of Peloponnese, well captured and mapped by the EO model (see map on the right).

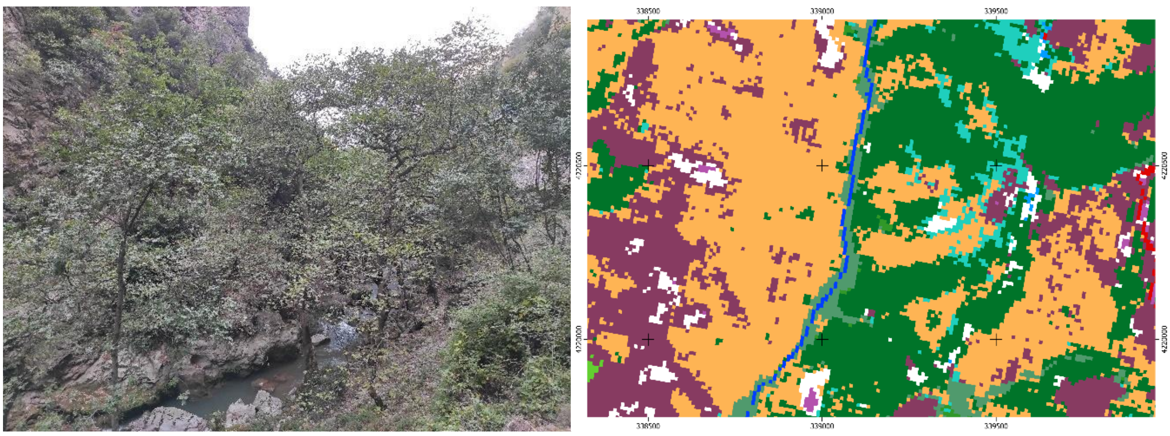


Figure 25: Ecosystems in the National Park of Chelmos-Vouraikos (inside the river gorge). Sclerophyllous vegetation, riparian forests and coniferous forests are well captured, however rocky cliffs are underestimated.

7.1.6. EUNIS Habitat Mapping for São Miguel

7.1.6.1. Training data preparation

The generated extensive training dataset prepared by local experts from São Miguel TS partner contained 167.688 points in total. Table 25 shows the total initial distribution of training points over all the EUNIS classes. To filter these points according to a proportional area-wise desired sample size per habitat class, the 'Coastal Zones Land Cover/Land Use 2018 (vector), Europe, 6-yearly' map of CLMS (Copernicus Land Monitoring Service) was used. This vector map contains delineated polygons classified into a EUNIS habitat at level 1, however the level 1 class N (Coastal habitats) is absent and rather classified into class M (Marine habitats). Therefore, the polygons around the coastline of São Miguel were manually reclassified into class N. According to the vector map of CLMS, level 1 class 'Q' (i.e., peatlands) is not present on São Miguel. However, local expert knowledge informed that some habitats can be classified into level 3 Q classes. Therefore, the desired sample size for class Q was set to 400. The total area of São Miguel without the Marine habitats contains 74.050,51 ha. A desired sample size per level 1 class (besides class Q) was estimated by deriving the fraction of each level 1 class on the island of São Miguel and then multiplying that fraction with the desired sample size of the total training dataset. The desired sample size of the total training dataset was set to 50.000 points, as this number was best suitable for this TS based on the target to contain abundant training points per habitat class but considering the island's area, the minimal distance between points and the computational constraints.

Table 25: Area distribution of EUNIS level 1 habitat classes mapped in the CLMS Coastal Zone map. Fraction per class was calculated by dividing the area per class by the total area on the island. The desired sample size was generated by multiplying the fraction by 50.000 (i.e., the desired total number of training points).

Class name	Total area of class [ha]	Fraction in island	Desired sample size
Urban - J	6120,09	0,0826	4132
Cropland - V	32211,67	0,4350	21750
Forest - T	20150,61	0,2721	13606
Grassland - R	8593,70	0,1161	5803
Shrubland - S	5051,04	0,0682	3411
Sparse_vegetation - U	471,24	0,0064	318
Water - C	872,38	0,0118	589
Coastal - N	580,12	0,0078	392

Like the methodology for Peloponnese, all level 3 classes for the same level 1 class were summed up and the fraction of each level 3 habitat class in its associated level 1 class was calculated. Next, this fraction was multiplied with the desired sample size per EUNIS2021 level 1 habitat class, hence a proportional area-wise amount of sample points for each EUNIS2021 level 3 class was found. This amount was raised to a minimum of 20 points if the resulting amount for the level 3 class was under 20. A random sample point selection was executed to be extracted for each level 3 class the desired sample size, considering a minimum distance of 50 meters between all selected points. During the point selection per class, points were selected if their dataset origin matched with the priority source dataset listed in Annex 13.5.2. The points for class R21 originated from either FI2024 or TerceiraLandUse as the latter was used to increase the total number of points for this class, which are considered equally trustworthy. Therefore, no priority data source was assigned to the point selection for class R21. For class C16, only 15 points were selected despite a desired sample size of a minimum of 20 points. This is most likely due to the minimum distance criterion that could not be achieved for this class. The final number of selected training points was 52.817 points.

Annex 13.5.2 also shows whether the training points for a specific level 3 class were used in the hierarchical habitat mapping at level 1, 2 or 3. At level 1, all training points were used. At level 2, there is only one ‘S’ class, so there was no need to generate a model to separate multiple S classes on level 2. At level 3, points from 'C23', 'J24', 'J32', 'J53', 'J62', 'N11', 'N21', 'N33', 'Q21', 'Q62', 'R1T', 'R21', 'U3C', 'U62', 'V1', 'V21', 'V31' and 'V52' were excluded since these classes are the only level 2/3 class within their associated level 2 class. The points of classes ‘S4x’, ‘S43’ and ‘T2x’ are an exception here. For ‘S’ there was only one distinguished class on level 3 (i.e. ‘S43’), but the points within ‘S4x’ are not to be classified into ‘S43’, neither can they be categorized into an existing EUNIS2021 level 3 class. Therefore, these points were gathered in a separate class ‘S4x’ representing the ‘non-Macaronesian and non-cultivated (exotic) heath of the Azores’, or simply, all naturally occurring non-S43 heathland, and a model must be trained to separate ‘S4x’ points from ‘S43’.

Class S4x was created as a placeholder to represent non-endemic, exotic heathlands found in São Miguel that could not be accurately classified within the existing EUNIS typology. While EUNIS level-2 class S4 (“Temperate shrub heath”) and level-3 class S43 (“Macaronesian heaths”) describe shrublands dominated by Ericaceae and endemic species, the S4x areas are composed predominantly of exotic or introduced species such as *Pittosporum undulatum*, *Hedychium gardnerianum*, and *Hydrangea macrophylla*. These species are not mentioned as diagnostic or constant in any existing EUNIS classes.

The training points for S4x were derived from the regional forest inventory (IF2024) dataset, which focuses on species presence rather than habitat-level interpretation. Unlike habitat 4050, which corresponds directly to S43 and is represented by highly accurate, field-validated data from the LifeIPNatura project, the S4x points represent a distinct type of heath with no direct Annex I to EUNIS crosswalk. The classification of these areas was inferred based on species composition (shrubs <5 m) and structural characteristics, but with greater uncertainty due to the lack of direct habitat mapping or validation.

Importantly, this floristic composition—dominated by aggressive, introduced species forming stable shrub-dominated systems—has emerged because of human-driven introduction and spread. As such, it may constitute a novel anthropogenic vegetation type unique to the Azores archipelago, not currently represented in existing European typologies.

Furthermore, even the level-2 S4 definition is only partially applicable, as it presumes the prominence of *Ericaceae* species, which are absent or rare in these exotic heaths. Thus, the introduction of class S4x ensures ecological accuracy by avoiding the misclassification of non-endemic shrublands as S43. It also prevents the inappropriate merging of fundamentally different vegetation types, which is crucial particularly for conservation of endemic habitats.

Considering points in ‘T2x’, a similar reasoning can be applied: since they cannot be categorized into an existing EUNIS2021 level 3 class, they are gathered into ‘T2x’, representing the ‘non-Laurophyllous and non-cultivated (exotic) broadleaved evergreen forest of the Azores’, or simply, all naturally occurring non-T23 broadleaved evergreen forest. For areas in which a level 1 class only contained one level 2 class, or a level 2 class contained only one level 3 class, the post-processing step eventually assigned the correct level 2 or 3 category.

The EUNIS class T231 represents the Azorean laurisilva forests, a habitat endemic to the Azores and well documented under habitat code 9360 in the Habitats Directive Annex I. These laurel forests are composed primarily of native and endemic broadleaved evergreen species such as *Laurus azorica*, and their classification into T231 is well supported by field-validated data from the LifeIPNatura dataset, specifically designed to map Annex I habitats and includes detailed ecological descriptions and diagnostic species, allowing a confident crosswalk to EUNIS level-3 class T231.

However, not all broadleaved evergreen forests in São Miguel fall within the definition of endemic laurel forest. There are additional forest areas—mapped using the IF2024 regional forest inventory—that are composed of other broadleaved evergreen species, including both native and introduced taxa, but which do not exhibit the ecological or floristic characteristics required for classification as T231. These include formations lacking the endemic species associated with Azorean laurisilva and may reflect secondary, altered, or novel forest types.

Because no existing EUNIS level-3 class adequately captures the structure and composition of these non-laurisilva evergreen forests, they could not be assigned to T231 or any other subclass. While they broadly align with EUNIS level-2 class T2 (“Broadleaved evergreen woodland”), this classification is too general and does not distinguish between endemic laurel forests and other evergreen forest formations. To address this gap, T2x placeholder class fits these broadleaved evergreen non-endemic woodlands. As with S4x, the floristic composition of T2x may reflect novel or regionally unique assemblages influenced by human introduction of non-native species. This approach ensures that endemic laurisilva habitats (T231) are not conflated with structurally similar but ecologically different forest types, maintaining both the ecological integrity of the classification and its relevance for conservation and policy support.

The sets of training data at levels 1, 2 and 3 were used to train the models to classify the study area of São Miguel for each of the EUNIS categories in Annex 13.5.2. The inference module applied the models on the full study area and created the probability maps per class trained in the model, for each pixel. In the first attempt to create a wall-to-wall habitat map, the class with the highest probability was chosen for each pixel. Section 4.1.6.2 describes how some important post-processing steps were implemented. In contrast to the habitat mapping for Peloponnese, it was necessary to implement post-processing in each level of the habitat mapping pipeline to create a more accurate map. After each post-processing step per level, the modelled map at the following level must be hierarchically merged into the associated prior level (i.e., after post-processing of level 2, the modelled level 3 map (before post-processing level 3) must be merged into the post-processed level 2 map).

7.1.6.2. Post-processing steps

The first version of the habitat mapping for São Miguel was revised in an intermediate step to create final products with higher realistic accuracy. The first version exhibited some recurring and relevant misclassifications that had to be addressed for the creation of the final maps. The most pressing artifacts included:

- Coastal areas were mostly mapped to class J (industrial habitats) instead of N (coastal habitats).
- A large wetland in the west of São Miguel was mapped to C13 (inland surface waters) instead of Q62 (Periodically exposed shore with stable, mesotrophic sediments with pioneer or ephemeral vegetation).
- Class Q12 (Blanket bog) disappeared almost completely from the map despite being covered by about 300 training points, in comparison with classes Q11 (Raised bog) and Q21 (Oceanic valley mire) that are rare classes (table 24) with very low amount of training points in the first place for which it can be expected that the model will not be able to classify those correctly due to lack of training material.
- Many forest areas of T1 (Broadleaved deciduous forest) and T2 (Broadleaved evergreen forest) were often mistaken for T3 (Coniferous Forest) since this forest type is dominating on São Miguel Island.

The most common issues for these misclassifications are due to the following 3 reasons:

1. The random sampling method did not select enough points in a certain area (this was applicable to the misclassification between C13 and Q62, and the disappearance of class Q12 as most of these training points were collected from Terceira Island.
2. The area-wise proportional method to select the amount of training points per class assigned a sample size to some dominating habitat classes that was too high relatively to other underrepresented habitat classes. This happened with T3M (dominating) and some T1/T2 rare classes, since the probability of T3M was almost always much higher.
3. The predicted class (on the resulting habitat map) was represented by a very similar feature set compared to the true class.

The first two issues are technical mistakes that could have been easily prevented with an adjusted sampling strategy that ensures the following:

- Verifying that the selected training points cover the entire study area.
- Assigning a higher sampling size for training points of rare classes while avoiding an overfitting of the classifier through overspilled sample sizes for dominating classes.

Unfortunately, due to time limitations, it was not possible to make a second run of the whole process of training data collection, feature extraction, features selection, model generation and inference. However, the most serious misclassifications described above were addressed with some post-processing steps, which included:

- The wetlands, whose training points were accidentally skipped for sampling on the west, were reclassified to EUNIS class Q at level 1. Then, the probability maps for the Q classes (at levels 2 and 3) were further used to assign the ultimate classification.
- The peatland bogs (Q11, Q12, Q21) that were excluded throughout the mapping due to low amount of training points or training point selection elsewhere on Terceira Island, were imprinted on the level 3 habitat map by prioritizing their classification within delineated polygons (drawn based on the fully validated Natura 2000 dataset).
- The confusion between forest types at EUNIS level 2 and the overestimation of the T3 coverage was addressed by comparing the habitat map at level 2 to the CLC+ Backbone map 2021 of CLMS. The CLC+ map classifies forest areas into 'Coniferous Forest' (T3), 'Broadleaved deciduous forest' (T1) and 'Broadleaved evergreen' (T2). It was checked which pixels on the habitat map responded to the condition: classified as T3 on the habitat map but not classified as T3 on the CLC+ map. For these conflicting pixels, it was checked if the probability map indicated a higher probability for T1 than T2 in pixels which CLC+ classified as T1. If so, these pixels were reclassified to T1. If not, these pixels were kept at their original classification of T3. A similar method was applied for classifying T2. Then, the probability maps for the T classes at level 3 were further used to assign the ultimate classification. A version was also created in which the probability condition (probs(T1)>probs(T2) or other way around) was not considered. However, posterior evaluation based on expert opinion decided that this method introduced too many positive errors, meaning that pixels were reclassified to T1 (or: T2) whereas they should rather be T2 or T3 (or: T1 or T3).

Lastly, a major artifact in the first version of the habitat map was the misclassification of J (urban area) for N (coastal habitats) close to the coast. This is most likely due to the third reason listed above, that is, some classes can have very similar feature sets (pavement in class J vs. bare rocks in class N). This shows the importance of identifying exclusive features that can effectively distinguish similar classes from each other, especially at level 1 (considering our hierarchical merge process). Therefore, future studies and research efforts should focus on the identification of exclusive features still lacking in the current feature pool. The misclassifications J-N were compensated by using the CLMS coastal zone map in vector format. By overlaying this map over Google Aerial imagery in QGIS, polygons corresponding to coastal habitats were selected. Afterwards, four scenarios were created:

1. all pixels within the polygons classified as J were reclassified to N;
2. all pixels within the polygons classified as J were reclassified to N if the probability of N was higher or equal to 1%;
3. all pixels within the polygons classified as J were reclassified to N if the probability of N was higher or equal to 10%;
4. all pixels within the polygons classified as J were reclassified to N if the probability of N was higher or equal to 35%.

Review by expert opinion from São Miguel TS partner concluded that the scenario (1) in which all J pixels are reclassified to N regardless of the probability of N was the most accurate and best alternative mapping to the original version. This methodology could serve as a blueprint for correcting EO-derived habitat maps in other regions.

7.1.6.3. Level 1 EUNIS Habitat map

For model training at level 1, 91 features were selected. The most important features were: 1) Distance to inland water; 2) Climate annual precipitation; 3) Altitude from DEM; 4) Number of growing days with T > 5°C; and 5) Deep contextual Feature #6 (Sentinel 2 data). The model performed well for all classes with an accuracy of 87,01% (Figure 26). The normalized confusion matrix included class U (sparsely vegetated habitats) while this L1 habitat class is missing on the habitat map of São Miguel (no area), since the model considers all training points listed in Annex 13.5.2, thus it is trained to classify each class of the training points. However, when the model is applied and the classification model cannot identify any pixels over the mapping area for which U is the best suiting class, it can be eventually lost. This matches the reality of São Miguel where few areas fit the habitat description of “sparsely vegetated” due to the temperate climate and fertile volcanic soils promoting the quick growth of dense vegetation.

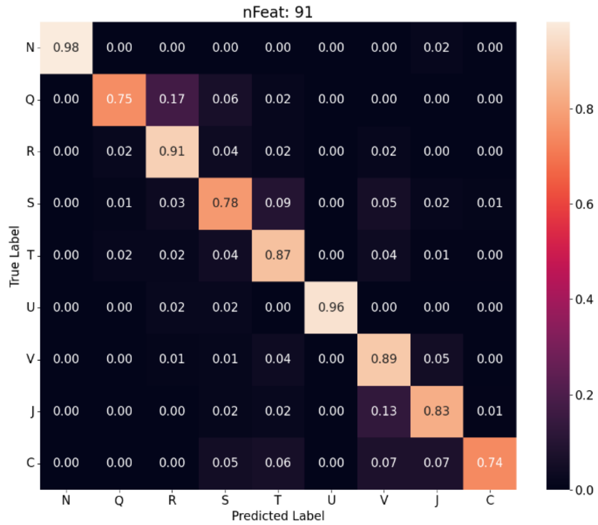


Figure 26: Normalized confusion matrix of winning Catboost model to map habitats at EUNIS L1 for São Miguel. The model accuracy contains 87.01%.

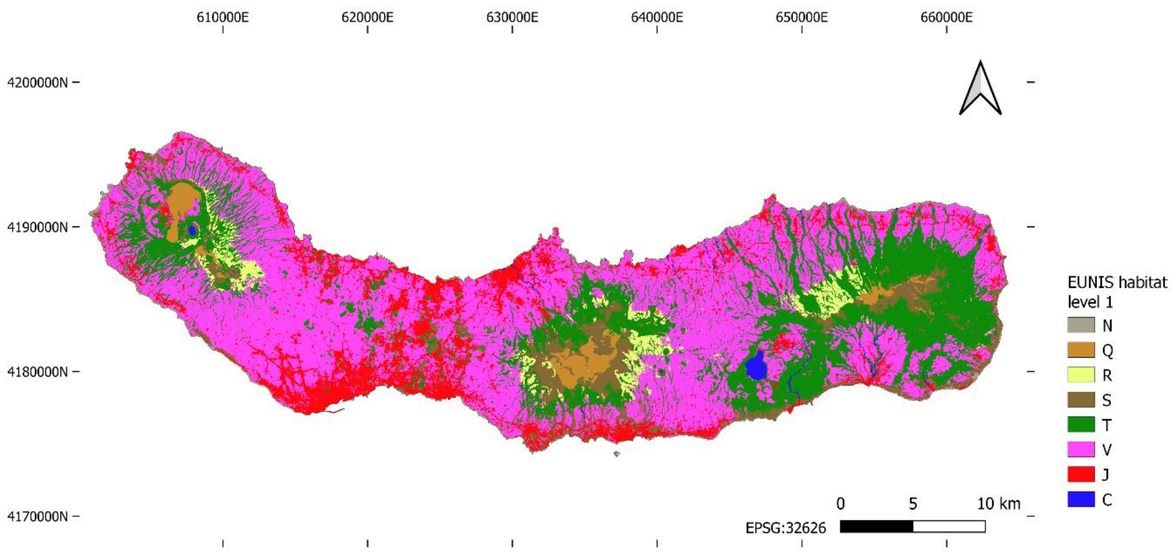


Figure 27: Predicted EUNIS level 1 habitat map of São Miguel. See Annex 13.4 for the legends of the habitat codes.

Like Peloponnese, São Miguel Island is majorly covered by class V (i.e., Vegetated man-made habitats) which includes most of the agricultural classes (Figure 28). The natural classes are located more towards the central axis of the island, and the Constructed, industrial and other artificial habitats (class J) cover mostly the boundaries at the coastal zones (Figure 27).

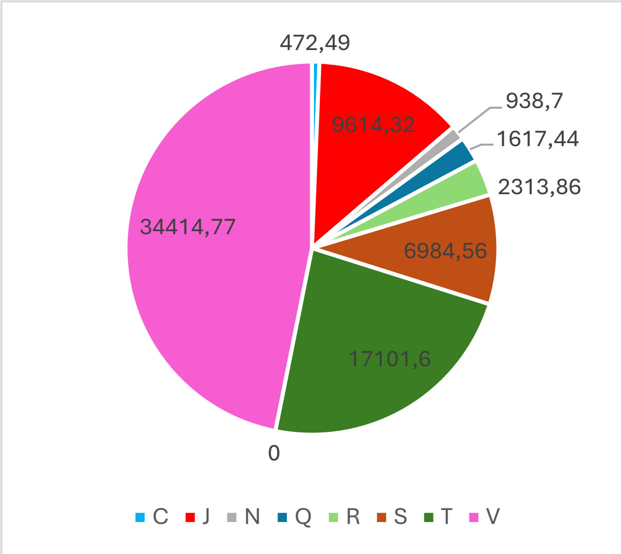


Figure 28: Area distribution [ha] of EUNIS level 1 habitat classes mapped within São Miguel. See Annex 9.4 to explain the habitat codes in the legend.

7.1.6.4. Level 2 EUNIS Habitat Map

At level 2, the model accuracy had satisfactory results (Table 26). Zooming in the normalized confusion matrix of the model for class N shows that N1 (Coastal dunes and sandy shores) is most difficult to be mapped correctly and often mistaken for other classes N2 (Coastal shingle) and N3 (Rock cliffs, ledges and shores, including the supralittoral) (Figure 29).

Table 26: Model accuracies of winning Catboost models generated to map EUNIS habitat level 2 classes within overarching EUNIS habitat level 1 class.

Overarching EUNIS habitat L1 class	Model accuracy [%] to map the associated L2 classes
C	100
J	88,93
N	83,93
Q	98,46
R	99,88
T	91,94
U	100,00
V	97,01

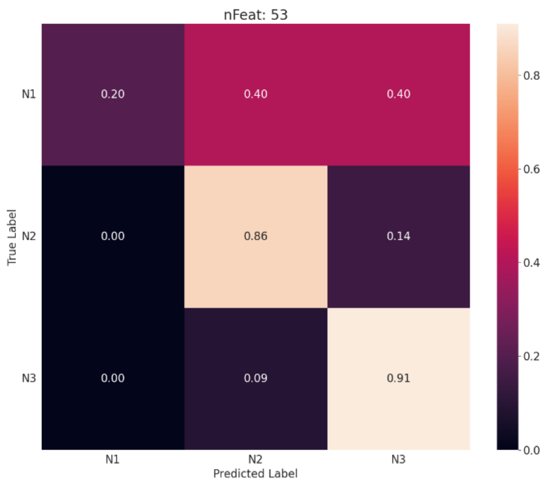


Figure 29: Normalized confusion matrix for mapping EUNIS level 2 habitat classes within EUNIS level 1 habitat N. The model accuracy contains 83.93%.

Table 27 gives an overview of the 10 most important features with predictive power in the model to differentiate level 2 classes within the associated level 1 class. Again, the contextual features are important to classify coastal habitats, and the RS derived products (e.g. B03-p50-10m, ndvi-ts2-10m, …) play a bigger role than in level 1 habitat modelling.

Table 27: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 1 classes. The list displays the features from top to bottom in order of importance with highest importance at the top.

Class-N	Class-T	Class-R
DEM-alt-20m	cfvo	gdd5
contextft_6	dist	nbr2-p50-20m
gdd5	gsp	contextft_6
DEM-slo-20m	contextft_1	gst
contextft_14	bio12	bdod
B03-p50-10m	vh_vv-p50-20m	anir-p10-20m
gsp	rep-p50-20m	ndre2-p10-20m
ndgi-p90-10m	contextft_2	ndvi-ts2-10m
contextft_2	cec	vppamp
contextft_11	DEM-alt-20m	contextft_5

Figure 30 illustrates the output of the habitat mapping of EUNIS habitat classes at level 2. The map, as well as the bar chart displayed in Figure 31, shows that, within the natural vegetation habitat classes, mostly coniferous forest (T3), broadleaved evergreen forest (T2) and temperate shrub heathland (S4) dominate the land cover. The area distribution shows that N2 (Coastal shingle), Q2 (Valley mires, poor fens and transition mires), R1 (natural ungrazed dry grasslands) and T1 (Broadleaved deciduous forest) can be considered as rare classes, which are in accordance with their limited sampling size of training data points.

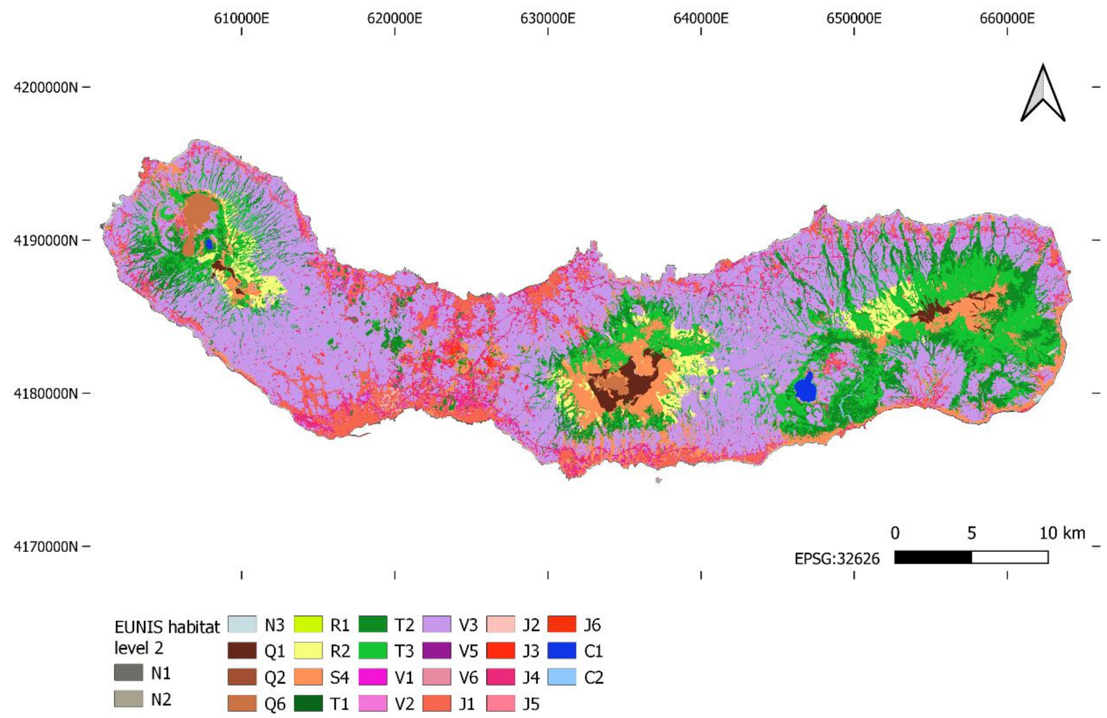


Figure 30: Predicted EUNIS level 2 habitat map of São Miguel. See Annex 9.4 to explain the habitat codes in the legend.

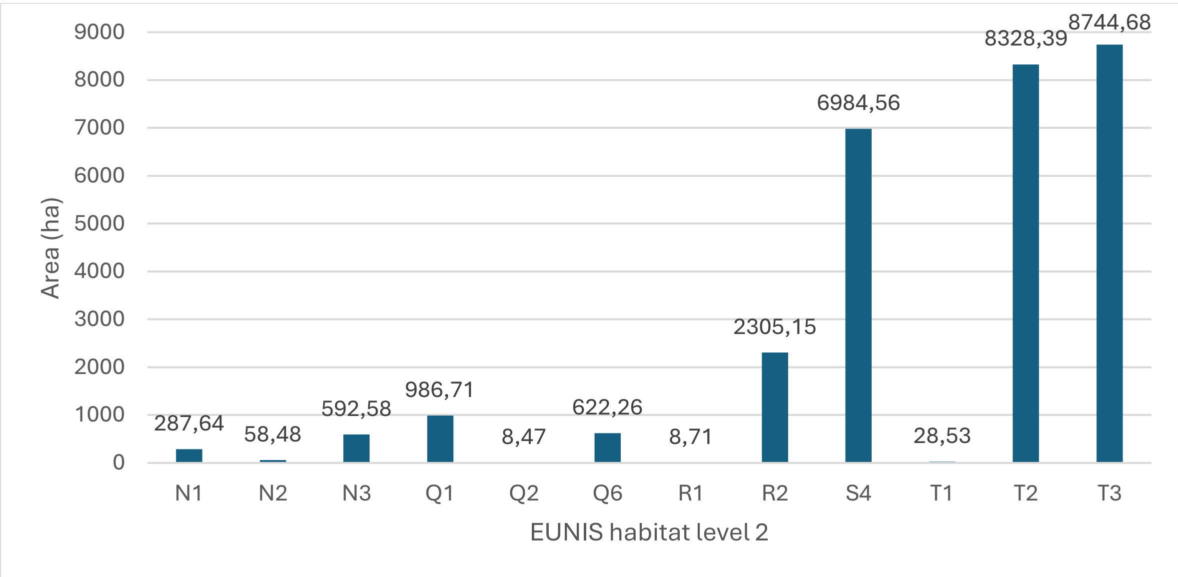


Figure 31: Area distribution [ha] of the natural vegetation EUNIS L2 habitat classes in São Miguel.

7.1.6.5. Level 3 EUNIS Habitat Map

At level 3, the model accuracies exhibit satisfactory results (Table 28). Zooming in the normalized confusion matrix of the model for class T2 shows that class T29 (Broadleaved evergreen plantation of non-site-native trees) is most difficult to be mapped correctly and often mistaken for the other classes like T2x (non-Laurophyllous and non-cultivated [exotic] broadleaved evergreen forest not fitting in T23 and not further differentiated to level 3) (Figure 32). This confusion was expected as these two classes share the same vegetation characteristics, with the main difference being that T29 is cultivated, not naturally occurring.

Table 28: Model accuracies of winning Catboost models generated to map EUNIS habitat L3 classes within overarching EUNIS habitat L2 class.

Overarching EUNIS habitat level 2 class	Model accuracy [%] to map the associated level 3 classes
C1	100,00
J1	89,12
J4	100,00
Q1	93,75
S4	99,52
T1	100,00
T2	97,63
T3	99,61
V6	100,00

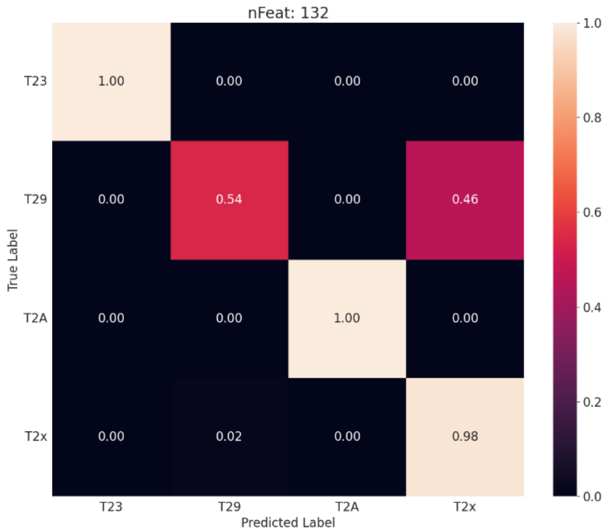


Figure 32: Normalized confusion matrix for mapping EUNIS level habitat classes within EUNIS level 2 habitat T2. The model accuracy contains 97.63%.

Table 29 shows an overview of the 10 most important features with predictive power in the model to differentiate the level 3 classes within the associated level 2 class. For level 3 classes within classes Q1 (Raised and blanket bogs) and T2 (Broadleaved evergreen forest), the top 10 are a combination of contextual features and biophysical parameters. For level 3 classes within T1 (Deciduous broadleaved forest), Sentinel-1 backscatter products are important.

Table 29: Top 10 features with highest importance within the Catboost model generated for three overarching EUNIS level 2 classes. Features ordered from top to bottom in order of importance.

Class-Q1	Class-T1	Class-T2
bio12	contextft_7	phh20
gsp	rep-p10-20m	cfvo
contextft_12	VV-iqr-20m	contextft_7
gdd5	gst	nbr2-p50-20m
DEM-alt-20m	phh20	contextft_12
soc	rvi-iqr-20m	contextft_3
contextft_4	gsp	contextft_1
contextft_13	vh_vv-iqr_20m	bio12
B06-p10-20m	gdd5	cec
contextft_15	lai10	lai01

Figure 33 shows the final habitat map of São Miguel at EUNIS level 3, with further zoom-ins showcased in Figure 34. At level 3, the EUNIS habitat classes S4x (all temperate shrub heathland not further differentiated to level 3), T2x (all broadleaved evergreen forest not further differentiated to level 3) and T3M (Coniferous plantation of non-site native trees) were the most prominent (Figure 35). R21 (Mesic permanent pasture of lowlands and mountains) follows with 2305,15 ha and is classified by the IUCN Red list of habitats as 'vulnerable'(Janssen and Rodwell 2016).

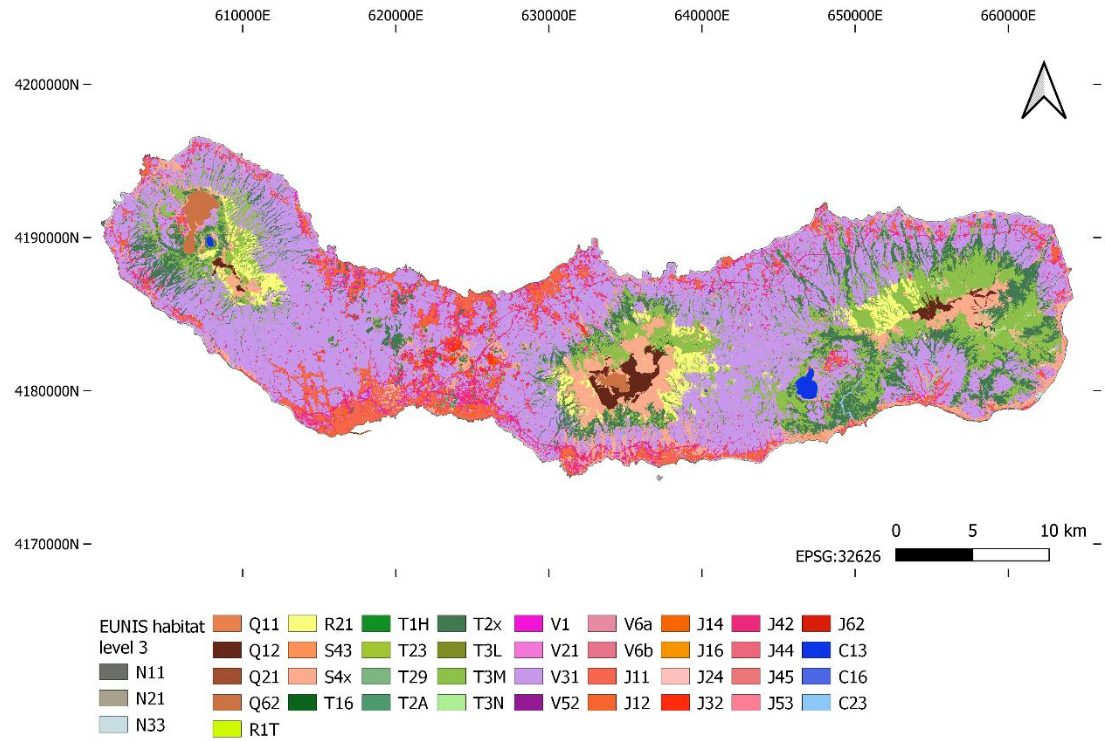


Figure 33: Predicted EUNIS level 3 habitat map of São Miguel. See Annex 9.4 to explain the habitat codes in the legend.

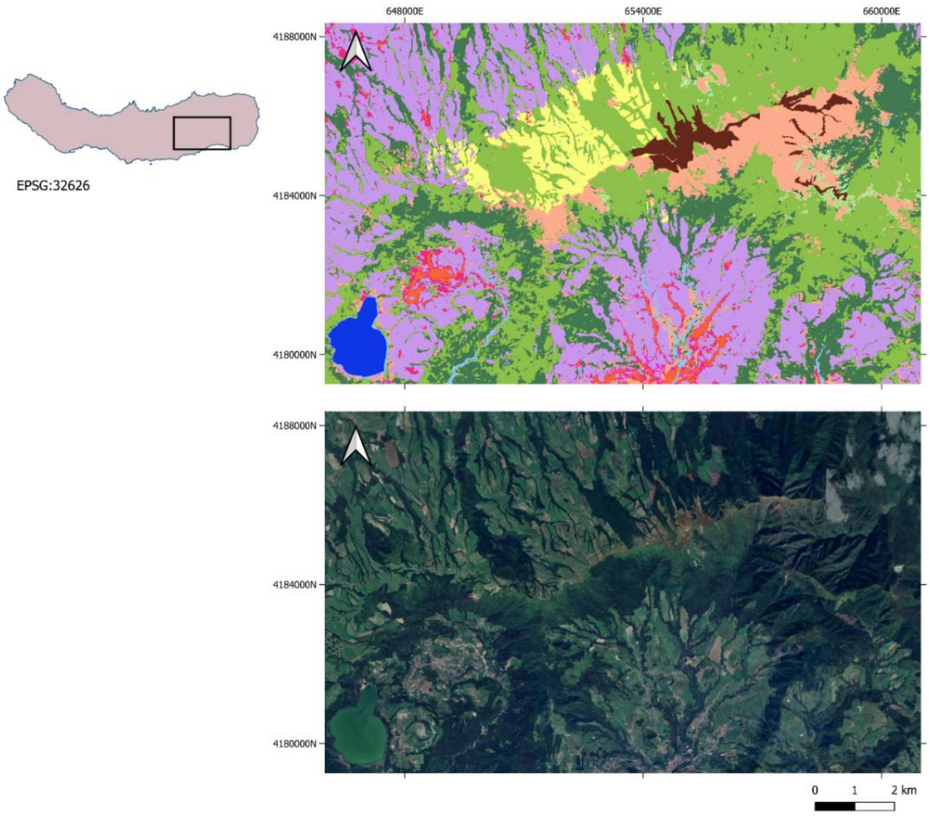


Figure 34: Zoom-in on area in East of São Miguel to visually check the correspondence of the predicted EUNIS level 3 habitat map with Google Aerial imagery. See Figure 33 to link the colours in the habitat map to a EUNIS habitat class.

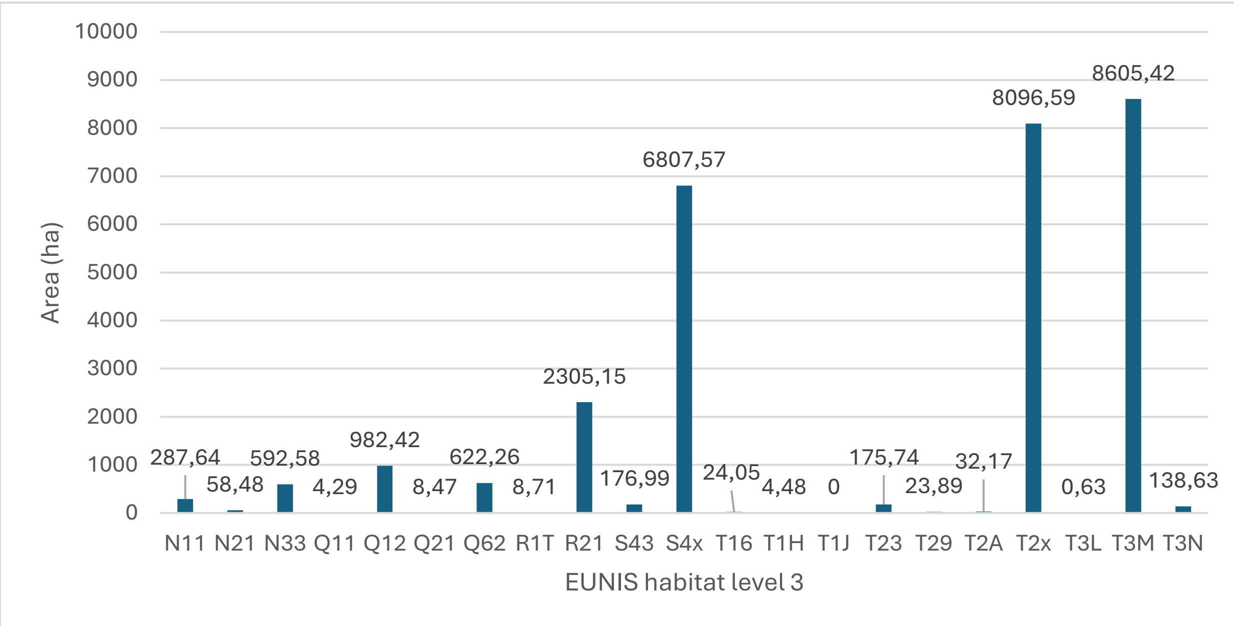


Figure 35: Area distribution [ha] of the natural vegetation EUNIS L3 habitat classes in São Miguel.

7.1.6.6. Validation

The analyses described above show that some classes, for which training data was collected and used in the training dataset, were not mapped in the final habitat maps for São Miguel. These include: T1J, U3C and U62. This can be considered an expected result. T1J and U3C only contained the minimum of 20 training points each in the final training dataset. Therefore, the classes might not be very easily distinguished from others since the low amount of training points is associated with less distinctive power and the trained model does not recognize them easily. T1J is a subclass of broadleaved deciduous forest (T1), which is known to be much less abundant than broadleaved evergreen (T2) and coniferous forests (T3). Meanwhile, U3C and U62 were only sourced within the Natura 2000 areas and not for the whole island, hence their training point datasets were boosted with points from Pico and Terceira Island, where these habitats cover larger extents than São Miguel.

Since the training data for executing the habitat mapping in São Miguel originally consisted of a much larger original dataset, which was filtered to obtain the training dataset, the remaining points not used for training could be used as an external validation set.

Map accuracy (rather than model accuracy) was checked by comparing the original validated classification of the external validation set with the predicted classification in the EUNIS level 1 habitat map (Figure 36, left). This map accuracy was 78,44%. The same approach was performed to check how the original classifications of the points compared to the level 1 classifications on the CLMS Coastal Zone map (Figure 36, right). This map accuracy was 66,67%. Thus, at EUNIS L1, the habitat map performs better than the existing CLMS product.

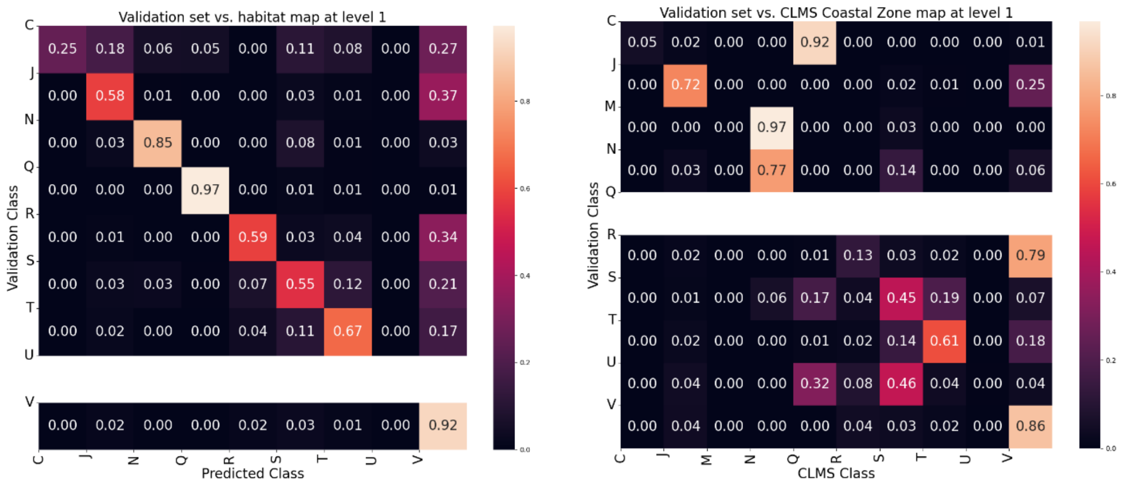


Figure 36: Normalized confusion matrix for [left] - Original classification of external validation points vs. the predicted classification in the habitat map at level 1. Class U was not mapped on the modelled habitat map. The map accuracy based on these validation points is 78,44%. [Right] Original classification of external validation points vs. the classification on the Coastal Zone map of CLMS (2018) at level 1. Class Q is not mapped on the CLMS Coastal Zone map. The map accuracy is 66,67%.

Map accuracy for EUNIS habitat level 2 was also checked by comparing the original validated classification of the external validation set with the predicted classification in the habitat map (Figure 37). Zooming in the same forest and heath classes (T1, T2, T3 and S4) that are used for creating the forest condition accounts, the normalized confusion matrix shows that the habitat map mistakes T1 often for S4. Also, misclassifications to T2 or T3 occur. T2 is mostly misclassified for T3 and V3. T3 is mapped with the best accuracy relative to the other forest types. Sometimes, confusion with V3 occurs. The same applies to S4.

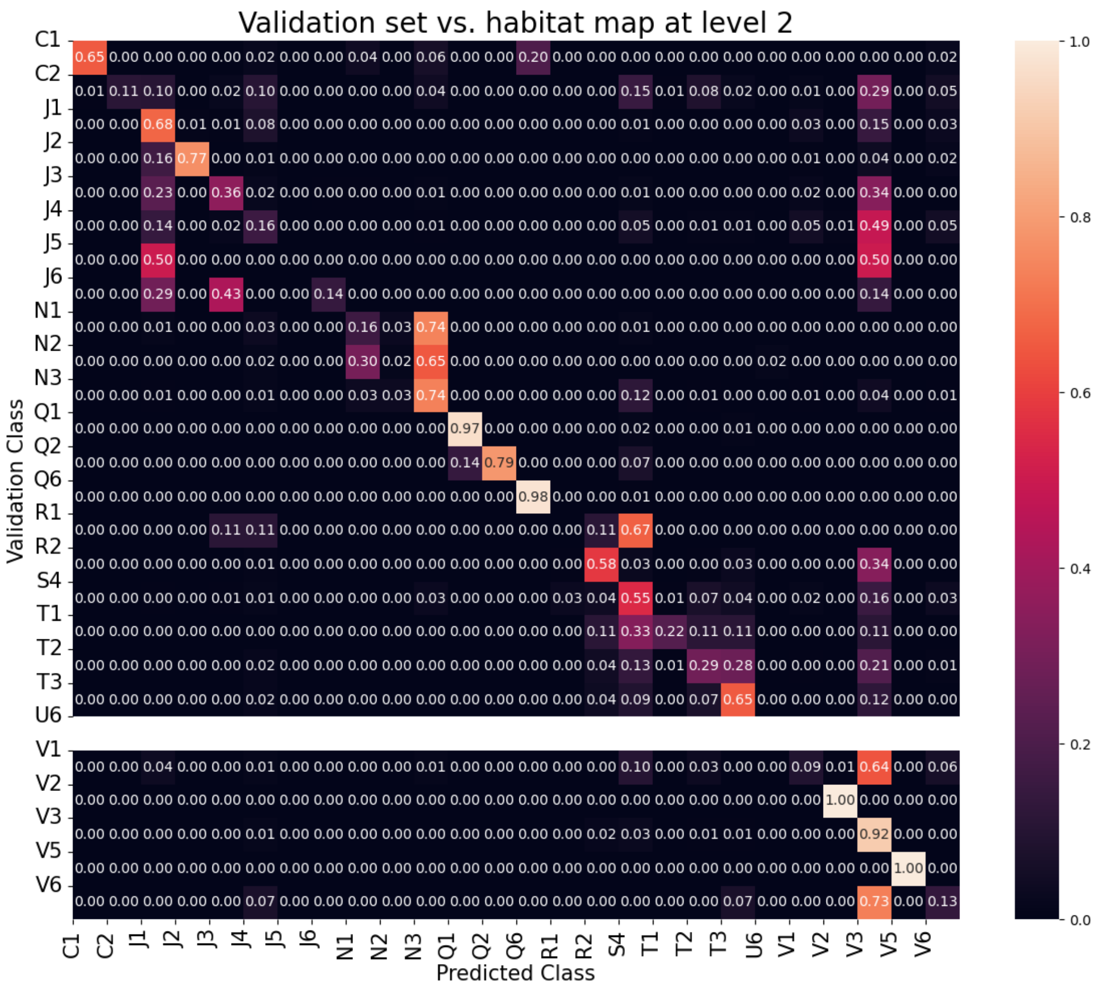
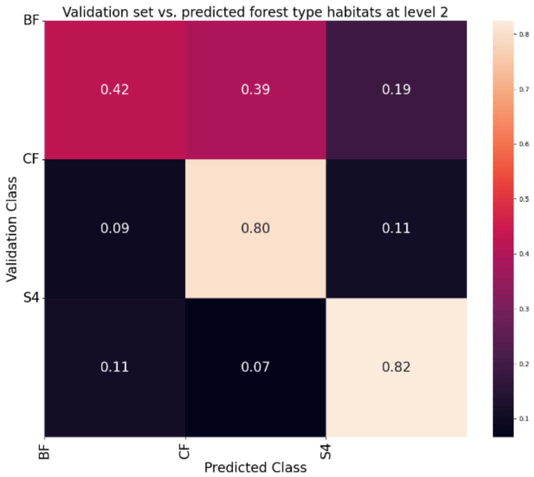


Figure 37: Normalized confusion matrix for the original classification of external validation points vs. the predicted classification in the habitat map at level 2.

In a second validation check, the HRL Forest Type Layer (FTY) of CLMS (2018) which classifies the forest areas was compared to the external validation points that indicate forest area (EEA, 2020a) (Figure 38, right). The CLMS FTY map only differentiated Broadleaved Forest (BF) from Coniferous Forest (CF), thus, it does not classify 'Broadleaved evergreen forest' (T2) and 'Broadleaved deciduous forest' (T1) separately. Therefore, the training points for T1 and T2 were merged on class BF and then a normalized confusion matrix was created. The map accuracy for mapping the forest classes BF and CF by the CLMS FTY layer based on the external validation points is 69,48%. Similarly, the mapped forest classes on the level 2 habitat map were translated to BF and CF, then a confusion matrix was created with the validation points as well (Figure 38, left). When only BF and CF are considered, the map accuracy is 76,58%.

When class S4 is also considered, the map accuracy for forest classes used in the forest condition account is 70,58%. Thus, the level 2 habitat map performs better than the CLMS FTY layer in predicting the distribution of broadleaved and coniferous forest in São Miguel.



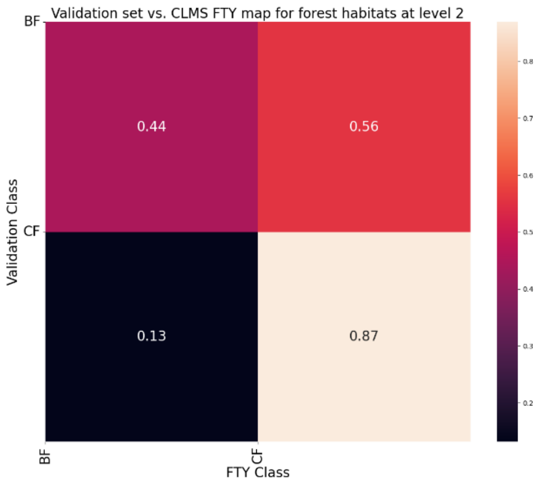


Figure 38: Normalized confusion matrix for [left] - original classification of external validation points for Broadleaved forest (BF, i.e., T1 and T2 merged), Coniferous forest (CF) and S4 points vs. the predicted classification of the forest classes in the level 2 habitat map, and [right] original classification of external validation points for Broadleaved forest (BF, i.e., T1 and T2 merged), CF and S4 points (map accuracy considering only BF and CF: 76,58%, map accuracy including all classes: 70,58%) vs. the classification on the CLMS FTY layer (2018) (map accuracy: 69,48%).

As previously mentioned, improving habitat mapping accuracy requires identifying exclusive features or indicators that can effectively distinguish similar classes from each other. This first run of the EUNIS habitat mapping, presented in this chapter, lays the groundwork for succeeding efforts by highlighting which classes are most often misclassified, guiding further research for finding or developing more decisive and context-dependent features to separate them. In addition to those described in Section 7.1.6.2 and which were addressed during the post-processing steps, Table 30 summarizes some other relevant misclassifications that occurred for São Miguel at the different hierarchical levels.

Table 30: Relevant misclassifications identified in the habitat mapping of São Miguel.

Misclassification	Description
J vs N	Built-up pavement in class J near the coast mistaken for coastal cliff bare rocks in class N.
Q12 vs R / S	Q12 bogs resemble ungrazed grasslands or shrubs nearby, possibly due to infiltration by exotic shrubs.
C23 vs J4.2	Narrow streams/creeks are difficult to map at 10 m, usually hidden by riparian vegetation and sometimes ephemeral, thus resembling the dirt/pavement of narrow roads through vegetation.
R1T vs S4x	Ungrazed grasslands in altitude with low frequency of moving events might grow herbaceous shrubs resembling S vegetation.
R21 vs V31	Semi-natural grasslands at medium altitude and artificially improved grasslands look similar when the latter are not under crop rotation.

J3.2 vs S4x / V31.	Active and inactive extractive sites were combined to increase sampling size, but inactive sites will have vegetation regrowth soon after.
J4.2 vs S / T / V	Secondary narrow roads sidelined by grass, shrubs or tree canopies can induce misclassifications at 10 m resolution.
V31 vs S4x	Many V31 areas rotate maize crops from May-September and during this period look like S vegetation.
V31 vs V21 / V52	When rotating crops, V31 resembles ornamental / plantations.
S vs T	Same species can simultaneously form T2 Forests or S4 Heaths depending on potential growth size and stunting influence of strong winds in altitude (as surveyed by the forest inventory). Forest Inventory classification is based on species dominance. In many areas, the mixture of vegetation from both classes is unavoidable (T class tree canopies with S class shrubs in the understory). Most S vegetation is aggregated in S4x at level 3, potentially creating a heterogenous grouping of a broad range of exotic or altered heathland types resembling habitats from other EUNIS classes. More accurate field validated data is needed for further differentiation within this virtual class.
T2x vs T3M	Stunted coniferous plantations might resemble T2. Further differentiation is needed within the heterogenous T2x virtual class.
C16 vs V31	Temporary lakes are few, small and ephemeral, looking blended into the grasslands when dry.
J1 subclasses	Stricter criteria needed to discriminate continuous/discontinuous urban areas.
Q62 vs C13	Q62 and C13 are both large surface bodies of water, except C13 are eutrophic while Q62 are mesotrophic.
V1 vs J1.2 / S4x	Many arable lands with permanent cultures are in built-up farms or house backyards, resembling areas of discontinuous urban fabric (J4.2). Their vegetation may look like class S shrubs.
V6a vs J1.2 / S4x	Tree orchards in farms and backyards resemble discontinuous urban fabric and their vegetation may look like class S shrubs.
V6b vs T / S	Forest reserves have many of the same species naturally occurring in the wild, except they are planted and managed by human intervention.

It is important to acknowledge the challenge posed by unique and locally relevant EUNIS classes whose mapping and classification is very hard due to their limited spatial extent. Two notorious examples in the current habitat map for São Miguel include the fumarolic fields (under class U62) or the only European tea plantations (under class V52). Class U62 is absent from the level 3 habitat map, while class V52 was prone to misclassification with other shrub vegetation or man-made cultivations. These are unique, differentiated, ecologically and culturally important classes represented within the EUNIS classification. Yet, due to their limited spatial extent and subtle spectral or structural signatures, these habitats are challenging to sample systematically and to detect through remote sensing. Consequently, despite their confirmed presence on the ground, these habitats are often underrepresented or omitted in spatial datasets. This highlights the need to reconcile the ecological importance of rare and spatially marginal habitat types with the constraints of automated classification, especially in diverse and heterogeneous landscapes of oceanic islands.

7.2. Ecosystem condition mapping

7.2.1. PEOPLE-EA approach for forest condition index

EU legislation requires ecosystem condition accounts to be produced every three years (European Commission 2022). For forest and woodland this requires reporting the mandatory ecosystem condition variables/indicators 'dead wood' (in m^3/ha) and 'tree cover density' (in %). There are some other recommended variables/indicators to report on, which include soil organic carbon content, forest productivity and several more. During the PEOPLE-EA (https://esa-people-ea.org/en) project, a method was established to create a forest condition index. This work follows up previous studies on forest condition monitoring conducted in Europe (Bruzón et al., 2023; Maes et al., 2023), but with a higher focus on the use of Earth Observation data sources. Besides, this section is closely linked to SELINA WP3 (Ecosystem condition mapping and assessment). WP3 subtasks aim to develop a minimal standardized set of indicators per ecosystem type and create a structured methodology for reference condition selection.

Although a project-wide methodology for variable and reference condition selection for multiple ecosystem types is being developed collaboratively within SELINA in the context of WP3, since it is currently in progress at the time of this report, this work in Task 5.2. has performed an exploratory analysis for forest ecosystems with application of the developed PEOPLE-EA method on the two TS, Peloponnese and São Miguel.

The PEOPLE-EA method follows the SEEA-EA framework, which proposes a stepwise approach:

1. Definition and selection of ecosystem condition variables with intention to cover each class of the Ecosystem Condition Typology (ECT) (overview in Figure 39).
2. Definition of the reference conditions and rescaling of the variables to ecosystem condition indicators which range between 0 and 1.
3. Aggregation of the partial indicators into a single overarching ecosystem condition index using specific indicator weights.

Group A: **Abiotic ecosystem characteristics**	Class A1. **Physical state characteristics**: physical descriptors of the abiotic components of the ecosystem (e.g. soil structure, water availability)
	Class A2. **Chemical state characteristics**: chemical composition of abiotic ecosystem compartments (e.g. soil nutrient levels, water quality, air pollutant concentrations)
Group B: **Biotic ecosystem characteristics**	Class B1. **Compositional state characteristics**: composition / diversity of ecological communities at a given location and time (e.g. presence/abundance of key species, diversity of relevant species groups)
	Class B2. **Structural state characteristics**: aggregate properties (e.g. mass, density) of the whole ecosystem or its main biotic components (e.g. total biomass, canopy coverage, annual maximum NDVI)
	Class B3. **Functional state characteristics**: summary statistics (e.g. frequency, intensity) of the biological, chemical, and physical interactions between the main ecosystem compartments (e.g. primary productivity, community age, disturbance frequency)
Group C: **Landscape level characteris-tics**	Class C1. **Landscape and seascape characteristics**: metrics describing mosaics of ecosystem types at coarse (landscape, seascape) spatial scales (e.g. landscape diversity, connectivity, fragmentation)

Figure 39: SEEA EA Ecosystem Condition Typology per group (abiotic, biotic and landscape level) and per associated classes explained (Vallecillo Rodriguez et al. 2022).

This Section explains how this approach was applied to create forest condition accounts for Peloponnese (2018-2022) and São Miguel (2018-2023).

Table 31 shows a list of proposed forest condition variables during PEOPLE-EA, covering all ECT classes (Bruelheide et al. 2024). First, forest condition variables, to be used to create the final forest condition index (FCI), are identified from this list. The variables composing the index aim to represent comprehensively all the components of ecosystems. The index is usually composed of a single variable for each ECT class, since the variables belonging to the same ECT class are usually highly correlated. The variables are selected considering their relevance, their direct relationship to forest conditions and the availability of data for their measurement. The table below gives an overview of these proposed variables, the associated ECT class and their temporal coverage. To generate long time series for some variables, multiple datasets had to be used. In that case, the datasets were always harmonized, and a resampling was applied if needed (to have consistency in spatial resolution and extent).

Table 31: Overview of forest condition variables.

Forest Condition Variable	ECT class	Dataset origin → applied algorithm	Temporal coverage & spatial resolution
Normalized Difference Water Index (NDWI) *Water content index*	A1 Abiotic: Physical state	GEE Landsat 7 & 8 (harmonization) Collection 1 Tier 1 32-Day	2000-2021, 30 m

		NDWI composite (U.S. Geological Survey 2021, 2022) - aggregation to three-annual average	
Soil Organic Carbon content (SOC) *Organic carbon content in the first 30 cm of the soil (topsoil)*	A2 Abiotic: Chemical state	2003 OCTOP: Topsoil Organic Carbon Content for Europe (Jones et al. 2005) 2014 LUCAS: Topsoil Soil Organic Carbon (LUCAS) for EU25 (de Brogniez et al. 2014) OCTOP for 2000-2013 & LUCAS from 2014 onwards	2003, 1 km 2014, 500 m
Threatened Forest Bird Species Diversity (TFBSD) *Assessment of bird population status (population sizes, trends for breeding and wintering populations, pressures, and threats for Special Protection Area trigger species)*	B1 Biotic: Compositional state	Population trend of bird species: datasets from Article 12, Birds Directive 2009/147/EC reporting (2008-2012) (EEA 2021b) - original dataset used: 2008 values for 2000-2011, 2012 for 2012 onwards	2000 & 2018, 5 km
AGB (Above Ground Biomass) *The stock in living or dead biomass aboveground (plant stem, leaves, head, spike, seeds and foliage)*	B2 Biotic: Structural state	ESA's Climate Change Initiative Biomass (Santoro and Cartus 2024) ESA Forest carbon monitoring (Forest Carbon Monitoring Consortium 2021) - ESA CCI dataset for 2000-2020, ESA FCM from 2021 onwards	2010, 100 m, 2017-2021, 20 m
Leaf Area Index (LAI) *The ratio of one-sided leaf is per unit ground area*	B2 Biotic: Structural state	Copernicus Leaf Area Index (European Commission Directorate-General Joint Research Centre (a) 2017, European Commission Directorate-General Joint Research Centre (b) 2017) - aggregation to annual average	1999-2014, 1 km, 2015-present, 300 m
Plant Phenology Index (PPI)	B2 Biotic: Structural state	Copernicus Plant Phenology Index Seasonal	2017-present, 10 m

Index is linearly related to green leaf area index, is used to track canopy green foliage dynamics (plant phenology)		Trajectories (EEA 2021a) - aggregation to annual average	
Tree Cover Density (TCD) *Proportional canopy coverage per satellite pixel in a range of 0 to 100%*	B2 Biotic: Structural state	Copernicus HRL tree cover density layer (EEA 2020b)- year 2012 for 2000-2014, year 2015 for 2015-2017, year 2018 for 2018 onwards	2012, 2015, 2018, 100 m
Net Primary Productivity (NPP) *The rate of accumulation of biomass or energy*	B3 Biotic: Functional state	Copernicus Dry Matter Productivity and Net Primary Production (European Commission Directorate-General Joint Research Centre (c) 2018, European Commission Directorate-General Joint Research Centre (d) 2018) - aggregation of annual sum, transfer of GDMP to NPP	1999-2014, 1 km 2015-present, 300 m
Fraction of Green Vegetation Cover (FCOVER) *The fraction of ground covered by green vegetation*	B3 Biotic: Functional state	Copernicus Fraction of green Vegetation Cover (European Commission Directorate-General Joint Research Centre (e) 2017, European Commission Directorate-General Joint Research Centre (f) 2017) - aggregation to annual average	1999-2014, 1 km 2015-present, 300 m
Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) *The fraction of incoming solar radiation in the spectrum 400-700 nm that is absorbed by vegetation canopy*	B3 Biotic: Functional state	Terrascope Sentinel-2 fAPAR (VITO 2024) - derived from ESA L2A products, fAPAR generation, aggregation to annual average	2016-present
Drought Severity (SD) *Consecutive occurrences of water deficiency*	B3 Biotic: Functional state	European Drought Observatory – Drought Indicator v3 (European Commission Joint Research Center: Drought Team 2024) - transfer	2012-2024

		from categorical to quantitative unit, aggregation to annual average	
Normalized Difference Vegetation Index (NDVI) *A measure of the amount and vigor of vegetation on the land surface*	B3 Biotic: Functional state	Copernicus Normalized Difference Vegetation Index (European Commission Directorate-General Joint Research Centre (g) 2016, European Commission Directorate-General Joint Research Centre (h) 2021)-aggregation to annual average	2000-2020, 1 km 2021-present, 300 m
Forest Connectivity Percentage (FCP) *The percentage of connected forest patches*	C1 Landscape and seascape at coarse scale	CORINE Land Cover (EEA 2020a) - GUIDOS Toolbox algorithm for connectivity for selected classes (Vogt and Riitters 2017)	2000, 2006, 2012, 2018 100 m
Landscape Naturalness (LN) *The degree of natural characteristics in the landscape (i.e., not human-made structures)*	C1 Landscape and seascape at coarse scale	CORINE Land Cover (EEA 2020a) - GUIDOS Toolbox algorithm for naturalness for selected classes (Vogt and Riitters 2017)	2000, 2006, 2012, 2018 100 m
Forest Fragmentation (FF) *The division of forest into smaller patches*	C1 Landscape and seascape at coarse scale	GEO BON Relative Magnitude of Fragmentation (RMF) (Naimi and Kissling 2022)	1992-2020 300 m

As can be derived from Table 31, the Ecosystem Condition Typology (ECT) classes contain variables that are not restricted to describing productivity but also cover biodiversity (e.g. Threatened Forest Bird Species Diversity) and ecosystem integrity (e.g. Forest Connectivity Percentage). Under the SEEA EA framework, ecosystem condition is defined as the quality of an ecosystem measured in terms of its abiotic and biotic characteristics. It reflects the ecological state of the ecosystem and is not defined as the optimal condition for delivering the maximum capacity of ecosystem services.

"NDWI", "SOC", "TFBSD", "AGB", "NPP", and "FCP" were selected as forest condition variables (Table 32). We aimed at selecting these variables to cover each ECT class, to involve Earth Observation datasets in the methodology and to have data availability both on spatial and temporal domain for the two test sites. Besides, we tried to avoid redundancy in the selected forest condition variables by selecting variables with different dataset origin (i.e., different satellites, different algorithms, etc.). This is important in order to ensure the interpretability of the final forest condition index. However, as the ECT classes itself do overlap to a certain extent, total avoidance of redundancy could not be achieved.

In the PEOPLE-EA approach, the variable ‘Soil Organic Carbon (SOC) is the only proposed ‘A2 Abiotic: chemical state’ variable and is derived from the OCTOP (Topsoil Organic Carbon Content for Europe, 2003) and LUCAS (Topsoil Soil Organic Carbon for EU25, 2014) archive. Unfortunately, these datasets do not cover São Miguel, so it was aimed to find an alternative dataset. Datasets describing the chemical state of forests at high spatial and temporal resolution are very scarce. The best option was to extract SOC data from ISRIC SoilGrids Database at 250m spatial resolution (SoilGrids 2025). This data is not annual, in fact no temporal extent is given for the data. However, the same data per forest type available in SoilGrids online database for each year can be used, as this will still influence the magnitude of the final forest condition index and is therefore useful.

Recently, an issue with the CLMS NPP dataset has been discovered. The product is derived from the CLMS GDMP (300 m) dataset, created by PROBA-V data from 2014 until June 2020, and from July 2020 onwards based upon Sentinel-3/OLCI data. The transition from PROBA-V to Sentinel-3 in mid-2020 has introduced artefacts in the data and has not yet been resolved. Therefore, further work has proceeded with the NPP dataset from MODIS at 500 m resolution (Running and Zhao 2021).

Each variable was rescaled to an indicator ranging from 0 to 1, with values getting closer to 1 as the forest condition improves. The method of rescaling variables to indicators requires identifying the 98th percentile (i.e., Upper Reference) of the variable’s values among forest with ideal condition (i.e., the reference area) and the 2nd percentile (i.e., Lower Reference) of the variable’s values among the remaining forest around the area of interest. The reference areas should contain the desired state of the variable and represent forests in good health. Generally, primary forest or protected forest areas are taken as reference areas since they have very little signs of human impact and remain highly tree-covered throughout time series. However, some if not many areas lack primary forest or protected areas. In that case, it is best to leave the selection of reference sites up to local expert knowledge.

These upper and lower reference values for rescaling the variables to indicators are obtained from the reference year which is normally the first year the data was available. Often the year 2000 is taken since this is far enough in time (and therefore with higher probability that the forests will have experienced fewer human impacts) and many datasets are available from this year onwards. However, it is important that all condition variables should have the same reference year. In the case that for one variable the reference year is only available from more recent datasets, that year should be used as reference year for all condition variables used for creating the index.

Rescaling of variables to indicators was done by the following calculation (Figure 40):

$$Xi = (X_{observed} – X_{LowerReference}) / (X_{UpperReference} – X_{LowerReference})$$

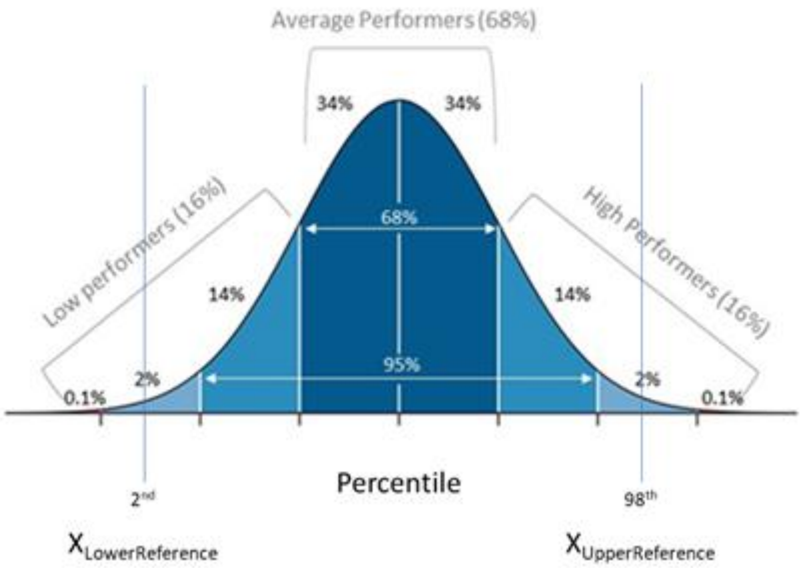


Figure 40: Visualization of extraction of 98th (i.e., Upper Reference) and 2nd percentile (i.e., Lower Reference) per condition variable for rescaling variables to indicators (Bruelheide et al. 2024).

The Forest Condition Index (FCI) quantifies forest health by summarizing the information of several indicators, each representing a different aspect (abiotic, biotic and landscape characteristics) of the forest condition. The weights for indicators are determined according to the following criteria:

- their spatial resolution (referring to pixel size in meters).
- their temporal extent (referring to the range of years covered by the dataset).
- their temporal resolution (referring to the number of observations per year).
- and the level in which the dataset is up to date (referring to the proximity of the latest year to the present).

Indicators are ranked according to each of these criteria based on expert opinion (i.e., a common understanding and not just used in this project only). Larger values of these criteria are associated with a better reliability in realistic representation of the ecological conditions and rank in a higher position (Table 32).

Table 32: Forest condition indicators ranked on 4 criteria (spatial resolution, temporal resolution, temporal frequency, and dataset quality). When data from more than one indicator is considered equal in one criterion, each indicator is attributed to the average of the positions they would represent (e.g. for spatial resolution NDWI, AGB and FCP in the table below).

Variable	Spatial resolution	Temporal resolution	Temporal frequency	Dataset quality	Total	Final weight
Normalized Difference Water Index (NDWI)	5	6	6	5.5	**22.5**	0.27
Soil Organic Carbon (SOC)	2	1	1	2	**6**	0.07
Threatened Forest Bird Species Diversity (TFBSD)	1	3.5	3	1	**8.5**	0.10
Above Ground Biomass (AGB)	5	3.5	3	4	**15.5**	0.18
Net Primary Productivity (NPP)	3	5	5	5.5	**18.5**	0.22
Forest Connectivity Percentage (FCP)	5	2	3	3	**13**	0.15
					84	

The overall Forest Condition Index (FCI) is derived by calculating the arithmetic average. In other words, by multiplying each condition indicator with its assigned weight and summing up the weighted indicators to get the FCI: FCI = (NDWI x 0.27) + (SOC x 0.07) + (TFBSD x 0.10) + (AGB x 0.18) + (NPP x 0.22) + (FCP x 0.15).

7.2.2. Peloponnese – ARIES for PEOPLE-EA

During the PEOPLE-EA project, the ARIES for PEOPLE-EA tool was developed (BC3 - VITO 2025) https://peopleea.integratedmodelling.org/modeler/#/login. The tool is now available for all users, providing an automatic approach to calculate the forest condition index for specified map boundaries or administrative regions (encoded in the tool), according to the above-described method. Besides, variable and indicator accounts can also be derived and visualized. Within the tool, forest condition indices can be calculated per forest type and per biogeographical region (EEA 2024). The forest types are classified based on the CLMS (Copernicus Land Monitoring Service) CORINE Land Cover layer (2018) (EEA 2020a). The selected forest types of CORINE include Broadleaved Forest, Coniferous Forest, Mixed Forest and Transitional woodland & shrub. The European biogeographical regions are divided by the EEA classification (Figure 41).

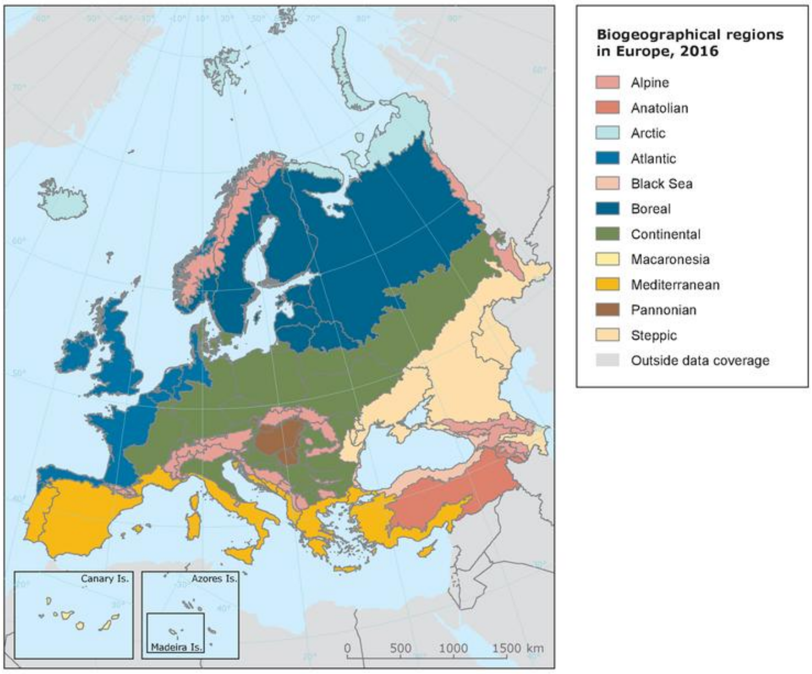


Figure 41: European biogeographical regions classified by the European Environmental Agency (EEA 2024).

According to the classification of the EEA Biogeographical regions, Peloponnese is situated within the Mediterranean zone. Reference sites are automatically derived by combining the classified forest areas, the biogeographical zone and the indication of area of interest. More specifically, in the PEOPLE-EA approach the reference sites per forest type and per biogeographical region are selected within the study area based on combining these datasets:

- European Primary Forest Database (EPFD) v2.0 (Sabatini et al. 2021),
- World Database on Protected Areas (WDPA) (UN WCMC 2025) and
- CLMS HRL Tree Cover Density (EEA 2020b).

The selected forest condition variables to compute the forest condition index for Peloponnese are the ones described Table 32: NDWI, SOC, TFBSD, NPP, AGB and FCP. The OCTOP and LUCAS databases cover Peloponesse so the SOC data could be derived from these databases. In the ARIES for PEOPLE-EA Explorer, this full process of rescaling the variables to indicators and calculating the forest condition index can be executed under the hood (i.e., computation is performed in the back end automatically when the user fills out the desired parameters in the front-end). In practice, users can choose to retrieve the mean condition variable account per forest type (per biogeographical zone), as well as calculate the mean condition indicator values or derive the mean forest condition indices (Figure 42). This can be done per year or for a full period.

Table 1. Net Primary Productivity Variable

Table 1. Net Primary Productivity Variable Account (in tC/ha/year)

	Transitional woodland scrub, Mediterranean	Broadleaf forest, Mediterranean	Mixed forest, Mediterranean	Coniferous forest, Mediterranean
Year 2018	865.23	1070.29	1058.54	922.19
Year 2019	867.86	1042.78	1034.49	917.08
Year 2020	854.37	1023.06	1019.82	886.91
Year 2021	730.98	874.24	869.20	772.09
Year 2022	802.15	981.81	950.65	842.06
Net change	-63.07	-88.48	-107.89	-80.13

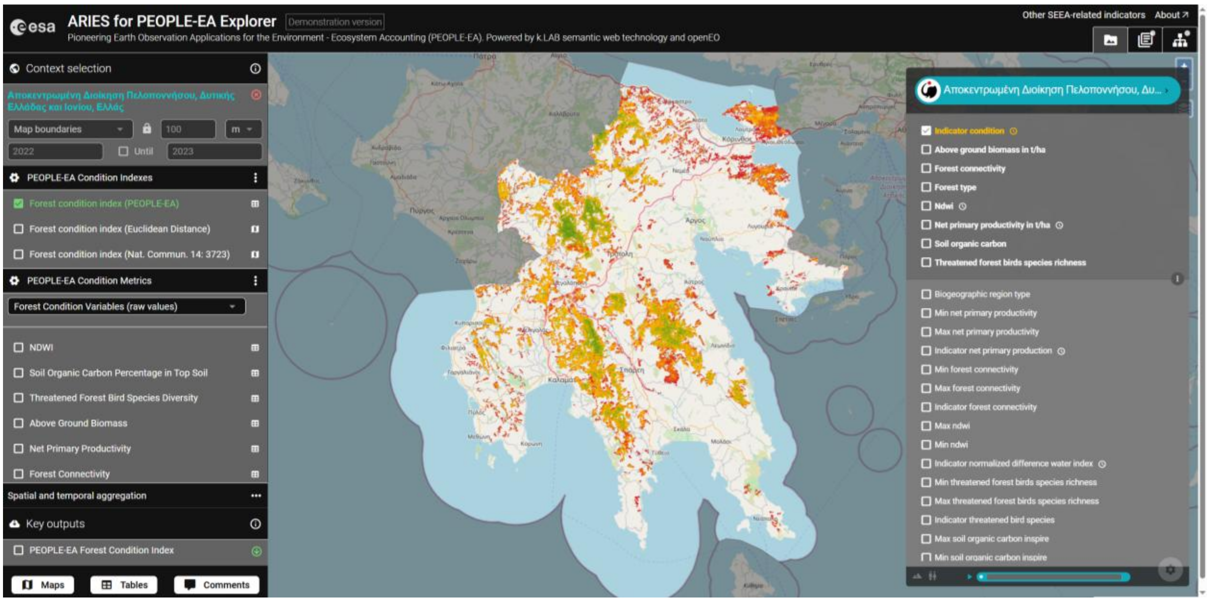
Table 1. Net Primary Productivity Indicator

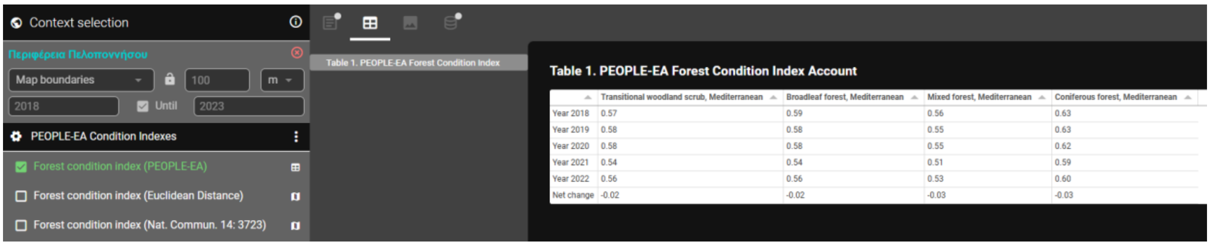
Table 1. Net Primary Productivity Indicator Account

	Transitional woodland scrub, Mediterranean	Broadleaf forest, Mediterranean	Mixed forest, Mediterranean	Coniferous forest, Mediterranean
Year 2018	0.75	0.83	0.75	0.72
Year 2019	0.76	0.81	0.73	0.72
Year 2020	0.75	0.78	0.71	0.69
Year 2021	0.62	0.61	0.57	0.57
Year 2022	0.69	0.74	0.65	0.64
Net change	-0.06	-0.10	-0.11	-0.08

Figure 42: Examples of output ARIES for PEOPLE-EA Explorer - Net Primary Productivity variable condition account [above] and Net Primary Productivity Indicator account [below].

Figure 43 illustrates how the forest condition index was generated per forest type for 2022 in Peloponnese. The output is a GeoTIFF with forest condition index calculated per 100 m x 100 m pixel and a table collecting the mean forest condition index per forest type per year.





	Transitional woodland scrub, Mediterranean	Broadleaf forest, Mediterranean	Mixed forest, Mediterranean	Coniferous forest, Mediterranean
Year 2018	0.57	0.59	0.56	0.63
Year 2019	0.58	0.58	0.55	0.63
Year 2020	0.58	0.58	0.55	0.62
Year 2021	0.54	0.54	0.51	0.59
Year 2022	0.56	0.56	0.53	0.60
Net change	-0.02	-0.02	-0.03	-0.03

Figure 43: Example of output ARIES for PEOPLE-EA Explorer, showing the spatial raster map of forest condition index for 2018 [above] and the tabular forest condition account for 2018-2022 per forest type [below].

7.2.3. São Miguel – Forest condition assessment

For São Miguel, the forest condition index FCI for the period 2018-2023 was not calculated by use of the ARIES for PEOPLE-EA tool. This is because not all input maps for ARIES currently cover the extent of the Azores islands. Therefore, a better option was to perform the PEOPLE-EA approach offline. Besides, the offline method provides the option to specify our own reference sites per forest type and to delineate our own accounting area. The accounting area was derived by the use of the EUNIS habitat level 2 map, generated as described in Section 4.1.2. From this map, the accounting area was delineated as all pixels classified to either T1 (broadleaved deciduous forest), T2 (broadleaved evergreen forest, T3 (coniferous forest) or S4 (temperate shrub heathland). The reference sites per forest type are located within each forest type on the habitat map and were selected based on expert review (Figure 44).

The selection of reference areas per forest class and for temperate shrub heathland was done by conjugating the most trustworthy field-validated source datasets, namely the Natura 2000 habitats dataset and the forest perimeter sub-dataset, to narrow down the habitat level 2 areas in the most ideal condition. The current version of the general island-wide forest inventory is highly incomplete in the validation of most of the attributes related to forest condition, most notably missing the status of degradation by abiotic factors (strong winds, landslides or tree cutting). From the outset, this limitation (which also limited the national-centric approach, as mentioned in Section 7.1.1.1) greatly reduced the pool of reference areas to the central and eastern sections of the island (Figure 44), where most of the N2000 habitats and forest perimeter areas are located. Since this forest condition assessment was derived from the habitat map L2, created only for São Miguel, no reference areas were selected from N2000 areas of other islands like Pico and Terceira to avoid error propagation, even though these islands have larger extents of potentially pristine Macaronesian endemic heaths (S43) and Macaronesian laurophyllous forest (T23) than São Miguel.

Whenever overlapping occurred between the two datasets, priority was given to selection of reference areas from the LIFE IP Natura 2000 over the forest perimeter for the same reasons listed in section 6.2.2, although any overlapping areas with clearly contradicting classifications were discarded. One additional reason for such predilection is that the forest perimeter dataset classifies any areas with over 10% of trees (over 5 m height) as a forest rather than a heathland, which neglects arborized bogs and arborized heaths with a relatively low % of tree cover but otherwise dominated by typical bogland and heathland species, respectively.

Initially, all vectorized areas of the forest perimeter were considered, while for the Natura 2000 dataset, only the habitats matching EUNIS T2 and S4 classes were considered. The former were filtered by species composition to isolate the areas better matching each of the desired classes: all areas classified as heathlands and dominated by heath species were assigned to S4; all areas classified as “Forest” and dominated by broadleaved deciduous tree species were assigned to T1; all areas classified as “Forest” and dominated by broadleaved evergreen tree species were assigned to T2; all areas classified as “Forest” and dominated by coniferous tree species were assigned to T3. All areas not fitting these groups or already covered by Natura 2000 S4 and T2 areas were excluded. From the remaining forest perimeter areas, only those better classified as “forest stand” in the “fragmentation” attribute were selected (over 0.5 ha of area and 20 m of width).

On a second step, the filtered polygons according to the criteria above were intersected with a standalone vectorized field-validated dataset of the areas influenced by abiotic factors to subtract the overlaps. The remaining polygons not affected by abiotic factors were intersected with the corresponding polygons of the same class in the habitat map level 2 to exclude areas with high likelihood of habitat misclassifications. At last, the remaining polygons were cleared of cropping artifacts and invalid geometries. A final visual confirmation with Google Earth Historic Imagery was done to remove any areas where the vegetation may have been disturbed very recently by tree cutting or slope movements.

Some pitfalls were unavoidable in this selection process. Due to the shortage of forest areas dominated by broadleaved deciduous species in São Miguel, almost no T1 areas became selectable, which would prevent the calculation of the FCI for this class. Therefore, all T1 areas from the general island-wide but incomplete and not fully validated forest inventory were included, if they were at least consistent with the habitat L2 map and free from abiotic factors. Likewise, due to the incompleteness of the island-wide forest inventory, both spontaneous (naturally occurring) and cultivated forests were considered alike without distinction because most of the forest within the perimeter dataset is cultivated. Furthermore, due to the widespread presence of invasive exotic species in the understory layers, it was not possible to only select reference areas representative of pure composition with autochthonous vegetation, as very few remain intact. Hence, the proposed FCI could not account for the aspects of forest degradation related to competition with exotic species in São Miguel.

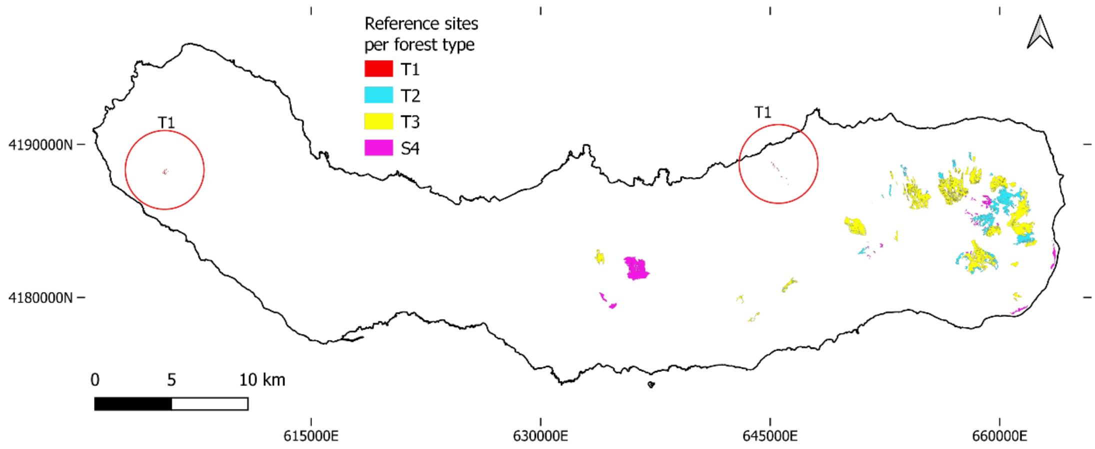


Figure 44: Locations of the forest reference sites per forest type and heathland in São Miguel.

The High-Resolution Layer Tree Cover Density (HRL-TCD) of CLMS (2018, 10 m) was used to filter the provided reference sites and aim for a more homogeneous set of reference sites per forest type considering their tree cover density (EEA 2020b) (Figure 45).

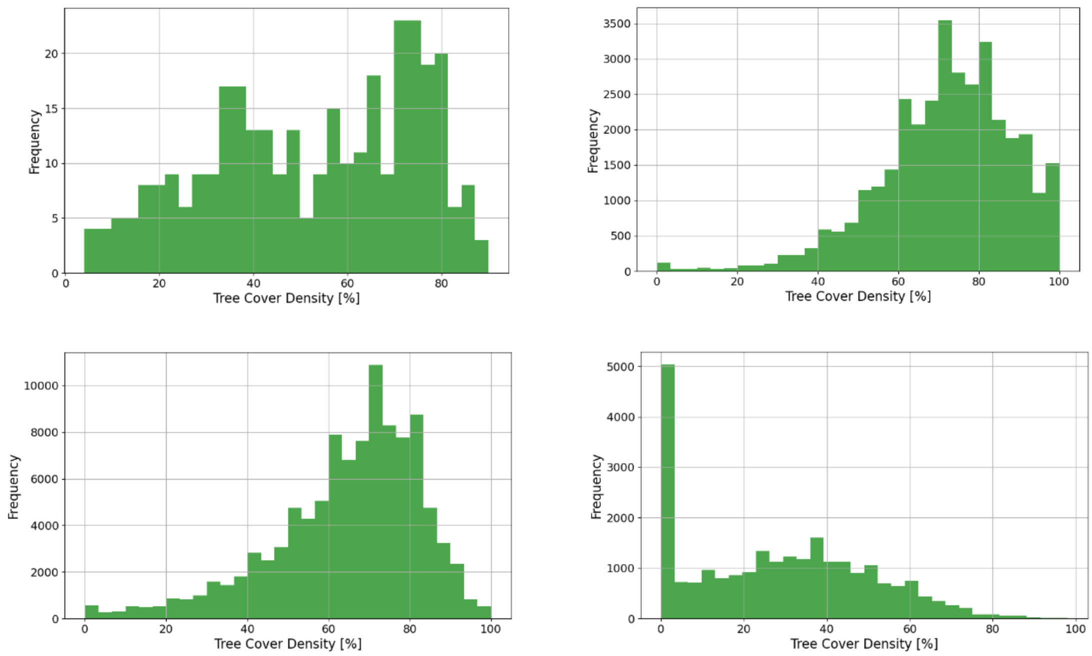


Figure 45: Distributions of tree cover density (TCD) within the provided forest reference sites per forest types: T1 (Broadleaved deciduous forest) [top-left], T2 (Broadleaved evergreen forest) [top-right], T3 (Coniferous Forest) [bottom-left] and S4 (Temperate shrub heathland) [bottom-right].

For classes T1, T2 and T3 (EUNIS habitats classified as forest), pixels within the provided reference sites were retained if their TCD was higher than or equal to 50% and lower than or equal to 100%. For class S4, reference sites with a broader range of TCD were decided to be allowed as reference sites since S4 classifies under EUNIS habitats for shrub-or heathland. Provided reference sites of S4 with a TCD lower than 20% were filtered out (Figure 46).

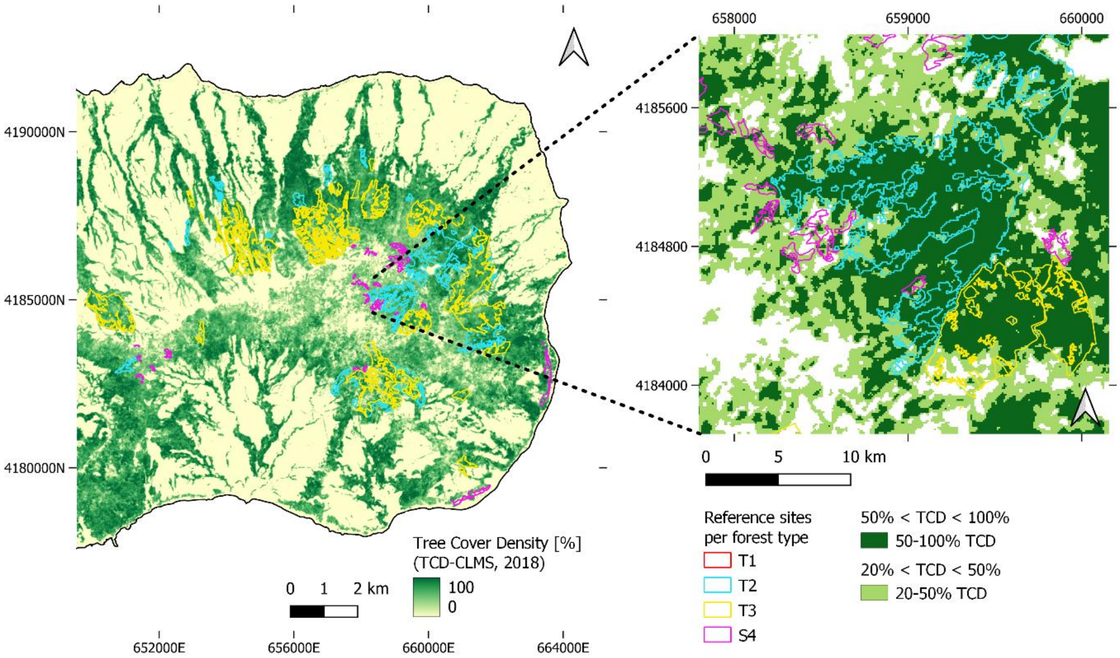


Figure 46: Illustration of the filtering process: filtering out areas from the provided forest reference sites where tree cover density is lower than 50% for T1, T2 and T3, and where tree cover density is lower than 20% for S4.

Table 33: Number of pixels within the provided forest reference sites after the filtering process illustrated in Figure 43.

EUNIS habitat L2	Forest type	Pixel count T1, T2 & T3 filtered for 50% < TCD < 100% S4 filtered for 20% < TCD < 100%
T1	Broadleaved deciduous forest	179
T2	Broadleaved evergreen forest	31502
T3	Coniferous forest	83817
S4	Temperate shrub heathland	15216

The pixels indicated in Table 33, after filtering based on TCD, were used for deriving the UpperReference (i.e., the 98th percentile of their distribution) for rescaling the condition variables to indicators. The EUNIS habitat map, developed at level 2, was used to indicate all area predicted to be either T1, T2, T3 or S4. All area for these classes that did not contain the reference sites (after filtering for their TCD), were used to derive the LowerReference (i.e., the 2nd percentile of their distribution) for rescaling. Then, all area used for UpperReference and LowerReference (i.e., the reference sites and all surrounding forest) were combined to create the accounting area, for which a forest condition index was generated (Figure 47).

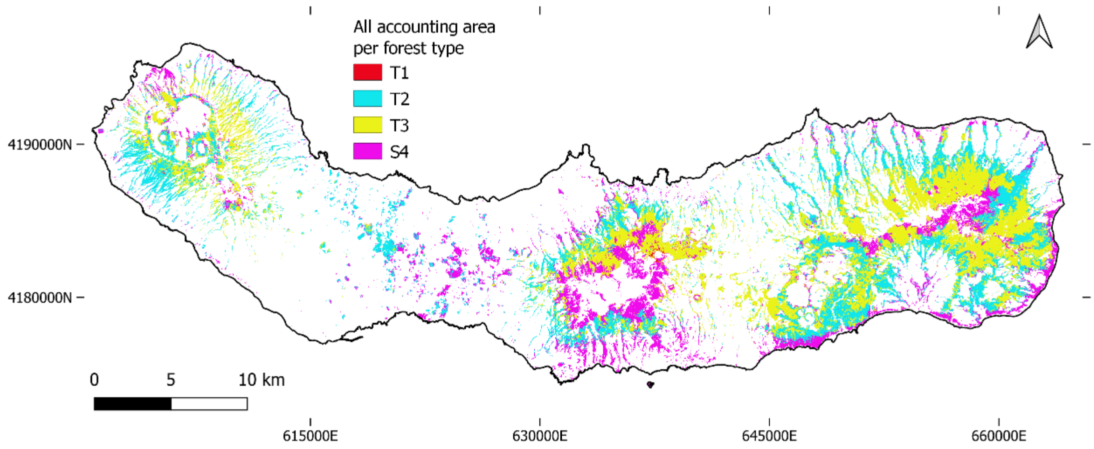


Figure 47: Illustration of all accounting areas per forest type: all area on São Miguel for which a forest condition index will be generated, comprising the areas used as forest reference sites and all remaining forest.

As described earlier, usually the year 2000 is taken as the 'reference year' to which all values of the forest condition variables are rescaled to retrieve forest condition indicator values. However, many datasets that qualify for the forest condition index calculation lack coverage or good mapping quality over the Azores. The AGB data only shows good quality starting from 2017. Therefore, 2017 was used as the reference year for all datasets.

The condition variables were rescaled to condition indicators based on the formula in Figure 40. The following tables (Table 34-Table 37) contain the yearly mean indicators per variable and per forest type. The evolution of condition indicator values should be approached with caution. It is important to understand that the evolution can only be analyzed per forest type and the value indicates the proximity towards the ideal condition (i.e., the closer to 1, the closer to the ideal condition) that was established in the year 2017. This means that if the condition variable in the reference year (i.e., 2017) was actually a very good reference, a low condition indicator value would suggest that the forest type has experienced a decrease in its condition relative to 2017 considering that variable. However, if the condition variable in the reference year was not a good reference, a high condition indicator value will indicate that the situation has not changed a lot. Thus, the raw forest condition variable accounts that are listed in section 13.6 of the Annex should be considered in analyzing the indicator accounts.

The yearly Net Primary Productivity (NPP) indicator shows that, in 2020, each forest type performed best (Table 34). After 2020, each forest drifts further away from reference NPP.

Table 34: Net primary productivity (NPP) condition indicator account per forest type for 2018-2023.

NPP	T1	T2	T3	S4
2018	0.35	0.43	0.42	0.49
2019	0.37	0.41	0.40	0.49
2020	0.44	0.45	0.45	0.52
2021	0.27	0.37	0.36	0.45
2022	0.29	0.34	0.32	0.41
2023	0.21	0.32	0.29	0.37

The Aboveground Biomass (AGB) indicator does not follow a similar pattern as the NPP indicator. This is because the NPP indicated the yearly net primary productivity (above and belowground) while the AGB origins from SAR data which measures the aboveground biomass stock, hence, not the yearly accumulation. For each forest type, the AGB indicator seems to remain quite stable through time (Table 35). It increases slightly between 2018 and 2023 with the best trend detected for Coniferous Forest (T3).

Table 35: Aboveground biomass (AGB) condition indicator account per forest type for 2018-2023.

AGB	T1	T2	T3	S4
2018	0.34	0.33	0.46	0.23
2019	0.37	0.33	0.49	0.24
2020	0.38	0.35	0.52	0.25
2021	0.38	0.35	0.51	0.25
2022	0.39	0.35	0.52	0.25
2023	0.39	0.35	0.52	0.25

The Normalized Difference Water Index (NDWI) is sensitive to changes in liquid water content of vegetation canopies. The NDWI indicator had small fluctuations over time, but these are quite negligible (Table 36). Note that the NDWI variable values were averaged per set of three years (e.g. NDWI of 2018 averaged over 2017, 2018 and 2019). This was done to smooth out seasonal effects throughout the years. Therefore, the NDWI will mostly become quite low when there are consecutive years with low water availability, since forests are not affected to a very high extent if there is only one year with drought events.

Table 36: Normalized difference water index (NDWI) condition indicator account per forest type for 2018-2023.

NDWI	T1	T2	T3	S4
2018	0.55	0.60	0.47	0.41
2019	0.58	0.61	0.48	0.43
2020	0.59	0.63	0.50	0.44
2021	0.52	0.63	0.50	0.42
2022	0.53	0.65	0.52	0.42
2023	0.56	0.65	0.53	0.42

Forest Connectivity data origins from CORINE Land Cover dataset, which is generated every 6 years (2000, 2006, 2012, 2018 and yet to be published 2024). Therefore, the FC indicator showed similar values per forest type 6 years in a row, from 2018 onwards (Table 37). The broadleaved deciduous forest (T1) is closest to the optimal condition for forest connectivity compared to the state in 2017. The other forest types showed a less desired state for forest connectivity compared to their reference condition.

Table 37: Forest connectivity (FC) condition indicator account per forest type for 2018-2023.

FC	T1	T2	T3	S4
2018	0.52	0.27	0.35	0.33
2019	0.52	0.27	0.35	0.33
2020	0.52	0.27	0.35	0.33
2021	0.52	0.27	0.35	0.33

2022	0.52	0.27	0.35	0.33
2023	0.52	0.27	0.35	0.33

Threatened Forest Bird Species Diversity data has been published only for 2008 and 2012. The data for 2012 is used from 2012 onwards. Therefore, the indicator values showed similar responses for the years between 2018 and 2023 per forest type (Table 38). For T2, T3 and S4, the TFBSD condition during 2018-2023 performed only half as good compared to 2017 (i.e., same values as 2012). Forest type T1 remained more stable in proximity to its ideal state.

Table 38: Threatened Forest Bird Species Diversity (TFBSD) condition indicator account per forest type for 2018-2023.

TFBSD	T1	T2	T3	S4
2018	0.91	0.52	0.61	0.52
2019	0.91	0.52	0.61	0.52
2020	0.91	0.52	0.61	0.52
2021	0.91	0.52	0.61	0.52
2022	0.91	0.52	0.61	0.52
2023	0.91	0.52	0.61	0.52

As explained earlier, the only available spatial soil organic carbon (SOC) raster data for São Miguel was retrievable from ISRIC SoilGrids, which does not contain annual maps but rather a continuously updated dataset. Hence, the SOC condition indicator remained a single value through time per forest type (Table 39) but still influences the magnitude of the final forest condition index. The values show that forest types T2 and T3 contain related SOC values far away from their reference condition.

Table 39: Soil organic carbon (SOC) condition indicator account per forest type for 2018-2023.

SOC	T1	T2	T3	S4
2018	0.89	0.36	0.46	0.69
2019	0.89	0.36	0.46	0.69
2020	0.89	0.36	0.46	0.69
2021	0.89	0.36	0.46	0.69
2022	0.89	0.36	0.46	0.69
2023	0.89	0.36	0.46	0.69

The condition indicators in Table 34 to 39 were used in the arithmetic average calculation described in section 7.2.1 to derive the yearly forest condition index (FCI) maps aggregating all selected indicators. This integration step is essential to:

- Provide a single, interpretable metric of overall forest condition per forest type.
- Aligning with the SEEA EA reporting structure.
- Enable meaningful comparisons across forest types and different years.

7.3. Forest Carbon Accounting

Forests are essential to the global carbon balance, functioning as both carbon sinks and carbon sources. While forests have long been studied for commercial use and land-use planning, carbon accounting is a relatively recent focus, driven by the urgent need to quantify carbon stocks, emissions, and removals to support climate change mitigation and meet greenhouse gas (GHG) reporting requirements (Calvo Buendia et al. 2019). Deforestation and forest degradation currently contribute approximately 10–12% of global anthropogenic emissions (IPCC 2007). At the same time, the forestry sector is recognized as having over 50% of the global mitigation potential among nature-based solutions (Griscom et al. 2017). This underscores the critical role of carbon accounting in forest ecosystems for understanding and managing their contributions to ES and climate regulation.

Accurate carbon accounting is not only essential for integrating forests into climate policies at local, national, and global levels, it enhances transparency in climate commitments and helps ensure that vital ecosystem services — such as biodiversity, water regulation, and soil health — are properly valued in environmental and economic planning.

Carbon accounting should be approached from different perspectives:

- **Stock accounting**: Establishing the terrestrial carbon stock of a given area as an average carbon stock per unit area for specific land use, cover type, or forest type at a certain time.
- **Carbon Fluxes accounting**: Measuring the amount of carbon emitted or sequestrated due to land or forest management practices, deforestation, or degradation. This is necessary to assess the scale of emissions from the forestry sector relative to other sectors.

In recent years, significant advancements have been made in carbon accounting methods; however, despite these improvements, the influence of intrinsic forest dynamics on carbon accounting is still not fully understood. Variables such as spatial forest heterogeneity, thinning practices, clear-cutting, deforestation, forest degradation, habitat restoration, harvest rotation cycles, and natural forest growth all have significant impacts on the estimation of carbon stocks and emissions over time (Fargione et al. 2018, Lutz et al. 2018, Pugh et al. 2019). Addressing these complexities is crucial for improving the reliability of carbon accounting frameworks and their role in climate change mitigation strategies.

In this Section, two distinct approaches are presented for assessing carbon stocks and carbon fluxes. The first approach is based on RS, using advanced methodologies integrating multiple satellite data sources to capture the spatial variability of carbon stocks and to directly estimate emissions from spatial datasets. The second approach involves estimating biomass changes over time by calculating Gross Primary Productivity (GPP) through the Light Use Efficiency (LUE) model. This method provides insights into carbon fluxes by linking vegetation productivity with environmental conditions. For both TS, examples are given of carbon stock mapping and carbon flux accounting using these complementary methodologies, highlighting their strengths and applications in ecosystem carbon assessments.

7.3.1. Carbon stock maps using remote sensing derived products

Carbon pools are ecosystem components that either sequester, store or release carbon, traditionally categorized into five main types: living above-ground biomass (AGB), living below-ground biomass (BGB), dead organic matter (DOM) in wood, DOM in litter, and soil organic matter (SOM). Stock accounting sums carbon pools at a single point in time, decisions on which pools to include depending on data availability, measurement costs, and the required level of accuracy (MacDicken 1997). Some key definitions are needed to clarify the methodology (Brown 1997, Watson 2009):

- **Forest Above-Ground Biomass (AGB):** Organic matter resulting from primary production through photosynthesis, minus consumption through respiration and harvest. The AGB carbon pool includes all living vegetation above the soil, such as stems, stumps, branches, bark, seeds, and foliage. Biomass assessments provide insights into forest structure and function, helping estimate timber, fuel, and fodder quantities (Brown 1997). Since approximately 50% of dry forest biomass is carbon (Westlake 1966), these assessments also serve as indicators of carbon stocks.
- **Below-Ground Biomass (BGB):** The BGB carbon pool consists of the biomass contained within live roots.
- **Carbon Flux Accounting:** A primary approach to emissions accounting, which estimates the net balance of carbon additions to and removals from a carbon pool.
- **Emissions Accounting:** Assesses the net greenhouse gas emissions to the atmosphere resulting from forest activities.
- **Stock Accounting:** Measures the amount of carbon stored in forest ecosystems at a single point in time.
- **Carbon Sink:** A carbon pool where more carbon flows in than out, such as forests that sequester carbon through tree growth and biological processes.
- **Carbon Source:** A carbon pool where more carbon flows out than in. Forests can act as a net source of carbon due to processes like decay, combustion, and respiration.

This Section presents the methods used for carbon stock and emissions accounting using remote sensing-derived products. We propose an unsupervised approach requiring little to no field data for calibration. This approach integrates various sources of optical and radar satellite data, primarily from space agency repositories, with most datasets available free of charge from the European Space Agency's (ESA) Copernicus platform.

The methodology used by SarVision for carbon accounting encompasses both carbon stock mapping and carbon flux calculations over time, considering structural changes in forest ecosystems. The process involves five key steps:

1. **Stratification of the landscape** into major structural vegetation structures/types.
2. **Analysis of forest structural changes** occurring during the emission calculation period, mainly Deforestation and degradation.
3. **Mapping of above-ground biomass (AGB)** for the stratified vegetation cover types.
4. **Carbon estimation** derived from the AGB map.
5. **Calculation of carbon flux** between two dates for each vegetation type.

Table 40 summarizes the steps and data used during the carbon flux mapping process. This approach (Figure 48) ensures a comprehensive assessment of forest carbon dynamics, since it combines information on vegetation structure and forest change into the calculation of both the carbon stock and the carbon flux.

Table 40: Summary of products, remote sensing data, map resolutions and reference dates. Information applies for both study sites.

Component	Product	RS data	Resolution (m)	# Maps	Reference year
Forest structural Mapping	Forest Structural Maps	Sentinel-2	10	2	2021 2023
Forest structural change	Yearly deforestation and degradation maps	Sentinel-1	10	3	2021 2022 2023
Stratified Biomass mapping	Use of SATVI vegetation index	Sentinel -2	10	2	2021 2023
	Use of h100 vegetation canopy height	GEDI Lidar	25m footprint		2020-2021-2022
Carbon Stock Mapping	Use of Biomass maps	-	10	2	2021 2023
Carbon Flux mapping	Use of carbon maps	-	10	1	Difference between years

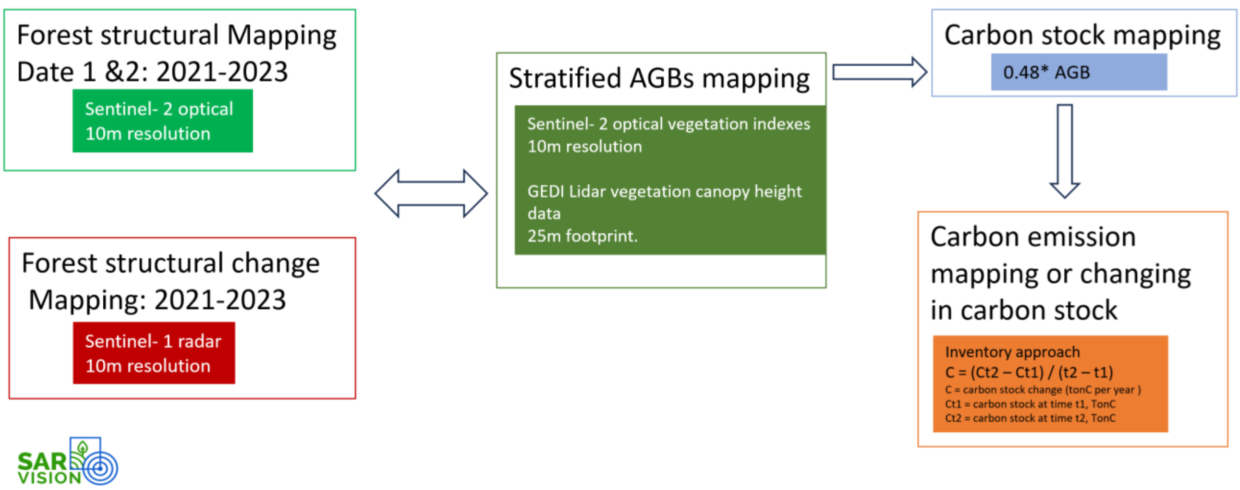


Figure 48: SarVision methodology for the unsupervised carbon stock and emission mapping.

The first step in carbon stock mapping is the stratification of the landscape into distinct vegetation structural types, which are directly correlated with above-ground biomass (AGB) levels. The High Carbon Stock (HCS) Vegetation Classification System (Figure 49) illustrates how accurately delineation of high/low density forests, open canopy forests, woodlands, shrublands, grasslands, wetlands and agricultural areas, is crucial for AGB mapping.

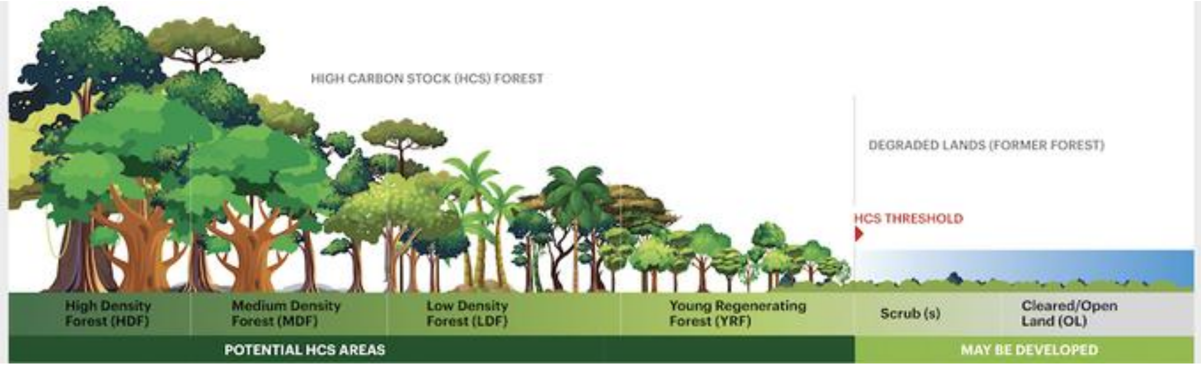


Figure 49: Biomass stratification system proposed by the High Carbon Stock (HCS) classification system. (https://highcarbonstock.org/what-is-the-high-carbon-stock-approach/)

The structural classification of vegetation into different structural types follows a combination of supervised and unsupervised classification of remote sensing images. This approach integrates optical and radar imagery, incorporating field data when available to enhance accuracy. Additionally, high-resolution data is utilized for both image classification and final map validation. The flowchart in Figure 50 outlines the sequential steps, including image preprocessing, image processing, classification, post-classification refinement, and validation (Hoekman et al. 2010b, 2010a). Both for the São Miguel and Peloponnese TS, forest structural maps were created, one for 2021 and for 2023.

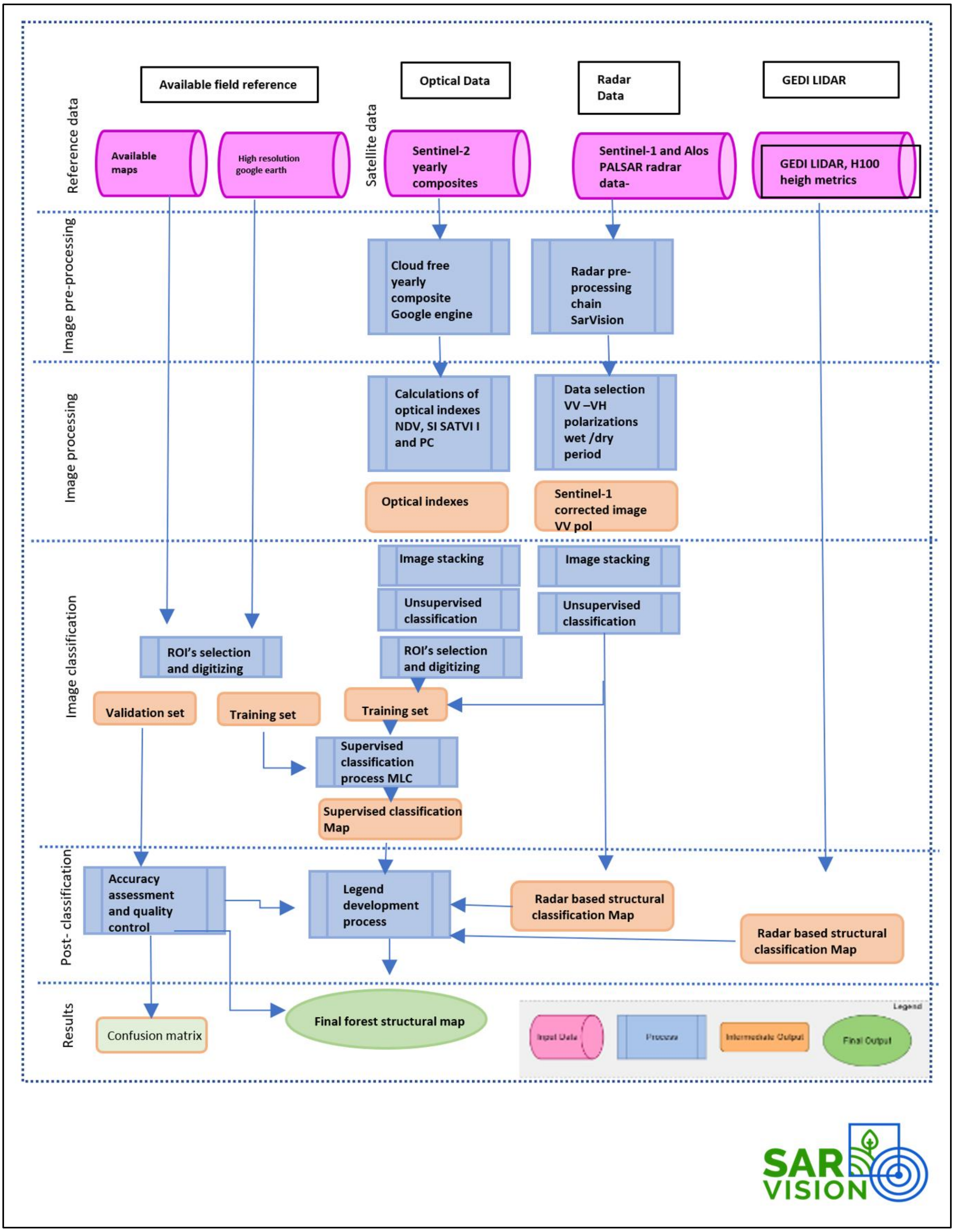


Figure 50: Flowchart of the forest structure mapping methodology.

The second step in carbon stock mapping is the evaluation of forest changes in the study area. This step is crucial for assessing forest structural changes caused by deforestation and forest degradation, which lead to variations in carbon stock over time. This analysis is conducted using the SarSentry algorithm, a state-of-the-art deforestation and forest degradation detection system based on the analysis of Sentinel-1 radar time series (Hoekman et al. 2020).

The output of this algorithm is a thematic map that identifies areas affected by deforestation and degradation. The SarSentry process involves (Figure 51):

- Downloading and preprocessing all available Sentinel-1 radar images for the study period (in this case, 2017–2023).
- Processing the radar data to generate thematic deforestation and degradation maps.
- Evaluating and refining these thematic products into final deforestation and degradation maps.

In some cases, applying the SarSentry algorithm is not straightforward, particularly in regions with strong seasonal vegetation changes. In such instances, the algorithm is applied to yearly radar composites, which help average out seasonal variations and focus solely on significant structural forest changes, such as deforestation or other types of forest loss.

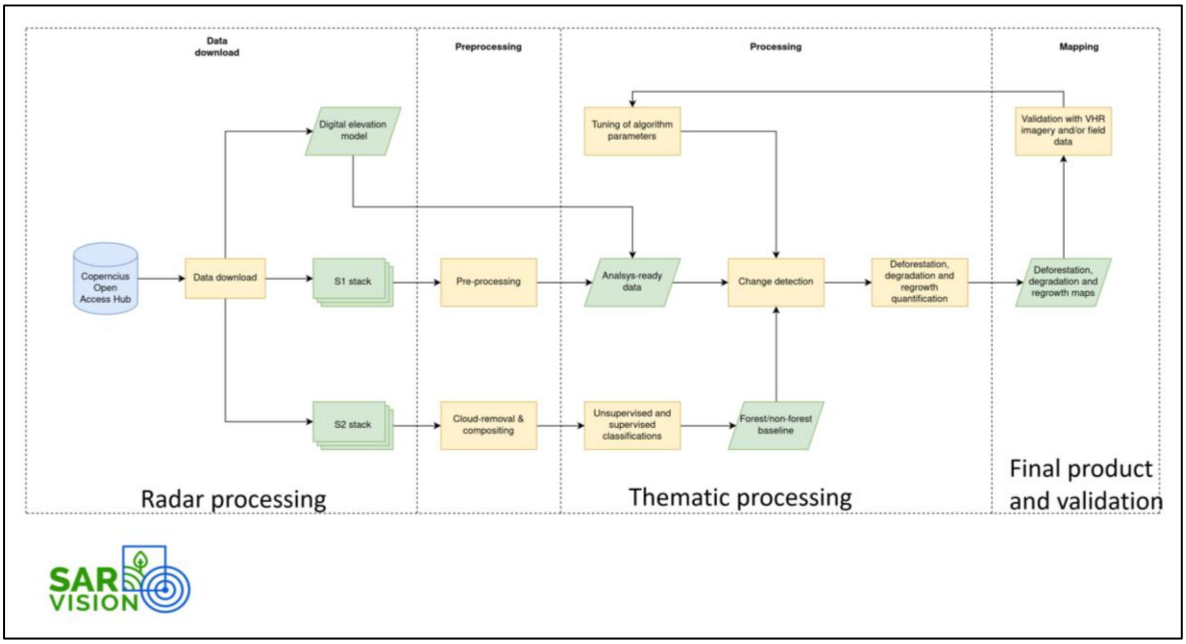


Figure 51: Main steps involved in the processing of Sentinel-1 radar images during the creation of thematical deforestation and forest degradation data using the SarSentry system.

The third stage in the carbon stock mapping process is the creation of biomass maps for each period. This process is complex and requires multiple data sources and several integration steps, including:

1. Input Data Sources:
    - Forest structural map (from Step 1).
    - Remote sensing images:
        - Optical vegetation index images (for areas with steep slopes).
        - L-Band ALOS PALSAR radar images (for flat terrain).
    - GEDI LiDAR vegetation canopy height data (point-based vegetation height measurements).
2. Data Integration Using the SarCarbon Algorithm (an inhouse biomass mapping algorithm used internally in SarVision):

- The SarCarbon algorithm integrates the point-based GEDI LiDAR data with raster-based remote sensing images, generating an intermediate raster-based canopy height map.
- This height map is then converted into biomass data using allometric equations that relate canopy height to biomass. (Hansen, 2012)

3. Validation:
    - The final biomass map is validated using field data collected for the corresponding year.

This approach ensures an accurate spatial representation of biomass distribution across the landscape. A visual representation of these steps can be seen in Figure 52.

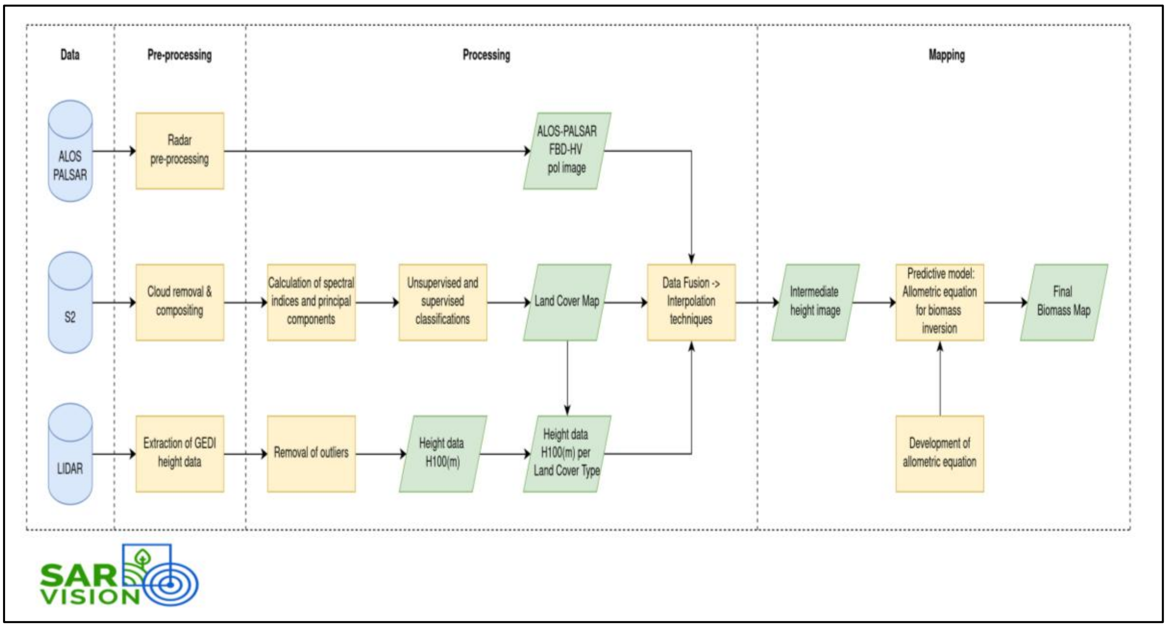


Figure 52: Steps for biomass mapping

For both TS, field data on vegetation heights, reference biomass estimates, and allometric equations, which are essential for calibrating the SarCarbon system in the biomass mapping process, were not available. To address this limitation, default aboveground biomass values from the IPCC Guidelines (Volume 4, Chapter 4, Tables 4.7–4.8) were utilized (Eggleston et al. 2006). Specifically, biomass values for Temperate Oceanic and continental European Natural and Planted Forests were applied to estimate forest biomass in the study areas. This approach provides standardized biomass estimates based on internationally recognized methodologies, ensuring comparability and consistency in the carbon accounting process.

The final biomass maps are raster tiff float maps at 10m resolution (grayscale), where each pixel value represents the calculated Above-Ground Biomass (AGB) in tons per hectare for that specific location. If the actual biomass value for a pixel needs to be retrieved, the data must be normalized to the pixel resolution. In cases where the biomass stock for an entire forest structural type needs to be calculated, the biomass values from all pixels within that class should be summed and corrected accordingly to ensure accurate estimates.

Carbon data can be easily derived from the biomass maps by multiplying each map by the conversion factor for the corresponding cover types (Table 4.3, IPCC Volume 4, Chapter 4) (Matthews 1993, Andreae and Merlet 2001, Gayoso et al. 2002, McGroddy et al. 2004, Eggleston et al. 2006). After generating the biomass maps for both 2021 and 2023 at each TS, the carbon maps were done. The final carbon maps are float raster maps (grayscale), where each pixel value represents the estimated tons of carbon per hectare (TonC/ha) for that area. Values should be normalized by pixel area to ensure accuracy.

Finally, for estimating carbon emissions by carbon stock changes, one of the IPCC-approved methodologies for flux accounting is the inventory approach (Penman et al. 2003). This method measures the difference in carbon stocks between two points in time, also known as periodic accounting or the stock-difference approach, according to the following formula:

$$\Delta C = \sum \quad (Ct2 - Ct1)/(\text{t2-t1})$$

C = carbon stock change, tonnes C per year
Ct1 = carbon stock at time t1, tonnes C
Ct2 = carbon stock at time t2, tonnes C

By measuring stock changes over time, this approach allows for the assessment of carbon fluxes across large areas with diverse species compositions and site conditions. It is a widely used method for estimating forest carbon dynamics, particularly in national greenhouse gas inventories and land use change assessments. For the two TS, the estimated carbon maps from 2021 were subtracted from the maps of 2023.

7.3.2. Carbon Flux mapping using GPP

The method applied by SarVision can estimate the carbon sequestration by subtracting the biomass stock at the end of 2021 from the biomass stock at the end of 2023. VITO adopts an alternative approach to defining carbon sequestration by estimating Gross Primary Productivity (GPP) using the Light Use Efficiency (LUE) concept. This methodology has also been implemented in the operational Copernicus Land Monitoring Service (CLMS), where GPP is derived at a 300 m resolution using Sentinel-3 data (EC directorate-general JRC). For this specific use case, the approach has been adopted to Sentinel-2 to allow providing 10 m resolution information on GPP. Additionally, the LUE term in the GPP estimation formula (see Section "Calculation of GDMP/DMP") has been re-calibrated for specific land cover types.

Since the exact methodology of the operational CLMS service was applied in this approach, we could not integrate the classified habitat maps of Peloponnese and São Miguel. Thus, the classification of land cover types followed the method of CLMS and not the classification of the habitat maps. Unfortunately, this implies that the results of the carbon accounts cannot be brought in direct relationship with the results of the habitat maps, the extent maps and the forest condition accounts for the two TSs. The calculation of annual GPP was only done for wetland, broadleaf forest, coniferous forest, evergreen forest and low woody vegetation. Table 41 highlights from which data source the corresponding classes were retrieved.

Table 41: Overview of datasets used to map the corresponding ecosystem type class.

Ecosystem Type Class	Data source
Wetland	Corine Land Cover (CLC) classification (EEA 2020a) & HRL water & wetness layer 2018 (EEA 2020c)
Broad leaf forest	HRL forest 2021 (EEA 2025)
Coniferous forest	HRL forest 2021 (EEA 2025)
Evergreen forest	CLC+ backbone 2021 (EEA, 2025)
Low woody vegetation	CLC+ backbone 2021 (EEA, 2024b)

The forest and wetland maps for the two study areas can be seen in Figure 53.

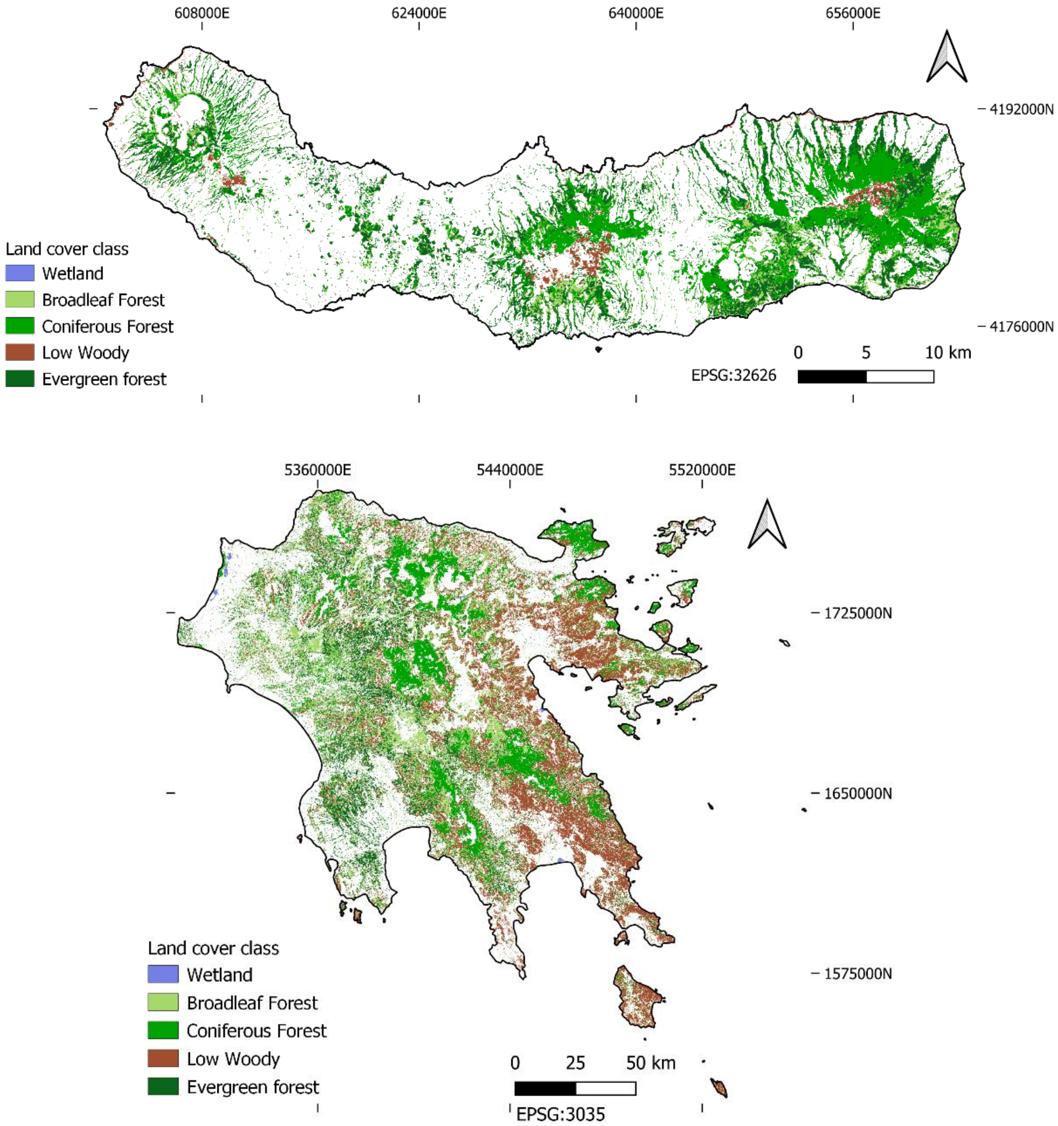


Figure 53: Ecosystem types for which the carbon sequestration between end of 2021 and end of 2023 are to be generated in São Miguel [above] and Peloponnese [below] (sources in Table 41).

The method below is used to go from annual GPP maps towards annual produced ANPP (aboveground carbon accumulation) maps, demonstrating how the carbon sequestration between end of 2021 and end of 2023 can be estimated with direct growth information instead of carbon stock data.

The Gross Dry Matter Productivity (GDMP) can be calculated according to the following Penman Monteith equation (Swinnen et al. 2021), where:

$$GDMP = R\ x\ \mathcal{E}c\ x\ fAPAR\ x\ \mathcal{E}LUEc\ x\ \mathcal{E}T\ x\ \mathcal{E}CO2$$
$$GDMP = fAPAR\ x\ \mathcal{E}LUEc\ x\ R\ x\ \mathcal{E}(T, CO2)$$
$$GDMP = fAPAR\ x\ \mathcal{E}LUEc\ x\ GDMPmax$$

- R the total shortwave incoming radiation (0.2 µm – 3.0 µm).
- $\mathcal{E}c$ the fraction of PAR (photosynthetic absorbed radiation, 0.4 µm – 0.7 µm).
- fAPAR the PAR-fraction absorbed (PA) by green vegetation.
- $\mathcal{E}LUEc$ the maximum light uses efficiency (DM = dry matter) under optimal conditions (land cover specific).
- $\mathcal{E}T$ the normalized temperature effect and $\mathcal{E}CO2$ normalized CO2 fertilization effect.

Dry Matter Productivity (DMP) can be calculated with the formula:

$$DMP = GDMP\ x\ CUE$$

where CUE (carbon use efficiency) taken as 0.5 as a constant factor specified in Table 3 in ATBD by (Swinnen et al. 2021).

In practice, meteorological inputs (R and T) are required and retrieved from AgERA5 based on ECMWF reanalysis (Boogaard et al. 2020). The meteorological inputs combined with the variable CO_2-level are input for the reduced Penman-Monteith model, with LUE of 1 to generate max GDMP (Lan and Keeling 2025):

$$GDMPmax = R\ x\ \mathcal{E}(T, CO2) = R\ x\ \mathcal{E}c\ x\ \mathcal{E}T\ x\ \mathcal{E}CO2$$

Daily GDMPmax values are averaged to get the mean dekad (10-daily or different for last dekad depending on the length of the month) GDMPmax, and this image is resampled (with bilinear interpolation) to the same resolution and extent of the fAPAR image, derived from Sentinel-2 (10m) (VITO 2024). Cloud masking and Whittaker smoothing have been applied to the fAPAR data to ensure continuous dekad information.

The biome-specific $\mathcal{E}LUEc$ (Light Use Efficiency coefficient) is calibrated specifically for wetland and forest types. This was done by retrieving the data for the flux tower Eddy Covariance (EC) data per land cover class within Europe. Since too little flux tower data was available for low woody vegetation and evergreen forests, the LUE term calibrated for coniferous forests was applied, as this land cover class is expected to exhibit the most similar productivity characteristics in these regions. In contrast, wetlands and broadleaved forests were assigned distinct LUE terms. In this method, the flux data is considered as a ground truth measurement of GPP. The $\mathcal{E}LUEc$ per class was optimized by calibrating the GPP model's output ($GDMP(10) = fAPAR(10)\ x\ \mathcal{E}LUEc\ x\ GDMPmax(10)$) against the flux tower GPP measurements. Table 42 shows the optimized $\mathcal{E}LUEc$ for the different land cover classes.

Table 42: Calibrated Light Use Efficiency coefficient per ecosystem type class.

Ecosystem type class	$\mathcal{E}LUEc$
Wetlands	1.80
Broadleaf forest	2.28
Coniferous forest	2.58
Evergreen forest	2.58
Low woody vegetation	2.58

When the optimized $\mathcal{E}LUEc$ is found, the GDMPmax can be multiplied with the LUE per land cover class and the 10-daily fAPAR that results in dekadal GDMP results. To get annual GDMP maps, the dekadal GDMP values must be multiplied by the number of days in each dekad and summed. Then, the GDMP values contain the unit of kgDM/ha/year. The annual GPP map is created by formula:

$$GPP = GDMP\ x\ 0.5\ x\ 0.1 = DMP\ x\ 0.1$$

with 0.1 the conversion factor to go from kgDM/ha/year to gC/m^2/year for GPP. Annual GPP maps for São Miguel and Peloponnese are given in Figure 54 and Figure 55, respectively.

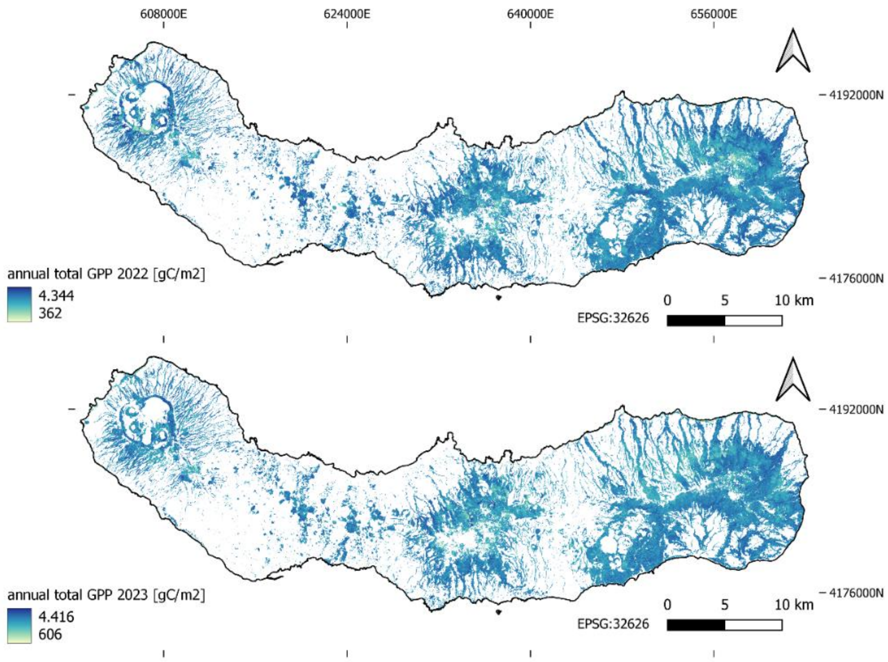


Figure 54: Annual Total GPP for 2022 and 2023 in São Miguel for classes: wetland, broadleaf forest, coniferous forest, low woody and evergreen forest.

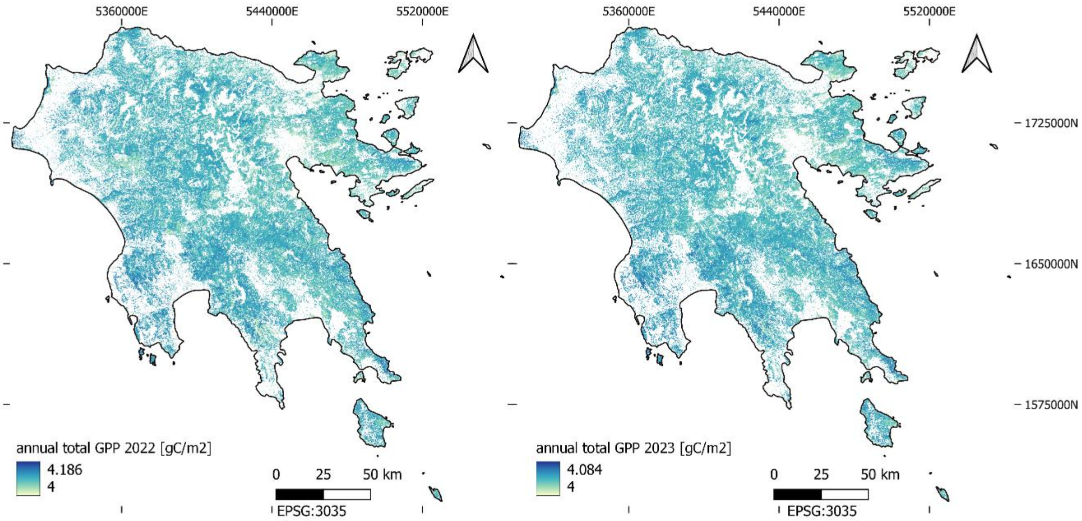


Figure 55: Annual Total GPP for 2022 and 2023 in Peloponnese for classes: wetland, broadleaf forest, coniferous forest, low woody and evergreen forest.

Next, the Net Primary Productivity (NPP, net increment in carbon accumulation):

$$NPP = GPP\ x\ 0.5$$

Aboveground net primary productivity (ANPP, net accumulation of carbon aboveground) and belowground net primary productivity (BNPP, net accumulation of carbon belowground) are derived by (IPCC Task Force on National Greenhouse Gas Inventories, 2003a):

$$ANPP = \frac{NPP}{1 + R}$$

$$BNPP = NPP - ANPP$$

with R the root-to-shoot (R:S) ratio, specific for the vegetation type under environmental conditions (IPCC Task Force on National Greenhouse Gas Inventories 2003). The IPCC has listed R values by broad categories in table 3A.1.8 of Annex 3A.1 for the IPCC Good Practice Guidance for LULUCF. The categories by IPCC in the table do not correspond one-to-one with the land cover classes for which the GPP was generated. The IPCC category 'Shrubland' can be taken for land cover class 'Low Woody'. Forest is divided into IPCC categories 'Other broadleaf forest' or 'Conifer Forest/plantation'. The R values can be retrieved following a decision rule depending on the aboveground biomass (AGB in ton/ha) for the vegetation in question. The generated AGB maps of SarVision were used to choose the best suitable R value per pixel. Lastly, the IPCC table only contains 'Tidal marsh' as a category approaching the 'Wetland' class. Since 'Tidal marsh' is a very specific wetland class, and the wetland ecosystem types in either São Miguel or Peloponnese resemble rather wetlands with very little rooting system (e.g. *Sphagnum* peats in blanket bogs), the R value for the land cover class 'Wetland' was decided on based on expert opinion. Table 43 shows R:S values per land cover class.

Table 43: Root-to-shoot ratio values used to derive aboveground carbon accumulation (ANPP) from net primary productivity (NPP) per ecosystem type class.

IPCC category	Land cover class	AGB [ton/ha]	(mean) R
Conifer forest/plantation	Coniferous Forest	< 50	0.46
		50-150	0.32
		> 150	0.23
Other broadleaved forest	Broadleaved Forest Evergreen Forest	< 75	0.43
		75-150	0.26
		> 150	0.24
Shrubland	Low Woody	/	2.83
/	Wetland	/	3.00

In practice, the NPP raster maps of 2022 and 2023 were converted into ANPP (& BNPP) by overlaying them with the ecosystem type classes and applying the decision rules for R by comparing to the AGB raster maps of 2021 and 2023, respectively, per land cover class. Lastly, the ANPP and BNPP maps still containing gC/m^2/year units, were converted to tonC/ha by multiplying by 0.01. Once the ANPP map of 2022 and 2023 are created, their values were summed pixel-wise to create the total accumulation of aboveground carbon in the biomass for period between end of 2021 and end of 2023. This resulting map was compared to the carbon sequestration map of SarVision.

8. Results

8.1. Ecosystem extent mapping

Ecosystem extent mapping is essential for understanding and managing ecosystem services, as it provides a spatial representation of different ecosystems and their distribution. By accurately mapping ecosystem extents, policymakers, conservationists, and land managers can assess and monitor ecosystem condition changes over time and identify areas of high ecological value. This information is crucial for sustainable land-use planning, biodiversity conservation, ecosystem restoration, climate change adaptation, and carbon sequestration efforts. Additionally, ecosystem extent mapping forms the base for ecosystem service accounts such as water regulation, soil fertility, and carbon storage, ensuring their protection and sustainable use for future generations.

8.1.1. National-Centric Approach

Both TS adhered to a nationally focused methodological framework. The primary distinction lays in the availability of national-level data for each site, which in turn influenced the integration of additional CLMS datasets. This Section presents the results derived from the methodology outlined in Section 7.1.1.

8.1.1.1. Peloponnese

The availability of national data for the Peloponnese was limited, with only a single dataset (N2K) partially covering the peninsula. Given this constraint, along with the region's comparably large size, multiple CLMS datasets were integrated to enhance the existing national data, ensuring the creation of a comprehensive ecosystem extent map aligned with the European Ecosystem Typology. To assess how each approach contributed to mapping towards the ETA, statistics were extracted at Level 1, as presented in Table 44.

By mapping the ecosystem extent only using national data, most of the area remained unmapped as seen under the ‘no data’ category. The ‘National only’ map cannot be directly compared to the other two mapping approaches due to its partial coverage.

The Level 1 ecosystem extent maps show similar overall results between the two mapping approaches, National & CLMS’ and ‘CLMS only.’ This similarity arises from the fact that the ‘National & CLMS’ map was generated by supplementing the national dataset with the ‘CLMS only’ map in areas where national data was absent or by using a combination of CLMS datasets to further refine and disaggregate national data into lower ETA classes. Both mapping approaches contain areas classified as "no crosswalk," which were filled with CLMS data that could not be directly linked to the ETA (mentioned in Section 4.1.1.2 section b). This category is more extensive in the ‘CLMS only’ approach, as the presence of national data allows for the allocation of these areas to an ETA class. The inability to assign these classes to the ETA stems from the fact that they originate from land cover datasets whose classifications do not align with the land use categories defined by the ETA.

Overall, the National & CLMS and CLMS only maps show very similar Level 1 ecosystem extents. However, class 5. 'Heathlands and shrub' are more expansive under the former approach. The National mapping classifies large areas as sclerophyllous vegetation, whereas CLMS designates these same areas as forest. This difference stems from variations in classification definitions, creating challenges in direct comparisons. The N2K map follows the EUNIS classification, which categorizes habitats, while CLMS data focuses on land use. As a result, mapping EUNIS classes to the ETA is not suitable for certain categories, making CLMS a better fit for these definitions.

Table 44: Results of the ecosystem extent delineation at level 1 of the European Ecosystem Typology using three mapping approaches for the Peloponnese TS.

European ecosystem typology (level 1)	National only (ha)	National & CLMS (ha)	CLMS only (ha)
1. Settlement and other artificial areas	3 545	74 705	75 115
2. Cropland	33 750	416 001	399 959
3. Grassland	435	117 322	133 401
4. Forest and woodlands	149 130	821 407	835 387
5. Heathlands and shrub	111 912	368 869	298 570
6. Sparsely vegetated ecosystems	2 934	6 387	5 297
7. Inland wetlands	1 403	2 153	2 018
8. Rivers and canals	11	3 385	3 733
9. Lakes and reservoirs	446	2 894	2 650
10. Marine inlets and transitional waters	1 483	1 647	2 078
11. Coastal beaches, dunes and wetlands	3 722	5 784	4 825
No data	1 825 313	-	-
No crosswalk	-	313 532	371 052
Total	**2 134 085**	**2 134 085**	**2 134 085**

Looking closely at the hierarchical levels that could be mapped (Table 45), the National only approach is obviously restricted spatially as mentioned before, with most area mapped to No data (85.5%). However, of the area mapped, most could be mapped down to level 3. The use of CLMS greatly improves the level of detail that can be crosslinked and mapped. The National & CLMS approach resulted in a higher proportion of classes mapped to level 3 (34%) than the CLMS only approach (24%). The CLMS & National approach also allowed for a bit more area to be crosswalked to the ETA (85%) in comparison to the CLMS only approach (83%).

Table 45: Results of ecosystem extent delineation to all levels of the European Ecosystem Typology using three mapping approaches for the Peloponnese TS.

	Level	ha	Percentage (%)
Step 1 National only	1	3 829.6	0.2
	2	73 833.2	3.5
	3	231 109.4	10.8
	No data	1 825 312.9	85.5
	Grand total	**2 134 085**	**100**
Step 2 National & CLMS	1	13 418.2	0.6
	2	1 092 423.2	51.2
	3	714 711.6	33.5
	No crosswalk	313 532.1	14.7
	Grand total	**2 134 085**	**100**
Step 3 CLMS only	1	33 908.6	1.6
	2	1 218 248.3	57.1
	3	510 875.8	23.9
	No crosswalk	371 052.3	17.4
	Grand total	**2 134 085**	**100**

The mapping outcomes for each of the three approaches discussed above are shown below. The 'National only' map, shown in Figure 1Figure 56, shows areas mapped using the N2K dataset and crosswalked to the ETA, while the black areas represent regions with no available data. Figure 57 illustrates the 'National & CLMS' map, where certain N2K-mapped areas were further refined using CLMS data — for example, class 2.3 (Permanent Crops) was disaggregated into 2.3.1 (Olives). However, since most N2K-mapped areas were already crosswalked to Level 3, only minimal additional CLMS data was required for further refinement. Instead, CLMS data proved particularly valuable in filling previously unmapped areas. Figure 58 presents the 'CLMS only' map which was developed based only on CLMS data. The legends of each of the figures show the ETA classes that were crosswalked. In Figure 57 and Figure 58, the classes that could not be crosslinked to the ETA have been subdivided into CLCplus Backbone, Priority Area Mapping and IMD, reflecting the datasets from which these classes were derived.

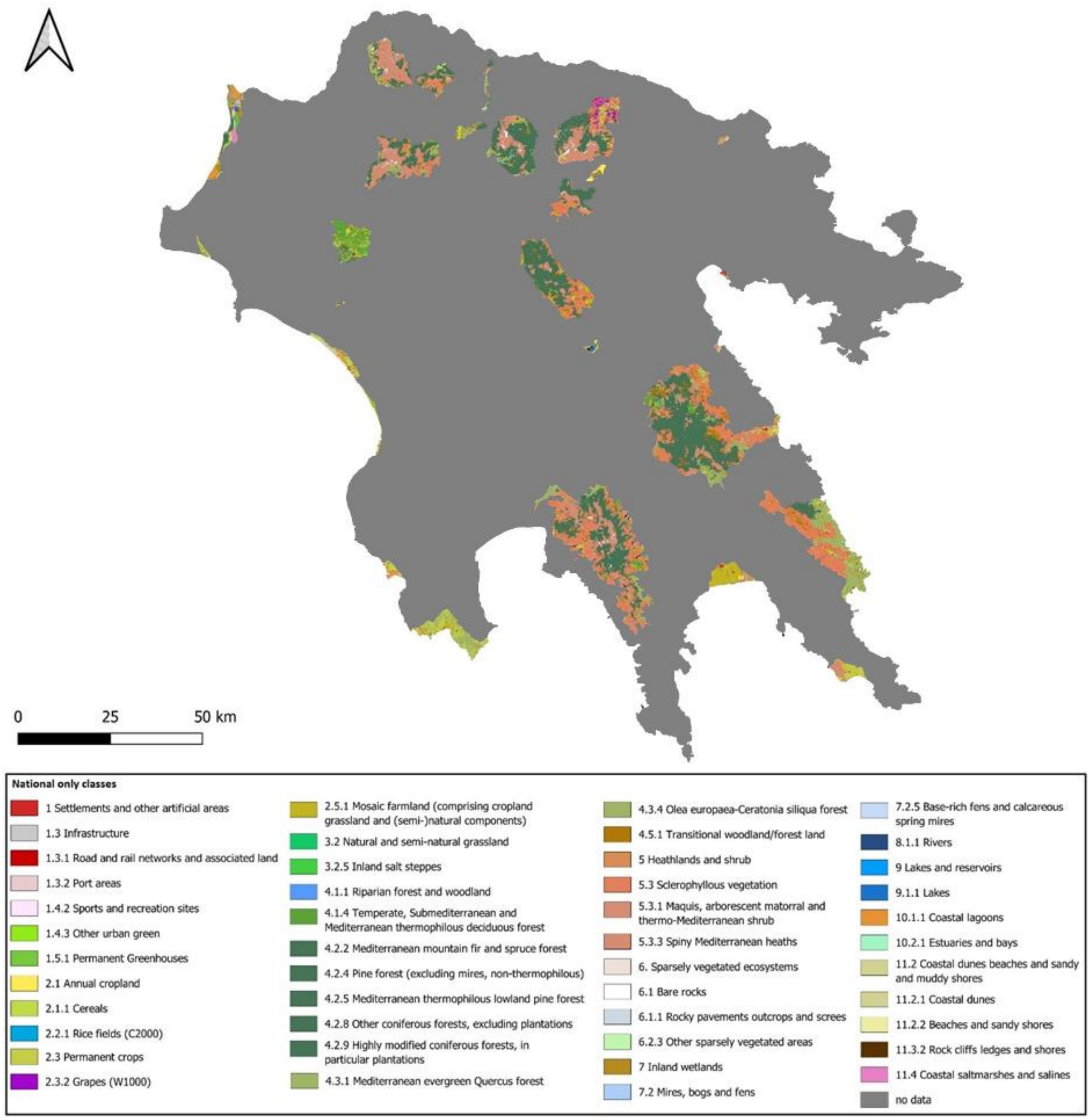


Figure 56: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using National data only (EPSG:3035).

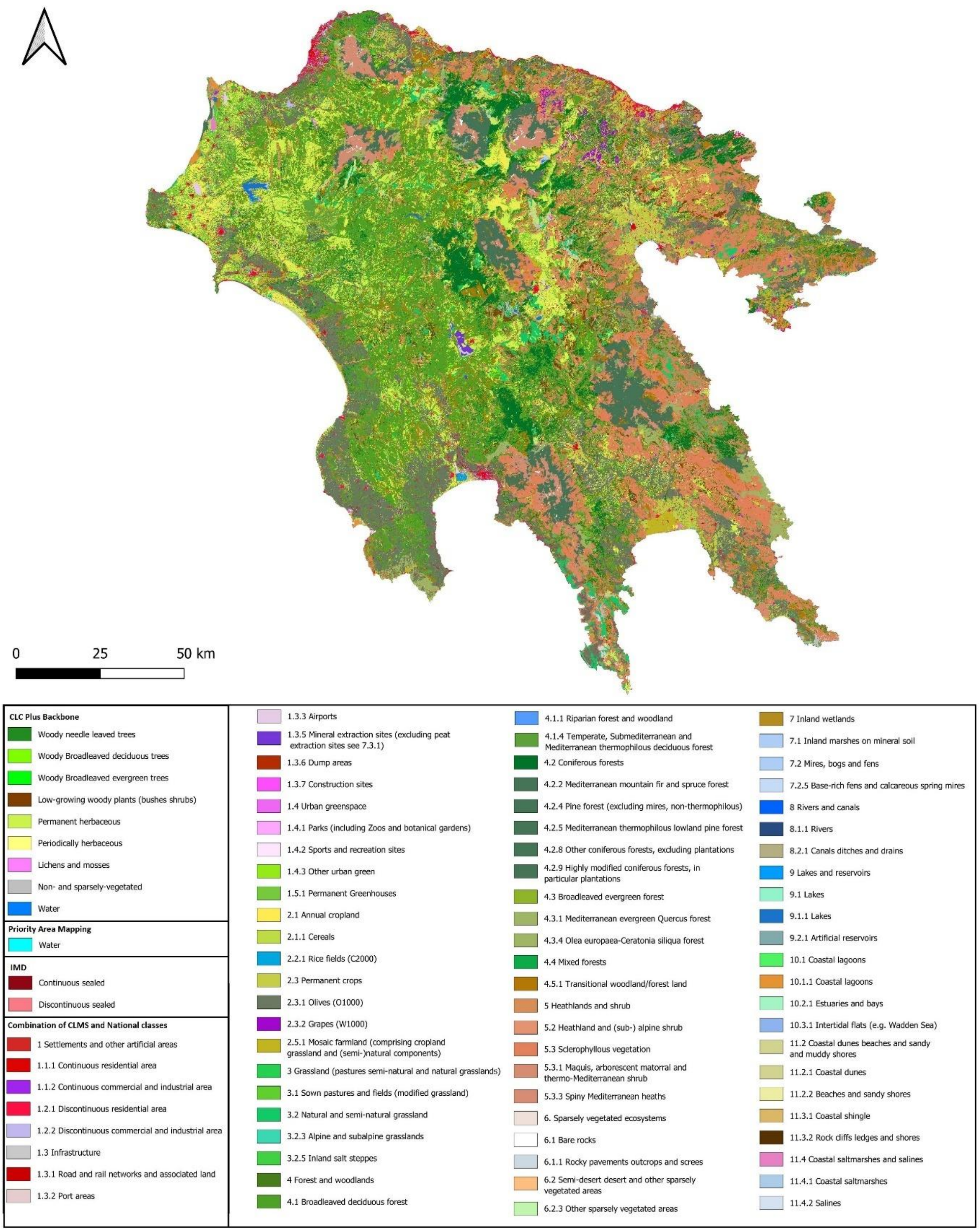


Figure 57: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using National data enhanced with CLMS data (EPSG:3035).

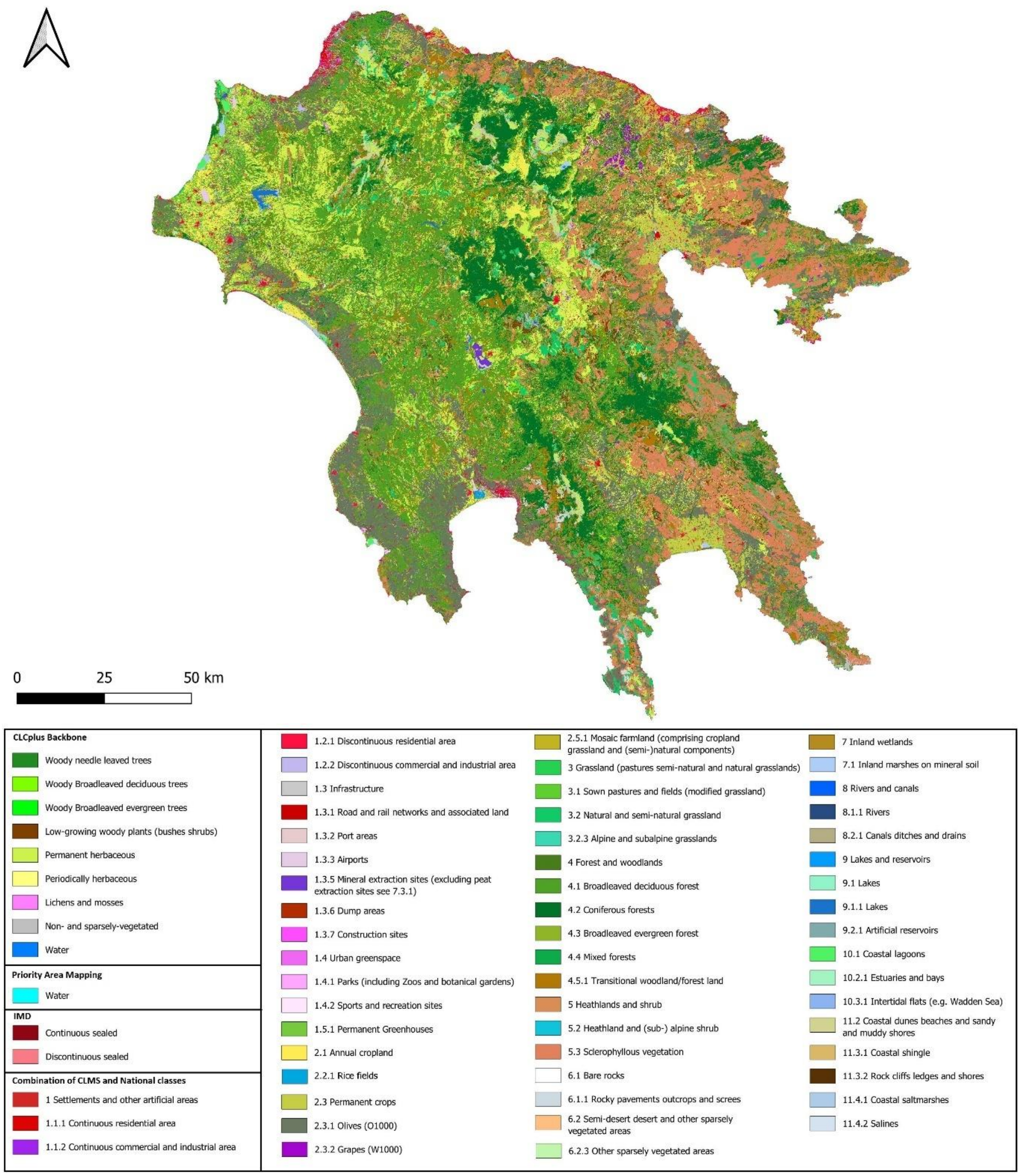


Figure 58: The European Ecosystem Typology ecosystem extent map for Peloponnese TS using CLMS data only (EPSG:3035).

8.1.1.2. São Miguel

The national reference data for São Miguel was of good quality, providing a reliable baseline for mapping, requiring less CLMS data than Peloponnese. To assess how each approach contributed to mapping towards the ETA, statistics were extracted at Level 1 (Table 46).

All three mapping approaches produced notably different class values at Level 1. The key differences between the 'National only' and 'National & CLMS' approaches were observed in the cropland and grassland classes. These classes were not accurately delineated under the national mapping, as highlighted by local expertise and visual assessments of satellite imagery. The integration of the CLCplus BB layer significantly improved their classification. Additionally, the CLMS data improved the classification of national data crosswalked to Level 1, particularly in the 6. Sparsely Vegetated Ecosystems (Level 2, 6.1 Bare Rocks). This class was further refined by distinguishing 6.1 Bare Rocks from 11.3.2 Rocks, Cliffs, and Shores, resulting in a more detailed and accurate representation.

When comparing the 'National' and 'CLMS only' maps, class 7 "Inland Wetlands Ecosystem" is absent in the latter. This group was mapped as a mix of heathlands and forests (both exotic and planted, as well as native and spontaneous) using CLMS data sources. A national expert confirmed it as an inland wetland, despite its appearance resembling heathland based on visual assessment. Additionally, many areas mapped as 5 "Heathlands and Shrub" in the CLMS only map were classified as 3 "Grassland" or 4 "Forest and Woodlands", likely due to differences in the definitions used by the two typologies. The CLMS only map also provided more detailed mapping of urban areas, as several roads and buildings were omitted in the national mapping.

Table 46: Shows the results of the ecosystem extent delineation at level 1 of the European Ecosystem Typology using three mapping approaches for the São Miguel TS.

European ecosystem typology (level 1)	National only (ha)	National & CLMS (ha)	CLMS only (ha)
1. Settlement and other artificial areas	4554	4621	6306
2. Cropland	8831	15526	11475
3. Grassland	35883	29120	29274
4. Forest and woodlands	21919	21919	20153
5. Heathlands and shrub	466	466	5119
6. Sparsely vegetated ecosystems	871	626	309
7. Inland wetlands	1054	1054	-
8. Rivers and canals	23	23	20
9. Lakes and reservoirs	831	831	852
11. Coastal beaches, dunes and wetlands	78	323	513
No data	-	-	90
No crosswalk	-	-	399
Total	**74510**	**74510**	**74510**

An in-depth analysis of the ecosystem extent delineation outcomes at the three levels of the European Ecosystem Typology is seen in Table 47. The aim of the delineation process is to map to a level as detailed as possible in the typology structure; level 3 in the case of the European Ecosystem Typology. The National & CLMS approach resulted in 28% of area mapped to level 3, where the National level approach reached 24%. Both these approaches map >96% of classes to a combined level 2 and 3. The CLMS only approach resulted in a 91% area mapped at levels 2 and 3. In comparison to the Peloponnese example, only a small area (0.66%) could not be crosswalked to the ETA using the CLMS only approach.

Table 47: Results of ecosystem extent delineation to all levels of the European Ecosystem Typology using three mapping approaches for the Sao Miguel TS.

Approaches	Level	Mapped area (ha)	Mapped area (%)
Step 1 National only	1	2 794.35	3.75
	2	53 535.31	71.85
	3	18 180.21	24.40
	Grand total	**74 509.87**	**100**
Step 2 National + CLMS	1	2 211.09	2.97
	2	51 687.68	69.37
	3	20 611.10	27.66
	Grand total	**74 509.87**	**100**
Step 3 CLMS only	1	3 665.3	8.14
	2	27 223.4	78.52
	3	60 666.9	12.68
	No crosswalk	1 455.9	0.66
	Grand total	**74 509.87**	**100**

The mapping outcomes for each of the three mapping approaches that have been discussed are presented below. The 'National only' (Figure 59) in comparison with the 'National & CLMS' (Figure 60) map clearly shows more detailed information on the level 2 class annual cropland. Figure 61 shows the mapping result using CLMS only. In comparison to the other maps, more areas have been mapped as heathlands and shrub in the center and around the eastern part of the island. These are mapped mostly as forest or wetland under the national mapping.

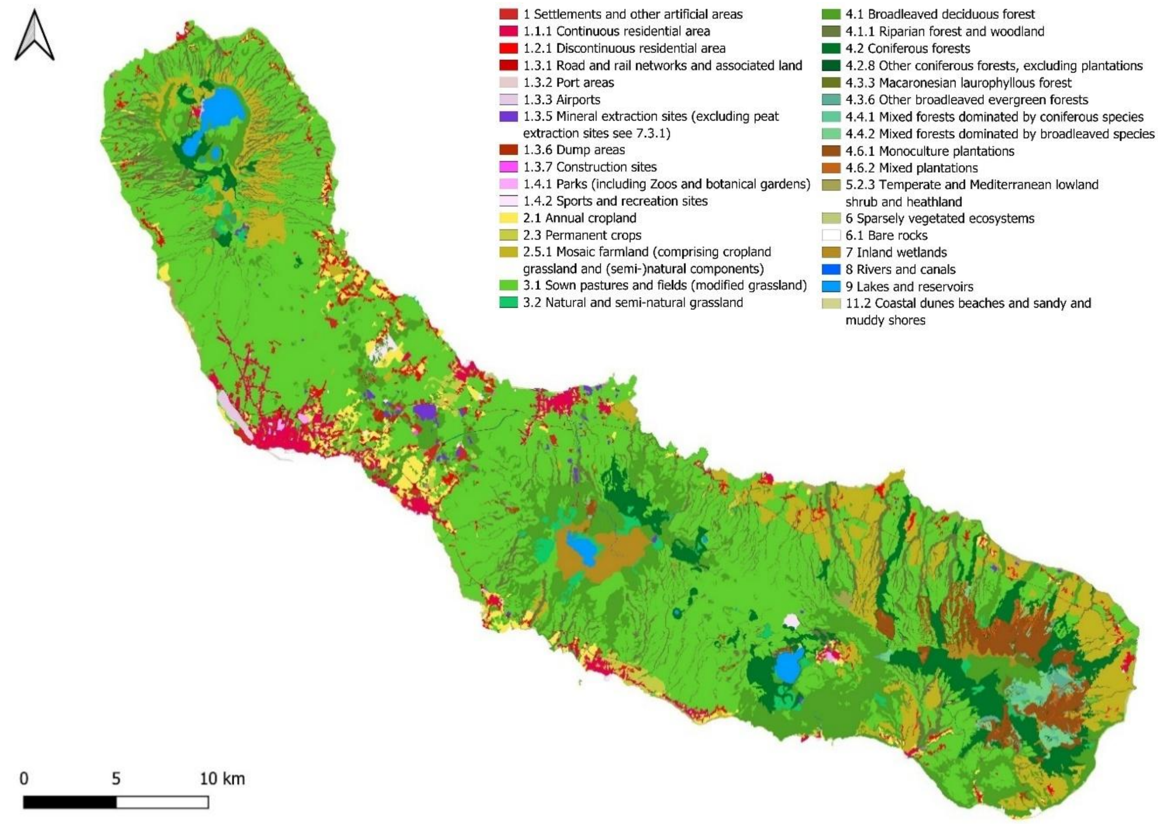


Figure 59: The ETA ecosystem extent map for São Miguel TS using National data only (EPSG:3035).

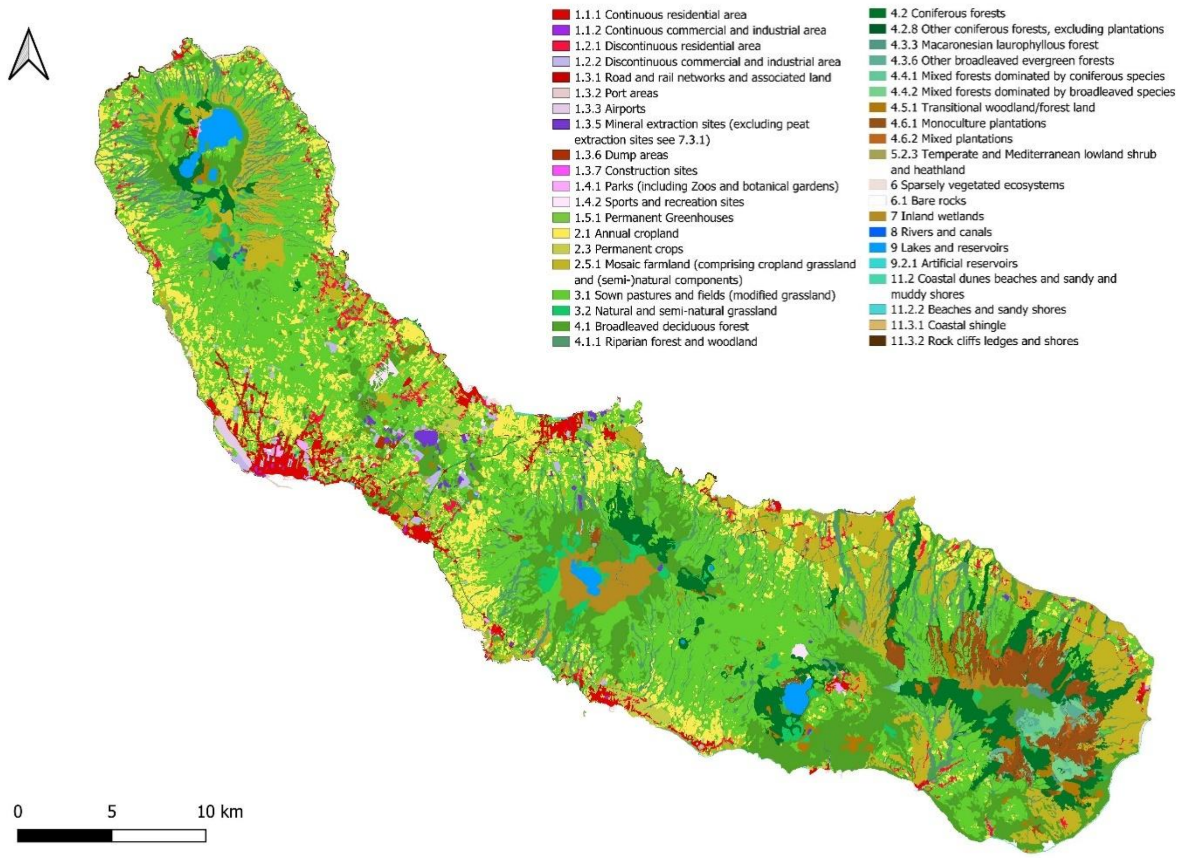


Figure 60: The ETA ecosystem extent map for São Miguel TS enhanced with CLMS data (EPSG:3035).

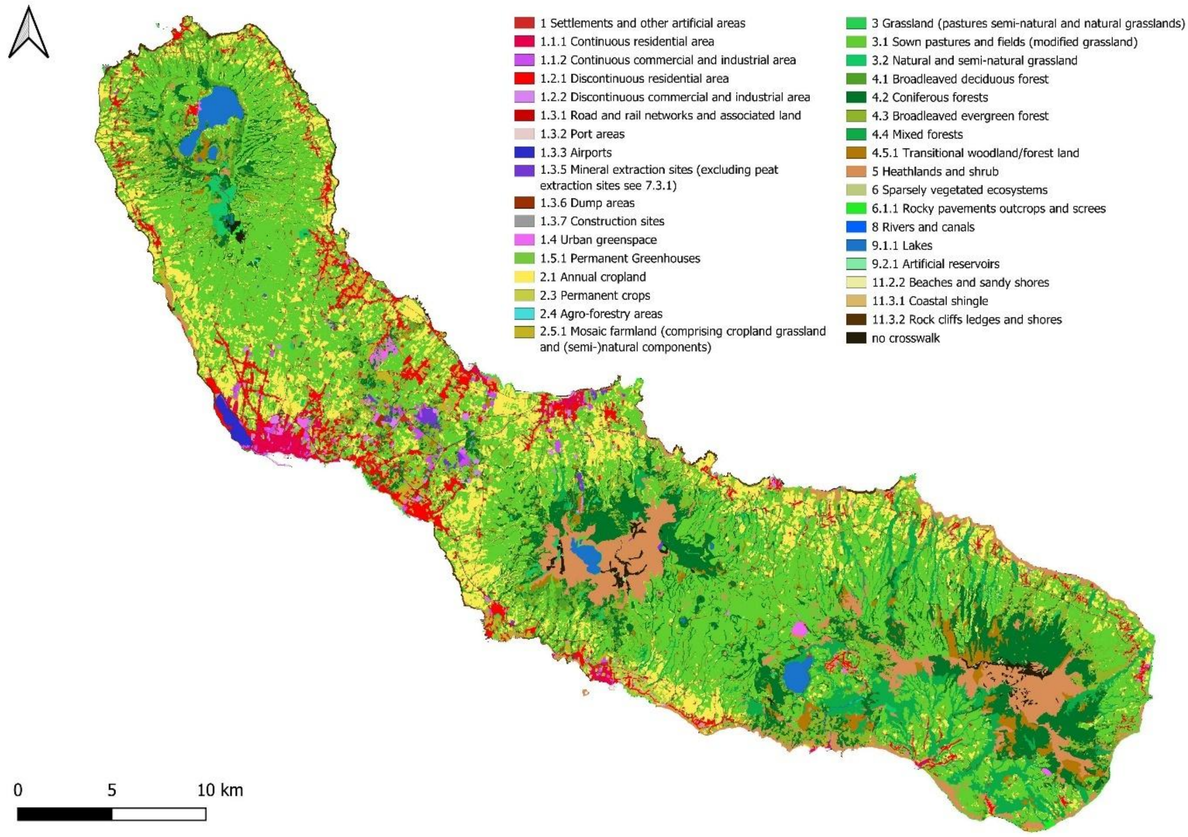


Figure 61: The ETA ecosystem extent map for São Miguel TS using CLMS data only (EPSG:3035).

8.1.2. Vegetation-centric approach

8.1.2.1. Peloponnese extent account

An ecosystem extent map for Peloponnese for the year 2020 was generated (Figure 62, Table 48) based on the EUNIS maps and according to the methodology as described before. As the cross-walking from EUNIS to Ecosystem Extent Typology is done according to the EU guidelines, the available EUNIS habitat classes and their mapping define largely the limitations of the ecosystem extent map. No further efforts were pursued, given the limited resources available, to map the anthropogenic classes (e.g. settlements and artificial areas) to a higher ecosystem type level.

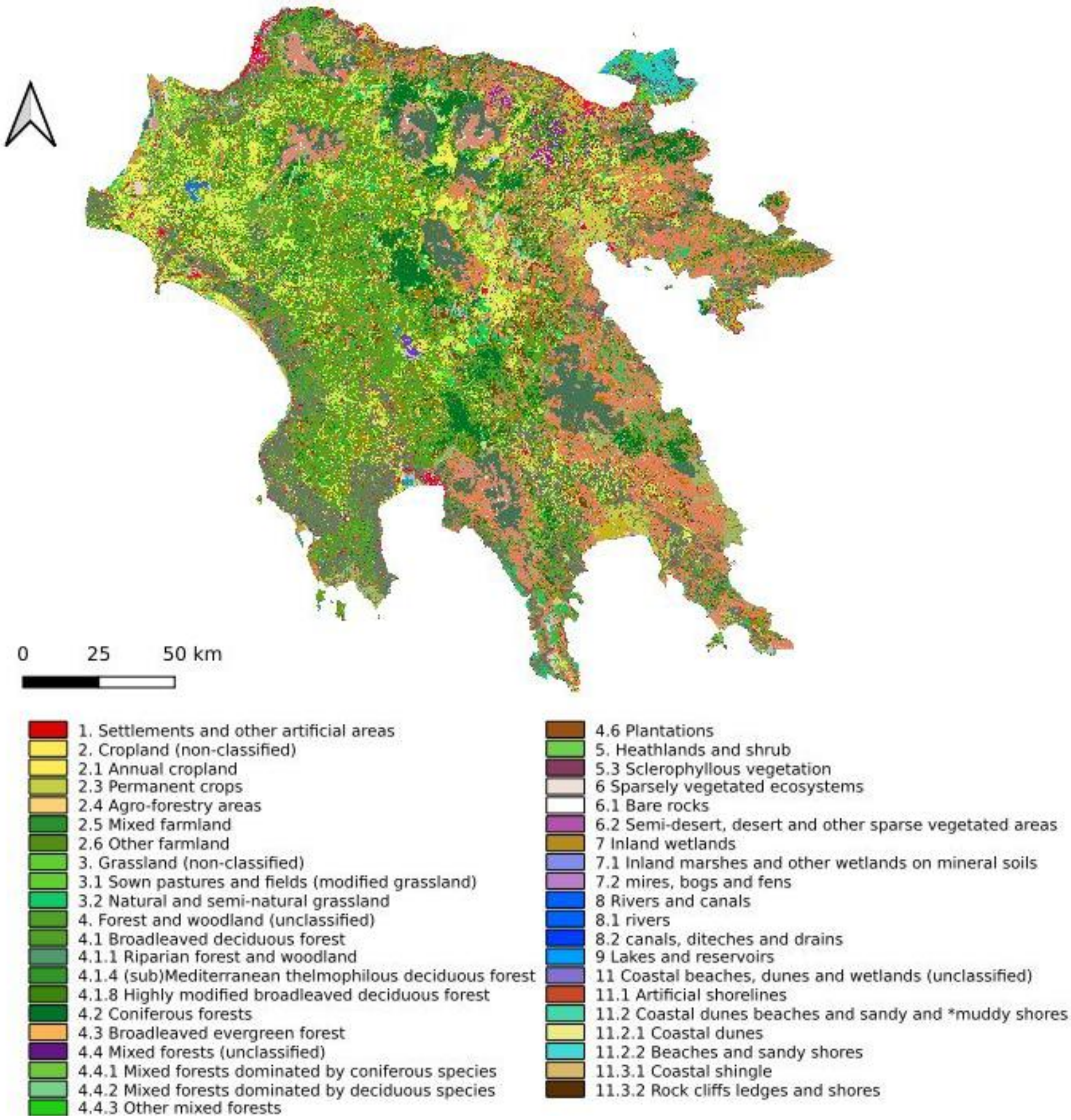


Figure 62: Ecosystem extent map 2020 for Peloponnese using the Sentinel / Vegetation centric approach (EPSG:3035).

Table 48: The Ecosystem Extent Account for Peloponnese for the year 2020.

L1	L2	L3	Ecosystem type	closing area (ha)	share of area
1			Settlements and other artificial areas	34707.30	1.6%
2			Cropland	78416.64	3.7%
	2.1		Annual cropland	75605.83	3.5%
	2.3		Permanent crops	112.22	0.0%
	2.5		Mixed farmland	2698.59	0.1%
3			Grassland	419845.67	19.6%
	3		Grassland - L2 unallocated	419440.56	19.6%
	3.2		Natural and seminatural grassland	405.11	0.0%
4			Forest and woodland	950681.22	44.4%

	4.0		Forest and woodland - L2 unallocated	265441.25	12.4%
	4.1		Broadleaved deciduous forest	97927.40	4.6%
		4.1.1	Riparian forest and woodland	41505.06	1.9%
		4.1.4	Temperate, Submediterranean and Mediterranean thermophilous deciduous fores	56422.34	2.6%
	4.2		Coniferous forest	248229.71	11.6%
		4.2.0	Coniferous forest - L2 unallocated	141651.80	6.6%
		4.2.4	Pine forest (excluding mires, non-thermophilous)	42758.90	2.0%
		4.2.5	Mediterranean thermophilous lowland pine forest	59340.46	2.8%
		4.2.8	Other coniferous forests, excluding plantations	1689.76	0.1%
		4.2.9	Highly modified coniferous forests, in particular plantations	2788.79	0.1%
	4.3		Broadleaved evergreen forest	339082.86	15.8%
		4.3.0	Broadleaved evergreen forest - L2 unallocated	101245.62	4.7%
		4.3.1	Mediterranean evergreen Quercus forest	122530.44	5.7%
		4.3.4	Olea europaea-Ceratonia siliqua forest	115306.80	5.4%
5			Heatland and shrub	580278.40	27.1%
	5.0		Heatland and shrub - L2 unallocated	91769.62	4.3%
	5.3		Sclerophyllous vegetation	488508.78	22.8%
6			Sparsely vegetated ecosystems	41729.98	1.9%
	6.0		Sparsely vegetated ecosystems - L2 unallocated	10008.70	0.5%
	6.1		Bare rocks	29529.21	1.4%
	6.2		Semi-desert, desert and other sparsely vegetated areas	2192.07	0.1%
7			Inland wetlands	1742.02	0.1%
	7.1		Inland marshes and other wetlands on mineral soil	1683.10	0.1%
	7.2		Mires, bogs and fens	58.92	0.0%
8			Rivers and Canals	5436.49	0.3%
	8.0		Rivers - L2 unallocated	4209.41	0.2%
	8.1		Rivers	203.04	0.0%
	8.2		Canals, ditches and drains	1024.04	0.0%
9			Lakes and Reservoirs	11822.36	0.6%
11			Coastal beaches, dunes and wetlands	17623.75	0.8%
	11.1		Artificial shorelines	8964.69	0.4%
	11.2		Coastal dunes, beaches and sandy and muddy shores	7452.00	0.3%
		11.2.0	Coastal dunes - L2 unallocated	629.62	0.0%
		11.2.1	Coastal dunes	4545.52	0.2%
		11.2.2	Beaches and sandy shores	1069.80	0.0%
		11.2.3	Muddy shores	1207.06	0.1%
			Totals	2142283	ha
				21422	km^2

8.1.2.2. São Miguel Extent account

Mapping São Miguel (Figure 63, Table 49) is very challenging from a remote sensing perspective, particularly in distinguishing the managed versus natural sites. Most of the island is managed and hence additional land use layers are required to distinguish some specific classes as well as mapping them into the EU ecosystem typology, such as:

- T3M (coniferous plantation of site-native trees) or T3N (coniferous plantation of site-native trees) are mapped to EU class 4.2.9 (highly modified coniferous forests, in particular plantations). They could potentially also be mapped to ETA 4.6.1 (monoculture or mixed plantations) but this would require additional land use information which was not available in the OpenStreetMap dataset.
- Mixed farmland (ETA 2.5.1) describes heterogeneous landscapes without one single dominant land cover that covers more than 50% of the Minimum Mapping Unit. It typically is a mix of arable land, permanent crops, pasture/grassland interspersed with areas such as hedges, ponds, etc. Therefore, in our map, most likely the grassland is over-estimated, and the cropland is under-estimated.

It should also be noted that the Settlements and other Artificial areas (ETA 1) were not further decomposed, since it was primarily focusing on mapping the (semi-)natural ecosystem types.

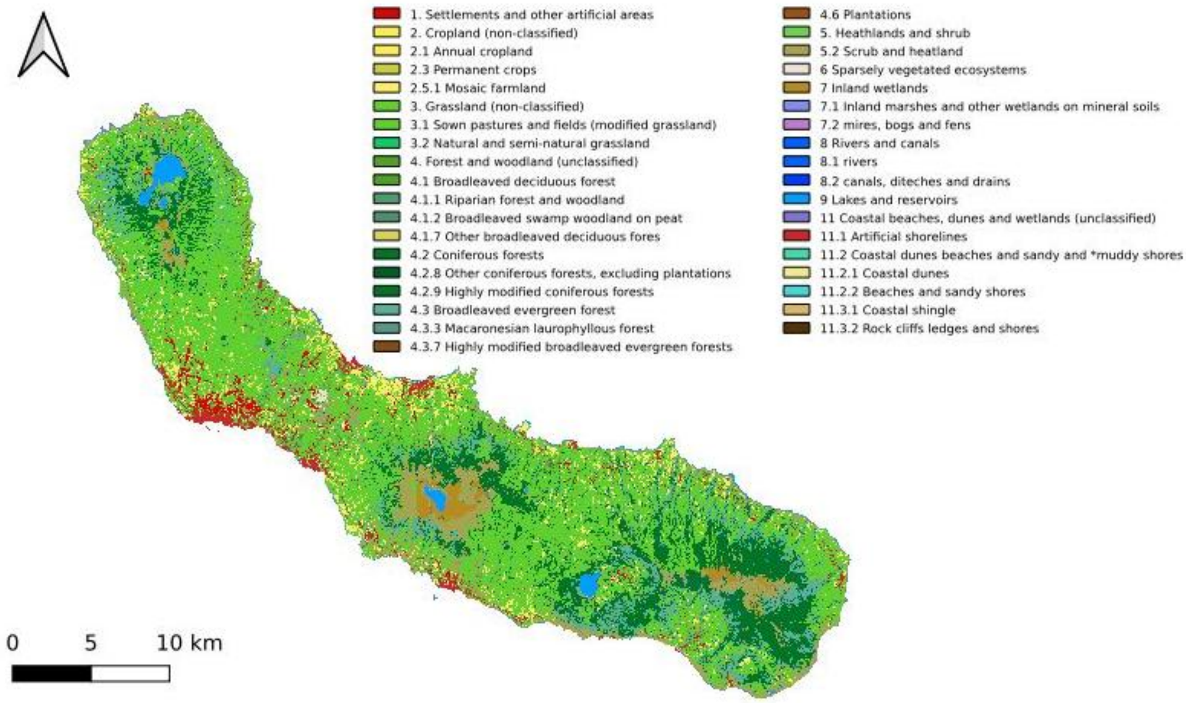


Figure 63: The European Ecosystem Typology ecosystem extent map for São Miguel TS using Sentinel data / vegetation centric approach (EPSG:3035).

Figure 64 shows a confidence of the extent map for São Miguel. It is based on the methodological rules as described above, with the definition that if both the EUNIS habitat map and the Land Cover/Land Use maps confirm the ecosystem type the confidence is highest (depicted with match_ALL label). If only one of the two maps confirm the ecosystem type, the one that prevailed for the selection (in our case mostly EUNIS) is depicted.

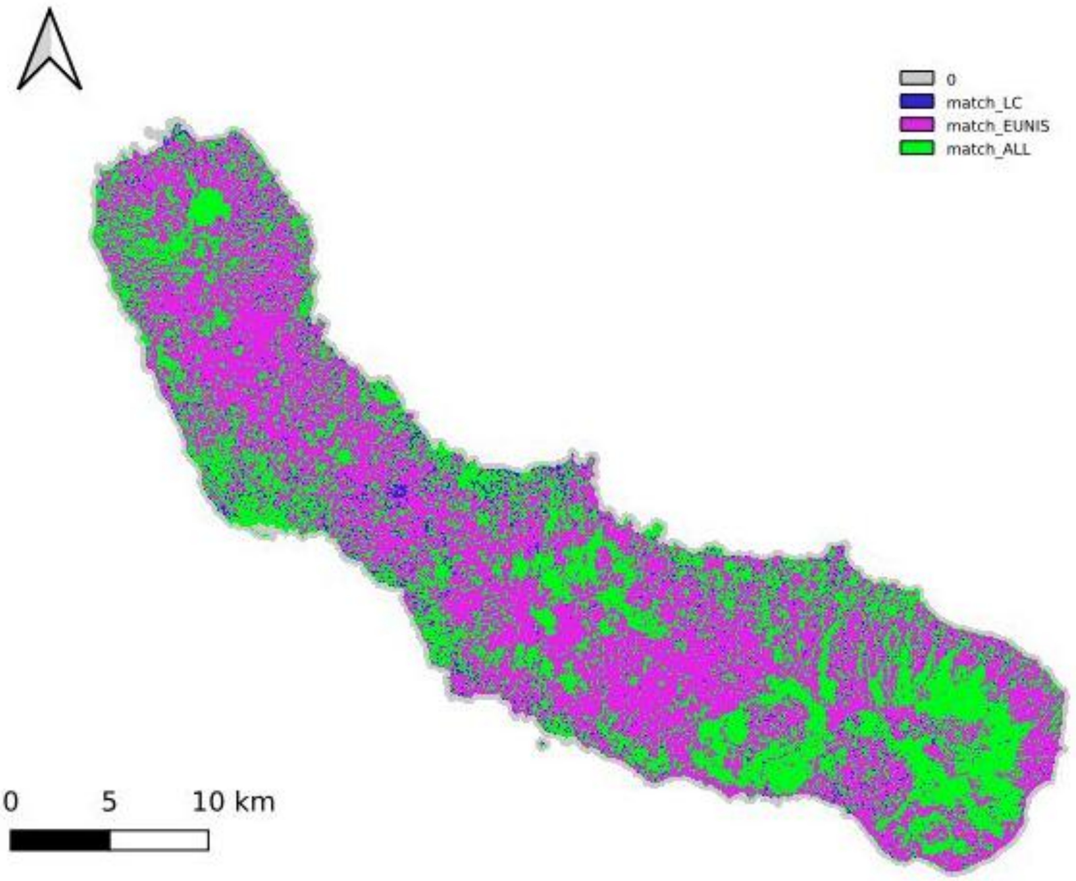


Figure 64: The European Ecosystem Typology ecosystem extent Quality Flag for São Miguel TS using Sentinel data / vegetation centric approach (EPSG:3035).

Table 49: The Ecosystem Extent Account for São Miguel for the year 2020.

L1	L2	L3	Ecosystem Type	closing area (ha)	Share of area
1			Settlements and other artificial areas	1303.44	1.8%
2			Cropland	4205.85	5.7%
	2.1		Annual cropland	525.18	0.7%
	2.3		Permanent crops	2.70	0.0%
	2.5		Mixed farmland	3677.97	4.9%
3			Grassland	36324.18	48.8%
	3		Grassland - L2 unallocated	34022.48	45.7%
	3.1		Sown pastures and other grass (modified grassland)	2301.70	3.1%
4			Forest and woodland	19604.89	26.3%

	4.0		Forest and woodland - L2 unallocated	1835.18	2.5%
	4.1		Broadleaved deciduous forest	27.07	0.0%
		4.1.2	Broadleaved swamp woodland on non-acid and acid peat	23.10	0.0%
		4.1.7	Other broadleaved deciduous forest, excluding highly-modified plantations	3.97	0.0%
	4.2		Coniferous forest	11762.57	15.8%
		4.2.8	Other coniferous forests, excluding plantations	0.61	0.0%
		4.2.9	Highly modified coniferous forests including stands of non-native trees species that have long been established in European ecosystems stands	11761.96	15.8%
	4.3		Broadleaved evergreen forest	5980.07	8.0%
		4.3.0	Broadleaved evergreen forest - L2 unallocated	5918.81	8.0%
		4.3.3	Macaronesian laurophyllous forest	35.39	0.0%
		4.3.7	Highly modified broadleaved evergreen forests including stands of non-native trees species that have long been established in European ecosystems stands	25.87	0.0%
5			Heatland and shrub	7221.47	9.7%
	5.0		Heatland and shrub - L2 unallocated	387.65	0.5%
	5.2		Heatland and (sub-)alpine shrubs	6833.82	9.2%
6			Sparsely vegetated ecosystems	702.12	0.9%
7			Inland wetlands	1013.87	1.4%
	7.0		Inland wetlands - L2 unallocated	991.16	1.3%
	7.1		Inland marshes on mineral soils	20.99	0.0%
	7.2		Mires, bogs and fens	1.72	0.0%
8			Rivers and Canals	608.74	0.8%
	8.0		Rivers and Canals - L2 unallocated	606.57	0.8%
	8.1		Rivers	0.24	0.0%
	8.2		Canals, ditches and drains	1.93	0.0%
9			Lakes and Reservoirs	1046.15	1.4%
11			Coastal beaches, dunes and wetlands	2381.59	3.2%
	11.0		L2 unallocated	573.13	0.8%
	11.1		Artificial shorelines	1697.03	2.3%
	11.2		Coastal dunes, beaches and sandy and muddy shores	111.43	0.1%
			Totals	74412.30	ha
				744.123	km^2

Within the limited time available, it was possible to create an ecosystem extent account up to Level 3 using the EUNIS habitat maps as an underlying dataset. The approach is scalable across the entire European continent, despite requiring further work to be conducted. The results of this experiment are a very valuable input for other projects to continue the work as is the ESA World Ecosystem Extent Dynamics project (https://esa-worldecosystems.org/en) which targets to create extent accounts at full European scale using this approach.

8.2. Ecosystem condition mapping

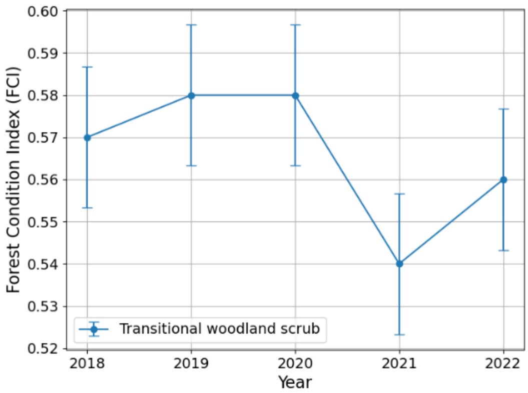
8.2.1. Peloponnese Forest Condition Index

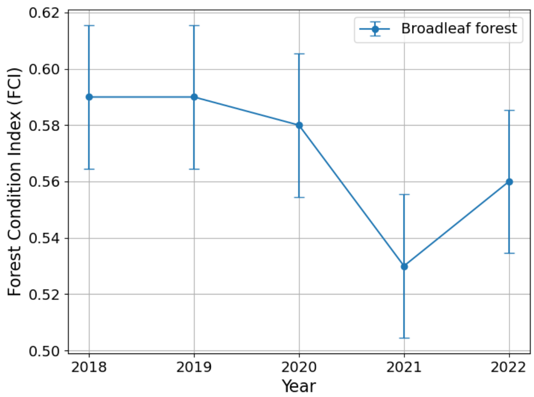
There is not a lot of differentiation in how close the different forest types on Peloponnese are to their ideal condition. Their mean forest condition indices fluctuate around 0.55 or 0.60 between 2018 and 2022 (Table 50).

Table 50: Forest condition account (mean forest condition indices per class) for forest types in Peloponnese for 2018-2022.

Year	Transitional woodland scrub, Mediterranean	Broadleaf forest, Mediterranean	Mixed forest, Mediterranean	Coniferous forest, Mediterranean
2018	0.57	0.59	0.56	0.63
2019	0.58	0.59	0.56	0.63
2020	0.58	0.58	0.55	0.62
2021	0.54	0.53	0.51	0.59
2022	0.56	0.56	0.53	0.60

The graphs below show that the mean forest condition index for each vegetation type seems to find a minimum in 2021 (Figure 65). As previously explained, the transition from PROBA-V to Sentinel-3/OLCI has introduced artefacts for products that are used to generate the CLMS NPP data. The ARIES for PEOPLE-EA Explorer still extracts NPP from this data source. Therefore, the forest condition indices were calculated with the CLMS NPP data. Hence, the drop in mean FCI in 2021 is most likely not a representation of reality but rather introduced by artefacts in the NPP data for period between 2020 and 2022.





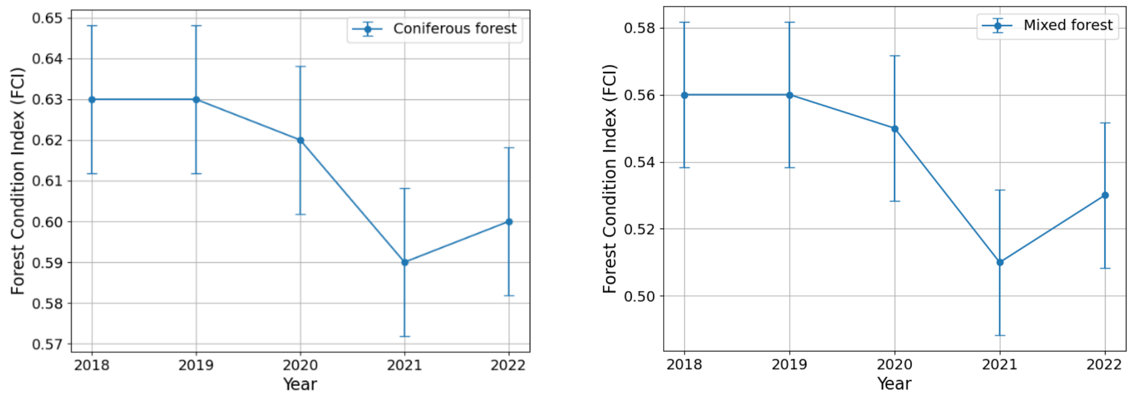


Figure 65: Evolution of mean forest condition index through time (2018-2022) per forest type in Peloponnese for Transitional woodland shrub, broadleaf forest, coniferous forest and mixed forest.

The FCI of 2022 is displayed in Figure 66 (other years behave similarly). The FCI is highest, or the forest condition is closest to its ideal/reference state, around the center of the peninsula.

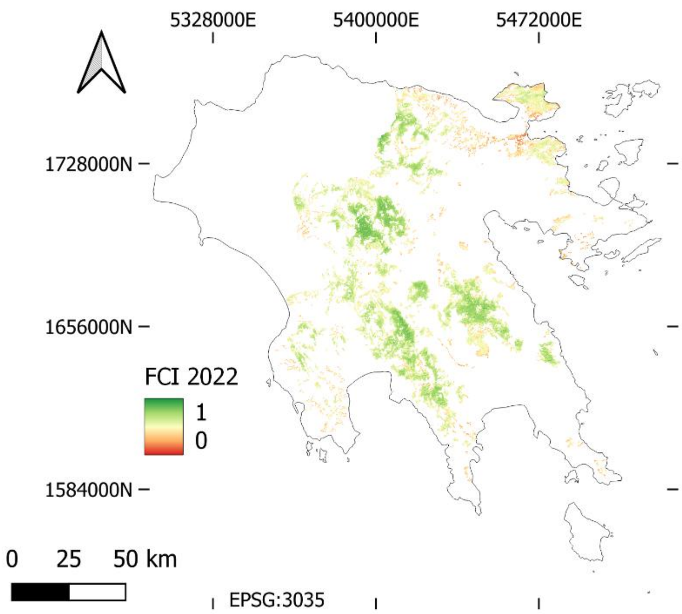


Figure 66: Spatial map of the forest condition index in 2022 for Peloponnese.

The net change in FCI between 2018 and 2022 is very slightly negative and the negative hotspots are mostly found at the edges of the forested areas (Figure 67). In general, the condition remains quite stable through time.

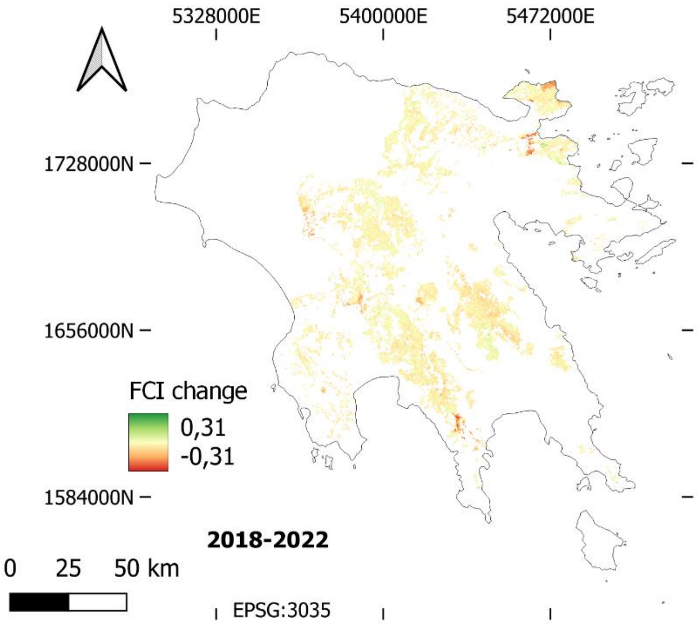


Figure 67: Net change map of forest condition index between 2018 and 2022 for Peloponnese. The redder the pixels are colored, the more a decline in forest condition index was observed. The greener the pixels are colored, the more an increase in forest condition.

The change maps per two consecutive years within the specified period indicate quite precisely where and to which extent the forest condition index increased or decreased in Peloponnese (Figure 68). There is quite a negative change between 2020 and 2021 over almost all forest area, but this decrease contains a maximum value of 0.15 in FCI. After 2021, almost the entire region experiences an increase in forest condition.

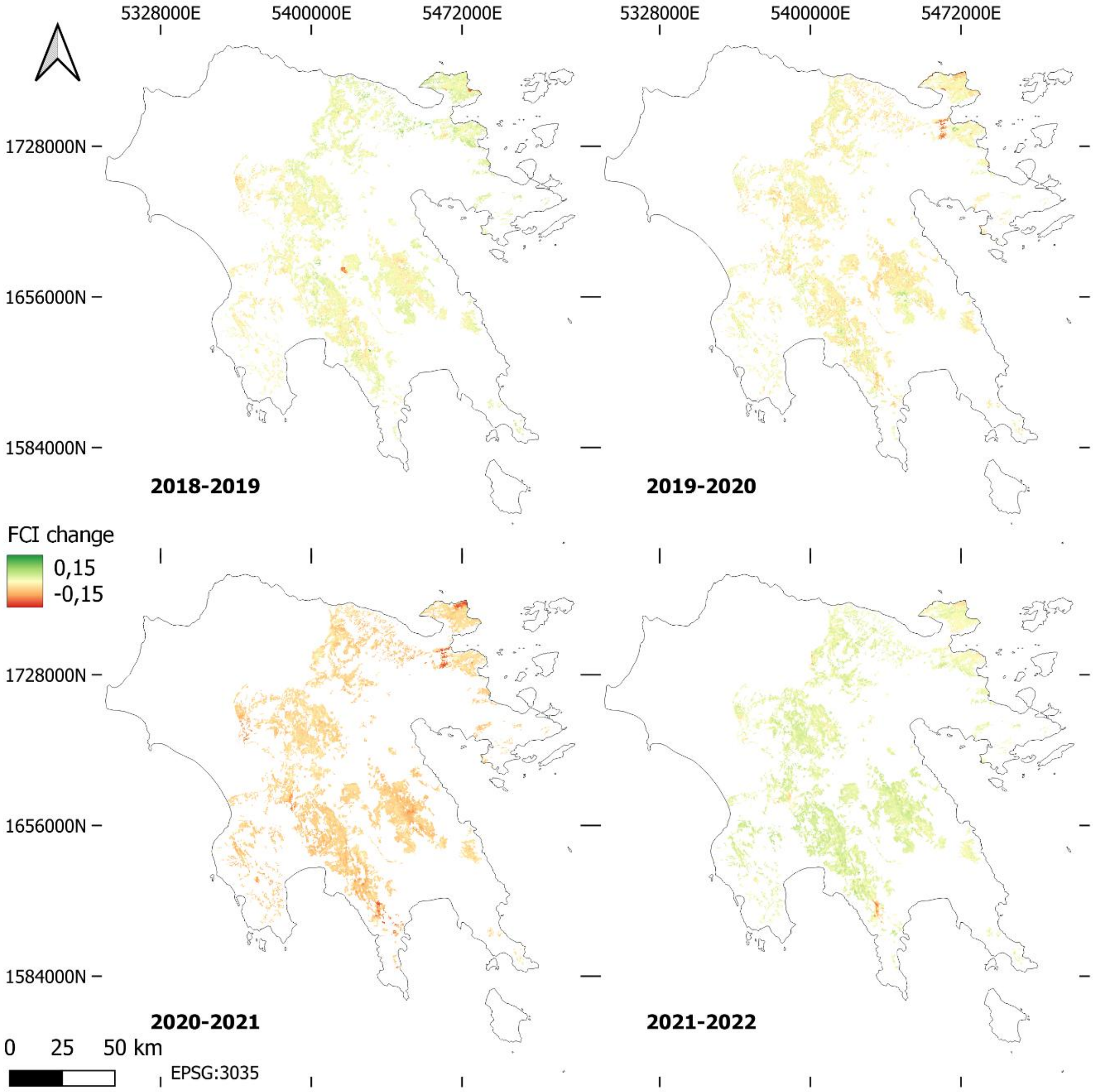


Figure 68: Change maps of forest condition index in pairs of consecutive years between 2018 and 2022 for Peloponnese.

The forest condition produced in the PEOPLE EA project was evaluated using field data from forest management studies (where available and spatially explicit), Habitat Directive monitoring (local conservation degree) sampling plots, MAES national project (LIFE IP 4 NATURA) field surveys and the relevant results for ecosystem condition, expert judgment on well-studied forests, and information from local authorities regarding forests and their status. Similar condition results are identified among the PEOPLE EA outcomes and the LIFE IP 4 NATURA (ecosystem condition), as well as the Habitat Directive monitoring (local conservation degree) results.

8.2.2. São Miguel Forest Condition Index

Forest condition for São Miguel comprised the period from 2018 until 2023 (Table 51). Forest types T1, T2 and T3 are fluctuating between 0.42 and 0.56 for 2018-2022. Forest type S4 is a bit further away from its ideal condition for the same period.

Table 51: Forest condition account (mean forest condition indices per class) for forest types (according to mapped EUNIS habitat level 2 classes on the habitat map) in Sao Miguel for period 2018-2023. T1: Broadleaved deciduous forest, T2: Broadleaved evergreen forest, T3: Coniferous Forest & S4: Temperate shrub heathland.

Year	T1	T2	T3	S4
2018	0.52	0.43	0.44	0.40
2019	0.54	0.43	0.45	0.40
2020	0.56	0.45	0.47	0.41
2021	0.50	0.43	0.45	0.40
2022	0.51	0.43	0.45	0.39
2023	0.50	0.42	0.44	0.38

All forest types experienced a maximum forest condition in 2020 (Figure 69). For T2, T3 and S4 it is followed by a mild decrease, while T1 undergoes a steeper decrease in 2021.

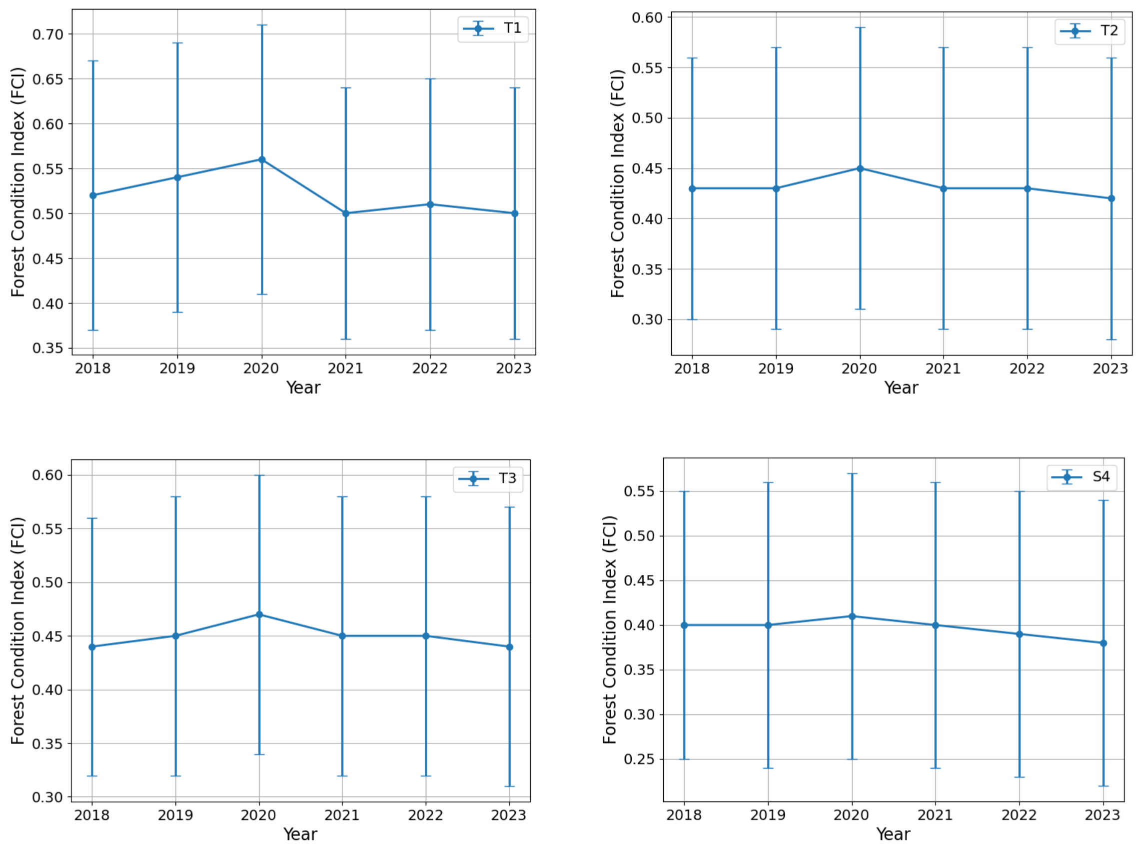


Figure 69: Evolution of mean forest condition index through time (2018-2023) per forest type in São Miguel: T1 Broadleaved deciduous forest, T2 Broadleaved evergreen forest, T3 Coniferous Forest and S4 Temperate shrub heathland.

The net change in forest condition index map (between 2018 and 2023) shows that the most decreases in forest condition index are located towards the central axis of the São Miguel (Figure 70). The maximum decrease in forest condition index contains an index of 0.35.

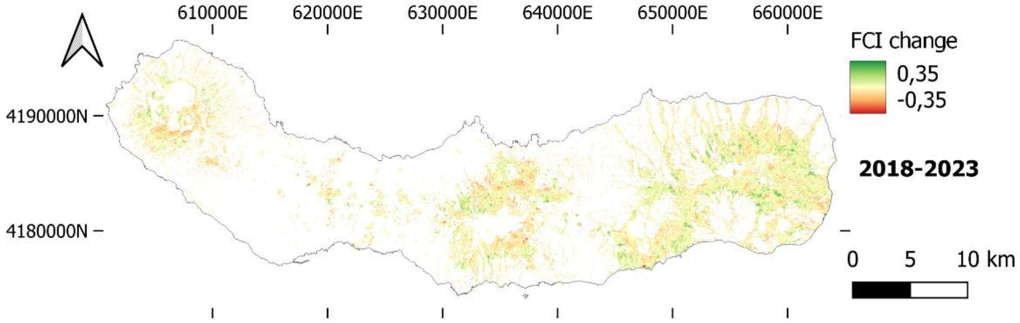


Figure 70: Net change map of forest condition index between 2018 and 2023 for Sao Miguel. The redder the pixels are colored, the more a decline in forest condition index was observed. The greener the pixels are colored, the more an increase in forest condition index was observed.

The yearly change maps disclose that quite extensive areas have experienced a sharp decline in forest condition index, mostly in the transition to 2021 and 2022 (Figure 71).

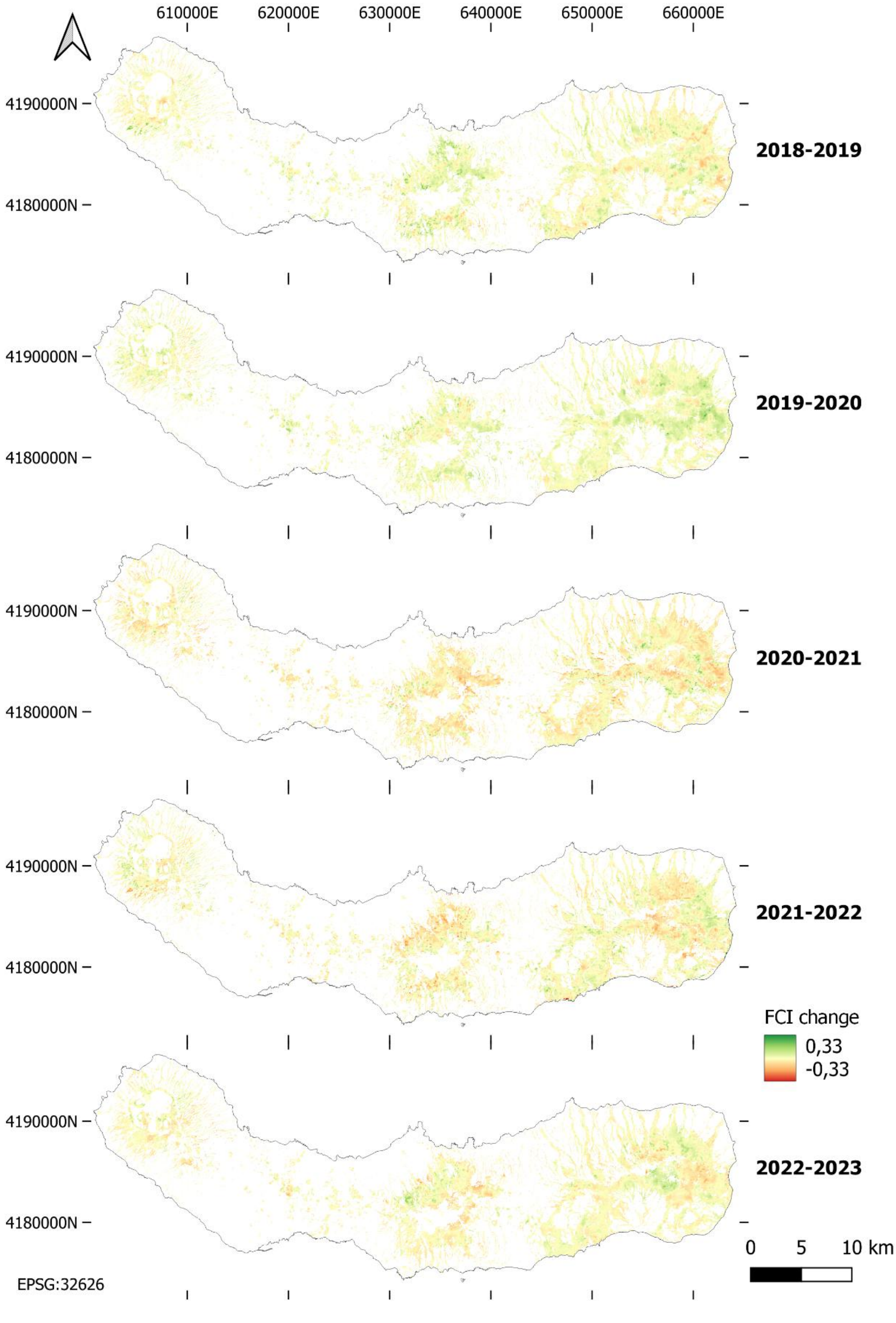


Figure 71: Change maps of forest condition index in consecutive years between 2018 and 2023 for São Miguel.

8.3. Forest Carbon Accounting

Here the results of carbon accounting using 2 methods are presented. The first method is based on the combination of RS-derived products (Section 7.3.1), while the second method is based on Gross Primary Productivity (GPP) estimates and maps (Section 7.3.2). The first method provides carbon emission estimates directly derived from biomass maps, created using an unsupervised approach that integrates data on forest structure (from forest structural maps), vegetation canopy height (obtained from GEDI LiDAR), and deforestation/degradation maps produced by the SarSentry change analysis. Carbon maps derived from these biomass maps are then used to calculate differences between two years, allowing for quantification of carbon fluxes. The second method utilizes GPP maps, which are converted into annual Net Primary Production (NPP), Aboveground Net Primary Production (ANPP), and Belowground Net Primary Production (BNPP) maps. The aboveground carbon sequestration for 2022 and 2023 is calculated by summing the ANPP values for both years.

8.3.1. Carbon stock maps using remote sensing-derived products

8.3.1.1. Results and validation of Forest structural maps

The process of carbon mapping begins with the creation of a high-quality Forest Structural Map (FSM) as a foundational step. Figure 72 presents the FSM created for São Miguel, covering 2017, 2021 and 2023, illustrating the spatial distribution of forest structure.

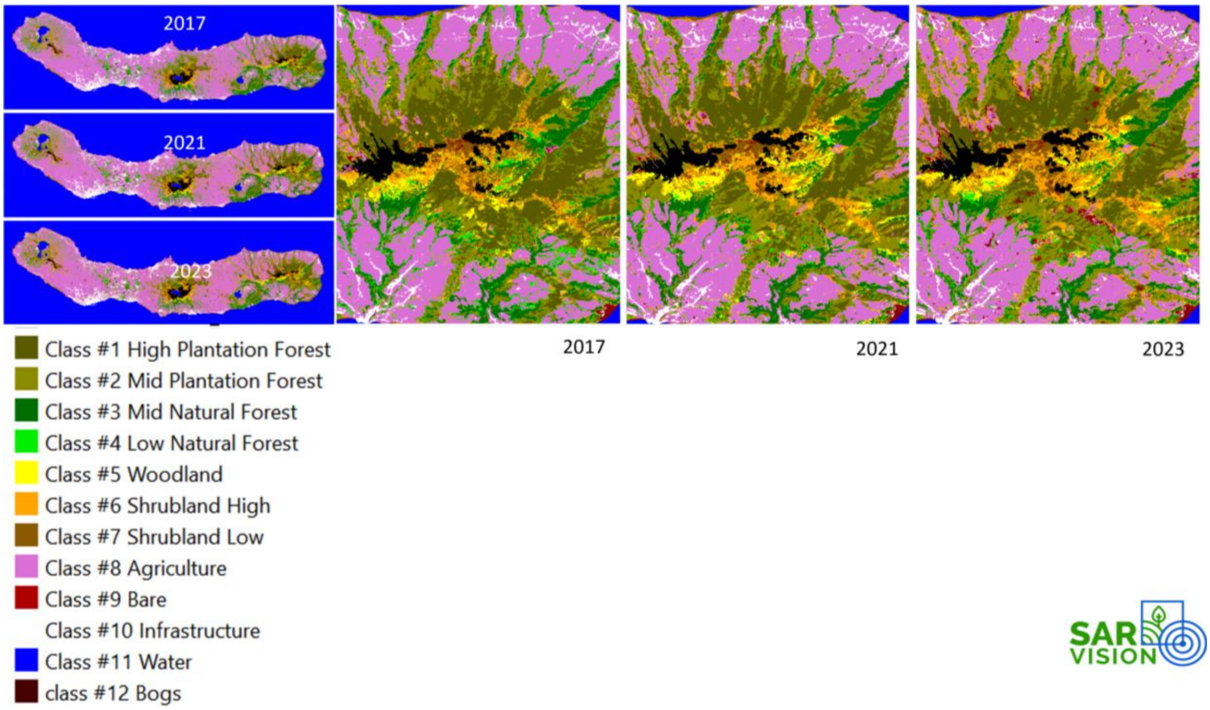
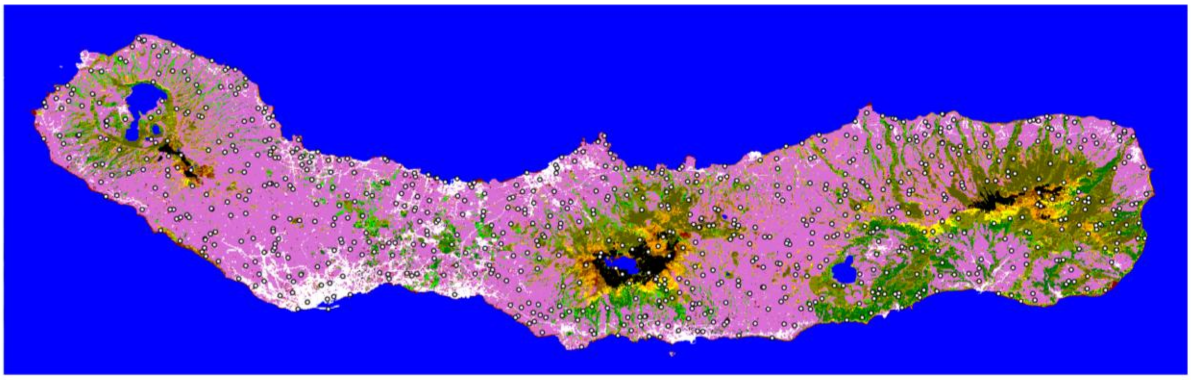


Figure 72: Forest structural maps of São Miguel for 2017, 2021 and 2023, the whole island (on the left), a detail over the east volcanic area where the Peatland bogs were mapped (on the right).

All maps were created using the same legend to facilitate easy comparison of changes between years. Most of these changes are linked to deforestation and forest degradation activities, although in some cases, landslides and slope movements, triggered by sudden rains, strong winds and seismic activity, have also contributed to the removal of vegetation from mountainous forested areas, creating temporary bare soil areas. In São Miguel, forests can be classified as coniferous or broadleaved, and they may be natural or planted. The maps identify four distinct forest types, which appear to be influenced by management practices (natural vs. planted) and exhibit different density levels and structural characteristics. Additionally, high and low shrublands and areas with tree cover below 10% were detected.

A significant portion of the island is covered by agricultural crops and grasslands, alongside areas designated for infrastructure and water bodies. One of the key improvements in this mapping effort is the detection and classification of highland peatlands or bogs, which play a crucial role in local water retention and belowground carbon accumulation and could otherwise be structurally mistaken for bare areas or low shrubland areas.



Overall Accurac												
Kappa Coefficie												
		1-2	3-4	5-6	7	9	8	11	10	Total	Commision erro	User Accuracy
1 High Forest	1-2	96.1	6.06	8.11	5.26	0	0.71	0	0	20.2	6.92	93.08
2 Low Forest	3-4	3.25	80.3	35.14	21.05	4.17	2.38	0	0	10.93	38.37	61.63
3 Shrubland h	5-6	0	4.55	27.03	10.53	4.17	0.48	0	0	2.29	44.44	55.56
4 Shrubland l	7	0.65	0	5.41	26.32	0	0.71	0	0	1.4	54.55	45.45
5 Bare	9	0	0	2.7	10.53	75	0.71	50	6.15	3.68	37.93	62.07
6 Crop high	8	0	9.09	21.62	26.32	8.33	94.76	50	6.15	53.88	6.13	93.87
7 Water	11	0	0	0	0	0	0	0	0	0	0	0
8 Infraestruc	10	0	0	0	0	8.33	0.24	0	87.69	7.62	5	95
	total	100	100	100	100	100	100	100	100	100		
	Ommision error	3.9	19.7	72.97	73.68	25	5.24	100	12.31			
	Producer accuracy	96.1	80.3	27.03	26.32	75	94.76	0	87.69			

Class #1 High Plantation Forest
Class #2 Mid Plantation Forest
Class #3 Mid Natural Forest
Class #4 Low Natural Forest
Class #5 Woodland
Class #6 Shrubland High
Class #7 Shrubland Low
Class #8 Agriculture
Class #9 Bare
Class #10 Infrastructure
Class #11 Water
class #12 Bogs

Figure 73: Confusion matrix presenting the validation results for the FSM 2023 in São Miguel. Map presents the location of randomly selected points used for map validation using high resolution Planet data as a reference. Overall accuracy was estimated in 87.5% with a calculated kappa coefficient of 0.8.

The validation of the 2023 map was conducted by inspecting high-resolution images over a set of randomly generated points, stratified across all map classes. A confusion matrix was created to show the proportion (in percentages) of correctly and incorrectly classified points for each class. The overall accuracy was calculated to be 87.5%, with a corresponding kappa coefficient of 0.8, indicating strong agreement between the classification and reference data. (Figure 73). 689 out of 767 points were confirmed to be classified correctly. Additionally, two types of classification errors were analysed:

- **Errors of Commission (False Positives)**, occur when a value is incorrectly classified into a class to which it does not belong, represented in the rows of the confusion matrix. These errors are related to user accuracy, which indicates the likelihood that a classified point belongs to the assigned class.
- **Errors of Omission (False Negatives)**, occurring when a value belongs to a class but is incorrectly assigned to another class, represented in the columns of the confusion matrix. These errors are related to producer accuracy, which measures how well class members were correctly classified.

Higher errors of commission and omission were observed in the shrubland classes, which is expected, as validating these structural classes using high-resolution images is inherently challenging. Differences between high or low shrublands are not always easy to differentiate. In contrast, both producer and user accuracies were found to be high for forest ecosystems, indicating greater reliability in their classification.

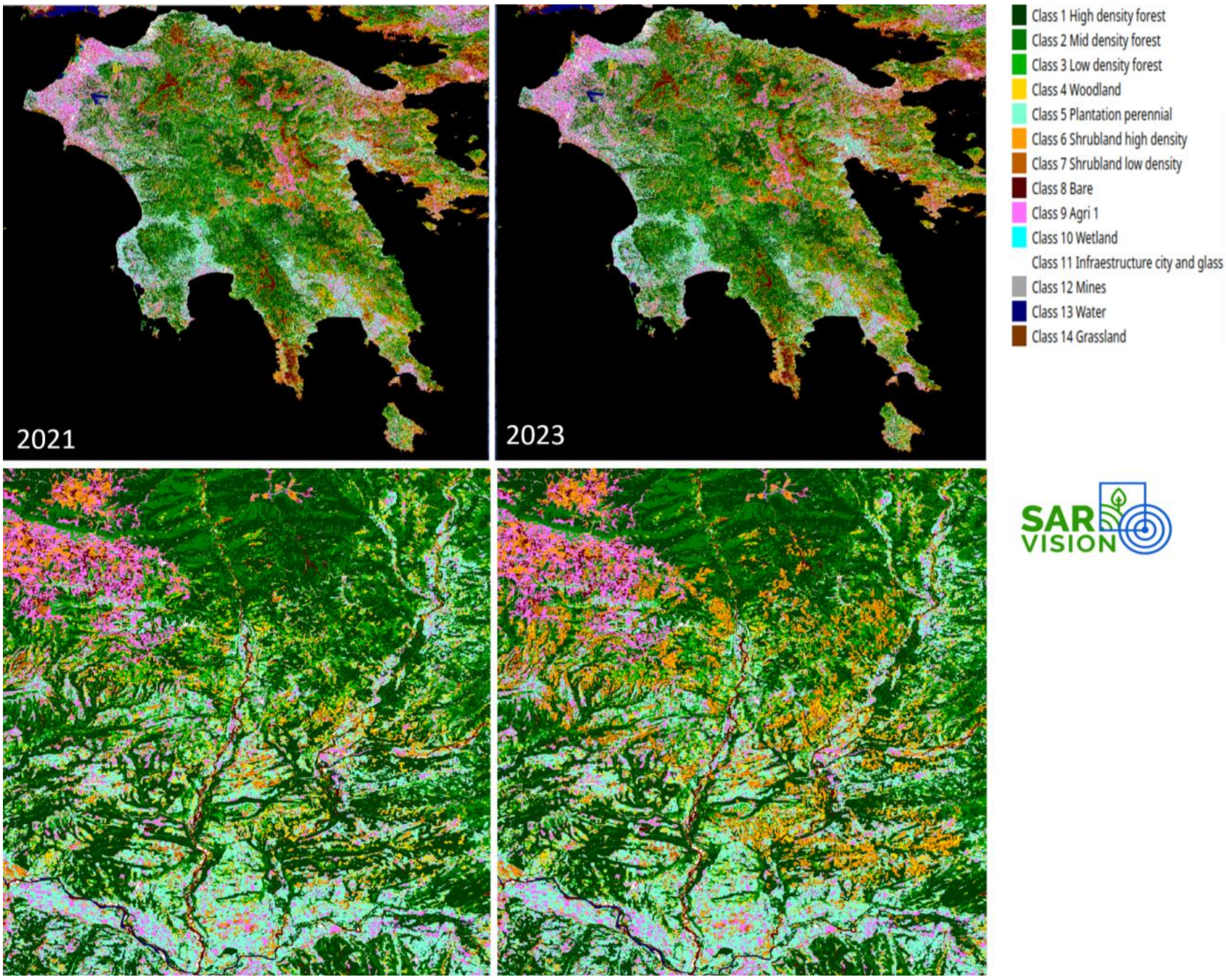
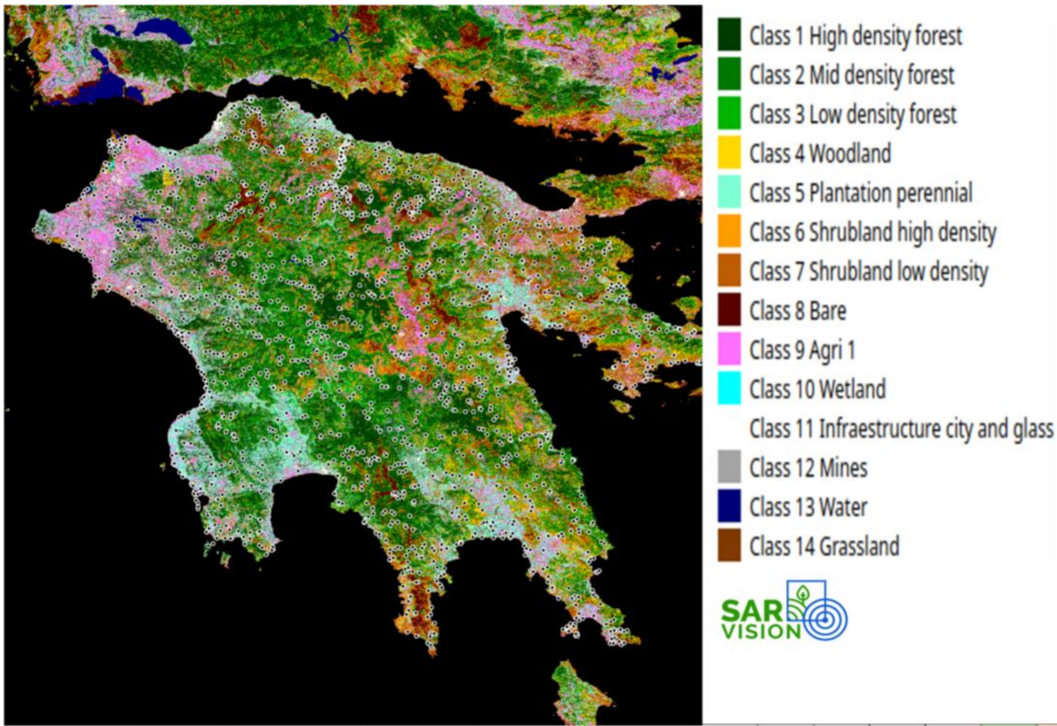


Figure 74: Forest Structural Maps of Peloponnesus 2021-2023. Structural differences can be seen specially the transformation of forest (green) into shrubland classes (orange). This forest loss and forest degradation were mainly attributed to fires, the whole area (at the top) and detail over an area of change between the two years (at the bottom).

For Peloponnesus, FSM were created for 2021 and 2023 (Figure 74). The forest structural types and ecosystems of Peloponnesus were more complex than those found in São Miguel. In this region, distinctions between forest types were primarily based on forest density, with classifications into high-, medium-, and low-density forests. Other natural vegetation structures included open woodlands and shrublands. Bare areas were identified at mountain peaks, which are occasionally covered by snow in winter, as well as in regions near mining activities and coastal areas. Additionally, tree plantations, particularly olive groves, were detected alongside extensive agricultural lands. Wetland areas were also identified in the northwest of the island. The mapping of Peloponnesus was more challenging, as reflected in the validation accuracy and confusion matrix. The overall accuracy was 82%, with a corresponding Kappa coefficient of 0.8. Out of 1827 randomly generated points, 1512 were correctly classified. Higher producer and user accuracy were observed, particularly for high- and medium-density forests. However, the highest classification errors were found in the shrubland and forest plantation classes. In contrast, classification accuracies for bare areas and low shrublands were high, indicating distinct vegetation density for these categories. On the other hand, woodlands, high shrublands, and forest plantations showed greater confusion with one another, suggesting lower mapping resolution for these classes (Figure 75).



		1	2	3	4	5	6	7	8	9	10	11	13	14	Total	commision	User Accuracy
High density forest	1	98.19	0	4.26	4.29	5.1	5.71	1.86	0.85	5.45	16.67	0	3.85	0	25.94	8.23	91.77
Mid forest	2	0	91.53	3.55	1.84	0.64	1.43	0	0	0.91	0	1.41	3.85	0	9.58	7.43	92.57
Low forest	3	0.23	3.95	83.69	7.36	9.55	17.14	4.19	0.85	6.36	0	1.41	0	8.33	10.13	36.22	63.78
Woodland	4	0	0	0.71	60.12	4.46	1.43	5.58	0.85	2.73	0	0	0	8.33	6.84	21.6	78.4
Plantation perennial	5	1.13	2.82	4.96	15.95	67.52	8.57	2.33	0.42	10	0	0	0	0	9.41	38.37	61.63
Shrubland high	6	0	1.13	0.71	4.91	6.37	48.57	1.86	0	2.73	0	0	0	0	3.39	45.16	54.84
Shrubland low	7	0	0	0.71	1.84	2.55	8.57	79.07	0.85	6.36	0	5.63	0	0	10.78	13.71	86.29
Bare	8	0	0	0	0	0.64	0	0	93.64	0	0	1.41	15.38	0	12.42	2.64	97.36
Agri	9	0.23	0	1.42	3.07	3.18	8.57	4.65	0.42	65.45	0	0	0	8.33	5.64	30.1	69.9
Wetland	10	0	0	0	0	0	0	0	0.42	0	66.67	0	0	0	0.27	20	80
Infraestructure	11	0	0.56	0	0.61	0	0	0.47	1.69	0	16.67	88.73	0	0	3.89	11.27	88.73
Water	13	0	0	0	0	0	0	0	0	0	0	1.41	76.92	0	1.15	4.76	95.24
Grassland	14	0.23	0	0	0	0	0	0	0	0	0	0	0	75	0.55	10	90
	total	100	100	100	100	100	100	100	100	100	100	100	100	100	100		
	Ommision error	1.81	8.47	16.31	39.88	32.48	51.43	20.93	6.36	34.55	33.33	11.27	23.08	25			
	Producer accuracy	98.19	91.53	83.69	60.12	67.52	48.57	79.07	93.64	65.45	66.67	88.73	76.92	75			

Figure 75: Confusion matrix presenting the validation results for the FSM 2023 in Peloponnese. Map presents the location of randomly selected points used for map validation using high resolution Planet data as a reference. Overall accuracy was estimated in 82% with a kappa coefficient of 0.8.

8.3.1.2. SarSentry Deforestation and degradation time series

The second step in carbon accounting is the accurate detection of vegetation structural changes in forest ecosystems. This is of great importance, as changes in forest cover represent a high potential for carbon emissions due to the high carbon levels in the stocks.

Forest change detection in this study was conducted using the SarSentry System, developed by SarVision (Hoekman et al. 2020). This system processes Sentinel-1 radar time series — acquired every 12 days globally, and every 6 days in European countries — to generate thematic maps that display deforestation and forest degradation within a predefined baseline forest mask.

For non-seasonal ecosystems, analyses are carried out every 12 days, while for seasonal ecosystems, they are performed on an annual basis to account for natural variations such as leaf fading, which can influence radar signal responses. Radar data enables detection of seasonal effects and forest structural changes, producing a sequence of maps showing forest areas remaining unchanged, and those affected by deforestation, degradation, or regrowth.

Each annual map includes a legend that classifies detected changes, with implications for carbon dynamics:

- Deforestation and degradation: Indicate a loss in biomass linked to carbon emissions.
- Regrowth: Indicates a gain in biomass, corresponding to carbon sequestration.

Deforestation detection involves temporal confirmation. The system identifies 3 classes:

- Early Deforestation (Red): Initial detections that require confirmation in future time.
- Deforestation (Yellow): Areas where deforestation has been confirmed.
- Deforestation 2 (dark Yellow): Areas previously deforested and confirmed, which may now show signs of regrowth.

Forest Degradation is detected through texture change analysis in radar imagery and is mapped at a coarser resolution (150 m) compared to deforestation (10 m). The 4 levels of degradation are:

- Levels 1–3: Represent canopy structure changes without converting the land to a non-forest class.
- Level 4: Reflects deeper impacts, such as internal structural degradation and biomass reduction.

Regrowth shows areas where the radar backscatter increases, often more visible in formerly non-forest areas. In established forest areas, a minimum increase of 2 dB in radar backscatter is generally required to indicate biomass gain in mature forests (i.e., forests in a climax state).

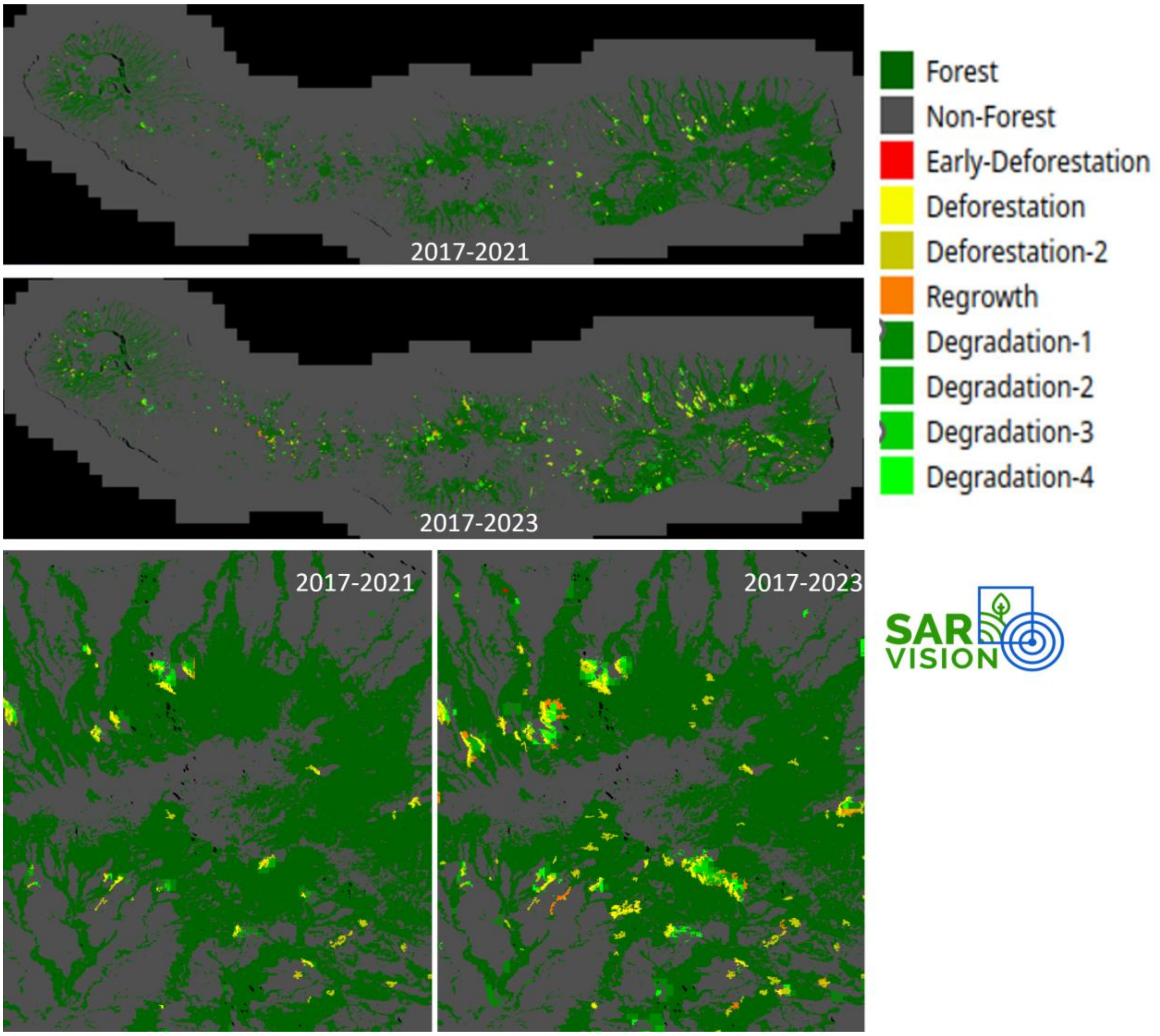


Figure 76: SarSentry results for São Miguel. Observation periods correspond to 2017-2021 and 2017-2023. SarSentry thematic products show the yearly forest change detections (deforestation, degradation and regrowth); the whole island (at the top) and detail over the east volcanic area where changes in the forest cover were detected (at the bottom).

The results for São Miguel are shown in Figure 76. An analysis of changes that occurred between the time periods 2017-2021 and 2017-2023 is presented. This data allows for tracking forest changes that can be associated with carbon fluxes. The use of detailed deforestation information over time helps with carbon accounting between two periods where biomass is mapped. In the case of deforestation, the assumption is that all carbon contained in the deforested area is being emitted. Estimating the gains or losses from degradation and regrowth depends on the development of locally calibrated models and requires input from field measurements and other local observation. Changes in the forest cover detected as deforestation can be associated with management and rotation practices or landslides on very steep slopes. Degradation processes might also be related to selective logging or other management practices.

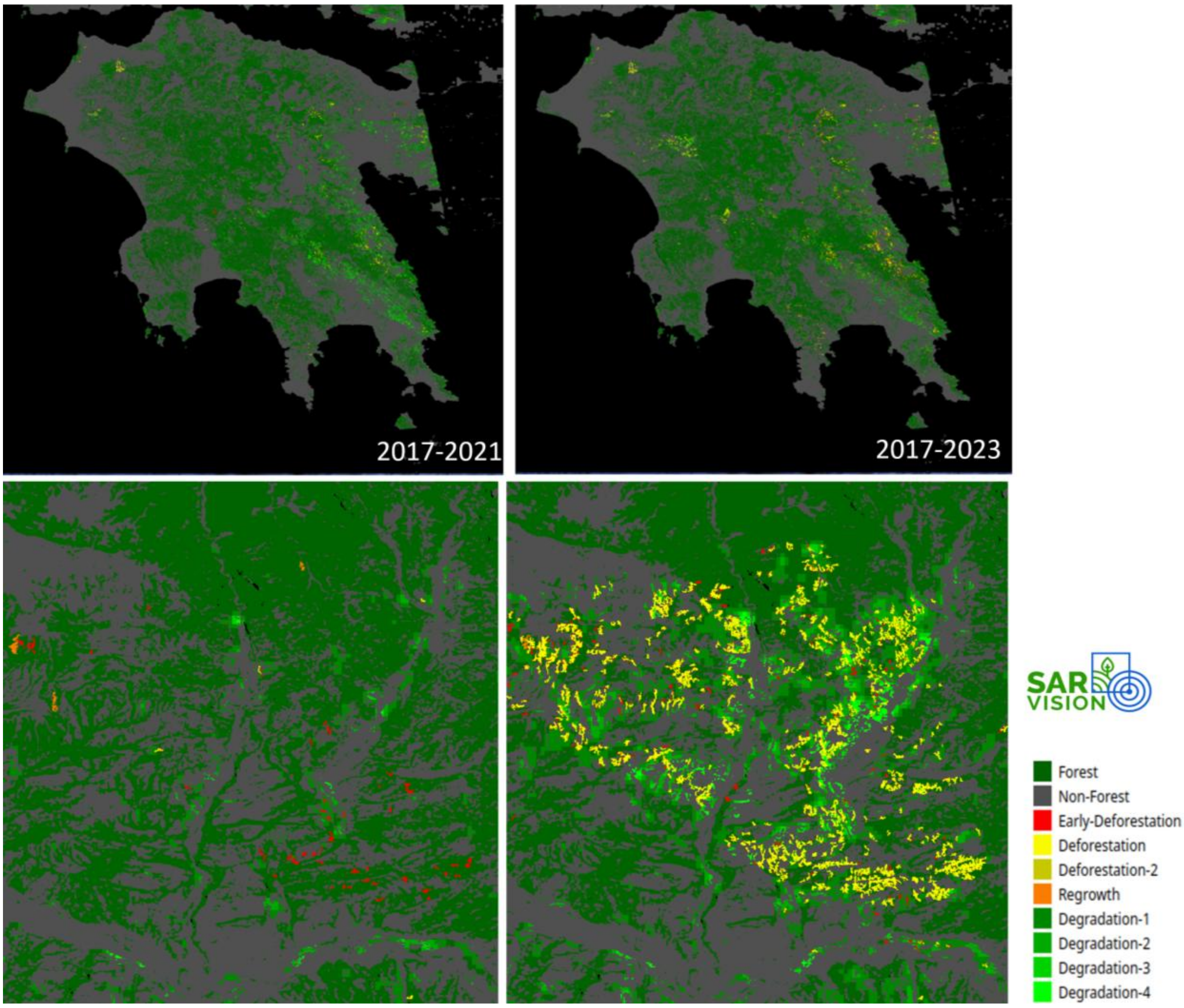


Figure 77: SarSentry thematic results for Peloponnesus for the 2017-2021 and the 2017-2023 periods; SarSentry thematic products show the yearly forest change detections (deforestation, degradation and regrowth); the whole peninsula (at the top) and detail over the central area where changes in the forest cover were detected (at the bottom).

For Peloponnesus, the SarSentry changes were also calculated for the period between 2017-2021 and 2017-2023. Figure 77 shows the results for the entire region, with a detailed view of an area where many changes were detected. In Peloponnesus, the changes in forest cover can be attributed to forest management, but most of the time they are associated with forest fires occurring during the summer.

8.3.1.3. Biomass Maps

The next step in carbon mapping is the production of biomass maps. This mapping process is carried out using the SarCarbon algorithm, developed by SarVision, it is a complex system that integrates raster data from the derived Shadow Index (SATVI), calculated using Sentinel-2 optical data for the year of mapping, the FSM map for the same year, and the SarSentry results for the same year, with canopy height elevation point data derived from the LiDAR GEDI system.

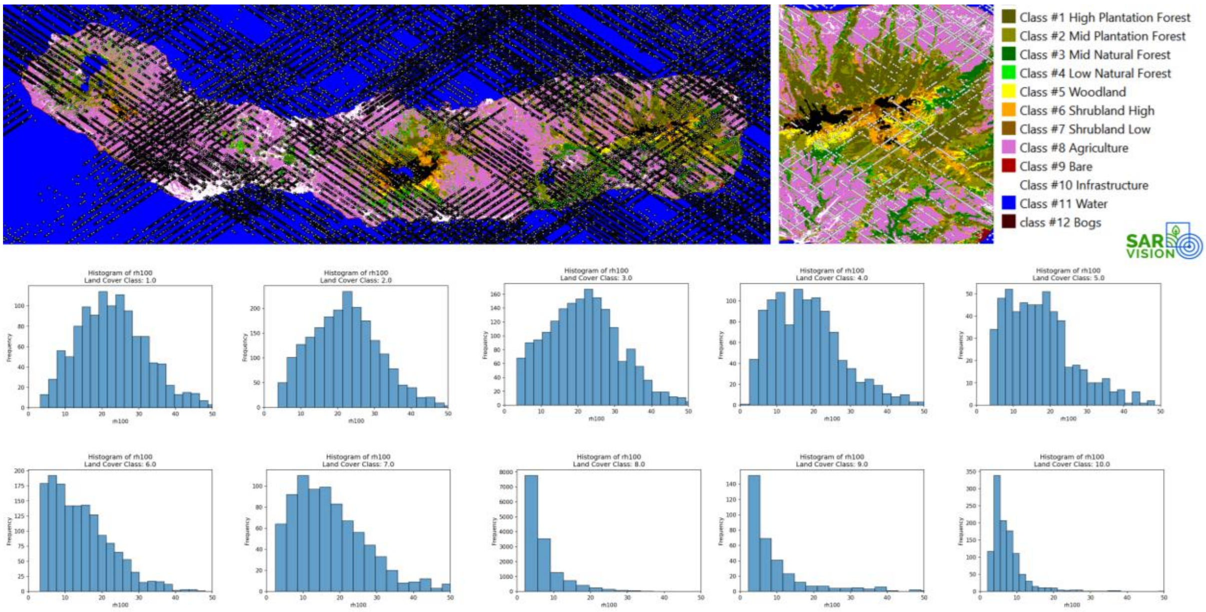


Figure 78: Location of GEDI footprints acquire over the area of São Miguel during years 2020, 2021 and 2022. In total, 57200 points with canopy height data (h100) were used for the SarCarbon analysis, covering most of the Island. Histograms showing the distribution of canopy height data for each of the forest structural classes of the structural maps are shown.

Figure 78 illustrates the distribution of all GEDI canopy height data downloaded for the island of São Miguel for the years 2020 and 2021. After data screening, a total of 57.200 points were used for analysis. The H100 variable from the GEDI database was utilized for canopy height calculation. Canopy height data in meters was extracted for all forest structural classes, and the distributions of canopy heights for each class were visualized using histograms, displaying the distribution of vegetation height ranging from 0 to 50 meters. The canopy height distributions per class are related to the SI index value and the correlation is used to create the intermediate raster canopy height map.

A pixel-based biomass classification model integrates reflectance data from the Sentinel-2 Optical Index with height distributions for each forest structural class. This process generates a temporal raster file of vegetation height, which is converted into biomass data using general allometric equations (Asner et al. 2012). Figure 79 shows the São Miguel biomass map.

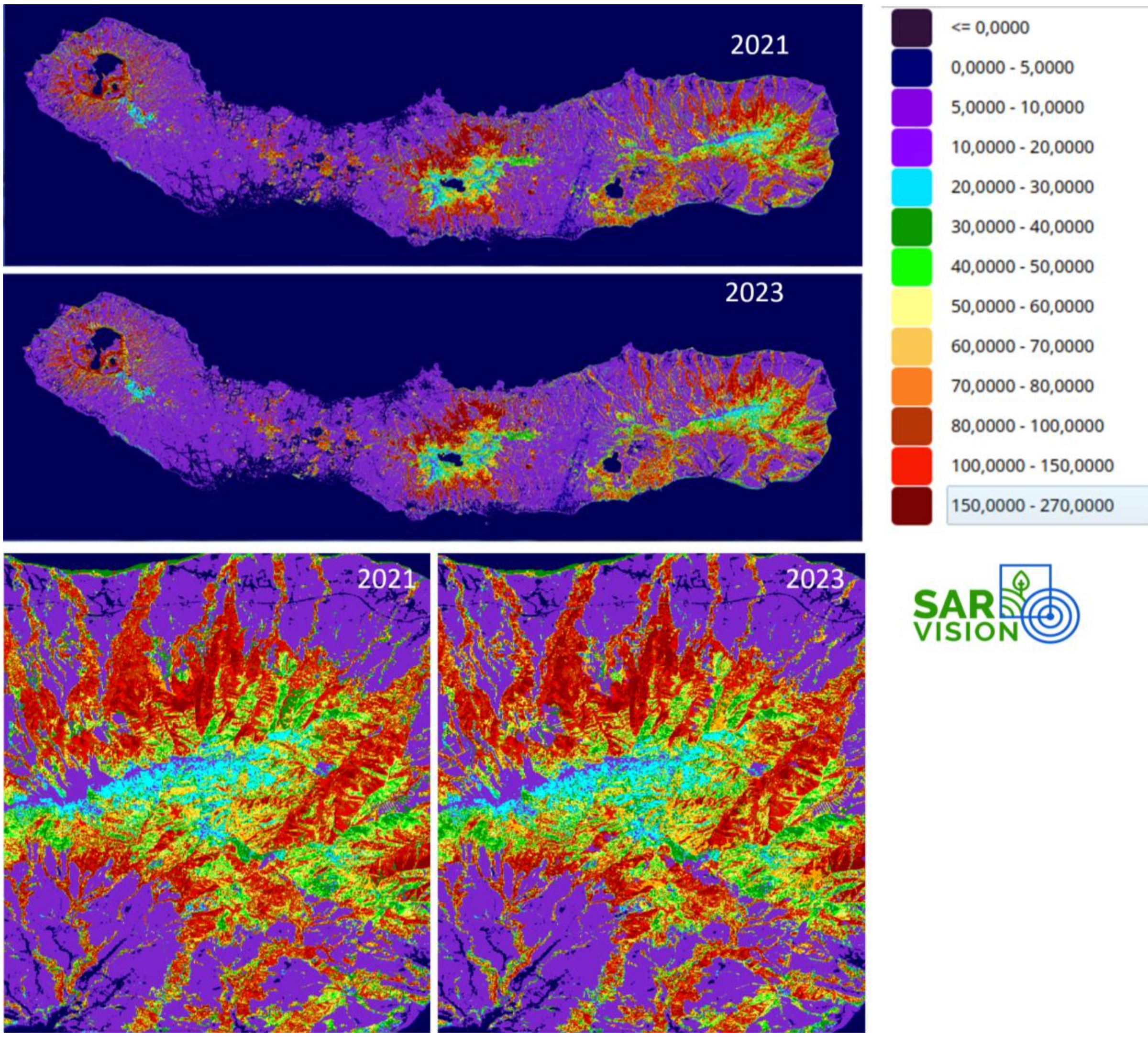


Figure 79: Final biomass maps for São Miguel. Data per pixel is expressed in tons /ha. The legend is aggregating the biomass values of a float map into classes. Values range from 5 ton/ha to a maximum of 270 ton/ha; the whole island (at the top) and a detail from the mountainous eastern sections are shown (at the bottom).

Biomass maps for 2021 and 2023 were calculated separately using different remote sensing data sources, including Sentinel-2 images and forest structural maps. However, the canopy height-derived data for each structural type is shared by both maps. The biomass maps are in float raster format; each pixel represents an estimated biomass value in tons/ha. Any calculations involving these maps should account for a correction factor due to the actual 10 m resolution. The estimated biomass values for São Miguel range from 0 to 270 tons/ha. Unfortunately, no field Aboveground Biomass (AGB) data was available to validate these maps. However, carbon estimates conducted on the island of Madeira, also in Macaronesia and sharing similar forest types, produced comparable results (Masseti and Gil 2020).

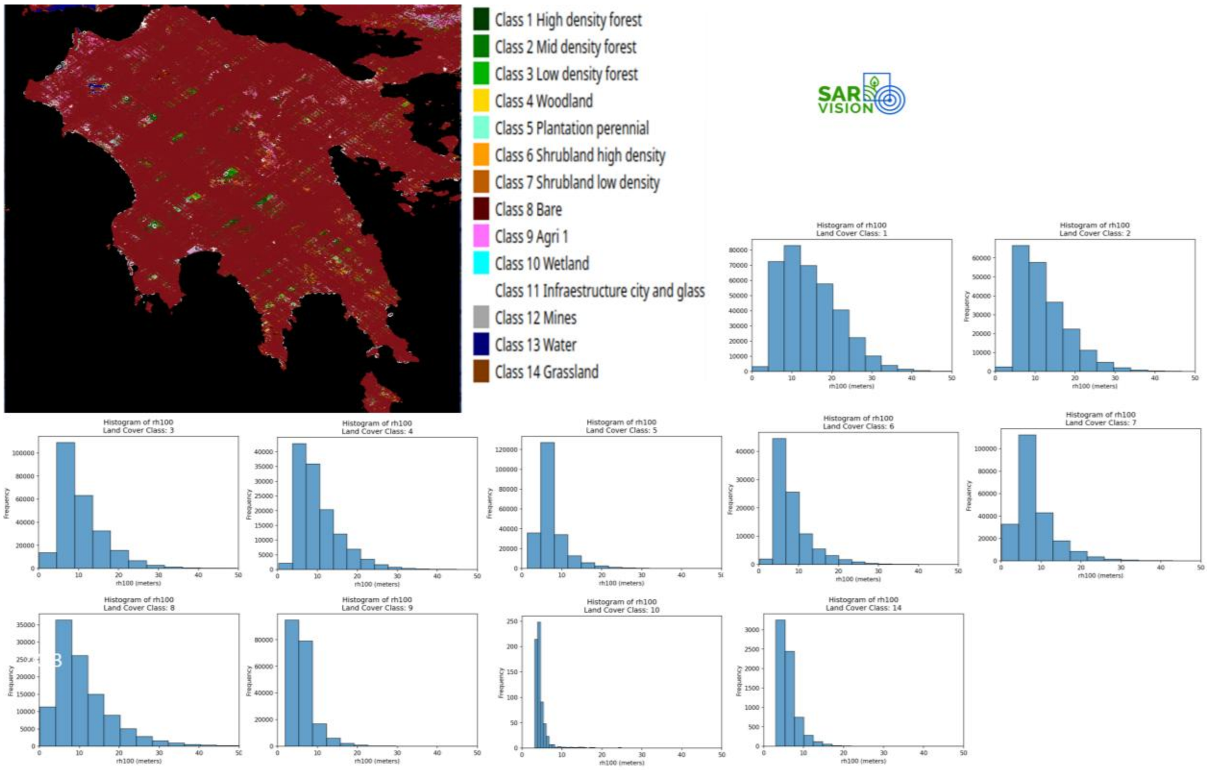


Figure 80: Location of GEDI footprints acquired over Peloponnese in 2019, 2020 and 2021. In total, 908.000 points with canopy height data (h100) were used for the SarCarbon analysis, covering most of the peninsula. Histograms showing the distribution of canopy height data for each of the forest structural classes of the structural maps are shown.

Figure 80 shows the GEDI coverage over Peloponnesus Peninsula. After data screening, a total of 908.000 points were used. Canopy height data (in meters) was calculated from the H100 variable and extracted for all vegetation structural types. Histograms were generated to illustrate canopy height distributions for each vegetation structural type. These distributions serve as inputs for the SarCarbon algorithm in defining biomass levels by vegetation type.

Biomass maps were produced for 2021 and 2023, showing estimated biomass levels across the region (Figure 81). Biomass values in Peloponnesus range from 0 to 270 tons per hectare, with the highest values observed in forested areas. Although no field-based AGB (Above-Ground Biomass) data is available for validation, the estimates align with values published in the 2006 IPCC Guidelines (Volume 4, Tables 4.7 and 4.8) (Eggleston et al. 2006).

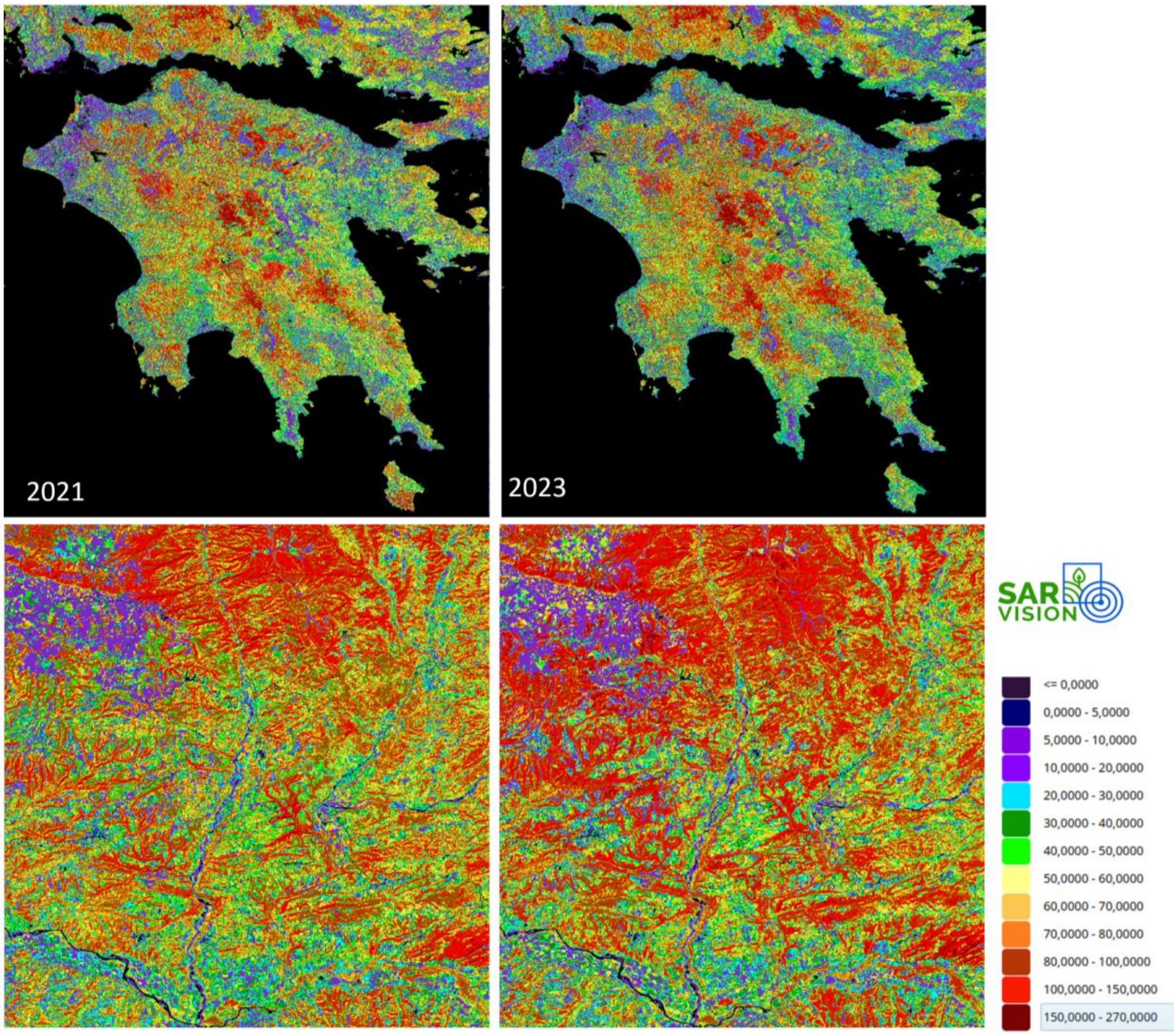


Figure 81: Final biomass maps estimated for Peloponnese. Data per pixel is expressed in tons/ha. The legend is aggregating the biomass values of a float map into ranges. Values range from 5 ton/ha to 270 ton/ha; the whole area (at the top) and detail of areas were the biomass changes between the years(at the bottom).

The proper validation of these maps should be done using field measured biomass estimated for the same study period. Unfortunately, this data was not available for the TS. For that reason, data published by IPCC is used as best practice guidelines for carbon estimates in different ecosystems around the globe. (IPCC guidelines, Vol 4, 2006 and 2019) (Eggleston et al. 2006, Calvo Buendia et al. 2019).

The first step in using this guideline is to define the ecosystem in question. For Peloponnesus and São Miguel, temperate continental forest and temperate oceanic forest respectively for European ecosystems are used as a reference, based on each TS geographical location and climatic classification (table 4.1 of the IPCC guideline, Vol. 4). As a validation proxy for the biomass map estimations in this project, we used published data from the same guidelines. For both temperate continental and oceanic natural forests, the reference biomass is approximately 120 tons per hectare. For temperate oceanic forest plantations, values are given for broadleaved (200 ton/ha) and coniferous (150-250 ton/ha) forest. For temperate continental forest values range form 30 to 200 ton/ha. (table 4.7 of the IPCC guideline, Vol. 4 and table 4.8 of the IPCC guideline, Vol. 4).

8.3.1.4. Carbon Maps

The carbon mapping process follows biomass mapping, as carbon is a fraction of the estimated AGB. Conversion fractions for different ecosystems are provided in the IPCC Guidelines (table 4.3 of the IPCC guideline, Vol. 4,) (2006–2019, Table 4.3) (Eggleston et al. 2006, Calvo Buendia et al. 2019). According to the IPCC table, average conversion factors for various forest types in temperate oceanic and continental ecosystems range between 0.47 and 0.51. A conversion factor of 0.48 was used for both TS.

Final carbon maps are shown for São Miguel (Figure 82) and Peloponnese (Figure 83). Carbon estimates for each TS range from 0-100 tons of carbon per hectare (ton C/ha). The maps are raster-based floating-point data, allowing carbon values (ton C/ha) to be read per pixel. When extracting data for carbon accounting, corrections are applied based on the data resolution.

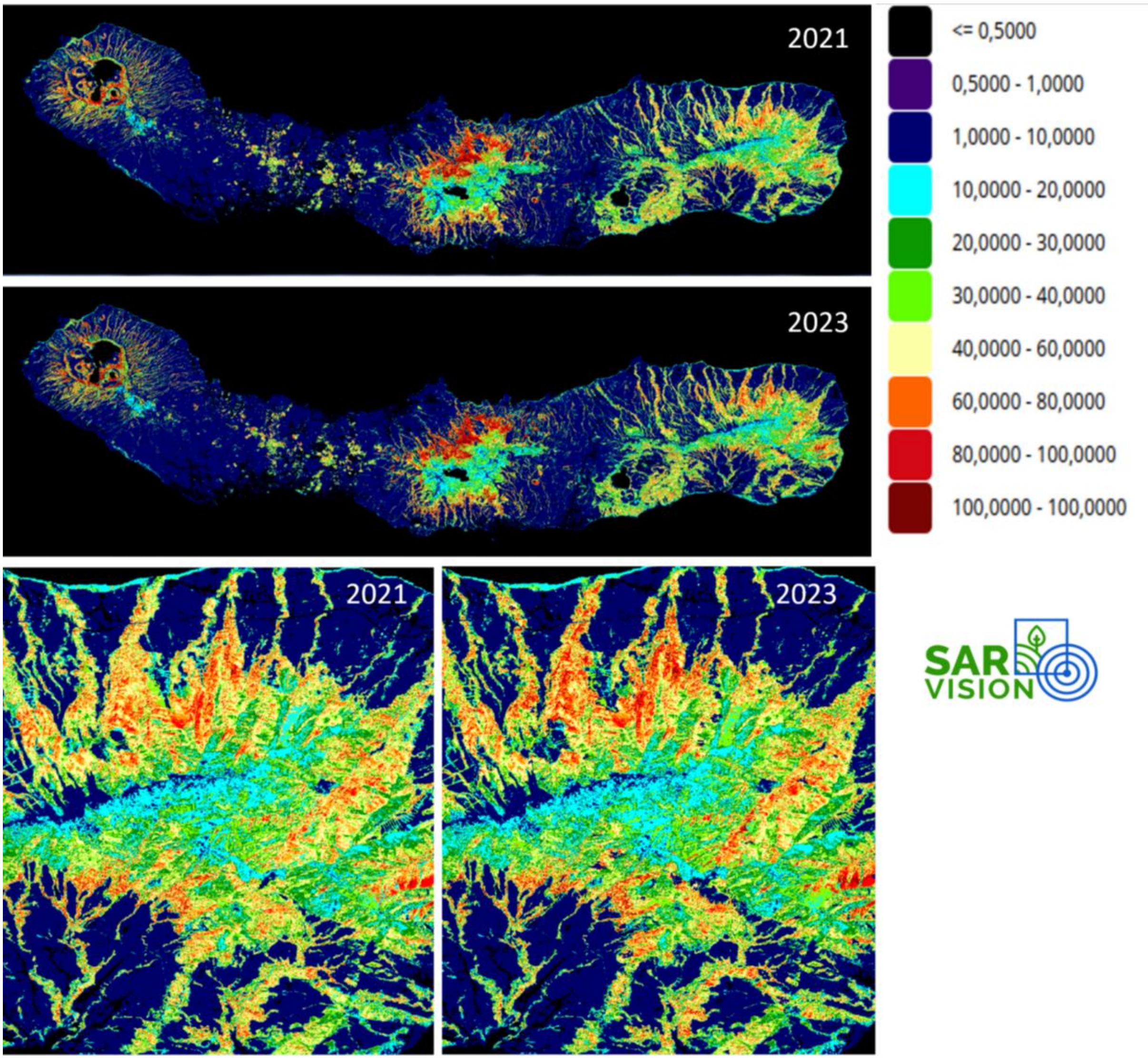


Figure 82: Final carbon maps for São Miguel. Data per pixel is expressed in tons C/ha. The legend is aggregating the carbons values of a float map into classes. Values range from 5 ton C/ha to a maximum of 100 ton C/ha; the whole island (at the top) and detail over a mountainous region of the island (at the bottom).

Additional carbon reference data was found for Madeira Island, also belonging in Macaronesia and where forest types occur like those from São Miguel. Reported carbon levels for Madeira were of the same magnitude (Masseti and Gil 2020).

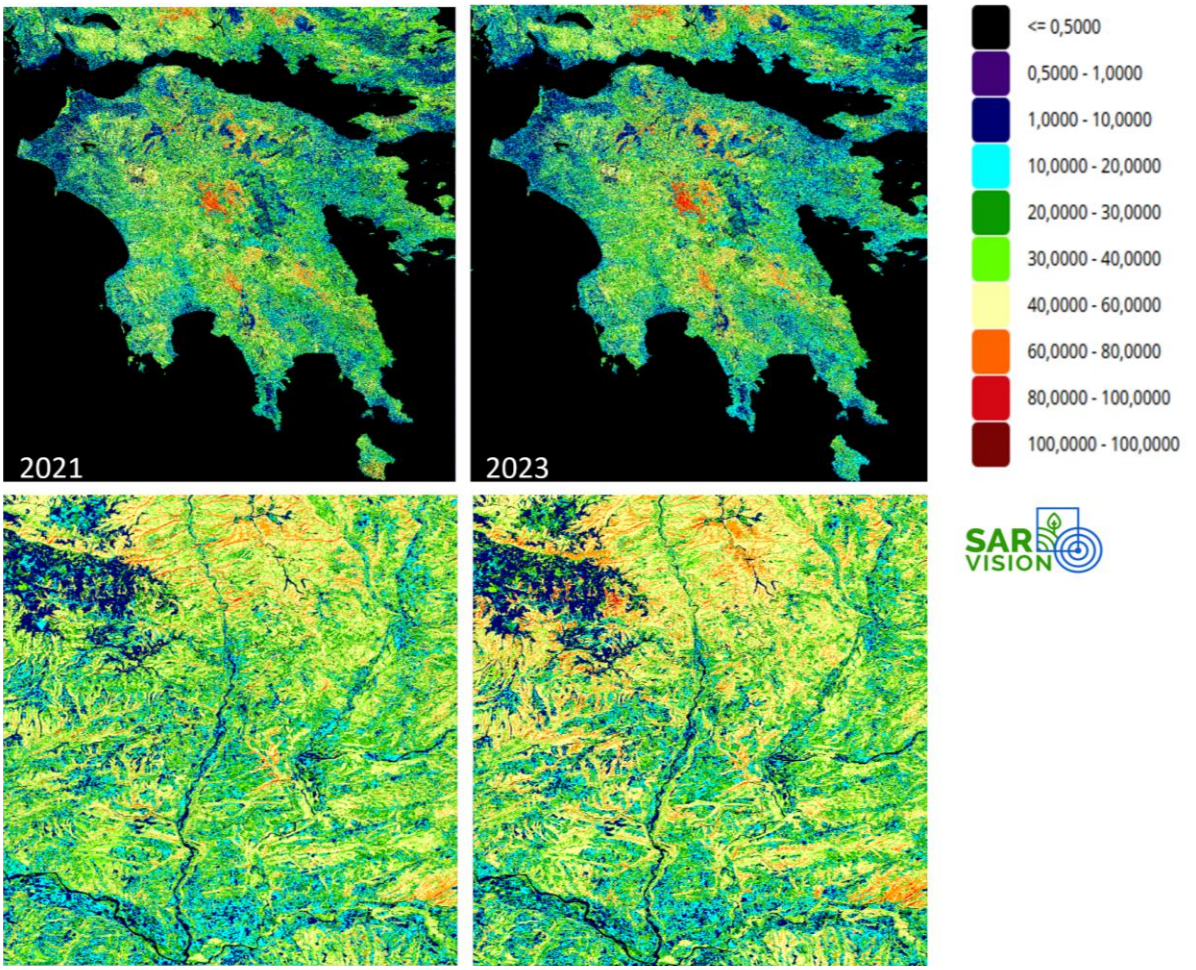


Figure 83: Final carbon maps for Peloponnese. Data per pixel is expressed in tons C/ha. The legend is aggregating the carbons values of a float map into classes. Values range from 5 ton C/ha to 100 ton C/ha; the whole area (at the top) and details over the central region where changes were observed (at the bottom).

8.3.1.5. Carbon emissions by carbon stock changes

For the calculation of carbon emissions, we used the equation presented in Section 7.3.1., for both TS, which calculates the difference between the carbon maps from the two available dates. The difference maps can be seen in Figure 84 for São Miguel and Figure 85 for Peloponnese. The carbon difference maps are expressed in tons of carbon per hectare (ton C/ha). The legend aggregates the carbon values from a floating-point raster map into distinct carbon classes. The estimated carbon differences between the 2023 and 2021 maps show both positive and negative values. Negative values indicate areas where carbon emissions were detected, while positive values suggest regions where carbon sequestration occurred. This information serves the calculation of the total carbon in ton C/ha emitted or sequestered for each pixel. For instance, in the case of the areas where deforestation was detected between the two years, it is expected that there is loss of biomass and a subsequent carbon emission due to the forest change. In the case of São Miguel (Figure 84) it is possible to see the red pixels where there are carbon emissions between 30 to 136 tons/ha in a period of two years. Most of the area appears in black where very small changes were detected. For the rest, pixels in orange and blue show fluxes (carbon emission or uptakes), corresponding to the published data for both ecosystems.

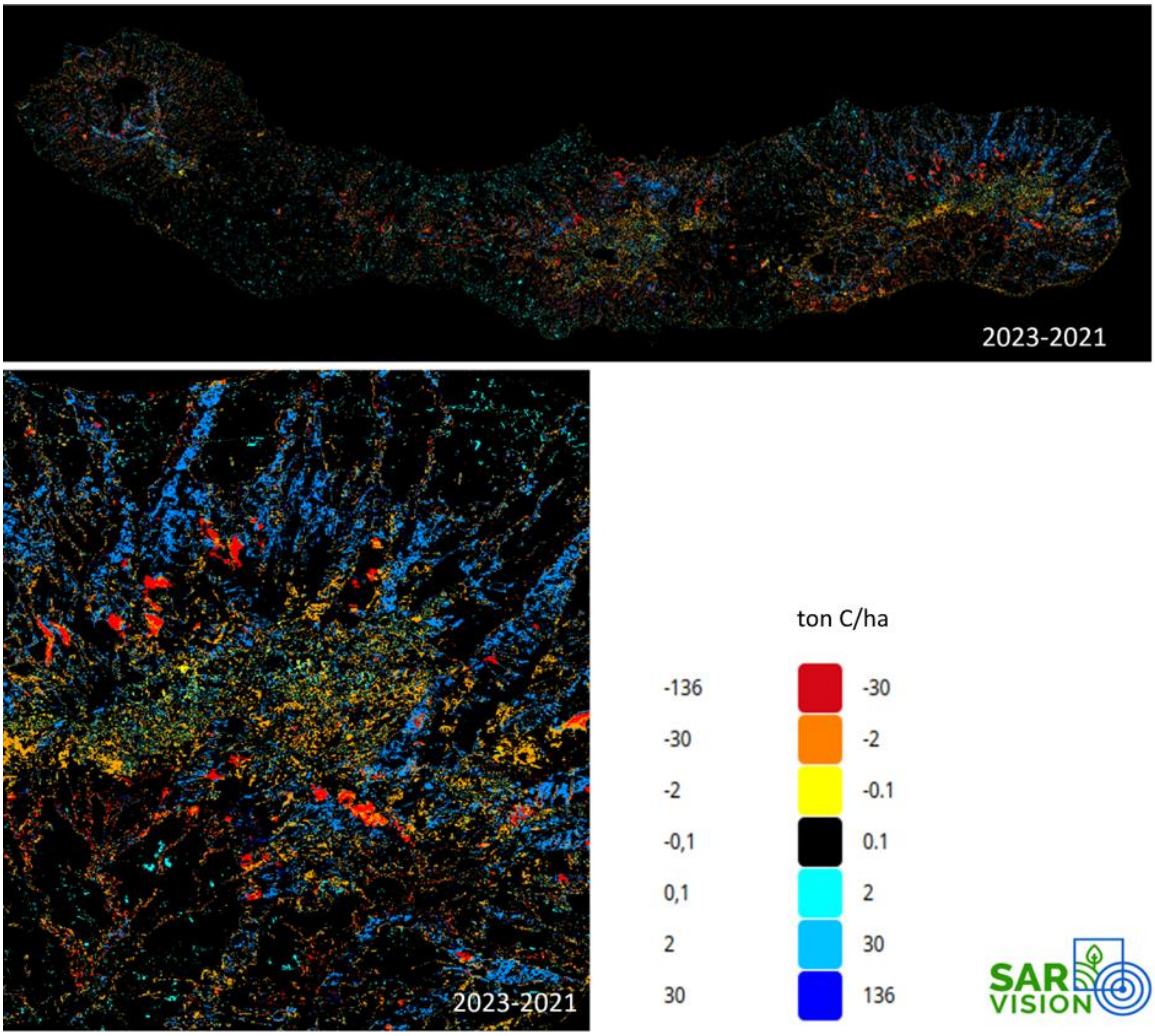


Figure 84: Carbon difference (flux) map for São Miguel expressing data in tons of carbon per hectare (ton C/ha). The legend aggregates the carbon values from a floating-point raster map into distinct carbon classes; the whole island (at the top) and detail at the east end showing red areas where carbon emissions are due to deforestation (at the bottom). This calculation represents the change detected in two years.

In the case of São Miguel, most of the high emission values (red) correspond to areas where deforestation was detected. In Peloponnesus, the areas with the highest carbon differences, or highest emissions, are primarily located in regions affected by reported forest fires.

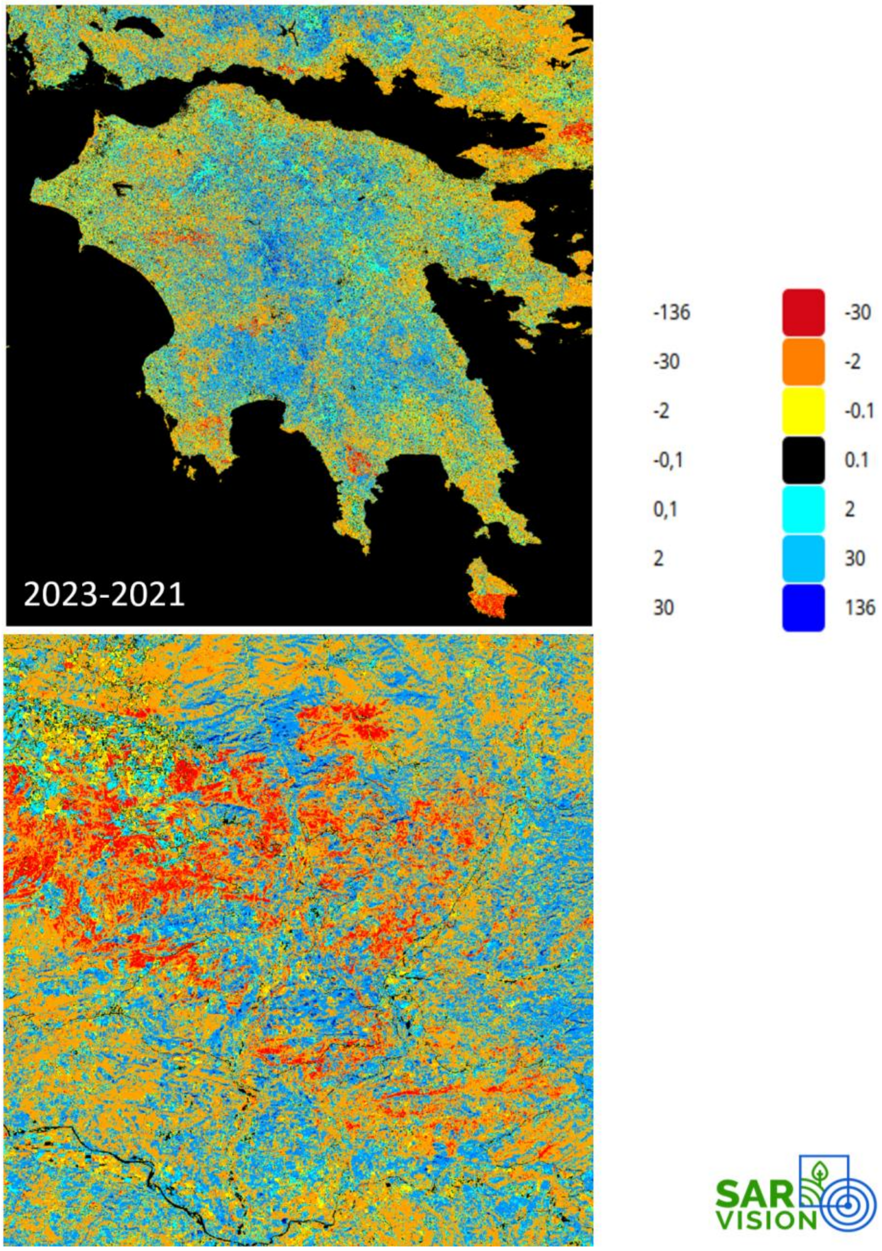


Figure 85: Carbon difference (flux) map for Peloponesse expressing data in tons of carbon per hectare (ton C/ha). The legend aggregates the carbon values from a floating-point raster map into distinct carbon classes; the whole peninsula (at the top) and detail at the central region showing red areas where carbon emissions are due to forest change (at the bottom). This calculation represents the change detected in two years.

There is no specific data available to validate carbon emissions per ecosystem for either TS. However, Table 4.9 from the IPCC guidelines provides a reference (table 4.9 of the IPCC guideline, Vol 4). For temperate oceanic forests, estimated biomass accumulation per year is approximately 2,3 tons per hectare and 4.0-7.5 for continental forest. This range aligns with the accumulated carbon values calculated from the C difference map for both TS.

8.3.2. Carbon Flux mapping using GPP

The total above-and belowground carbon sequestration (ANPP and BNPP respectively) within 2022 and 2023 for São Miguel (Figure 86) and Peloponnese (Figure 87) are shown. The total aboveground carbon sequestration is much higher in São Miguel than in Peloponnese. The hotspots of belowground carbon sequestration in São Miguel seem to be oriented around the habitats classified to EUNIS level 1 Q class, corresponding mainly to the blanket bogs.

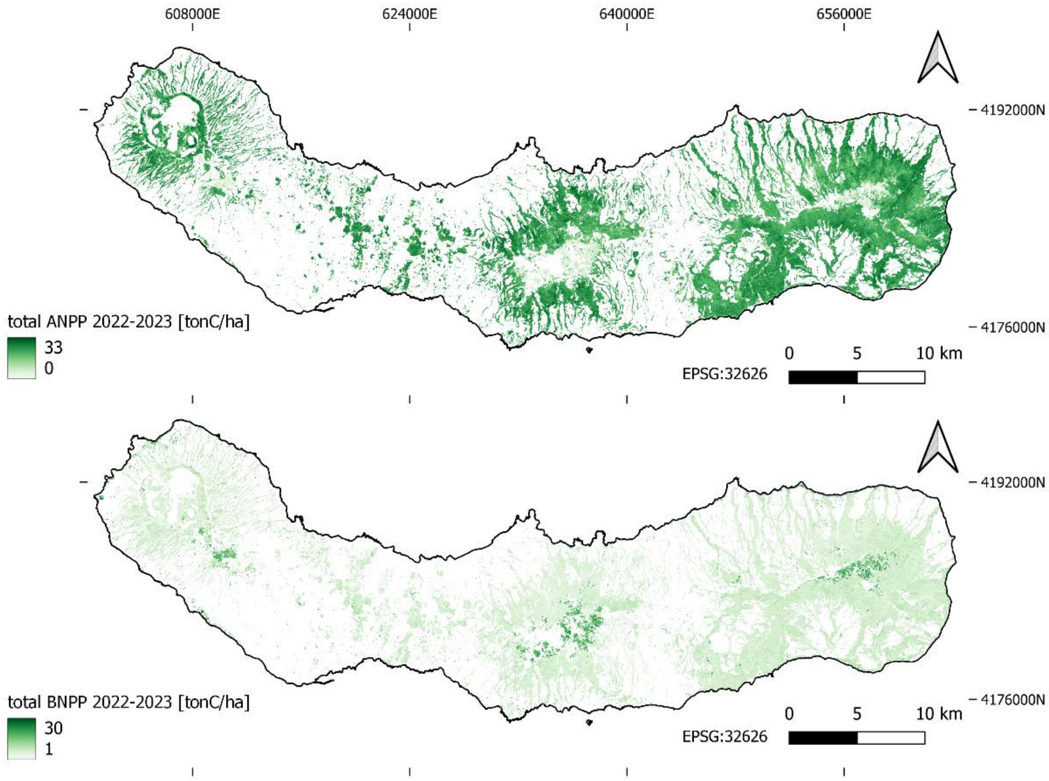


Figure 86: Total aboveground and belowground carbon sequestration for São Miguel by summing the aboveground carbon accumulation (ANPP) in 2022-2023 and belowground carbon accumulation (BNPP) in 2022-2023, respectively.

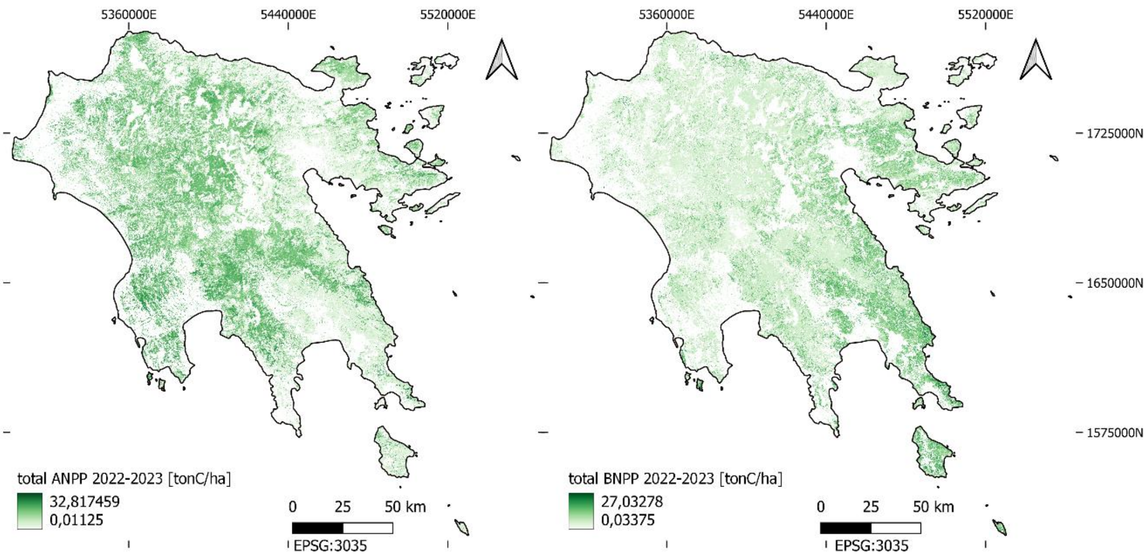


Figure 87: Total aboveground and belowground carbon sequestration for Peloponnese by summing the aboveground carbon accumulation (ANPP) in 2022-2023 and belowground carbon accumulation (BNPP) in 2022-2023, respectively.

Table 52 shows the mean total ANPP and BNPP for 2022 and 2023 summed. The table does not display a value for the 'Wetland' class in São Miguel, as the land cover map that was used to generate the GPP maps did not contain areas classified as 'Wetland' within the island.

Table 52: Overview of the mean total aboveground and belowground carbon sequestration (ANPP and BNPP respectively) for 2022 and 2023 summed, per ecosystem type class in São Miguel and Peloponnese.

TS - Class	Mean total ANPP 2022-2023 [tonC/ha]	Mean total BNPP 2022-2023 [tonC/ha]
São Miguel – Wetland	/	/
São Miguel – Broadleaf forest	21.05	6.68
São Miguel – Coniferous forest	21.53	6.26
São Miguel – Low Woody	5.69	17.94
São Miguel – Evergreen forest	23.63	7.19
Peloponnese – Wetland	2.57	7.72
Peloponnese – Broadleaf forest	15.19	5.18
Peloponnese – Coniferous forest	16.38	5.23
Peloponnese – Low Woody	4.55	12.87
Peloponnese – Evergreen forest	18.51	5.64

9. Discussion

In this Section, the national-centric approach for improving ecosystem extent mapping using CLMS data is addressed, followed by highlighting how São Miguel TS demonstrated the complementarity of forest structure mapping with the vegetation-centric habitat mapping to enhance classification accuracy. The results of the forest condition index for São Miguel TS and its limitations for future improvement are also further explored. Also, the differences between biomass and flux-based carbon accounting are discussed, highlighting how the approaches can be complementary and potentially lead to improved accuracy when integrated. Finally, some remarks on the potential of using RS products alone (particularly through unsupervised methods) to produce cost-effective datasets that support scalable and consistent carbon accounting efforts.

9.1. Ecosystem extent – National-centric approach

The aim of national-centric approach was to assess the potential of European Copernicus Land Monitoring Service data and products for the mapping of ecosystem extents. In the two test areas, two different starting situations existed:

- São Miguel, with a rather complete coverage of the island with two national databases (land use and forest inventory) as well as a complete coverage with CLMS data.
- Peloponnesus, with a much more limited amount of national data (only related to protected area, no full coverage) and with a mix of different CLMS data sets, none of them provide full coverage of the territory.

Some key lessons can be drawn from this work. Seamless, wall-to-wall data should be the basis for any extent mapping. A wall-to-wall coverage of the data is needed to ensure a homogeneous coverage and quality of the information and to avoid patches with different levels of information depth. The results of Peloponnesus are an example of such patchiness, as the resulting extent data has been sourced from national protected area data and various CLMS data covering different parts of the territory. At the same time, the combination of CLMS data and national data can provide additional value, for example by reaching a greater depth for individual classes of the extent typology, as demonstrated in the case of São Miguel.

However, land use data, which is mainly needed for assessing ecosystem extents, is not available from CLMS in a wall-to-wall manner. The so-called CLMS Priority Area Monitoring data offers some datasets—such as Urban Atlas, Coastal Zones, or Riparian Zones—that provide land use information, but not in a seamless manner. In contrast, the highest potential of land cover data, that is, the so-called CLMS High Resolution Layers, is provided for forested areas, where at least a differentiation by leaf types is possible, and by the new crop type layer, which allows for the differentiation of several crop types. This last layer is particularly useful, as even many national datasets do not provide such information.

Overall, in Europe, CLMS data can provide a first approximation of ecosystem extent based on open and freely available data, but with limitations regarding the biological background and actual use of most habitats or ecosystems, providing information mainly on ETA level 1.

9.2. Ecosystem extent – Complementing the vegetation-centric approach with forest structure mapping

Misclassifications between EUNIS habitat classes that share floristic (species composition) and physiognomic (structure and form) similarities were shown as a limitation of the automated habitat mapping, especially when relying predominantly on spectral information derived from optical remote sensing. Several systematic misclassifications remain, despite post-processing corrections. A notorious example is the misclassification between Heathlands (S) and (T) Forests. This issue is especially relevant in the context of the Azores, where vegetation types occur in complex and intermingled mosaic formations, with overlapping species composition and structural variability driven by environmental gradients such as elevation and wind exposure. The same dominant woody species can form T2 forests under favorable growth conditions or S4 heathlands when stunted by harsh climatic or edaphic factors, such as strong wind regimes at higher altitudes. This overlap complicates the classification process, as it challenges the reliability of spectral reflectance-based differentiation between classes, especially in transition zones where vegetation occurs in early successional stages. Thus, leveraging structural and compositional approaches may help particularly in such places where the same plant communities express themselves differently due to topographic, climatic, or disturbance-related gradients.

Even though the forest structure map is primarily concerned with quantifying canopy height and density, assigning classes based on vertical structure alone, while the EUNIS habitat classification encompasses a broader ecological scope, integrating species composition, habitat functions, and environmental conditions; complementary radar-based forest structure mapping, capable of capturing vertical vegetation structure, is a valid strategy for addressing these classification ambiguities in vegetated habitat classes. By using structural features like tree height and canopy density — key ecological criteria in forest inventories and the EUNIS classification, such as the 5-meter height threshold that separates forests from lower vegetation like heathlands — radar-based forest structure mapping strengthens the vegetation-based approach applied to both TS.

This was empirically observed by overlapping SarVision forest structure maps with VITO EUNIS habitat maps in São Miguel. These demonstrated a good level of spatial and boundary agreement with each other, particularly in the mountainous and forested areas of São Miguel, where vegetation is more structurally stratified. However, the forest structure maps clearly improved the delineation between low-growing shrubland and true woodland (Figure 88), while also discriminating the areas showing cultivation patterns from those occurring naturally, providing a more a nuanced mapping at the same spatial resolution.

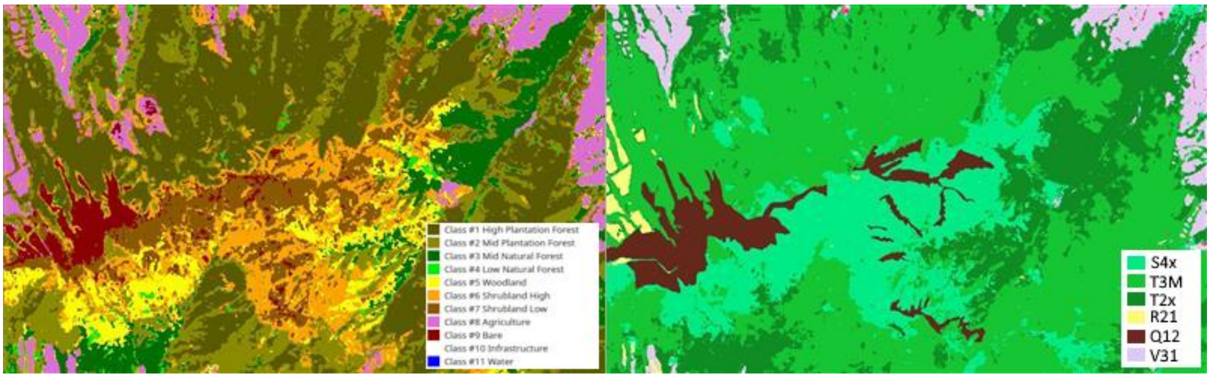


Figure 88: Side comparison of the forest structure map [left] and the EUNIS habitat map [right] in the same mountainous and forested area of São Miguel.

Structural mapping also enhances the temporal resolution of ecosystem monitoring by capturing short-term vegetation dynamics influencing the current biophysical extent, condition and carbon stocks and fluxes. Traditional land use or forest inventory datasets may continue to classify an area as “forest”, based on long-term management intentions or planned reforestation, while failing to capture abrupt and immediate changes in canopy height and tree density caused by disturbances such as tree harvesting, tree loss due to strong winds or slope movements, thus missing the transient stages of bareness, delayed replanting or early successional regrowth. In this perspective, complementing habitat mapping with forest structural mapping can further avoid misclassifications caused by the confusion between current/transient states and long term intended land use.

This complementary approach also highlights the broader methodological point of the importance of identifying and integrating exclusive features that are capable of separating classes with overlapping spectral or compositional characteristics. The first iteration of the EUNIS habitat mapping presented here serves as a baseline diagnostic, identifying which classes are most frequently confused and where additional or alternative features could contribute to more effective differentiation, while reinforcing that accurate and consistent field validation remains critical to improving this process.

9.3. Ecosystem condition mapping – Forest Condition Index

A more detailed discussion is done on the forest condition account created for São Miguel, since the forest condition account for Peloponnese indicated that only little fluctuation in the forest condition index was present. Besides, the main dip in forest condition index per forest type that was observed in 2021 is with high probability due to an artefact in the NPP data by the transition from PROBA-V to Sentinel-3/OLCI.

The condition indicators used for creating the forest condition accounts for forest types in São Miguel show that the most fluctuating condition indicator is the Net Primary Productivity (NPP). Aboveground biomass (AGB) and Normalized Difference Water Index (NDWI) remain quite stable through time and forest connectivity (FC) and threatened forest bird species diversity (TFBSD) remains the same for the period because they do not have updated data for after 2018 and 2012, respectively. Besides, the ranking of the condition indicators based on the specified criteria and expert knowledge also assign the second highest weight to the NPP condition indicator to be used in the calculation of the forest condition index. Hence, the NPP indicator has a high influence on the final FCI.

The graphs in Figure 89 show the evolution of the mean NPP indicator and the mean FCI per forest type through time. For each forest type (T1, T2, T3) as well as for temperate shrub heathland (S4), the NPP indicator decreases with index values between 0.07 and 0.17 from 2020 to 2021. Similarly to Peloponnese, a dip in the forest condition index in 2021 is observed. However, in contrast to Peloponnese, the forest condition account for São Miguel was created with the MODIS NPP dataset, to avoid the known issue of the CLMS NPP dataset. Therefore, this decrease cannot simply be assigned to satellite sensor artefacts.

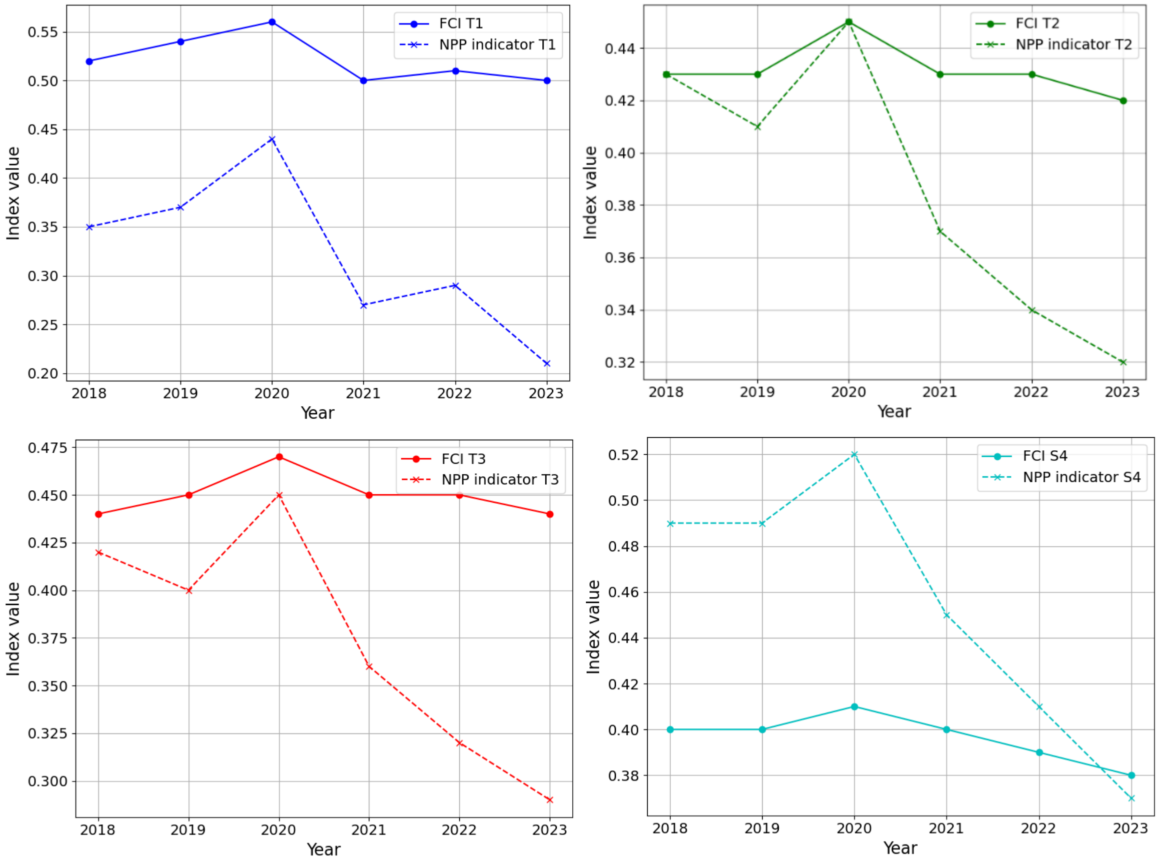


Figure 89: Evolution of the mean Net Primary Productivity (NPP) condition indicator plotted with the evolution of the mean forest condition index through time per forest type in Sáo Miguel for T1 Broadleaved deciduous forest, T2 Broadleaved evergreen forest, T3 Coniferous forest and S4 Temperate shrub heathland.

Net Primary Productivity in forests is much influenced by climate variables and elevation (Mehmood et al. 2024). Elevation gradients alter microclimatic conditions such as temperature and moisture availability, thus, affecting the NPP profiles over short geographical ranges (Chen and Zhang 2023). Lower elevations generally would support higher NPP since they are related to higher temperatures and water availability (Tao et al. 2022). At higher elevations, cooler temperatures occur, and the atmospheric pressure is lower, which leads to reduced NPP (Xu et al. 2024). Besides, at high altitudes the increased wind exposure can have strong effects on the vegetation’s productivity and the landforms (Zianis et al. 2005).

European forests have experienced record-breaking dry conditions during summer periods within the last decade. More specifically, the 2018-2020 drought and the summer drought in 2022 stand out. In 2018, insufficient precipitation and heatwaves were recorded (Zscheischler and Fischer 2020). Studies found that this led to wide-spread premature leaf senescence with consequences of unprecedented drought-induced tree mortality across various species and reduced tree growth (Bose et al. 2020, Schnabel et al. 2024). During 2019 and 2020, the exceptionally dry conditions even persisted and the tree stress responses in 2019 were even more pronounced compared to 2018 (Schnabel et al. 2024). This emphasizes the effect of consecutive and compound dry years compared to isolated dry years (Sachsenmaier et al. 2024). While the 2018-2020 drought was more pronounced in Northern and Central Europe, the summer drought of 2022 affected the southern regions of Europe by exceptionally pronounced atmospheric and soil dryness. A study by (Gharun et al. 2024) found that the negative impact on forests of the 2022 summer drought, indicated by declined GOSIF (global OCO-2 solar-induced fluorescence) was significantly greater than in 2018, despite the atmospheric and soil drought scores being lower in 2018.

Regional meteorological and udometric data further confirms the trend towards drier (Figure 90) and hotter (Figure 91) summers in the North Atlantic region in recent years, reflecting the rise of air temperature and change in rainfall patterns driven by climate change. The 2018 summer stands out as the driest since at least 2016, causing a short-term summer "drought" that was still remarkably impactful for agricultural crops and livestock due to the surface water shortage, causing major financial losses for producers.

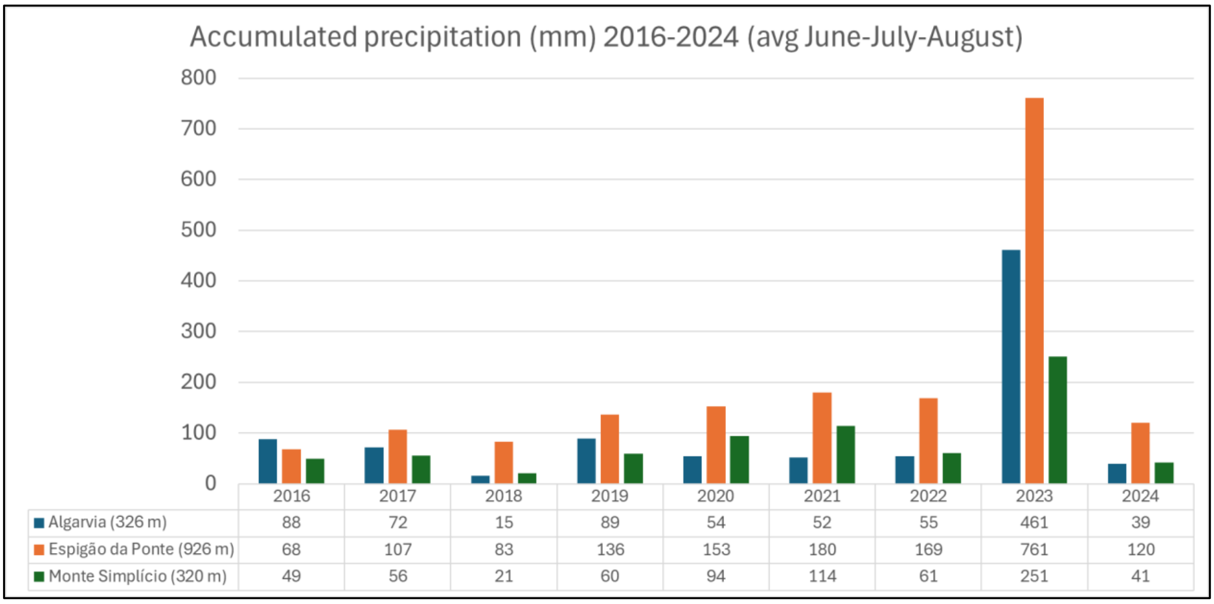


Figure 90: Average accumulated precipitation in summer months from 2016 to 2024 in São Miguel Island. Data gathered from the Hydrometeorological Network of the Azores (https://redehidro.ambiente.azores.gov.pt/).

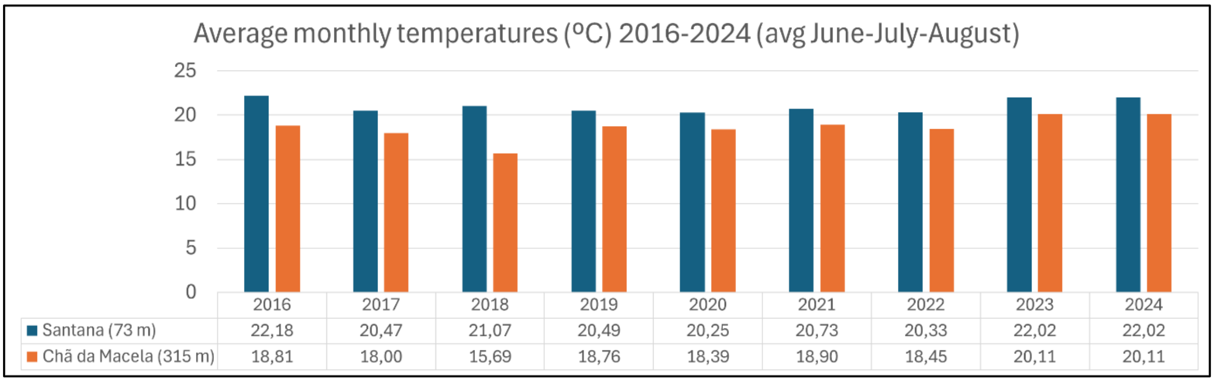


Figure 91: Average monthly temperatures in summer months from 2016 to 2024 in São Miguel Island. Data gathered from the Hydrometeorological Network of the Azores (https://redehidro.ambiente.azores.gov.pt/).

This suggests that consecutive drier years from 2018 to 2020 might have affected the forest condition to decrease abruptly in 2021, due to a loss in forest stability and resilience. Figure 92 shows that the decrease in forest condition index from 2020 to 2021 was highest at high elevations following the mountain ridges. The summer drought in 2022 had persisted the negative impacts on forest growth. Despite increasing precipitation rates in 2023, it is evident that forest need more time to recover and to be able to show positive responses in growth rates. Hence, the forest condition index will remain low in 2023.

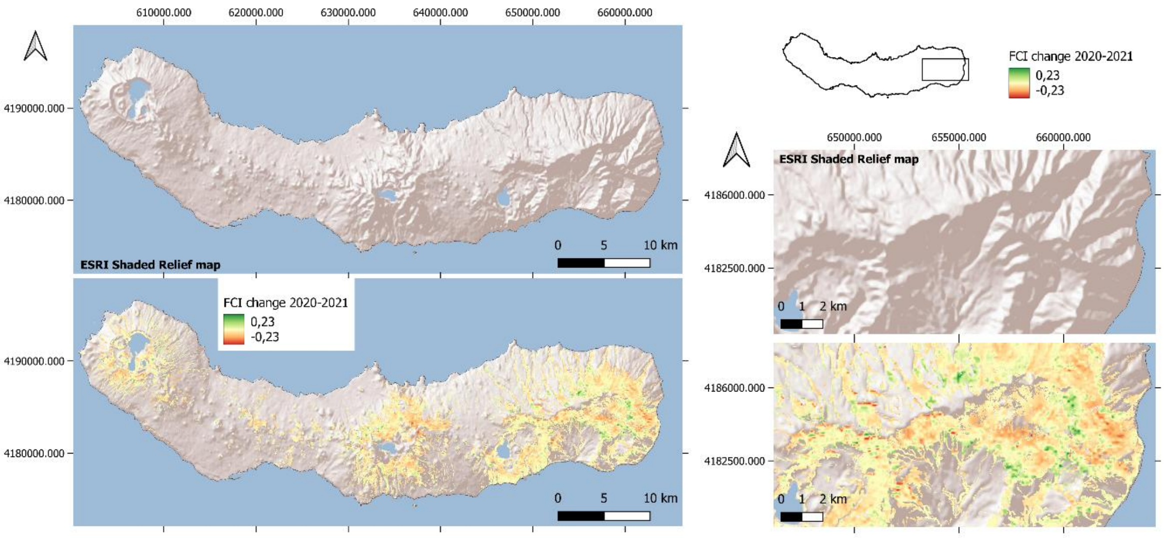


Figure 92: [Left] Forest condition index change map for 2020-2021 (bottom left) overlaid on orthographic relief image (ESRI, bottom top)) for São Miguel. [Right] Zoom-in on area in the East of São Miguel that demonstrates that most of the decline in forest condition index from 2020 to 2021 happened on higher elevations, following the mountain ridges.

It is essential to point out some limitations in the approach to calculating the forest condition accounts, based on the approach that was established in PEOPLE-EA.

First, obtaining high-resolution (both spatial and temporal) datasets that cover all ECT (Ecosystem Condition Typology) classes is a difficult task. To create a full ecosystem condition account, there is a need for spatial and temporal maps. These maps often, if not mostly, origin from models or remote sensing data sources, which have their own typical limitations in accuracy. The maps will have different coordinate reference systems, spatial resolution and temporal extent, which must be harmonized. Within this process of harmonization, error propagation occurs, which is quite hard to account for.

For the short period of 2018-2023, some datasets do not offer revised versions and therefore the most recent maps must be reused for the following years to be able to generate a condition account. This is the case for ‘Forest Connectivity’, ‘Threatened Forest Bird Species Diversity’ and ‘Soil Organic Carbon’. Therefore, these indicators cannot have a significant influence on the evolution of the forest condition index through time as their values do not change over time. However, they impact the level of magnitude in which the FCI occurs.

Besides limitations of data availability, there is also the issue with accuracy of existing datasets. As described before, the commonly used NPP dataset of CLMS was found to contain sensor artefacts due to the transition of PROBA-V to Sentinel-3/OLCI in 2020-2021. It is important to consider the complexities of satellite-derived datasets and be cautious when interpreting the results.

Second, in the forest condition account for São Miguel TS, the generated EUNIS level 2 habitat map was used to extract all forest area per forest type. The validation of the habitat map indicated that at level 2, some forest areas are misclassified in their forest type. Generally, the distribution of coniferous forests is overestimated, and broadleaved deciduous forest is hard to map since it is less predominant on the island. Thus, it is important to acknowledge that the misclassifications in habitat mapping of the EUNIS level 2 forest types may cause error propagation in the forest condition accounts per forest type, despite matching the reference areas per type to the map based on field validated areas. More specifically, this means that we might have rescaled variables within forest type A to indicators based on UpperReference and LowerReference data from forest type B, since forest type B was misclassified for forest type A in our habitat map. However, a small analysis was performed to compare mapping of the forest types with the mapping by existing CLMS products. The map accuracy of the level 2 habitat map (computed by use of the external validation points) was better than the map accuracy of the CLMS FTY layer. Thus, it may be concluded that by using the level 2 habitat map as a base to identify the areas for which forest condition accounts would be generated, the error propagation is inferior to the CLMS FTY.

Thirdly, the selection of reference sites is the main bottleneck within the condition accounting approach. The rescaling of the variables to indicators depends on the UpperReference and LowerReference, which are subjected to the criteria underlying the choice of which area is considered as ‘reference forest’ and which is not. The concept of ‘reference forest’ has many nuances and the understanding of it can be highly subjective. Questions come up like ‘Can we only consider pristine or primary forest as reference forest?’, ‘Is primary forest still present here?’, or ‘Can Forest containing exotic invasive species be considered as reference forest?’. The latter question was a particularly interesting one in the context of São Miguel TS. Usually, invaded forests are considered disturbed and in worse condition by deviating from the typical composition of pre-existing native forests. Although undesirable from a conservation standpoint, under climate change projections, maintaining forest extent and integrity may require a transition to accepting exotic species better suited to irregular climatic conditions.

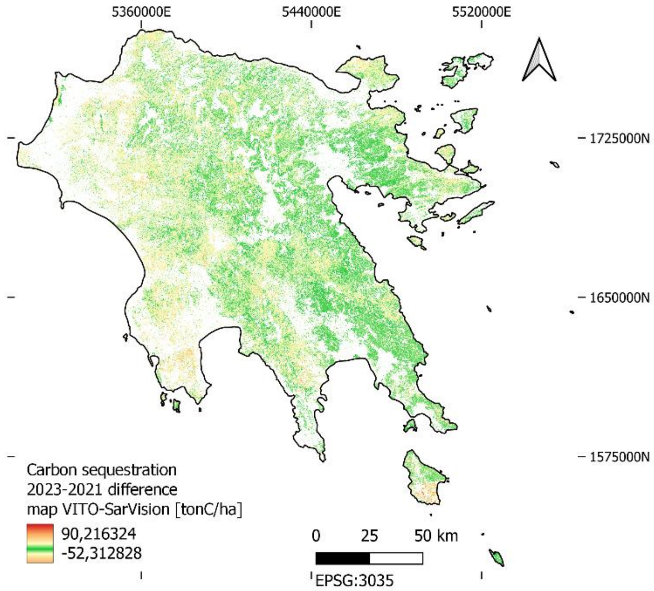
This process of compositional shift is clearly at a very advanced stage in the case of São Miguel due to historical human action, which is why it was not possible to isolate ‘reference forest’ areas free of exotic invasion in São Miguel or assign a weight to their influence as a variable in the overall FCI. We recommend this topic to be further explored to improve future forest condition assessments.

Lastly, the PEOPLE-EA method is experimental and is not yet fine-tuned to local scales. In principle, it is a decent method to create comparable accounts within Europe, but European biogeographical zones may contain a lot of variation that needs to be accounted for when identifying the appropriate variables to be used in the forest condition accounting. We refer to the Algorithm Theoretical Baseline Document (ATBD) of the Ecosystem Forest Condition of PEOPLE-EA (Deliverable 7) for further information (Bruelheide et al. 2024).

9.4. Carbon flux mapping

For estimating the aboveground carbon sequestration on Peloponnese and São Miguel, this task applied two different approaches. On the one hand, SarVision generated aboveground biomass maps with radar data, from which aboveground carbon stocks can be derived by multiplying with a factor of 0.48 and then the aboveground carbon sequestration can be calculated by subtracting the carbon stock at the end of 2021 from the carbon stock at the end of 2023. On the other hand, VITO generated annual GPP maps, which can be transferred into annual NPP, ANPP and BNPP maps by formulas and root-to-shoot ratios per ecosystem type class, then summing the ANPP values of 2022 and 2023 to derive the aboveground carbon sequestration.

Figure 93 shows a generated difference map to check how well the two approaches correspond to calculate aboveground carbon sequestration for Peloponnese and São Miguel. The difference map is generated by subtracting the carbon sequestration results of SarVision from the carbon sequestration results of VITO. The greener the pixels are colored, the better the results of VITO match with the results of SarVision. The redder the pixels are colored, the bigger the difference in their results.



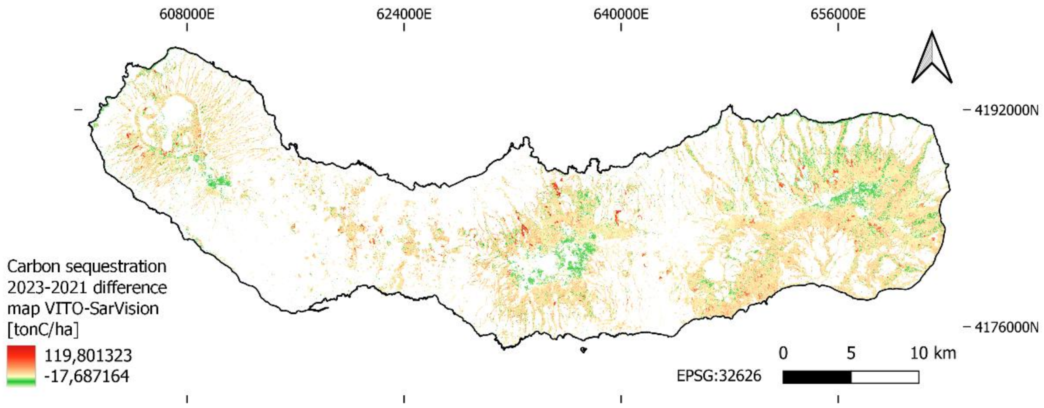


Figure 93: Carbon sequestration difference map between VITO and SarVision approaches.

The difference between the two approaches seems very pronounced in some locations. Figure 94 zooms in to one of these locations in São Miguel.

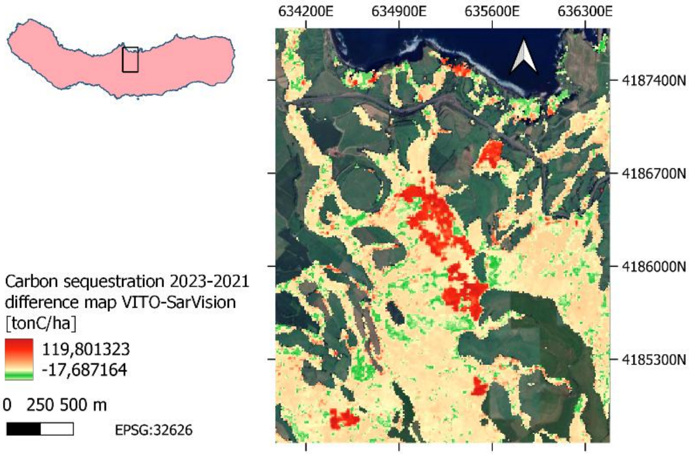
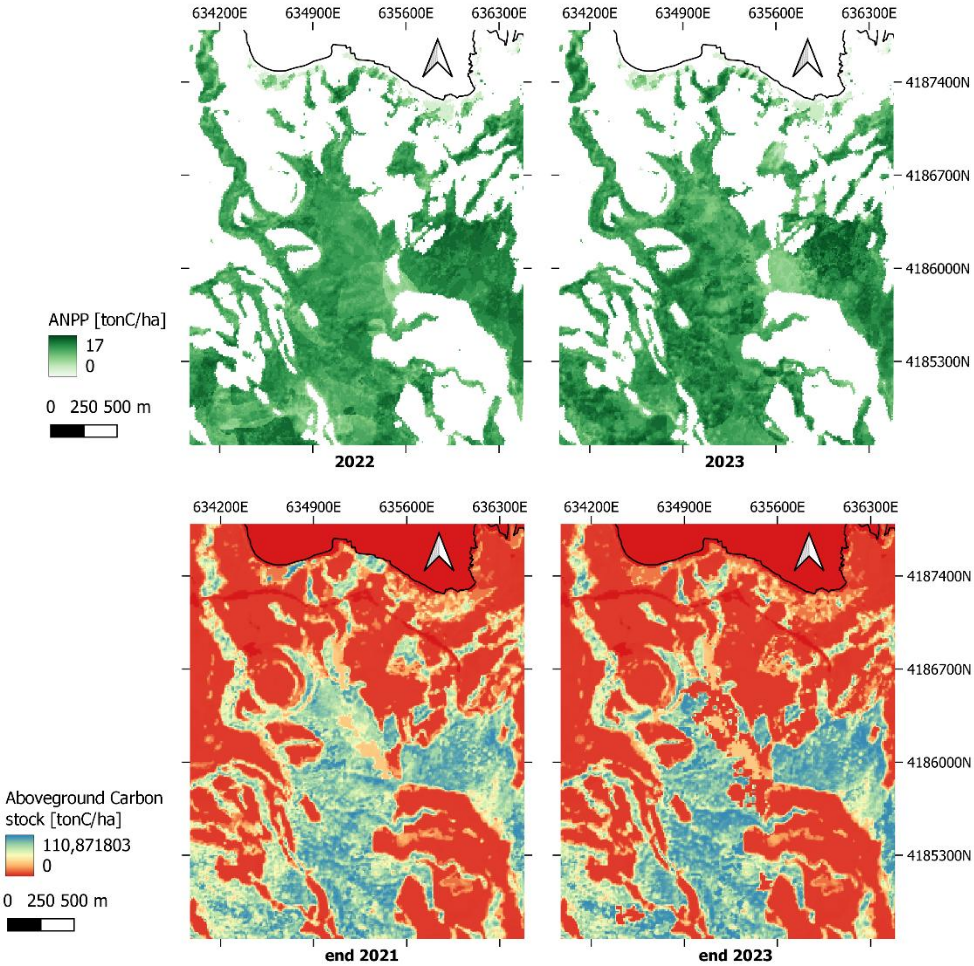


Figure 94: Zoom-in to an area in São Miguel where the difference map indicated high divergence for carbon sequestration between SarVision and VITO approaches.

Figure 95 shows that annual ANPP calculated by VITO is positive in 2022 and 2023, suggesting that each year more aboveground carbon is sequestrated. Meanwhile, in the SarVision aboveground carbon stock maps (calculated for the end of 2021 and end of 2023), the carbon stock decreases abruptly between 2021 and 2023, matching a deforestation event somewhere during this period.



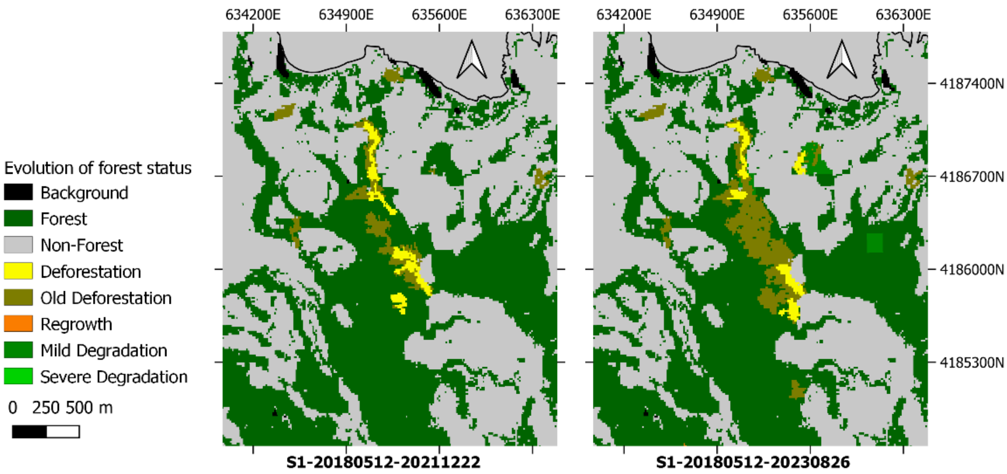


Figure 95: Example of divergence in carbon sequestration results between SarVision and VITO. [Top] VITO ANPP results show aboveground increment in 2022 and 2023. [Middle] SarVision carbon stock maps showing an abrupt stock decrease. [Bottom] Forest degradation maps attribute this abrupt decline in carbon stock is due to (old) deforestations events.

This divergence in carbon sequestration stems from the fundamental methodological differences between the VITO and SarVision approaches, namely, a flux-based vs. stock-based representation of terrestrial carbon dynamics. The VITO ANPP estimates are derived from a modified light use efficiency (LUE) model, where Gross Dry Matter Productivity (GDMP) is calculated as a function of incoming solar radiation (R), absorbed photosynthetically active radiation (fAPAR), and biome-specific LUE coefficients (ƐLUEc), further modulated by meteorological variables (temperature, CO_2 fertilization) and surface greenness derived from Sentinel-2 observations. This method does not directly reference existing biomass, but instead models carbon fluxes as a function of photosynthetic activity. As such, even sparse or early successional regrowth (e.g., grasses, herbs, or low shrubs), such as following a deforestation event, can produce significant flux values, particularly in years like 2023, which experienced increased precipitation and favorable growing conditions, leading to a modelled increase in carbon sequestration despite minimal structural biomass recovery.

Any deforestation events causing abrupt loss in biomass in late 2021, prior to the start of the ANPP time series, might cause the flux-based VITO model to not incorporate the pre-2022 carbon stock baselines and not account for the ensuing biomass loss. Instead, it primarily quantifies the productivity of vegetation present from 2022 onward. Validation using historical high-resolution aerial imagery from Google Earth further confirms significant forest removal occurring during 2021 in the example of Figure 95 and other conflicting areas, supporting the interpretation of a pre-existing carbon stock loss not visible in the flux-based estimates. Moreover, because the model uses land cover–specific ƐLUEc coefficients that were originally calibrated for forest types such as coniferous forests, the application of these coefficients to areas that, post-deforestation, are dominated by grasses or shrubs may lead to overestimation of productivity. In effect, the model assumes forest-like photosynthetic efficiency even though the vegetation structure and physiology no longer match those assumptions, akin to a land use change that is not reflected in the LUE parameterization.

In contrast, the SarVision approach estimates carbon stocks directly by quantifying standing aboveground biomass (AGB) using remote sensing techniques and allometric models, then converts these into carbon stocks using a fixed coefficient (0.48). Carbon loss is inferred from observed reductions in biomass associated with deforestation or degradation. The SarVision stock-based maps show a pronounced decrease in carbon between 2021 and 2022, attributed to forest clearance resulting in permanent loss of woody biomass. Since this approach integrates past structural losses and does not assume regrowth until biomass accumulation is observable, the carbon stock remains low post-disturbance, even if photosynthetic activity (as measured by the flux model) temporarily increases.

Thus, the discrepancy can be attributed to the VITO model interpreting increased fAPAR and favorable climatic conditions as enhanced productivity (ANPP) for biomass increment, while the SarVision method registers a continued deficit in carbon stock due to limited regrowth. In other words, the VITO flux-based model reflects potential carbon uptake capacity driven by current vegetation activity and environmental drivers, whereas the SarVision model reflects actual storage of carbon in biomass. The co-location of high ANPP and low carbon stock in 2023 highlights the distinction between carbon flux and stock, emphasizing the importance of interpreting these datasets in the context of disturbance history and land cover transitions.

9.5. Carbon stock mapping using remote sensing-derived products

For both TS, a sequential series of maps were created for carbon flux accounting (Figure 96). Each of these intermediate maps plays a key role in estimating carbon stocks and carbon flux through different methods.

- The first map in the series is the forest structural map, which serves as the base for delineating different forest structures or, when possible, forest types.
- The second map provides information on forest cover changes caused by deforestation and degradation processes.
- The third map presents biomass estimations in tons per hectare, stratified using canopy height data derived from GEDI LiDAR for different structural classes.
- Finally, the carbon-derived map, created as a fraction of the biomass map, enables the calculation of carbon emissions by comparing data between two dates.

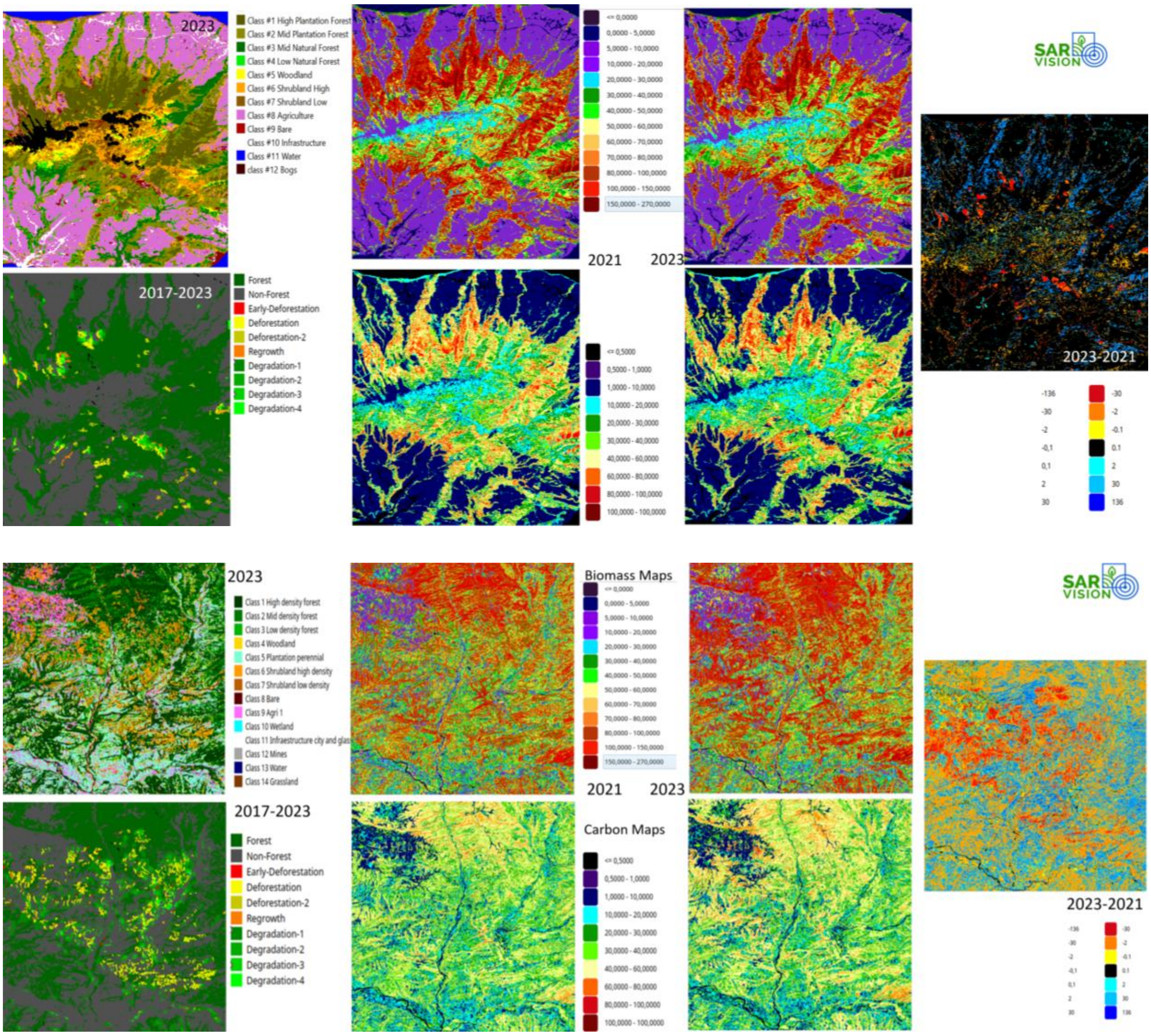


Figure 96: Details of the series of maps using remote sensing derive products created for São Miguel [above] and for Peloponnesus [below]., necessary for the carbon stock and carbon flux accounting for a certain area.

These maps were created without the direct use of field reference data for calibration or validation. Instead, they rely primarily on unsupervised methods designed to integrate diverse types (optical, radar and Lidar) of high-resolution remote sensing data, which is freely available to the public, to generate the best possible spatial data to support carbon accounting. Even so, these maps showed good correspondence with published reference data. Therefore, a remote sensing-based approach should be considered a viable method for carbon accounting in areas where field data is unavailable or non-existent. This has significant implications for carbon accounting methods, particularly in terms of the investments and resources required to meet reporting objectives. Carbon estimates that highly rely on field data are extremely expensive and often fail to capture spatial and temporal variations in biomass and carbon levels within a single ecosystem.

In this approach, the different data products support one another by being correlated and linked at the landscape or forest structural class level. Canopy height distributions, derived from LiDAR data (e.g., GEDI), are used as a proxy for vegetation structure in the field. By applying probability density functions to describe and link both the canopy height data and remote sensing variables (such as spectral indices or vegetation indices) within each structural vegetation type, it becomes possible to map fine-scale variations within each class. This approach also facilitates consistent monitoring of forest changes, such as deforestation and degradation, enabling reliable carbon accounting on an annual basis over a defined monitoring period. Additionally, updating baseline data (e.g., every five years) is straightforward and cost-effective, making this method suitable for long-term ecosystem monitoring, carbon flux accounting and reporting.

9.6. Links with other SELINA work packages

9.6.1. SELINA WP3

There is a close connection between the work ongoing in WP3 (Ecosystem condition mapping and assessment) and the work presented in this deliverable for task 2 in WP5, considering the forest condition account (see section 7.2). In WP3 the tasks aim to develop a minimal standardized set of indicators per ecosystem type and to create a structured methodology for reference condition selection. As previously stated, although a project-wide methodology for variable and reference condition selection for multiple ecosystem types is being developed collaboratively within SELINA in the context of WP3, the work in T5.2 focused on forest, which has attempted to create forest condition accounts based on the PEOPLE-EA approach for TS. The created forest condition accounts can serve as prototypes that can be further analyzed and optimized in future efforts and developments.

For instance, in the context of São Miguel TS, according to the gap identified in section 9.3., there is a demand for developing forest condition indicators, primarily based on compositional state characteristics, able to integrate the compositional impact of invasive alien species over the condition of native habitats. In regions like the Azores, it is usual for invaded forests to perform just as well or even outperform native forests in the most frequently used and mandatory reporting indicators, weakening their informative power. This was also lacking in the PEOPLE-EA approach and could not be further explored due to time and methodological constraints and the need for ground-truthing field surveys, since compositional characteristics usually require much more field input and localized expertise.

Nevertheless, Task 5.2. work on São Miguel TS and the expertise of the local partner has been feeding into Task 3.2. by contributing to the process of selection of the minimum set of key indicators for forest ecosystems, aiding in the literature review and subsequently collaborating on scoring and selecting a minimum set of key forest condition indicators within the forest ET group. Due to asynchronization between the progress of Task 3.2 and 5.2., coupled with the limitations in dataset availability for São Miguel, explained in section 7.2.3, it was not possible to apply the same set of minimum indicators, chosen by Task 3.2. ET group, to the integrated FCI applied in this report, or evaluate whether this minimum set would be suitable specifically in the context of this TS. Still, the authors maintain that accounting for invasiveness in the Azores remains an issue for both PEOPLE-EA and SELINA minimum set.

9.6.2. SELINA WP4

The work of Task 5.2. has interconnected with WP4 particularly through São Miguel TS, which is also amongst the TS contributing to the Air Quality and Climate Regulation ES group formed within WP4. São Miguel TS is represented in D4.2 Diagnostic of ES model decision-support capabilities by being selected as a best-case practice for the work carried out to model carbon sequestration in the context of Task 5.2, although Peloponnese TS was also included in the ongoing discussions. The D4.1 checklist questions of the Terms of Reference for ES assessment were applied on the actual ES model applications of the case study of São Miguel TS for the various designated diagnostic topics.

Both Peloponnese and São Miguel TS are also contributing towards D4.3. Guidance on enabling ES models for decision-support by co-designing the Terms of Reference for Air quality and carbon sequestration ES Assessment. This is being done through testing the various templates for each of the modules (Frame & Scoping, Methodology and Evaluation criteria) meant to provide guidance to decision-makers commissioning ecosystem services assessment by supporting the writing of the Terms of Reference underlying a call for tender. São Miguel TS will likewise be the exemplary case study for testing the final version of the full Terms of Reference template in the Air quality & Climate regulation ES group. It is relevant to note that this guidance is aligned with the stages of the SELINA framework for integrated ecosystem assessment (FIEA) under development by WP6.

9.7. Uptake by SELINA Demonstration Projects and Compendium of Guidance (CoG)

As T5.2 has shown potential in calculating yearly forest condition indices spatially over a study area, it would be useful to test this approach in different areas and at different scales. Currently, there is an intention to apply the method on a SELINA demonstration project (DP) in Belgium, namely Bosland national park. Here, the aim is to evaluate different methods of forest condition accounting that are described in this document. First, the ARIES for PEOPLE-EA tool will be used to do a preliminary study. Hence, the reference areas are not chosen based on expert knowledge but by the ones encoded in the tool. In a second run, the offline method will be applied in which we use the same forest condition variables as in the ARIES tool, but choosing the reference areas based on expert knowledge. Lastly, a final round will explore possibilities of integrating different forest condition variables in the method, based on stakeholder involvement. Note that these plans are still under discussion and have not yet completely been set in stone.

Parallelly, SarSentry is an algorithm ready to be implemented for the analysis of Sentinel-1 radar time series. In Europe, Sentinel-1 data is available every six days, making it suitable for both the analysis of historical deforestation (starting from 2016) and the near real-time monitoring of changes in closed, dense forests. The outputs from this algorithm provide valuable information on forest condition and can significantly refine estimates of carbon emissions resulting from deforestation and degradation. For this reason, SarSentry could be of interest for implementation in one of the SELINA demonstration projects in Spain, where there is interest; however, the partner in question has requested implementation at the national scale, which would require more resources than are currently available within the SELINA project budget, since these algorithms are not open-source in the present and must be executed by experienced professionals to ensure reliable and accurate results.

Regarding the contribution to the CoG, Task 5.2 has demonstrated innovations that should certainly be considered as potential candidates for inclusion in the CoG. However, this will require further discussion within the SELINA consortium.

In addition, Task 5.2 has demonstrated how remote sensing-based products and advanced modelling techniques can be effectively applied to Copernicus data, including imagery acquired by ESA satellites and made available through the Copernicus platform. This approach fosters independence in data production from non-European sensors and empowers European countries to design and implement their own monitoring strategies using EU-developed technology. For example, the use of Sentinel-1 and Sentinel-2 imagery for mapping ecosystem extent and condition directly supports the EU Biodiversity Strategy for 2030, the Nature Restoration Law, and SEEA EA reporting. These capabilities are critical for generating high-resolution, harmonized datasets that inform evidence-based environmental policy and enable effective long-term ecosystem monitoring at both national and EU scales.

The specific case of São Miguel TS. where high-resolution Sentinel-1 and Sentinel-2 data were used to map ecosystem extent and forest condition, enabled precise carbon stock and degradation assessments. These methods directly support national obligations under the LULUCF Regulation, as well as broader goals of the EU Biodiversity Strategy for 2030 and the Nature Restoration Law. Such applications provide scalable tools for evidence-based environmental decision-making and strengthen the EU's capacity for long-term, sovereign ecosystem monitoring.

9.8. Contribution to policy making aligning with the SELINA Framework

The D5.2 report, "Enhancing the spatial and temporal resolution of ecosystem accounts using satellite data," directly supports and advances several core objectives of European environmental policy, particularly in the context of the EU Biodiversity Strategy for 2030, the European Green Deal, and the implementation of the System of Environmental-Economic Accounting - Ecosystem Accounting (SEEA EA) framework.

By integrating remote sensing technologies and state of the art spatial modelling with ecosystem accounting practice, this report contributes to building harmonized, high-resolution data systems that are essential for tracking ecosystem extent, condition, and services across Europe. These innovations align closely with the goals of the MAES (Mapping and Assessment of Ecosystems and their Services) initiative and provide foundational tools for reporting on natural capital and ecosystem degradation.

For instance, by integrating over 150 environmental predictors and deploying advanced machine learning models such as CatBoost and Convolutional Neural Networks (CNNs), these workflows enable high-resolution (10 m), scalable, and repeatable mapping of habitats and ecosystem types. This technological advancement is critical for operationalizing the European

Ecosystem Typology (ETA) and delivering consistent, harmonized data across Member States, which is essential for implementing the System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA).

Furthermore, the deployment of the ARIES for PEOPLE-EA tool for Forest Condition Index (FCI) mapping allows year-by-year tracking of ecosystem health across biogeographical zones. This dynamic monitoring capability provides the granularity and frequency needed for responsive policymaking, ensuring that degradation trends, restoration outcomes, and climate-related impacts can be quickly identified and addressed.

Updated deforestation and degradation data like the one produced by the SarSentry System significantly enhance the effectiveness and policy relevance of the SEEA EA by:

Improving Accuracy of Ecosystem Extent and Condition Accounts
Deforestation and degradation directly affect the size, structure, and health of ecosystems. Regularly updated data from radar satellite monitoring:

- Provides near-real-time tracking of forest cover changes.
- Helps detect subtle forms of degradation (e.g., thinning, selective logging, fires).
- Enhances spatial resolution and temporal consistency, enabling more precise ecosystem extent and condition assessments.

Enabling Timely Policy Response
With frequent updates (e.g., bimonthly, annual or seasonal), policymakers can:

- Identify hotspots of forest loss and degradation.
- Prioritize restoration or protection efforts.
- Adapt land use planning and conservation strategies in line with EU policies like the Nature Restoration Law and the Biodiversity Strategy for 2030.

Enhancing Carbon Sequestration Accounting
Degradation and deforestation significantly impact carbon stocks and fluxes. Updated data improves:

- Carbon stock maps by showing where biomass has been lost or gained.
- Carbon flux estimates, which are critical for climate policy.

Supporting Ecosystem Service Valuation
Changes in forest condition affect ES and accurate degradation data allows for more realistic valuation of lost or diminished ES

Increasing Integration and Comparability

- Up-to-date, geospatial deforestation/degradation data (e.g., from Sentinel-1/2, GEDI, or SarSentry) ensures:
- Consistency with international datasets and standards.
- Easier integration with national accounts, enabling harmonized SEEA EA implementation across EU Member States.

Unsupervised biomass mapping allows for scalable, cost-effective, and regular monitoring of biomass across diverse landscapes without the need for extensive field calibration, which is especially valuable for remote or under-sampled areas in the EU. This supports the development of comprehensive, EU-wide carbon accounts, aligned with the SEEA EA.

Allometric models using canopy height (e.g., from GEDI LiDAR, Sentinel data) improve the accuracy of biomass and carbon stock assessments by linking structural vegetation characteristics with carbon content. This precision is crucial for:

- Greenhouse gas inventories,
- Evaluating the effectiveness of nature-based climate solutions,
- Monitoring progress towards carbon neutrality goals under the European Green Deal.

These innovations collectively enhance the EU's ability to meet its targets on biodiversity conservation, ecosystem restoration, and integration of natural capital into economic decision-making, support the transition toward evidence-based sustainable decisions about the natural capital, offering enhanced insights with increased spatial and temporal resolution into the state and dynamics of European ecosystems.

10. Recommendations

This Section offers recommendations for general readers interested in the application of remote sensing and advanced modelling products for the System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA) and carbon accounts. The focus is on identifying cost-effective methods to support ecosystem extent accounting, ecosystem condition accounting and carbon flux accounting.

The choice of methods should consider the availability of national datasets and field validation data, which may lead to variations in results across regions. The applicability of the methods developed in this project will also depend on user requirements, particularly regarding spatial resolution, temporal monitoring frequency, data updating needs and available resources.

As RS technologies and advanced spatial modelling methodologies continue to evolve, these innovations will lead to improved resolution, accuracy, and monitoring capabilities. Furthermore, the integration of Artificial Intelligence (AI) is expected to play a significant role in enhancing data integration and automation, making ecosystem and carbon accounting more efficient and scalable.

Establishing an accurate SEEA EA baseline for a country may initially be costly. However, by combining this effort with appropriate analysis, advance spatial modelling, and semi-automated remote sensing-derived spatial products, the accounting process can eventually become more cost-effective and more efficient. The key question then becomes: How can users effectively integrate these methods to maximize benefits while minimizing costs?

To support technical recommendations, this Section presents the recommendations from the user's point of view. Authors from São Miguel and Peloponnesus TS share their perspective on data use and how it contributes to improve their ecosystem accounting needs.

10.1. Ecosystem extent accounts

Table 53 displays a comparison between the national-centric and vegetation-centric approaches used for creating the ecosystem extent accounts. First, direct validation or evaluation of which approach works most successfully is not possible as there is no validation data. However, it is possible to set up recommendations guiding the choice on which approach is best suitable considering the available resources and the intention.

Table 53: Comparison between the national-centric and vegetation-centric approaches for ecosystem extent accounts.

National-centric approach	Vegetation-centric approach
Focus on all ecosystem types.	Focus on EUNIS Level 3 habitat (natural) types.
Ecosystem extent mapping using available data (national data and/or CLMS data).	Generate EUNIS Level 3 habitat maps based on combining national field-validated input data and established habitat modelling workflow.
Crosswalk national data and CLMS data to EU typology.	Crosswalk EUNIS habitat maps into Ecosystem extent accounts following EU typology.
Create baseline for ecosystem extents, establish level of detail that can be achieved with openly available data. Limited scalability (time and space).	Create ecosystem extents based on modelling procedure with input of accurate and field-validated habitat type information, with habitat map serving as an extended information source accompanying the extent map. Fully scalable approach (time and space).

The national-centric approach is highly dependent on the datasets that are freely available: national datasets and useful CLMS layers for the study area. Besides, it is important that these datasets cover the study area in a wall-to-wall manner, otherwise this approach might lead to a patchy structure in mapping. The addition of open and freely available CLMS data has shown to add additional value to certain ecosystem extent classes.

CLMS data alone can provide a first approximation of ecosystem extent based on open and freely available data, but with limitations to the biological background and use of most habitats or ecosystems. The mapping methods in the national-centric approach also potentially contain many “no crosswalk” areas where CLMS data is not directly linked to the ETA. Besides, the crosswalk can be restricted to levels 1 or 2 of ETA. The vegetation-centric approach is based on crosswalking the EUNIS habitat maps with additional layers (e.g. Land use datasets) to ecosystem extent maps in ETA. Since the ETA is derived from EUNIS at level 3, the crosswalk is easier, and the intermediate habitat maps provide additional useful information next to the extent maps. However, the input requirement to execute the vegetation-centric approach can be significantly higher than in the national-centric approach.

The creation of EUNIS habitat maps demands very precise information on each possible EUNIS habitat type occurring in the study area and each class should contain a set of geolocated classified training points. The habitat modelling approach is highly dependent on the input of the training data, the input features that are used and how advanced post-processing is executed. As described earlier, the identification of features missing in the current feature set that can be useful to differentiate certain habitat classes from each other, is key. This shows that there is a huge link between field expertise and RS analysts. It is important that experts from these two fields work closely with each other and learn from each other. For field experts it is important to know what kind of information is required from the field, for RS analysts it is important to know which adjustments in the model can increase the predictive power based on the ecological background of the habitat classes.

As the vegetation-centric approach focuses mostly on the natural vegetation classes within the EUNIS typology, the level of refinement within the non-natural or non-vegetated areas (e.g. urban area, water) is often lower than in the national-centric approach. However, the vegetation-centric approach is a very flexible approach as it is based on machine learning and AI-algorithms that require as input simply a training dataset of point locations with their habitat classification, a feature set and a habitat typology. Therefore, one can provide input to the model containing any classes that need to be mapped if enough points per class are provided, containing an accurate geolocation, minimum distance between points and confidence in classification. It is also important to note that the current EUNIS2021 typology is still under revision and therefore not yet complete. As seen earlier, it can often be hard to classify all habitat types in a study area like Peloponnese or Sao Miguel to a discrete habitat category in EUNIS. National authorities often have their own classification system, and a direct crosswalk is not always easy or some important habitat information can be lost in the crosswalk. This is something to consider when validating the habitat maps.

In general, the vegetation-centric approach requires the set up and execution of an intensive field campaign to collect training data if not yet available. Hence, enough financial budget and available time is needed. The national-centric approach has proven to perform well with existing and freely available datasets given the data is highly dependent on hot-spots (which worked well for Sao Miguel) and focuses on non-natural classes too. However, with the national-centric approach it will be more difficult to reach more detailed levels in ETA for the natural classes. Besides, the national-centric approach is limited in the extent on which it can be applied since a new crosswalk with the national data has to be established. The vegetation-centric approach is scalable since the integration of training data for each EUNIS habitat class can fill up a database on which models are trained, and these models can be applied on areas within similar environmental and biogeographical conditions (e.g. biogeographical zones of EEA). The habitat maps, as an intermediate result for ecosystem extent mapping, are linked to species information and can serve as additional information for authorities considering conservation or restoration interventions.

The current habitat mapping and extent mapping pipeline by VITO is still a work in progress. Ongoing work must address current limitations such as better selection of training points within the collected field-validated set, the addition of missing input features, more advanced post-processing steps, etc. In the ESA-WEED project (https://esa-worldecosystems.org/en) this will be explored and improved.

10.2. Forest Condition Accounts

The applied method for forest condition accounting is an experimental method established in the PEOPLE-EA project. Further work should explore improvements such as a better choice of reference areas and the choice of condition variables per ECT class. One could argue that the choice should not be the same for each country or study area. However, if the set of ecosystem condition variables would not be consistent between study areas, the comparative potential would be lost. Another important factor is the integration of invasive or exotic species as a factor in the forest condition accounting and how this exactly should be integrated.

It is important to emphasize again that the forest condition index is not an indicator of productivity only and it goes beyond the detection of forest change (deforestation or forest degradation). Since it aims to consider variables for each ECT class, it looks further than only productivity or biomass. Therefore, it should be interpreted as how the condition of the forest is relative to a reference and desired state, looking at factors such as biodiversity, landscape connectivity, soil status and productivity all together.

10.3. Carbon accounts

Table 54 displays a comparison between the methods of SarVision (left) and VITO (right) to derive the aboveground carbon sequestration.

Table 54: Comparison between annual carbon stock maps and annual carbon sequestration (uptake maps).

Annual carbon stock maps	Annual carbon sequestration (uptake) maps
Based on SAR data.	Based on optical data, applying the Penman-Monteith equation.
Land cover types classified with combination of SAR and optical data.	Land cover types classified with optical data only.
Sequestration is calculated by subtracting carbon stocks from each other	Sequestration is calculated by summing 10-daily ANPP’s with each other.
Aboveground carbon stock does not provide information on belowground carbon.	GPP is converted into aboveground carbon sequestration and belowground carbon sequestration.

The two methods are compared in the table above, still they are inherently very different methods and should rather be looked at complementary than excluding another. The results have shown clearly that an annual accurate land cover type map is important to generate more reliable carbon accounts. The method of SarVision integrates radar data and therefore provides a lot of structural information of the land cover types. The deforestation and degradation time series provide information about the land cover type changes. The land cover type change is taken implicitly into account in the VITO approach and therefore an overestimation of the aboveground carbon sequestration is made. However, with the integration of annual land cover type maps (e.g. annual habitat maps) in the VITO approach, this limitation can be easily addressed. Besides, the generation of the GPP maps also allows to derive belowground carbon sequestration which is a very important factor within carbon accounts that should not be forgotten.

10.4. Recommendation on the use of new remote sensing technology

As remote sensing systems continue to evolve — reducing errors and enhancing spatial resolution and accuracy — the potential to develop robust, spatially explicit, and policy-relevant ecosystem accounts grow significantly. The high costs typically associated with traditional land survey methods and field validation campaigns contrast sharply with the cost-effectiveness and scalability of remote sensing-based data production.

Optical, radar, and LiDAR systems complement each other in observing the Earth by capturing radiometric and structural differences. These technologies allow for improved differentiation of land cover classes and provide valuable biophysical information such as vegetation density and height. Together, they enable the creation of biomass and carbon maps, often in unsupervised or semi-supervised ways, reducing the need for extensive field resources while increasing data production efficiency for reporting purposes.

The integration of forest structural change products also supports improved monitoring of carbon fluxes, enhancing the overall efficiency and effectiveness of carbon accounting and reporting systems. However, further research and development are needed in remote sensing to calibrate inversion models and adapt methodologies to local environmental conditions. Combining remote sensing with advanced modelling techniques, such as Gross Primary Productivity (GPP) estimation, can lead to more accurate and region-specific results.

Ongoing advancements in spatial and temporal resolution, together with innovations in remote sensing and spatial modelling, are paving the way for long-term monitoring strategies. These strategies support the consistent tracking of ecosystem extent, condition, and services, making ecosystem accounts more accessible, reliable, and relevant for robust, evidence-based scientific analysis feeding into policymaking.

10.5. Recommendation Regarding the use of ETA typology

The use of multiple datasets for ecosystem extent mapping requires a significant effort in the crosswalking process to align classifications with the European Ecosystem Typology (ETA). Despite the complexity, this step is essential to ensure the integration of existing datasets and their harmonization with new remote sensing–derived products that offer enhanced structural and biophysical information.

The ETA serves as a robust framework for organizing spatially explicit data from remote sensing, particularly in the following applications:

- Land cover and land use maps are used to classify or monitor changes in ecosystem types over time.
- Biophysical characteristics — such as vegetation height, density, and biomass — derived from LiDAR, radar, and optical sensors can be linked to structural indicators for each ecosystem class.
- Carbon accounting and ecosystem condition assessments can be calculated per ecosystem type, for example by using radar-derived forest degradation within the “forest” classes as defined in ETA Level 3.

Through this integration, the ETA supports a standardized and policy-relevant approach to ecosystem accounting, enabling better alignment between remote sensing innovations and the SEEA EA framework.

10.6. Recommendations from Peloponnese test site (UPATRAS)

For the TS of Peloponnese, the applied methodology, revealed that complex and diverse ecosystem types, as the ones present in the region (e.g. inland marshes, fens, different types of evergreen forests an) can be adequately captured and mapped in a standartised way, even with quite limited reference data. Moreover, mapping ecosystem condition revealed that monitoring of ecosystems health is possible via EO data and technics and can support local and regional conservation and management needs, something that was missing from Greece until now. For instance, the intra-annual documentation of deforestation and regrowth is a novel input that can be further developed and supported via field documentation from local authorities’ data and can be integrated into early warning systems for ecosystem degradation or for identifying illegal logging (especially for firewood) or clear cutting, issues in the daily practice of the Forest and Nature Protection authorities.

The proposed methodological framework can be also tested for large scale (local) studies to support applied practice, especially for monitoring ecosystems and their services at the local level, with focus on areas that have been already identified as vulnerable via the outcomes of the present study (e.g. at areas with fragile ecosystems or sites with severe degradation or reforestation, that need special measures and action for their protection, in order to properly recover, e.g. protect natural regeneration after a wildfires).
Regarding the models, integration of up-to-date climate data from local meteorological stations is also crucial and can better support the temporal scale of habitat type mapping,

however applicable in practice only for the local or regional scale. Other data that can be integrated into the models include vegetation diseases and insect infestations that affect phenology in extensive areas (e.g. for the TS region, riparian vegetation with *Platanus orientalis* is currently affected by the fungus *Ceratocystis platani* that grows in the wood of trees and attacks the roots, trunk and the branches). To achieve this, close cooperation with local authorities and scientific institutions is required, something that has already been achieved and supported by using the results of this study as a high-quality material to communicate that science-based decisions are feasible by using already available national and open-source datasets (filed data and EO) and with quite limited time and human resources. This study can be also suggested to be integrated in the regional monitoring framework for forests and combine / compare outcomes with the ongoing national forest inventory project, as well as to the Natura 2000 documentation and survey scheme of the Natural Environment and Climate Change Agency

10.7. Recommendation from São Miguel test site (FGF)

Given São Miguel's unique geological properties, complex land-use dynamics, and highly vulnerable ecological conditions shaped by both natural and anthropogenic pressures, the integration of biomass-based and flux-based carbon accounting approaches provides a critical improvement in the capacity to deliver robust, responsive, and policy-relevant ecosystem accounts with high degree of transferability to the whole Azores archipelago and other similarly affected Macaronesia archipelagos. The island's fragmented and rapidly evolving forested landscapes, where exotic plantations, degraded laurel remnants, recovering forests, and highly invasive species coexist, presents considerable challenges for accurate and current monitoring of their extent, condition and capacity for carbon sequestration.

Biomass-based carbon accounting, rooted in structural attributes such as canopy height and cover derived from radar and LiDAR, provides a robust estimate of carbon stock. This is essential in São Miguel, where static classifications from forest inventories and habitat maps, whose updating is traditionally costly and time-consuming, fail to capture with enough timeliness the rapidly changing natural and human-mediated vegetation dynamics. Meanwhile, flux-based carbon accounting complements this by capturing short-term carbon uptake of even low-stature, early successional vegetation, whose associated biomass regrowth and respective carbon stocks may take significantly longer to become detectable through structural or stock-based approaches, decreasing the lag between disturbances occurring and observing their effects on forests and heathlands.

This dual approach is especially important for application on the forest managed areas and the Natura 2000 sites of São Miguel. In the context of forest managed areas, particularly those certified for sustainable forestry practices, integrating flux-based and biomass-based carbon accounting allows forest managers to track not only the structural development of tree stands over time, but also the immediate carbon uptake dynamics during early replanting phases or post-harvest rotations. This also supports meaningful comparisons of the performance and productivity of native vs. exotic forests in terms of their carbon sequestration potential, thus enabling the assessment of trade-offs between timber forestry and ecological restoration of native forests. While exotic plantations are often favored for their fast growth and timber yield, native forests usually provide a broader array of ecosystem services and enriched biodiversity. The combined use of biomass stock-based and flux-based carbon accounting allows a more accurate comparison of how closely native forests match exotic plantations in terms of carbon sequestration, potentially strengthening the case for prioritizing their restoration. This evidence-based comparison reinforces long-term sustainability, promoting a shift towards forest management strategies that favor ecosystem multifunctionality over single-service optimization typically sought in monoculture plantations. In parallel, for Natura 2000 protected areas, while structural and stock-based carbon accounting reveals long-term changes in forest structure and habitat extent, crucial for assessing degradation or recovery, flux-based models aid in detecting short-term vegetation responses following restoration interventions or removal of exotic species removal.

The combined application of biomass stock-based and carbon flux-based carbon accounting in São Miguel has effectively tackled the initially expected outcomes of SELINA Task 5.2 by delivering an exemplary demonstration of enhancement of the spatial and temporal resolution of forest carbon accounts using satellite data to capture vegetation dynamics and detect changes in near real time. The robustness of carbon estimates was strengthened through integration with field-validated data in intermediary forest structure, habitat mapping, ecosystem extent and forest condition assessments, all while following standardization and alignment with the ETA. This is of particular importance for addressing the challenges faced by EU outermost regions, requiring higher-resolution, temporally updated assessments to support adaptive forest management and biodiversity conservation consistent with EU policy goals. However, transferability of the methodologies applied in this TS to more extensive EU continental contexts with significant biogeographical differences should continue to be tested and assessed in view of the encouraging outputs in this report.

11. Acknowledgements

All authors are thankful to the valuable contributions of the internal reviewers for the improvement of this report:

- Lars Hein (Wageningen University).
- Nicolas Grondard (Wageningen University).
- Benjamin Burkhard (Leibniz University Hannover).
- Fernando Santos-Martín (Rey Juan Carlos University).

The authors from Fundação Gaspar Frutuoso (FGF) team of the University of the Azores, responsible for São Miguel TS, are thankful to the following regional stakeholders:

- DRRFOT – Regional Directorate of Resources and Spatial Planning of the Government of the Azores, for providing the georreferenced Forest Inventory 2024 dataset, as well as UAV ortophotomaps of São Miguel and further field support. FGF are particularly thankful to the DRROT personnel directly involved in this partnership – Lourdes Penil, Vasco Medeiros and João Pacheco.
- SRAAC - Regional Secretariat for the Environment and Climate Action of the Government of the Azores for providing the Mapping of Terrestrial Habitats of the Natura 2000 Network in the Azores under the LIFE IP Azores Natura Project (LIFE17 IPE/PT/000010). FGF are particularly thankful to LIFE IP Azores Natura Project Coordinator, Diana Pereira, the scientific team led by professor Eduardo Dias and SRAAC personnel, Raquel Medeiros and Ana Moreira.

FGF also acknowledge the technical contributions of former team member Diogo Cabral in validating forest cover and forest structure intermediate maps in São Miguel TS.

12. References

Andreae MO, Merlet P (2001) Emission of trace gases and aerosols from biomass burning. Global Biogeochemical Cycles 15: 955–966.

Asner GP, Mascaro J, Muller-Landau HC, Vieilledent G, Vaudry R, Rasamoelina M, Hall JS, van Breuge M (2012) Universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia 168: 1147–1160. https://doi.org/10.1007/s00442-011-2165

BC3 - VITO (2025) PEOPLE Application. Available from: https://esa-people-ea.org/en/results/aries4people-application.

Boogaard H, Schubert J, De Wit A, Lazebnik J, Hutjes R, Van der Grijn G (2020) Agrometeorological indicators from 1979 to present derived from reanalysis. https://doi.org/10.24381/cds.6c68c9bb

Bose AK, Gessler A, Bolte A, Bottero A, Buras A, Cailleret M, Camarero JJ, Haeni M, Hereş A-M, Hevia A, Lévesque M, Linares JC, Martinez-Vilalta J, Matías L, Menzel A, Sánchez-Salguero R, Saurer M, Vennetier M, Ziche D, Rigling A (2020) Growth and resilience responses of Scots pine to extreme droughts across Europe depend on predrought growth conditions. Global Change Biology 26: 4521–4537. https://doi.org/10.1111/gcb.15153

de Brogniez D, Ballabio C, Stevens A, Jones RJA, Montanarella L, van Wesemael B (2014) A map of the topsoil organic carbon content of Europe generated by a generalized additive model. European Journal of Soil Science 66: 121–134. https://doi.org/10.1111/ejss.12193

Brown S (1997) Estimating Biomass and Biomass Change of Tropical Forests: a Primer. FAO, Rome, Italy

Bruelheide H, Jandt U, Marmol A, Fernandez N, Smets B, Buchhorn M, Giagnacovo L, Milli G, Jiménez-Alfaro B, Alvarez-Martinez JM (2024) D5.2 Past-to-present EBV modelled datasets and status indicator for selected terrestrial habitats in the Habitats Directive. Martin Luther University Halle-Wittenberg (MLU); Flemish Institute for Technological Research (VITO); University of Oviedo (UniOvi); Europa Biodiversity Observation Network: integrating data streams to support policy (EUROPABON) https://doi.org/10.3897/arphapreprints.e128158

Bruzón AG, Arrogante-Funes P, Santos-Martín F (2023) Modelling and testing forest ecosystems condition account. Journal of Environmental Management 345: 118676. https://doi.org/10.1016/j.jenvman.2023.118676

Calvo Buendia E, Tanabe K, Kranjc A, Baasansuren J, Fukuda M, Ngarize S, Osako A, Pyrozhenko Y, Shermanau P, Federici S (2019) 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. IPCC, Switzerland

Chen X, Zhang Y (2023) Impacts of climate, phenology, elevation and their interactions on the net primary productivity of vegetation in Yunnan, China under global warming. Ecological Indicators 154: 110533. https://doi.org/10.1016/j.ecolind.2023.110533

DeFries RS, Houghton RA, Hansen MC, Field CB, Skole D (2010) Carbon emissions from tropical deforestation and forest degradation: A critical review of the carbon fluxes associated with land-use change and forestry activities. Environmental Research Letters 5: 033001. https://doi.org/10.1088/1748-9326/5/3/033001

Dias E, Pereira D (2022) CARTOGRAFIA DOS HABITATS TERRESTRES DA REDE NATURA 2000 DOS AÇORES NO ÂMBITO DO PROJETO LIFE IP AZORES NATURA (LIFE17 IPE/PT/000010) – 2022. Available from: https://portal.azores.gov.pt/web/drpm/life-ip-azores-natura [upon request] (April 16, 2025).

Dorogush AV, Ershov V, Gulin A (2017) Workshop on ML Systems at NIPS.

DRAAC (2018) 2018 Land Use Map – Azores Territorial Planning Portal - Regional directorate of the environment and climatic action of the Azorean Government. Available from: https://ot.azores.gov.pt/COSA-2018.aspx (April 16, 2025).

DRRFOT (2024) 2nd Forest Inventory of the Autonomous Region of the Azores (IFRAA2) - Regional Directorate for Forest Resources and Spatial Planning.

EEA (2020a) CORINE Land Cover 2018 (raster 100 m), Europe, 6-yearly - version 2020_20u1, May 2020.

EEA (2020b) Tree Cover Density 2018 (raster 10 m), Europe, 3-yearly, Sep. 2020. Available from: https://doi.org/10.2909/486f77da-d605-423e-93a9-680760ab6791.

EEA (2020c) Water and Wetness 2018 (raster 10m), Europe, 3-yearly - version 2, Nov. 2020. Available from: https://land.copernicus.eu/en/products/high-resolution-layer-water-and-wetness/water-and-wetness-status-2018.

EEA (2021a) High Resolution Vegetation Phenology and Productivity: PPI Trajectories (10m) version 1 revision 1, Sep. 2021. Available from: https://land.copernicus.eu/en/products/vegetation/high-resolution-plant-phenology-index-seasonal-trajectories.

EEA (2021b) Population trend of bird species: datasets from Article 12, Birds Directive 2009/147/EC reporting (2008-2012) - PUBLIC VERSION - Jan. 2021. Available from: https://sdi.eea.europa.eu/catalogue/srv/api/records/7c2dd14f-60b6-4009-aca8-5d20300479a9.

EEA (2024) Biogeographical regions in Europe. Available from: https://www.eea.europa.eu/en/analysis/maps-and-charts/biogeographical-regions-in-europe-2?activeTab=570bee2d-1316-48cf-adde-4b640f92119b.

EEA (2025) Forest Type 2021 (raster 10m), Europe, 3-yearly.

Eggleston HS, Buendia L, Miwa K, Ngara T, Tanabe K (2006) 2006 IPCC Guidelines for National Greenhouse Gas Inventories. IGES, Japan

European Commission (2022) Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) No 691/2011 as regards introducing new environmental economic accounts modules. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2022:329:FIN (April 16, 2025).

European Commission Directorate-General Joint Research Centre (a) (2017) Leaf Area Index 1999-2020 (raster 1 km), global, 10-daily - version 1. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_lai_1km_v1_10daily.

European Commission Directorate-General Joint Research Centre (b) (2017) Leaf Area Index 2014-present (raster 300 m), global, 10-daily - version 1. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_lai_300m_v1_10daily.

European Commission Directorate-General Joint Research Centre (c) (2018) Dry Matter Productivity 1999-2020 (raster 1 km), global, 10-daily - version 2. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_dmp_1km_v2_10daily.

European Commission Directorate-General Joint Research Centre (d) (2018) Dry Matter Productivity 2014-present (raster 300 m), global, 10-daily - version 1. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_dmp_300m_v1_10daily.

European Commission Directorate-General Joint Research Centre (e) (2017) Fraction of Vegetation Cover 1999-2020 (raster 1 km), global, 10-daily - version 1. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_fcover_1km_v1_10daily.

European Commission Directorate-General Joint Research Centre (f) (2017) Fraction of Vegetation Cover 2014-present (raster 300 m), global, 10-daily - version 1. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_fcover_300m_v1_10daily.

European Commission Directorate-General Joint Research Centre (g) (2016) Normalised Difference Vegetation Index 1998-2020 (raster 1 km), global, 10-daily - version 2. Available from: https://globalland.vito.be/geonetwork/srv/api/records/clms_global_ndvi_1km_v2_10daily.

European Commission Directorate-General Joint Research Centre (h) (2021) Normalised Difference Vegetation Index 2020-present (raster 300 m), global, 10-daily - version 2. Available from:

https://globalland.vito.be/geonetwork/srv/api/records/clms_global_ndvi_300m_v2_10daily.

European Commission Joint Research Center: Drought Team (2024) European Drought Observatory.

EUROSTAT (2024) Guidance note on ecosystem extent accounts – draft updated version 2024. EUROSTAT. Guidance note

Fargione J, Tilman D, Hale B (2018) The role of land use and land-use change in forest carbon dynamics. Nature Sustainability 1: 301–310. https://doi.org/10.1038/s41893-018-0064-7

Forest Carbon Monitoring Consortium (2021) Forest Carbon Monitoring (FCM) Product Portal. Available from: https://www.forestcarbonplatform.org/.

Gayoso J, Guerra J, Alarcón D (2002) Contenido de carbono y funciones de biomasa en especies natives y exóticas. Universidad Austral de Chile, Valdivia, Chile

Gharun M, Shekhar A, Xiao J, Li X, Buchmann N (2024) Effect of the 2022 summer drought across forest types in Europe. Biogeosciences 21: 5481–5494. https://doi.org/10.5194/bg-21-5481-2024

Griscom BW, Adams J, Ellis PW, Houghton RA, Lomax G, Miteva DA, Schlesinger WH, Shoch D, Siikamäki JV, Smith P, Woodbury P, Zganjar C, Blackman A, Campari J, Conant RT, Delgado C, Elias P, Gopalakrishna T, Hamsik MR, Fargione J (2017) Natural climate solutions. Proceedings of the National Academy of Sciences 114: 11645–11650. https://doi.org/10.1073/pnas.1710465114

Hellenic Ministry of Environment and Energy (2016). Conservation status assessment of habitat types and species for terrestrial areas protected under the «Natura 2000» network at a national scale (unpublished data, available upon request).

Hoekman D, Kooij B, Quiñones M, Vellekoop S, Carolita I, Budhiman S, Arief R, Roswintiarti O (2020) Wide-Area Near-Real-Time Monitoring of Tropical Forest Degradation and Deforestation Using Sentinel-1. Remote Sensing 12: 3263. https://doi.org/10.3390/rs12193263

Hoekman DH, Vissers MAM, Wielaard NJ (2010a) PALSAR wide-area mapping methodology and map validation of Borneo. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 3: 605–617.

Hoekman DH, Vissers M, Tran T (2010b) Unsupervised full-polarimetric SAR data segmentation as a tool for classification of agricultural areas. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

IPCC (2007) Climate Change 2007: Mitigation of Climate Change. Contribution of Working Group III to the Fourth Assessment Report of the IPCC. Cambridge University Press Available from: https://www.ipcc.ch/report/ar4/wg3/.

IPCC Task Force on National Greenhouse Gas Inventories (2003) Good Practice Guidance for Land Use, Land Use Change and Forestry (LULUCF). Chapter 3: LUCF Sector Good Practice Guidance, Annex 3A.1 Biomass Default Tables for Section 3.2 Forest Land. Available from: https://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_files/Chp3/Anx_3A_1_Data_Tables.pdf.

Janssen JAM, Rodwell JS (2016) European Red List of Habitats: Part 2. Terrestrial and freshwater habitats. IUCN; Alterra; European Commission; Naturebureau UK; Publications Office of the European Union, Luxembourg. https://doi.org/10.2779/091372

Jones RJA, Hiederer R, Rusco E, Loveland PJ, Montanarella L (2005) Estimating organic carbon in the soils of Europe for policy support. European Journal of Soil Science 56: 655–671.

Kokkoris I, Eliadou E, Verde N, Mallinis G, Dimopoulos P (2020) Map of the Ecosystem types of Greece within the Natura 2000 SACs (scale 1:5,000).

Kokkoris, I. P., Kokkinos, V., Michos, E., Kalogeropoulos, R., Charalambides, M., Kounelis, A., Iliadou, E., Diamntidis, C.K., Mallinis, G., Bouras, C., Dimopoulos, P. (2021). MAES_GR: A Web-Based, Spatially Enabled Field Survey Platform for the MAES Implementation in Greece. *Land*, *10*(4), 381.

Lan X, Keeling R (2025) NOAA/GML Trends and Scripps CO2 Data. Available from: http://gml.noaa.gov/ccgg/trends/.

Lutz JA, Larson AJ, Swanson ME, Ziegler JM (2018) A critical review of the effects of forest thinning and disturbance on carbon sequestration. Ecological Applications 28: 2162–2174. https://doi.org/10.1002/eap.1794

MacDicken K (1997) A Guide to Monitoring Carbon Storage in Forestry and Agroforestry Projects. Winrock International, 1611 N. Kent St., Suite 600, Arlington, VA 22209, USA.

Maes J, Teller A, Erhard M, Conde S, Vallecillo RS, Barredo CJI, Paracchini M-L, Abdul MD, Trombetti M, Vigiak O, Zulian G, Addamo A, Grizzetti B, Somma F, Hagyo A, Vogt P, Polce C, Jones A, Marin A, Ivits E, Mauri A, Rega C, Czucz B, Ceccherini G, Pisoni E, Ceglar A, De PP, Cerrani I, Meroni M, Caudullo G, Lugato E, Vogt J, Spinoni J, Cammalleri C, Bastrup-Birk A, San-Miguel-Ayanz J, San RS, Kristensen P, Christiansen T, Zal N, De RA, De JCA, Pistocchi A, Del BAI, Tsiamis K, Gervasini E, Deriu I, La NA, Abad VR, Vizzarri M, Camia A, Robert N, Kakoulaki G, Garcia BE, Panagos P, Ballabio C, Scarpa S, Montanarella L, Orgiazzi A, Fernandez UO, Santos-Martín F (2020) Mapping and Assessment of Ecosystems and their Services: An EU ecosystem assessment. JRC Publications Repository. https://doi.org/10.2760/757183

Maes J, Bruzón AG, Barredo JI, Vallecillo S, Vogt P, Marí Rivero I, Santos-Martín F (2023) Accounting for forest condition in Europe based on an international statistical standard. Nature Communications 14: 3723. https://doi.org/10.1038/s41467-023-39434-0

Masseti A, Gil A (2020) Mapping and assessing land cover/land use and aboveground carbon stocks rapid changes in small oceanic islands' terrestrial ecosystems: A case study of Madeira Island, Portugal (2009–2011). Remote Sensing of Environment 239: 111625.

Matthews G (1993) The carbon content of trees. UK Forestry Commission, Edinburgh, UK

McGroddy ME, Daufresne T, Hedin LO (2004) Scaling of C:N:P stoichiometry in forests worldwide: Implications of terrestrial Redfield-type ratios. Ecology 85: 2390–2401.

Mehmood K, Anees SA, Rehman A, Rehman NU, Muhammad S, Shahzad F, Liu Q, Alharbi SA, Alfarraj S, Ansari MJ, Khan WR (2024) Assessment of climatic influences on net primary productivity along elevation gradients in temperate ecoregions. Trees, Forests and People 18. Available from: https://doi.org/10.1016/j.tfp.2024.100657.

Naimi B, Kissling WD (2022) Relative Magnitude of Fragmentation (RMF). Group on Earth Observation (GEO) Biodiversity Network (BON) Available from: https://portal.geobon.org/ebv-detail?id=4.

Penman J, Gytarsky M, Hiraishi T, Kruger D, Pipatti R, Buendia L, Miwa K, Ngara T, Tanabe K, Wagner F (2003) Good Practice Guidance for Land Use, land-Use Change and Forestry. Intergovernmental Panel on Climate Change (IPCC), IPCC/IGES, Hayama, Japan

Posit team (2024) RStudio: Integrated Development Environment for R. Available from: http://www.posit.co/.

Pugh TAM, Arneth A, Kautz M, Poulter B, Smith B (2019) Important role of forest disturbances in the global biomass turnover and carbon sinks. Nature Geoscience 12: 730–735. https://doi.org/10.1038/s41561-019-0427-2

QGIS Development Team (2024) QGIS Geographic Information System (Version 3.30.2). Open Source Geospatial Foundation.

Running S, Zhao M (2021) MODIS/Terra Net Primary Production Gap-Filled Yearly L4 Global 500m SIN Grid V061 [Data set]. Available from: https://doi.org/10.5067/MODIS/MOD17A3HGF.061.

Sabatini FM, Bluhm H, Kun Z, et al. (2021) European primary forest database v2.0. Scientific Data 8: 220. https://doi.org/10.1038/s41597-021-00988-7

Sachsenmaier L, Schnabel F, Dietrich P, Eisenhauer N, Ferlian O, Quosh J, Richter R, Wirth C (2024) Forest growth resistance and resilience to the 2018–2020 drought depend on tree diversity and mycorrhizal type. Journal of Ecology 112: 1787–1803. https://doi.org/10.1111/1365-2745.14360

Santoro M, Cartus O (2024) ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2015, 2016, 2017, 2018, 2019, 2020 and 2021, v5.01. https://doi.org/10.5285/bf535053562141c6bb7ad831f5998d77

Schnabel F, Barry KE, Eckhardt S, Guillemot J, Geilmann H, Kahl A, Moossen H, Bauhus J, Wirth C (2024) Neighbourhood species richness and drought-tolerance traits modulate tree growth and δC responses to drought. Plant Biology 26: 330–345. https://doi.org/10.1111/plb.13611

SoilGrids (2025) Soil Organic Carbon stock. Available from: https://soilgrids.org/.

Swinnen E, Toté C, Van Hoolst R (2021) Copernicus Global Land Operations “Vegetation and Energy” - CGLOPS-1: Algorithm Theoretical Basis Document. Issue 1.30. Flemish Institute for Technological Research (VITO) Available from: https://land.copernicus.eu/en/products/vegetation/gross-primary-production-v1-0-300m.

Tao J, Xie Y, Wang W, Zhu J, Zhang Y, Zhang X (2022) Elevational Gradient of Climate-Driving Effects on Cropland Ecosystem Net Primary Productivity in Alpine Region of the Southwest China. Remote Sensing 14: 3069. https://doi.org/10.3390/rs14133069

UN WCMC (2025) Protected Areas (WDPA). Available from: https://www.protectedplanet.net/en/thematic-areas/wdpa?tab=WDPA.

U.S. Geological Survey (2021) Landsat 7 Collection 1 Tier 1 32-Day NDWI Composite.

U.S. Geological Survey (2022) Landsat 8 Collection 1 Tier 1 32-Day NDWI Composite.

Vallecillo Rodriguez S, Maes J, Teller A, Babi Almenar J, Barredo Cano JI, Trombetti M, Abdul Malak D, Paracchini M, Carré A, Addamo A, Czucz B, Zulian G, Marando F, Erhard M, Liquete Garcia MDC, Romao C, Polce C, Pardo Valle A, Jones A, Zurbaran-Nucci M, Nocita M, Vysna V, De Jesus Cardoso A, Gervasini E, Magliozzi C, Baritz R, Barbero M, Andre V, Kokkoris IP, Dimopoulos P, Kovacevic V, Gumbert A (2022) EU-wide methodology to map and assess ecosystem condition. Publications Office of the European Union, Luxembourg https://doi.org/10.2760/13048

VITO (2024) Sentinel-2 Fraction Absorbed Photosynthetically Radiation (FAPAR) (tiles) - V2. Available from: https://docs.terrascope.be/DataProducts/Sentinel-2/ProductsOverview.html.

Vogt P, Riitters K (2017) GuidosToolbox: universal digital image object analysis. European Journal of Remote Sensing 50: 352–361. https://doi.org/10.1080/22797254.2017.1330650

Watson C (2009) Forest Carbon Accounting Overview Principles. Available from: https://www.undp.org/publications/forest-carbon-accounting-overview-principles.

Westlake DF (1966) The biomass and productivity of glyceria maxima: I. Seasonal changes in biomass. Journal of Ecology 54: 745–753.

Xu B, Feng Z, Chen Y, Zhou Y, Shao Y, Wang Z (2024) Assessing the Distribution and Driving Effects of Net Primary Productivity along an Elevation Gradient in Subtropical Regions of China. Forests 15: 340. https://doi.org/10.3390/f15020340

Zianis D, Muukkonen P, Mäkipää R, Mencuccini M (2005) Biomass and stem volume equations for tree species in Europe. Silva Fennica Monographs 2005: 1–63. https://doi.org/10.14214/sf.sfm4

Zscheischler J, Fischer EM (2020) The record-breaking compound hot and dry 2018 growing season in Germany. Weather and Climate Extremes 29: 100270. https://doi.org/10.1016/j.wace.2020.100270

SELINA

SCIENCE FOR EVIDENCE-BASED AND SUSTAINABLE DECISIONS ABOUT NATURAL CAPITAL

https://project-selina.eu/

13. Annex

13.1. Peloponnese Crosswalk

13.1.1. Peloponnese CLMS to ETA crosswalk

Overview of the datasets that were used to map the Peninsula of Peloponnese using only Copernicus Land Monitoring (CLMS) data crosswalked towards the European Ecosystem Typology (ETA). A supporting document with detailed descriptions of data sets can be provided upon request to the authors.

Subsection	Description of Subsection	Columns	Description
A1.1. CLMS Crosswalk	This subsection provides an overview of the mapping hierarchy and combination of CLMS data sets to map the peninsula of Peloponnese towards the ETA.	hierarchy	To map each of the ETA classes that occur in the peninsula of Peloponnese a hierarchical framework for classifying specific classes was developed to systematically populate a blank map of Peloponnese with various CLMS datasets. The first class mapped according to the hierarchy would be 1 and the last 170.
		dataset	This column lists all the datasets used to map Peloponnese based solely on CLMS data. Some datasets are either combinations or extractions of those marked in columns C to J and were exported as separate datasets, as further detailed in the script "01_data_pre_processing". The hierarchy of these datasets was determined based on their relevance. For instance, the Urban Atlas is expected to map urban areas most accurately since this is its primary focus. In contrast, Riparian Zones and Coastal Zones may be less precise in these areas, as their focus is on riparian and coastal regions, respectively.
		ua	See description of data set in support document.
		cz	See description in support document.
		rz	See description of data set in support document.
		clc	See description of data set in support document.

		clcbb	See description of data set in support document.
		IMD	See description of data set in support document.
		FTY_10m	See description of data set in support document.
		FTY_100m	See description of data set in support document.
		EUHydro	See description of data set in support document.
		crosswalked_code	These are the codes used for the exported TIF files. The codes are linked to the crosswalked ETA class.
		crosswalked_eta_class	These are the ETA class names linked to the crosswalked_code. Most data could be crosswalked to the ETA with the exception of crosswalked_codes 1 to 10 which were filled by clcbb data if no other suitable data was available, 11 filled by the priority area mapping class "water" which cannot be crosswalked to any ETA class and 13 and 14 which are areas that are continuously or discontinuously sealed based on the IMD layer.
A1.2. CLMS Typologies	This subsection provides an overview of the typologies and their class names of each of the used CLMS data sets along with their class codes (at different levels)	dataset	Input data set.
		class_level	The thematic class level of each of the data sets (if they are split up into different levels).
		class_name	The names of each of the classes.
		column_name	The column names given to each class at different levels (only relevant for the vector data sets).
		class_code	The class code as numbered in the original dataset.
		original_data_format	The original format of the data that was downloaded from the CLMS portal.
A1.3. European Ecosystem Typology	This subsection provides an overview of the European Ecosystem Typology at all levels of mapping	eta_code	Code used for the European Typology.
		class_level	The thematic class level of each of the data sets (if they are split up into different levels).
		class_name	The names of each of the classes.

13.1.1.1. CLMS Crosswalk

hierarchy	dataset	ua	cz	rz	clc	clcbb	IMD	FTY_ 10m	FTY_ 100m	EU Hydro	crosswalked_ code	crosswalked_eta_class
1	rpz			2120							29	1.5.1 Permanent Greenhouses
2	cz		21200								29	1.5.1 Permanent Greenhouses
3	ua	11300					80-100				16	1.1.2 Continuous commercial and industrial area
4	ua	11300					1-79				18	1.2.2 Discontinuous commercial and industrial area
5	ua	11100									15	1.1.1 Continuous residential area
6	ua	11210									17	1.2.1 Discontinuous residential area
7	ua	11220									17	1.2.1 Discontinuous residential area
8	ua	11230									17	1.2.1 Discontinuous residential area
9	ua	11240									17	1.2.1 Discontinuous residential area
10	ua	11300									17	1.2.1 Discontinuous residential area
11	ua	12210									20	1.3.1 Road and rail networks and associated land
12	ua	12220									20	1.3.1 Road and rail networks and associated land
13	ua	12230									20	1.3.1 Road and rail networks and associated land
14	ua	12300									21	1.3.2 Port areas
15	ua	12400									22	1.3.3 Airports
16	ua	13100									19	1.3 Infrastructure
17	ua	13300									25	1.3.7 Construction sites
18	ua	14100									27	1.4.1 Parks (including Zoos and botanical gardens)
19	ua	14200									28	1.4.2 Sports and recreation sites
20	rpz			1111							15	1.1.1 Continuous residential area
21	rpz			1112							17	1.2.1 Discontinuous residential area
22	rpz			1113							17	1.2.1 Discontinuous residential area
23	rpz_IMD_80_100			1120			80-100				16	1.1.2 Continuous commercial and industrial area
24	rpz_IMD_0_79			1120			1-79				18	1.2.2 Discontinuous commercial and industrial area

25	rpz			1210							20	1.3.1 Road and rail networks and associated land
26	rpz			1220							20	1.3.1 Road and rail networks and associated land
27	rpz			1230							21	1.3.2 Port areas
28	rpz			1240							22	1.3.3 Airports
29	rpz			1310							19	1.3 Infrastructure
30	rpz			1400							26	1.4 Urban greenspace
31	cz		11110								15	1.1.1 Continuous residential area
32	cz		11120								17	1.2.1 Discontinuous residential area
33	cz		11130								17	1.2.1 Discontinuous residential area
34	cz		12100				80-100				16	1.1.2 Continuous commercial and industrial area
35	cz		122								18	1.2.2 Discontinuous commercial and industrial area
36	cz		12100								20	1.3.1 Road and rail networks and associated land
37	cz		12200								20	1.3.1 Road and rail networks and associated land
38	cz		12310								21	1.3.2 Port areas
39	cz		12320								21	1.3.2 Port areas
40	cz		12330								21	1.3.2 Port areas
41	cz		12350								21	1.3.2 Port areas
42	cz		12360								21	1.3.2 Port areas
43	cz		12370								21	1.3.2 Port areas
44	cz		12400								22	1.3.3 Airports
45	cz		13110								23	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
46	cz		13120								24	1.3.6 Dump areas
47	cz		13130								25	1.3.7 Construction sites
48	cz		13200								25	1.3.7 Construction sites
49	cz		14000								26	1.4 Urban greenspace
50	rpz			8240							23	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
51	cz		82400								23	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)

52	clc_imd				1,2		80-100				15	1.1.1 Continuous residential area
53	clc_imd				1,2		1-79				17	1.2.1 Discontinuous residential area
54	clc_imd				3		80-100				16	1.1.2 Continuous commercial and industrial area
55	clc_imd				3		1-79				18	1.2.2 Discontinuous commercial and industrial area
56	clc_imd_clcbb				4	1	1-100				20	1.3.1 Road and rail networks and associated land
57	clc				5						21	1.3.2 Port areas
58	clc				6						22	1.3.3 Airports
59	clc				7						23	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
60	clc				8						24	1.3.6 Dump areas
61	clc				9						25	1.3.7 Construction sites
62	clc				11						28	1.4.2 Sports and recreation sites
63	IMD_1_79						1-79				14	Discontinuous sealed
64	IMD_80_100						80-100				13	Continuous sealed
65	rpz			1320							12	1. Settlements and other artificial areas
66	ua	13400									12	1. Settlements and other artificial areas
67	clcbb					1					12	1. Settlements and other artificial areas
68	rpz			8110							56	8.1.1 Rivers
69	rpz			8120							57	8.2.1 Canals, ditches and drains
70	rpz		82200	8130							56	8.1.1 Rivers
71	rpz		82300	8210							60	9.1.1 Lakes
72	rpz			8220							61	9.2.1 Artificial reservoirs
73	rpz			8230							61	9.2.1 Artificial reservoirs
74	rpz			8310							62	10.1 Coastal lagoons
75	rpz			8320							63	10.2.1 Estuaries and bays
76	rpz			7110							54	7.1 Inland marshes on mineral soil
77	rpz			7210							68	11.4.1 Coastal saltmarshes
78	rpz			7220							69	11.4.2 Salines
79	rpz_CLC_river_ bank			6220		10					56	8.1.1 Rivers

80	cz		71100								54	7.1 Inland marshes on mineral soil
81	cz		72100								68	11.4.1 Coastal saltmarshes
82	cz		72200								69	11.4.2 Salines
83	cz		72300								64	10.3.1 Intertidal flats (e.g., Wadden Sea)
84	cz		81100								56	8.1.1 Rivers
85	cz		81200								57	8.2.1 Canals, ditches and drains
86	cz		81300								56	8.1.1 Rivers
87	cz		82100								60	9.1.1 Lakes
88	cz		82200								61	9.2.1 Artificial reservoirs
89	cz		82300								61	9.2.1 Artificial reservoirs
90	cz		83100								62	10.1 Coastal lagoons
91	cz		83200								63	10.2.1 Estuaries and bays
92	cz_CLC_river_ bank		62200			10					56	8.1.1 Rivers
93	clc				35						54	7.1 Inland marshes on mineral soil
94	clc				37						68	11.4.1 Coastal saltmarshes
95	clc				38						69	11.4.2 Salines
96	clc				42						62	10.1 Coastal lagoons
97	rpz			7000							53	7. Inland wetlands
98	EUHydro									1	59	9.1 Lakes
99	clc_clcbb				40	10					55	8. Rivers and canals
100	clc_clcbb				41	10					58	9. Lakes and reservoirs
101	rpz			8000							11	Water
102	ua	50000									11	Water
103	rpz			2200							32	2.3 Permanent crops
104	rpz			2210							32	2.3 Permanent crops
105	rpz			2220							33	2.3.1 Olives (O1000)
106	cz		22100								32	2.3 Permanent crops
107	cz		22200								33	2.3.1 Olives (O1000)

108	clc				14						31	2.2.1 Rice fields (C2000)
109	clc_clcbb				15	5					34	2.3.2 Grapes (W1000)
110	clc_clcbb				16	3,4,5					32	2.3 Permanent crops
111	clc_clcbb				17	4					33	2.3.1 Olives (O1000)
112	cz		34000								45	4.5.1 Transitional woodland/forest land
113	rpz			3400							45	4.5.1 Transitional woodland/forest land
114	cz		36000								45	4.5.1 Transitional woodland/forest land
126	fty_CLCBB_evergreen_forest					4		1,2			43	4.3 Broadleaved evergreen forest
130	rpz			5300							48	5.3 Sclerophyllous vegetation
115	FTY_clcbb					5, 6, 7, 9		1,2			45	4.5.1 Transitional woodland/forest land
127	fty							1			41	4.1 Broadleaved deciduous forest
128	fty							2			42	4.2 Coniferous forests
131	cz		53000								48	5.3 Sclerophyllous vegetation
116	clc_clcbb				29	5,6,7,9					45	4.5.1 Transitional woodland/forest land
129	rpz			3000							40	4. Forest and woodlands
132	clc_clcbb				27	5					47	5.2 Heathland and (sub-) alpine shrub
133	clc_clcbb				28	5					48	5.3 Sclerophyllous vegetation
117	rpz			3310							44	4.4 Mixed forests
134	rpz			5000							46	5. Heathlands and shrub
118	cz		33100								44	4.4 Mixed forests
135	cz		52000								46	5. Heathlands and shrub
136	cz		23100								35	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
137	cz		23200								35	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
138	cz		23300								35	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
119	FTY_100m_clipped								3		44	4.4 Mixed forests
139	rpz			2310							35	2.5.1 Mosaic farmland (comprising cropland, grassland and

												(semi-)natural components)
140	rpz			2320							35	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
141	rpz			2330							35	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
120	cz_CLCBB_forest		31100, 32100			3					41	4.1 Broadleaved deciduous forest
121	cz_CLCBB_forest		31100, 32100			2					42	4.2 Coniferous forests
122	cz_CLCBB_forest		31100, 32100			4					43	4.3 Broadleaved evergreen forest
123	rpz_CLCBB_forest		31100, 32100			3					41	4.1 Broadleaved deciduous forest
124	rpz_CLCBB_forest		31100, 32100			2					42	4.2 Coniferous forests
125	rpz_CLCBB_forest		31100, 32100			4					43	4.3 Broadleaved evergreen forest
142	cz_CLCBB_arab_grass		21100, 41000			6					37	3.1 Sown pastures and fields (modified grassland)
143	cz_CLCBB_arab_grass		21100, 41000			7					30	2.1 Annual cropland
144	rpz_CLCBB_arab_grass		21100, 41000			6					37	3.1 Sown pastures and fields (modified grassland)
145	rpz_CLCBB_arab_grass		21100, 41000			7					30	2.1 Annual cropland
146	clc_clcbb				12,13	7					30	2.1 Annual cropland
147	cz		42100								38	3.2 Natural and semi-natural grassland
148	cz		42200								39	3.2.3 Alpine and subalpine grasslands
149	rpz			4210							38	3.2 Natural and semi-natural grassland
150	clc_clcbb				26	6					38	3.2 Natural and semi-natural grassland
151	rpz			4000							36	3. Grassland (pastures, semi-natural and natural grasslands)
152	cz		62111								65	11.2 Coastal dunes, beaches and sandy and muddy shores
153	cz		62112								66	11.3.1 Coastal shingle
154	cz		62120								52	6.2.3 Other sparsely vegetated areas
155	cz		62200								51	6.2 Semi-desert, desert and other sparsely vegetated areas

156	cz		63110								50	6.1.1 Rocky pavements, outcrops, and screes
157	cz		63120								67	11.3.2 Rock cliffs, ledges and shores
158	rpz			6210							65	11.2 Coastal dunes, beaches and sandy and muddy shores
159	rpz			6220							51	6.2 Semi-desert, desert and other sparsely vegetated areas
160	rpz			6320							49	6.1 Bare rocks
161	clcbb					1					1	1: Sealed
162	clcbb					2					2	2: Woody – needle leaved trees
163	clcbb					3					3	3: Woody – Broadleaved deciduous trees
164	clcbb					4					4	4: Woody – Broadleaved evergreen trees
165	clcbb					5					5	5: Low-growing woody plants (bushes, shrubs)
166	clcbb					6					6	6: Permanent herbaceous
167	clcbb					7					7	7: Periodically herbaceous
168	clcbb					8					8	8: Lichens and mosses
169	clcbb					9					9	9: Non- and sparsely-vegetated
170	clcbb					10					10	10: Water

```
# R Script – Peloponnese data pre-processing
# Excel files used as inputs for the scripts are too large to reasonably fit this report and are available upon request to the authors.

---
title: "CLMS based mapping of Peloponnese"
output: html_notebook
---


```{r rasterize National coastline dataset }

library(terra)
library(here)
library(sf)

# Read the Coastline polygon of Peloponnese
vector_data <- vect("path/Peloponnese_coastline_reprojected_polygon.gpkg", layer="Peloponnese_coastline_reprojected_polygon")
# Use the Forest type layer as a reference raster to rasterize the Coastline polygon of Peloponnese
reference_raster <- rast("path/FTY_2018_010m_03035_V1_0.tif")

# Apply a 10 m buffer to the vector_data so that the beaches of Peloponnese are included in the mapping
buffered_vector <- buffer(vector_data, width = 10)  # Buffer by 10 meters

# Export the buffered area of the coastline as a GeoPackage
writeVector(buffered_vector, "path/coastline_buffered_10m.gpkg", layer = "coastline_buffered_10m",overwrite=TRUE)

# Set the resolution of to 10m to rasterize the vector_data
resolution <- 10

# Function to round extent outwards to the nearest multiple of resolution
round_extent_outwards <- function(extent, res) {
  xmin <- floor(extent[1] / res) * res
  xmax <- ceiling(extent[2] / res) * res
  ymin <- floor(extent[3] / res) * res
  ymax <- ceiling(extent[4] / res) * res
  return(ext(xmin, xmax, ymin, ymax))
}

# Apply the rounding function to original extent
rounded_extent <- round_extent_outwards(ext(buffered_vector), resolution)
```
```

```
# Create a raster template with the rounded extent and resolution
raster_template <- rast(
  extent = rounded_extent,  # Use the adjusted extent
  crs = reference_raster,  # Define the CRS; replace with appropriate CRS if needed
  resolution = resolution  # Set the resolution to 10 meters
)

# Rasterize the vector layer to the template, setting the value to 1
rasterized_vector <- rasterize(buffered_vector, raster_template, field = 1, background = NA)

#Export the raster boundary
writeRaster(rasterized_vector,   here("path/boundary_10m.tif"),   filetype   =   "GTiff",   overwrite   =   TRUE,      datatype='INT1U',
gdal=c("COMPRESS=LZW"))

```

```{r Mask all raster CLMS datasets to above created raster}

library(terra)
library(here)

#Load the boundary reference raster to clip the CLMS datasets
reference_raster <- rast("path/boundary_10m.tif")

#Load the CLMS datasets
#High Resolution Layer Imperviousness 2018
IMD <- rast("path/IMD_2018_010m_03035_V2_0/IMD_2018_010m_03035_V2_0.tif")
#High Resolution Layer Forest Type Layer 10m resolution 2018
FTY <- rast("path/FTY_2018_010m_03035_V1_0/FTY_2018_010m_03035_V1_0.tif")
# CLCplus Backbone 2018
CLCPlusBB <- rast("path/CLMS_CLCplus_RASTER_2018_010m_eu_03035_V1_1.tif")
#High Resolution Layer Forest Type Layer 100m resolution 2018
fty_100 <- rast("path/FTY_2018_100m_eu_03035_V1_0.tif")
#CORINE Land Cover 2018
CLC <- rast("path/U2018_CLC2018_V2020_20u1.tif")

############Clip rasters##############

####CLC###

# Crop both rasters to the intersection extent
```

```
cropped_CLC <- crop(CLC, reference_raster)
plot(cropped_CLC)

# Resample the cropped IMD raster to match the resolution of the reference raster
resampled_CLC <- resample(cropped_CLC, reference_raster, method = "near")


# Mask the resampled raster based on the reference raster
masked_raster <- mask(resampled_CLC, reference_raster)

ext(masked_raster)

# Save the clipped raster
writeRaster(masked_raster, here("path/CLC_clipped.tif"), filetype = "GTiff", overwrite = TRUE)

####IMD###

# Crop both rasters to the intersection extent
cropped_IMD <- crop(IMD, reference_raster)
plot(cropped_IMD)

# Resample the cropped IMD raster to match the resolution of the reference raster
resampled_IMD <- resample(cropped_IMD, reference_raster, method = "near")


# Mask the resampled raster based on the reference raster
masked_raster <- mask(resampled_IMD, reference_raster)

ext(masked_raster)

# Save the clipped raster
writeRaster(masked_raster, here("path/IMD_clipped.tif"), filetype = "GTiff", overwrite = TRUE)

####FTY 10m###

# Crop both rasters to the intersection extent
cropped_FTY <- crop(FTY, reference_raster)

# Resample the cropped IMD raster to match the resolution of the reference raster
resampled_FTY <- resample(cropped_FTY, reference_raster, method = "near")

# Mask the resampled raster based on the reference raster
```

```
masked_raster <- mask(resampled_FTY, reference_raster)

# Save the clipped raster
writeRaster(masked_raster, here("path/FTY_clipped.tif"), filetype = "GTiff", overwrite = TRUE)


###CLCPlusBB###

# Crop both rasters to the intersection extent
cropped_CLCPlusBB <- crop(CLCPlusBB, reference_raster)

# Resample the cropped IMD raster to match the resolution of the reference raster
resampled_CLCPlusBB <- resample(cropped_CLCPlusBB, reference_raster, method = "near")

# Mask the resampled raster based on the reference raster
masked_raster_CLCPlusBB <- mask(resampled_CLCPlusBB, reference_raster)

# Save the clipped raster
writeRaster(masked_raster_CLCPlusBB, here("path/CLCPlusBB_clipped.tif"), filetype = "GTiff", overwrite = TRUE)

##### FTY 100m #####

# Crop both rasters to the intersection extent
cropped_fty_100 <- crop(fty_100, reference_raster)

# Resample the cropped IMD raster to match the resolution of the reference raster
resampled_fty_100 <- resample(cropped_fty_100, reference_raster, method = "near")

# Mask the resampled raster based on the reference raster
masked_raster_fty_100 <- mask(resampled_fty_100, reference_raster)

# Save the clipped raster
writeRaster(masked_raster_fty_100, here("path/FTY_100m_clipped.tif"), filetype = "GTiff", overwrite = TRUE)

` ` `

` ` `{r split IMD into continuous and discontinuous}
library(terra)
library(here)

#Load clipped IMD layer
IMD <- rast("path/IMD_clipped.tif")
```

````
# Split the raster into two based on the specified value ranges

# Raster 1: Values between 1 and 70
raster_1_79 <- ifel(IMD >= 1 & IMD <= 79, 1, NA)

# Raster 2: Values between 80 and 100
raster_80_100 <- ifel(IMD >= 80 & IMD <= 100, 1, NA)


# Save the clipped raster
writeRaster(raster_1_79, here("path/IMD_1_79.tif"), filetype = "GTiff", overwrite = TRUE)

# Save the clipped raster
writeRaster(raster_80_100, here("path/IMD_80_100.tif"), filetype = "GTiff", overwrite = TRUE)

```

```{r add IMD column to Coastal Zones (cz)}

# Install all packages
install.packages("exactextractr")

library(sf)
library(exactextractr)
library(dplyr)
library(here)
library(here)
library(terra)

# Load the raster dataset (IMD)
imd_raster <- raster("path/IMD_clipped.tif")

# Load the vector layer (cz)
cz <- st_read(here("path/CZ_2018.gdb"), layer="CZ_2018")

# Calculate the zonal median statistics for the entire dataset
medians <- exact_extract(imd_raster, cz, 'median')

# Add the median values as a new column in the original vector layer
cz$IMD_median <- medians
````

````
# Create the code_and_IMD column based on the conditions
cz <- cz %>%
  mutate(code_and_IMD = case_when(
    CODE_3_18 == 112 & IMD_median >= 80 ~ 112,
    CODE_3_18 == 112 & (IMD_median < 80 | is.na(IMD_median)) ~ 122,
    TRUE ~ CODE_5_18  # Fill remaining cases with CODE_5_18
  ))


# Write the output to a new GPKG
st_write(cz, "path/cz_and_IMD.gpkg", filetype = "GPKG", overwrite=TRUE)

```


```{r add IMD column to Riparian Zones (rz)}

# Install all packages
install.packages("exactextractr")

library(sf)
library(exactextractr)
library(dplyr)
library(here)
library(here)
library(terra)

# Load the raster dataset (IMD)
imd_raster <- raster("path/IMD_clipped.tif")

# Load the vector layer (rz)
rz <- st_read(here("path/rpz_DU002A.shp"))

# Calculate the zonal median statistics for the entire dataset
medians <- exact_extract(imd_raster, rz, 'median')

# Add the median values as a new column in the original vector layer
rz$IMD_median <- medians

# Create the code_and_IMD column based on the conditions
rz <- rz %>%
  mutate(code_and_IMD = case_when(
    CODE_3_18 == 112 & IMD_median >= 80 ~ 112,
````

```
    CODE_3_18 == 112 & (IMD_median < 80 | is.na(IMD_median)) ~ 122,
    TRUE ~ CODE_4_18  # Fill remaining cases with CODE_4_18
  ))


# If you want to write the output to a new shapefile
st_write(rz, "path/rpz_and_IMD.gpkg", filetype = "GPKG", overwrite=TRUE)

```

```{r add IMD column to ua}

install.packages("exactextractr")

library(sf)
library(exactextractr)
library(dplyr)
library(here)
library(here)
library(terra)

# Load the raster dataset (IMD)
imd_raster <- raster("path/IMD_clipped.tif")

vector_data <- st_read(here("path/coastline_buffered_10m.gpkg"), layer="coastline_buffered_10m")

# Load the vector layer Urban Atlas (ua)
ua <- st_read(here("path/UrbanAtlasBBox.gpkg"), layer="UA2018_FUA788")

# Get the bounding box (extent) of the vector_data
extent_vector_data <- st_bbox(vector_data)

# Subset the ua layer based on the extent of vector_data
ua_subset <- ua[st_intersects(ua, vector_data, sparse = FALSE), ]
plot (ua)


# Calculate the zonal median statistics for the entire dataset
medians <- exact_extract(imd_raster, ua, 'median')

# Add the median values as a new column in the original vector layer
ua$IMD_median <- medians
```

````
# Create the code_and_IMD column based on the conditions
ua <- ua %>%
  mutate(code_and_IMD = case_when(
    code_2018 == 12100 & IMD_median >= 80 ~ 112,
    code_2018 == 12100 & (IMD_median < 80 | is.na(IMD_median)) ~ 122,
    TRUE ~ as.numeric(code_2018)  # Fill remaining cases with code_2018 as numeric
  ))

# If you want to write the output to a new shapefile
st_write(ua, "path/ua_and_IMD.gpkg", filetype = "GPKG", overwrite=TRUE)


```

```{r create raster EUHydro Inland Water}

#Load reclassified vector layer
EU_IW <- vect(here("path/EU-Hydro.gdb"), layer="InlandWater")

# Load reference raster
reference_raster <- rast("path/boundary_10m.tif")

# Rasterize the vector layer with 10m resolution and snap it to the reference raster grid
EUHydro_raster_10m <- rasterize(EU_IW , rast(EU_IW , resolution = res(reference_raster), crs = crs(reference_raster)), field = 1)
EUHydro_aligned_raster <- resample(EUHydro_raster_10m, reference_raster, method = "near")

EUHydro_aligned_raster <-mask(EUHydro_aligned_raster,reference_raster)

# Save the raster with 10m resolution
writeRaster(EUHydro_aligned_raster, here("path/EUHydro_raster_10m.tif"), filetype = "GTiff", overwrite = TRUE, datatype='INT1U',
gdal=c("COMPRESS=LZW"))



```

```{r create raster of coastal zones}
library(terra)
library(dplyr)
library(here)
library(readr)
````

```
library(raster)


# Read the boundary in vector format
vector_data <- vect("path/coastline_buffered_10m.gpkg", layer="coastline_buffered_10m")

# Read coastal zones vector dataset with IMD integrated
cz <- vect(here("path/cz_and_IMD.gpkg"),extent=ext(vector_data), layer="cz_and_IMD")


# what is the layer of cz: cz @ptr[["layer"]]
# what are the column names of cz:
#cz @ptr[["names"]]


###########Extract the unique values of each dataset, reclassify them in order and export these as csv##############################
#create reclassification table of cz
cz_reclass_table <- data.frame(original_values = sort(unique(cz$code_and_IMD))) %>% mutate(reclassified_values = rank(original_values))

write_csv(cz_reclass_table, here("path/cz_unique_values.csv"))


###############Join the new class codes to the existing local component vector layers#################

#Create a dataframe of the coastal zone layer
cz_df <- as.data.frame(cz)

# Perform the join to map old values to new values
cz_df_with_reclass <- left_join(cz_df, cz_reclass_table, by = c("code_and_IMD" = "original_values"))

# Update the SpatVector with the new class codes
cz$reclassified_values <- cz_df_with_reclass$reclassified_values
cz$reclassified_values <- as.integer(cz$reclassified_values)

writeVector(cz, "path/cz_reclassified_v2.gpkg", overwrite=TRUE)

################Rasterize cz vector################

#Load reclassified vector layer
cz_recl <- vect(here("path/cz_reclassified_v2.gpkg"),extent=ext(vector_data), layer="cz_reclassified_v2")

# Load reference raster
```

```
reference_raster <- rast("path/Reprojected/boundary_10m.tif")

# Rasterize the vector layer with 10m resolution and snap it to the reference raster grid
cz_raster_10m  <-  rasterize(cz_recl  ,  rast(cz_recl  ,  resolution  =  res(reference_raster),  crs  =  crs(reference_raster)),  field  =
"reclassified_values")
cz_aligned_raster <- resample(cz_raster_10m, reference_raster, method = "near")

cz_aligned_raster <-mask(cz_aligned_raster,reference_raster)

# Save the raster with 10m resolution
writeRaster(cz_aligned_raster,   here("path/cz_raster_10m_v2.tif"),   filetype   =   "GTiff",   overwrite   =   TRUE,      datatype='INT1U',
gdal=c("COMPRESS=LZW"))


```

```{r create raster of riparian zones}

library(terra)
library(dplyr)
library(here)
library(readr)
library(raster)


# Read boundary vector data
vector_data <- vect("path/coastline_buffered_10m.gpkg", layer="coastline_buffered_10m")

rpz <- vect(
  here("path/rpz_and_IMD.gpkg"),
  extent=ext(vector_data), layer= "rpz_and_IMD")
# what are the column names of rpz:
rpz@ptr[["names"]]


###########Extract the unique values of each dataset, reclassify them in order and export these as csv##############################
#create reclassification table of rpz

rpz_reclass_table <- data.frame(original_values = sort(unique(rpz$code_and_IMD))) %>% mutate(reclassified_values = rank(original_values))

write_csv(rpz_reclass_table, here("path/rpz_unique_values.csv"))
```

````
###############Join the new class codes to the existing local component vector layers#################

#Create a dataframe of the coastal zone layer
rpz_df <- as.data.frame(rpz)

# Perform the join to map old values to new values
rpz_df_with_reclass <- left_join(rpz_df, rpz_reclass_table, by = c("code_and_IMD" = "original_values"))

# Update the SpatVector with the new class codes
rpz$reclassified_values <- rpz_df_with_reclass$reclassified_values
rpz$reclassified_values <- as.integer(rpz$reclassified_values)

writeVector(rpz, "path/rpz_reclassified_v2.gpkg", overwrite=TRUE)

################Rasterize rz vector################

#Load reclassified vector layer
rpz_recl <- vect(here("path/rpz_reclassified_v2.gpkg"),extent=ext(vector_data), layer="rpz_reclassified_v2")

# Load reference raster
reference_raster <- rast("path/boundary_10m.tif")

# Rasterize the vector layer with 10m resolution and snap it to the reference raster grid
rpz_raster_10m <- rasterize(rpz_recl , rast(rpz_recl , resolution = res(reference_raster), crs = crs(reference_raster)), field =
"reclassified_values")
rpz_aligned_raster <- resample(rpz_raster_10m, reference_raster, method = "near")

rpz_aligned_raster <-mask(rpz_aligned_raster,reference_raster)

# Save the raster with 10m resolution
writeRaster(rpz_aligned_raster, here("path/rpz_raster_10m_v2.tif"), filetype = "GTiff", overwrite = TRUE, datatype='INT1U',
gdal=c("COMPRESS=LZW"))


```

```{r create raster of urban atlas}

library(terra)
library(dplyr)
library(here)
library(readr)
````

```
library(raster)


# Read boundary vector data
vector_data <- vect("path/coastline_buffered_10m.gpkg", layer="coastline_buffered_10m")

ua <- vect(
  here("path/ua_and_IMD.gpkg"),
  extent=ext(vector_data), layer="ua_and_IMD")
# what are the column names of ua:
ua @ptr[["layer"]]
ua@ptr[["names"]]


###########Extract the unique values of each dataset, reclassify them in order and export these as csv##############################
#create reclassification table of ua
ua_reclass_table <- data.frame(original_values = sort(unique(ua$code_and_IMD))) %>% mutate(reclassified_values = rank(original_values))

write_csv(ua_reclass_table, here("path/ua_unique_values.csv"))

###############Join the new class codes to the existing local component vector layers#################

#Create a dataframe of the coastal zone layer
ua_df <- as.data.frame(ua)

# Perform the join to map old values to new values
ua_df_with_reclass <- left_join(ua_df, ua_reclass_table, by = c("code_and_IMD" = "original_values"))

# Update the SpatVector with the new class codes
ua$reclassified_values <- ua_df_with_reclass$reclassified_values
ua$reclassified_values <- as.integer(ua$reclassified_values)

writeVector(ua, "path/ua_reclassified_v2.gpkg", overwrite=TRUE)

################Rasterize ua vector################

#Load reclassified vector layer
ua_recl <- vect(here("path/ua_reclassified_v2.gpkg"),extent=ext(vector_data), layer="ua_reclassified_v2")

# Load reference raster
reference_raster <- rast("path/boundary_10m.tif")
```

````
# Rasterize the vector layer with 10m resolution and snap it to the reference raster grid
ua_raster_10m <- rasterize(ua_recl , rast(ua_recl , resolution = res(reference_raster), crs = crs(reference_raster)), field =
"reclassified_values")
ua_aligned_raster <- resample(ua_raster_10m, reference_raster, method = "near")

ua_aligned_raster <-mask(ua_aligned_raster,reference_raster)

# Save the raster with 10m resolution
writeRaster(ua_aligned_raster, here("path/ua_raster_10m_v2.tif"), filetype = "GTiff", overwrite = TRUE, datatype='INT1U',
gdal=c("COMPRESS=LZW"))


```

```{r Reclassify IMD to continuous and discontinuous}

# Load library
library(terra)
library(here)

#Load IMD
IMD <- rast("path/IMD_2018_010m_03035_V2_0.tif")

# Reclassify values between 1 and 79 to 1
IMD_reclassified <- ifel(IMD >= 1 & IMD <= 79, 1, IMD)


# Save the clipped raster
writeRaster(IMD_reclassified, here("path/reclassified_IMD_1_79.tif"), filetype = "GTiff", overwrite = TRUE)

```

```{r cz and rpz River Bank class and CLC+BB 10 = 811 ETA}

library(terra)

rpz <- rast("path/rpz_raster_10m_v2.tif")
cz <- rast("path/cz_raster_10m_v2.tif")
clcbb <- rast ("path/CLCPlusBB_clipped.tif")

# Create a raster dataset where river banks from rz (6220) and Water (10) from CLC+BB overlap
# Reclassify the raster based on the conditions
````

````
rpz_CLC_river_bank <- ifel(clcbb == 10 & rpz == 37, 1, 0)

# Export the reclassified raster
writeRaster(rpz_CLC_river_bank, "path/rz_CLC_river_bank.tif", overwrite = TRUE,
            datatype = "INT1U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))

############ Create a raster dataset where river banks from cz (62200) and Water (10) from CLC+BB overlap

# Reclassify the raster based on the conditions
cz_CLC_river_bank <- ifel(clcbb == 10 & cz == 43, 1, 0)

# Export the reclassified raster
writeRaster(cz_CLC_river_bank, "path/cz_CLC_river_bank.tif", overwrite = TRUE,
            datatype = "INT1U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))

```

```{r Transitional forest - FTY and CLC+BB}

#Create a transitional forest layer from the FTY and CLC+BB layers

library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
fty <- rast ("path/FTY_clipped.tif")

# Reclassify the raster based on the conditions
fty_CLCBB_transitional_forest <- ifel(clcbb %in% c(5, 6, 7, 9) & fty %in% c(1, 2), 1, 0)

# Export the reclassified raster
writeRaster(fty_CLCBB_transitional_forest, "path/fty_CLCBB_transitional_forest.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))

```

```{r rpz and clcbb forest reclassification}
# Load necessary library
library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
````

```
rpz <- rast("path/rpz_raster_10m_v2.tif")

# Reclassify forest using CLCBB inside of riparian zones forest boundaries as these were mapped very inconsistently
rpz_CLCBB_forest <- ifel(rpz == 22, 410,
                     ifel(rpz == 23, 410,
                     ifel(rpz == 24, 420,
                     ifel(rpz == 25, 420, 0))))

# Apply the reclassification based on the ruleset
rpz_CLCBB_forest <- ifel(clcbb == 2 & rpz == 22, 420, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 3 & rpz == 22, 410, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 4 & rpz == 22, 430, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 2 & rpz == 23, 420, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 3 & rpz == 23, 410, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 4 & rpz == 23, 430, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 2 & rpz == 24, 420, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 3 & rpz == 24, 410, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 4 & rpz == 24, 430, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 2 & rpz == 25, 420, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 3 & rpz == 25, 410, rpz_CLCBB_forest)
rpz_CLCBB_forest <- ifel(clcbb == 4 & rpz == 25, 430, rpz_CLCBB_forest)

# Export the reclassified raster
writeRaster(rpz_CLCBB_forest, "path/rpz_CLCBB_forest.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))
\`\`\`

\`\`\`{r cz and clcbb forest reclassification}
# Load necessary library
library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
cz <- rast("path/cz_raster_10m_v2.tif")

# Reclassify forest using CLCBB inside of coastal zones forest boundaries as these were mapped very inconsistently
cz_clcbb_forest <- ifel(cz == 27, 410,
                     ifel(cz == 28, 420, 0))

# Apply the reclassification based on the ruleset
cz_clcbb_forest <- ifel(clcbb == 2 & cz == 27, 420, cz_clcbb_forest)
cz_clcbb_forest <- ifel(clcbb == 3 & cz == 27, 410, cz_clcbb_forest)
```

````
cz_clcbb_forest <- ifel(clcbb == 4 & cz == 27, 430, cz_clcbb_forest)
cz_clcbb_forest <- ifel(clcbb == 2 & cz == 28, 420, cz_clcbb_forest)
cz_clcbb_forest <- ifel(clcbb == 3 & cz == 28, 410, cz_clcbb_forest)
cz_clcbb_forest <- ifel(clcbb == 4 & cz == 28, 430, cz_clcbb_forest)

# Export the reclassified raster
writeRaster(cz_clcbb_forest, "path/cz_CLCBB_forest.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))

```

```{r FTY & CLCBB - evergreen forest only }
# Load necessary library
library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
fty <- rast ("path/FTY_clipped.tif")

# Map evergreen forest that falls within the forest type 10m layer
fty_CLCBB_transitional_forest <- ifel(clcbb %in% c(4) & fty %in% c(1, 2), 1, 0)

# Export the reclassified raster
writeRaster(fty_CLCBB_transitional_forest, "path/fty_CLCBB_evergreen_forest.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))


```

```{r Grassland and arable land split using clcbb - Coastal Zones (cz) }

# Load necessary library
library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
cz <- rast("path/cz_raster_10m_v2.tif")

# Reclassify coastal zones arable land and pasture using CLCBB as this data is more accurate that the cz data
cz_clcbb_arab_gras <- ifel(cz == 20, 210,
                     ifel(cz == 33, 310, 0))
````

````
# Apply the reclassification based on the ruleset
cz_clcbb_arab_gras <- ifel(clcbb == 6 & cz == 20, 310, cz_clcbb_arab_gras)
cz_clcbb_arab_gras <- ifel(clcbb == 7 & cz == 20, 210, cz_clcbb_arab_gras)
cz_clcbb_arab_gras <- ifel(clcbb == 6 & cz == 33, 310, cz_clcbb_arab_gras)
cz_clcbb_arab_gras <- ifel(clcbb == 7 & cz == 33, 210, cz_clcbb_arab_gras)

# Export the reclassified raster
writeRaster(cz_clcbb_arab_gras, "path/cz_CLCBB_arab_grass.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))


```

```{r Grassland and arable land split using clcbb - Riparian Zones (rpz)}

# Load necessary library
library(terra)

#Load datasets
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
rpz <- rast("path/rpz_raster_10m_v2.tif")

# Reclassify riparian zones arable land and pasture using CLCBB as this data is more accurate that the rpz data
rpz_clcbb_arab_gras <- ifel(rpz == 13, 210,
                     ifel(rpz == 30, 310, 0))

# Apply the reclassification based on the ruleset
rpz_clcbb_arab_gras <- ifel(clcbb == 6 & rpz == 13, 310, rpz_clcbb_arab_gras)
rpz_clcbb_arab_gras <- ifel(clcbb == 7 & rpz == 13, 210, rpz_clcbb_arab_gras)
rpz_clcbb_arab_gras <- ifel(clcbb == 6 & rpz == 30, 310, rpz_clcbb_arab_gras)
rpz_clcbb_arab_gras <- ifel(clcbb == 7 & rpz == 30, 210, rpz_clcbb_arab_gras)

# Export the reclassified raster
writeRaster(rpz_clcbb_arab_gras, "path/rpz_CLCBB_arab_grass.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))


```

```{r clc_imd - Combination of Corine and IMD  to discriminate continuous and discontinuous settlements}

# Load necessary library
````

````
library(terra)

# Load the raster datasets
CLC <- rast("path/CLC_clipped.tif")
IMD <- rast("path/IMD_clipped.tif")
boundary <- rast("path/boundary_10m.tif")

# Initialize the new raster based on the boundary raster's extent and resolution
new_raster <- rast(boundary)

values(new_raster) <- 0  # Initialize with NA

# Rule 1: IMD in category 80-100 & CLC 1 or 2 -> new_code 111
new_raster <- ifel(IMD %in% c(80:100) & CLC %in% c(1, 2), 111, new_raster)

# Rule 2: IMD in category 1-79 & CLC 1 or 2 -> new_code 121
new_raster <- ifel(IMD %in% c(1:79) & CLC %in% c(1, 2), 121, new_raster)

# Rule 3: IMD in category 80-100 & CLC 3 -> new_code 112
new_raster <- ifel(IMD %in% c(80:100) & CLC == 3, 112, new_raster)

# Rule 4: IMD in category 1-79 & CLC 3 -> new_code 122
new_raster <- ifel(IMD %in% c(1:79) & CLC == 3, 122, new_raster)


# Export the reclassified raster
writeRaster(new_raster, "path/clc_imd.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))



```

```{r clc_imd_clcbb}

library(terra)

CLC <- rast("path/CLC_clipped.tif")
IMD <- rast("path/IMD_clipped.tif")
clc_bb <- rast("path/CLCPlusBB_clipped.tif")
boundary <- rast("path/boundary_10m.tif")
````

````
# Initialize the new raster based on the boundary raster's extent and resolution
new_raster <- rast(boundary)
values(new_raster) <- 0  # Initialize with NA

# Select all sealed area from IMD and clc_bb
IMD_clc_bb <- ifel(IMD %in% c(1:100) | clc_bb == 1, 131, new_raster)

#Reclassify the sealed areas that overlap with CLC (Roads and railways) as this class under the ETA
IMD_clc_bb_CLC <- ifel(IMD_clc_bb == 131 & CLC == 4, 131, 0)


# Export the reclassified raster
writeRaster(IMD_clc_bb_CLC, "path/clc_imd_clcbb.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))

```

```{r CLCBB and CLC class combinations}

# Load necessary library
library(terra)

# Load the raster datasets
clc_bb <- rast("path/CLCPlusBB_clipped.tif")
CLC <- rast("path/CLC_clipped.tif")
boundary <- rast("path/boundary_10m.tif")

# Initialize the new raster based on the boundary raster's extent and resolution and reclassify all classes to ETA classes based on their c
new_raster <- rast(boundary)
values(new_raster) <- 0  # Initialize with NA

# Rule
new_raster <- ifel(clc_bb == 7 & CLC == 12, 210, new_raster)
new_raster <- ifel(clc_bb == 7 & CLC == 13, 210, new_raster)
new_raster <- ifel(clc_bb == 5 & CLC == 15, 232, new_raster)
new_raster <- ifel(clc_bb %in% c(3,4,5) & CLC == 16, 230, new_raster)
new_raster <- ifel(clc_bb == 4 & CLC == 17, 231, new_raster)
new_raster <- ifel(clc_bb == 6 & CLC == 26, 320, new_raster)
new_raster <- ifel(clc_bb == 5 & CLC == 27, 520, new_raster)
new_raster <- ifel(clc_bb == 5 & CLC == 28, 530, new_raster)
new_raster <- ifel(clc_bb %in% c(5,6,7,9) & CLC == 29, 451, new_raster)
new_raster <- ifel(clc_bb == 10 & CLC == 40, 800, new_raster)
````

````r
new_raster <- ifel(clc_bb == 10 & CLC == 41, 900, new_raster)
new_raster <- ifel(IMD %in% c(1:79) & CLC %in% c(1, 2), 121, new_raster)
new_raster <- ifel(IMD %in% c(0:80) & CLC == 3, 112, new_raster)
new_raster <- ifel(IMD %in% c(1:79) & CLC == 3, 122, new_raster)


# Export the reclassified raster
writeRaster(new_raster, "path/clc_clcbb.tif", overwrite = TRUE,
            datatype = "INT2U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))



```
````
````

```
# R Script – CLMS only approach
# Excel files used as inputs for the scripts are too large to reasonably fit this report and are available upon request to the authors.

---
title: "CLMS only approach"
output: html_notebook
---

```{r load library and datasets}
library(terra)
library(readxl)

# Load the pre-processed raster datasets
# Load the boundary which will be used
boundary <- rast("path/boundary_10m.tif")

# Load datasets where rasters are consistent with the 'dataset' names in the Excel file
rpz <- rast("path/rpz_raster_10m_v2.tif")
cz <- rast("path/cz_raster_10m_v2.tif")
ua <- rast("path/ua_raster_10m_v2.tif")
IMD_1_79 <- rast("path/IMD_1_79.tif")
IMD_80_100 <- rast("path/IMD_80_100.tif")
EUHydro <- rast("path/EUHydro_raster_10m.tif")
clcbb <- rast ("path/CLCPlusBB_clipped.tif")
rpz_CLC_river_bank <- rast("path/rz_CLC_river_bank.tif")
cz_CLC_river_bank <- rast("path/cz_CLC_river_bank.tif")
FTY_clcbb <- rast("path/fty_CLCBB_transitional_forest.tif")
FTY_100m_clipped <- rast("path/FTY_100m_clipped.tif")
cz_CLCBB_forest<- rast("path/cz_CLCBB_forest.tif")
rpz_CLCBB_forest<- rast("path/rpz_CLCBB_forest.tif")
fty_CLCBB_evergreen_forest<- rast("path/fty_CLCBB_evergreen_forest.tif")
fty <- rast("path/FTY_clipped.tif")
cz_CLCBB_arab_grass <- rast("path/cz_CLCBB_arab_grass.tif")
rpz_CLCBB_arab_grass <- rast("path/rpz_CLCBB_arab_grass.tif")
clc_imd <- rast("path/clc_imd.tif")
clc_imd_clcbb  <- rast("path/clc_imd_clcbb.tif")
clc <- rast("path/CLC_clipped.tif")
clc_clcbb <- rast("path/clc_clcbb.tif")

```
```

````
```{r CLMS based mapping}
# Load the hierarchy that has been written to map Peloponnese according to only CLMS data
df <- read_excel("path/hierarchy.xlsx")

# Create a blank raster

blank_raster <- rast(ncol = ncol(boundary), nrow = nrow(boundary),
                     ext = ext(boundary), crs = crs(boundary), vals = 0)

# Start timing
start_time <- Sys.time()

  # Loop through each unique priority level
  for (priority in unique(df$priority)) {

    # Filter the dataframe for the current priority level
    priority_df <- df[df$priority == priority, ]

    # Create a temporary raster to store intermediate results for this priority level
    temp_raster <- blank_raster

    # Process rasters according to the current priority level
    for (i in 1:nrow(priority_df)) {
      current_dataset <- priority_df$dataset[i]
      current_codelist <- as.numeric(priority_df$codelist[i])
      current_reclas_val <- as.numeric(priority_df$reclas_val[i])

      # Get the raster layer (Assuming these layers are already loaded in the environment)
      raster_layer <- get(current_dataset)

      # Update the temporary raster with the current layer
      temp_raster <- ifel(temp_raster == 0 & raster_layer == current_codelist, current_reclas_val, temp_raster)
    }

    # Merge the temporary raster with the blank raster
    blank_raster <- ifel(blank_raster == 0, temp_raster, blank_raster)
  }

  # Save the final raster incrementally with compression
  writeRaster(blank_raster, "path/CLMS_only.tif", overwrite = TRUE,
              datatype = "INT1U", gdal = c("COMPRESS=LZW", "BIGTIFF=YES"))
````
````

````
# End timing
end_time <- Sys.time()

# Calculate time taken
time_taken <- end_time - start_time
time_in_minutes <- as.numeric(time_taken, units = "mins")
time_in_hours <- as.numeric(time_taken, units = "hours")

# Print the time taken
cat("Time taken:\n")
cat(time_in_minutes, "minutes\n")
cat(time_in_hours, "hours\n")
```
````
````

13.1.1.2. CLMS Typologies

dataset	class_level	class_name	column_name	class_code	original_data_format
ua	3	Continuous Urban fabric (S.L. > 80%)	code_2018	11100	vector
ua	4	Discontinuous Dense Urban Fabric (S.L.: 50% - 80%)	code_2018	11210	vector
ua	4	Discontinuous Medium Density Urban Fabric (S.L.: 30% - 50%)	code_2018	11220	vector
ua	4	Discontinuous Low Density Urban Fabric (S.L.: 10% - 30%)	code_2018	11230	vector
ua	4	Discontinuous very low density urban fabric (S.L. < 10%)	code_2018	11240	vector
ua	3	Isolated Structures	code_2018	11300	vector
ua	3	Industrial, commercial, public, military and private units	code_2018	12100	vector
ua	4	Fast transit roads and associated land	code_2018	12210	vector
ua	4	Other roads and associated land	code_2018	12220	vector
ua	4	Railways and associated land	code_2018	12230	vector
ua	3	Port areas	code_2018	12300	vector
ua	3	Airports	code_2018	12400	vector
ua	3	Mineral extraction and dump sites	code_2018	13100	vector
ua	3	Construction sites	code_2018	13300	vector
ua	3	Land without current use	code_2018	13400	vector
ua	3	Green urban areas	code_2018	14100	vector
ua	3	Sports and leisure facilities	code_2018	14200	vector
ua	2	Arable land (annual crops)	code_2018	21000	vector
ua	2	Permanent crops	code_2018	22000	vector
ua	2	Pastures	code_2018	23000	vector
ua	2	Complex and mixed cultivation patterns	code_2018	24000	vector
ua	2	Forests	code_2018	31000	vector
ua	2	Herbaceous vegetation associations	code_2018	32000	vector
ua	2	Open spaces with little or no vegetations	code_2018	33000	vector
ua	1	Wetlands	code_2018	40000	vector

ua	1	Water	code_2018	50000	vector
rpz	1	1 Urban	CODE_1_18	1000	vector
rpz	1	2 Cropland	CODE_1_18	2000	vector
rpz	1	3 Woodland and forest	CODE_1_18	3000	vector
rpz	1	4 Grassland	CODE_1_18	4000	vector
rpz	1	5 Heathland and scrub	CODE_1_18	5000	vector
rpz	1	6 Open spaces with little or no vegetation	CODE_1_18	6000	vector
rpz	1	7 Wetland	CODE_1_18	7000	vector
rpz	1	8 Water	CODE_1_18	8000	vector
rpz	2	1.1 Urban fabric, industrial, commercial, public, military and private units	CODE_2_18	1100	vector
rpz	2	1.2 Transport infrastructure	CODE_2_18	1200	vector
rpz	2	1.3 Mineral extraction, dump and construction sites, land without current use	CODE_2_18	1300	vector
rpz	2	1.4 Green urban, sports and leisure facilities	CODE_2_18	1400	vector
rpz	2	2.1 Arable land	CODE_2_18	2100	vector
rpz	2	2.2 Permanent crops	CODE_2_18	2200	vector
rpz	2	2.3 Heterogeneous agricultural area	CODE_2_18	2300	vector
rpz	2	3.1 Broadleaved forest	CODE_2_18	3100	vector
rpz	2	3.2 Coniferous forest	CODE_2_18	3200	vector
rpz	2	3.3 Mixed forest	CODE_2_18	3300	vector
rpz	2	3.4 Transitional woodland and scrub	CODE_2_18	3400	vector
rpz	2	3.5 Lines of trees and scrub	CODE_2_18	3500	vector
rpz	2	3.6 Damaged forest	CODE_2_18	3600	vector
rpz	2	4.1 Managed grassland	CODE_2_18	4100	vector
rpz	2	4.2 Natural & semi-natural grassland	CODE_2_18	4200	vector
rpz	2	5.1 Heathland and moorland	CODE_2_18	5100	vector
rpz	2	5.2 Alpine scrub land	CODE_2_18	5200	vector
rpz	2	5.3 Sclerophyllous scrubs	CODE_2_18	5300	vector
rpz	2	6.1 Sparsely vegetated areas	CODE_2_18	6100	vector
rpz	2	6.2 Beaches, dunes, river banks	CODE_2_18	6200	vector

rpz	2	6.3 Bare rocks, burnt areas, glaciers and perpetual snow	CODE_2_18	6300	vector
rpz	2	7.1 Inland wetlands	CODE_2_18	7100	vector
rpz	2	7.2 Coastal wetlands	CODE_2_18	7200	vector
rpz	2	8.1 Water courses	CODE_2_18	8100	vector
rpz	2	8.2 Lakes and reservoirs	CODE_2_18	8200	vector
rpz	2	8.3 Transitional waters	CODE_2_18	8300	vector
rpz	2	8.4 Sea and ocean	CODE_2_18	8400	vector
rpz	3	1.1.1 Urban fabric (predominantly public and private units)	CODE_3_18	1110	vector
rpz	3	1.1.2 Industrial, commercial and military units	CODE_3_18	1120	vector
rpz	3	1.2.1 Road networks and associated land	CODE_3_18	1210	vector
rpz	3	1.2.2 Railways and associated land	CODE_3_18	1220	vector
rpz	3	1.2.3 Port areas and associated land	CODE_3_18	1230	vector
rpz	3	1.2.4 Airports and associated land	CODE_3_18	1240	vector
rpz	3	1.3.1 Mineral extraction, dump and construction sites	CODE_3_18	1310	vector
rpz	3	1.3.2 Land without current use	CODE_3_18	1320	vector
rpz	3	2.1.1 Arable irrigated and non-irrigated land	CODE_3_18	2110	vector
rpz	3	2.1.2 Greenhouses	CODE_3_18	2120	vector
rpz	3	2.2.1 Vineyards, fruit trees and berry plantations	CODE_3_18	2210	vector
rpz	3	2.2.2 Olive groves	CODE_3_18	2220	vector
rpz	3	2.3.1 Annual crops associated with permanent crops	CODE_3_18	2310	vector
rpz	3	2.3.2 Complex cultivation patterns	CODE_3_18	2320	vector
rpz	3	2.3.3 Land principally occupied by agriculture with significant areas of natural vegetation	CODE_3_18	2330	vector
rpz	3	2.3.4 Agro-forestry	CODE_3_18	2340	vector
rpz	3	3.1.1 Natural & seminatural broadleaved forest	CODE_3_18	3110	vector
rpz	3	3.1.2 Highly artificial broadleaved plantations	CODE_3_18	3120	vector
rpz	3	3.2.1 Natural & seminatural coniferous forest	CODE_3_18	3210	vector
rpz	3	3.2.2 Highly artificial coniferous plantations	CODE_3_18	3220	vector
rpz	3	3.3.1 Natural & seminatural mixed forest	CODE_3_18	3310	vector
rpz	3	3.3.2 Highly artificial mixed plantations	CODE_3_18	3320	vector

rpz	3	4.2.1 Semi-natural grassland	CODE_3_18	4210	vector
rpz	3	4.2.2 Alpine and sub-alpine natural grassland	CODE_3_18	4220	vector
rpz	3	6.2.1 Beaches and dunes	CODE_3_18	6210	vector
rpz	3	6.2.2 River banks	CODE_3_18	6220	vector
rpz	3	6.3.1 Bare rocks, outcrops, cliffs	CODE_3_18	6310	vector
rpz	3	6.3.2 Burnt areas (except burnt forest)	CODE_3_18	6320	vector
rpz	3	6.3.3 Glaciers and perpetual snow	CODE_3_18	6330	vector
rpz	3	7.1.1 Inland marshes	CODE_3_18	7110	vector
rpz	3	7.1.2 Peat bogs	CODE_3_18	7120	vector
rpz	3	7.2.1 Salt marshes	CODE_3_18	7210	vector
rpz	3	7.2.2 Salines	CODE_3_18	7220	vector
rpz	3	7.2.3 Intertidal flats	CODE_3_18	7230	vector
rpz	3	8.1.1 Natural & seminatural water courses	CODE_3_18	8110	vector
rpz	3	8.1.2 Highly modified water courses and canals	CODE_3_18	8120	vector
rpz	3	8.1.3 Seasonally connected water courses (oxbows)	CODE_3_18	8130	vector
rpz	3	8.2.1 Natural lakes	CODE_3_18	8210	vector
rpz	3	8.2.2 Reservoirs	CODE_3_18	8220	vector
rpz	3	8.2.3 Aquaculture ponds	CODE_3_18	8230	vector
rpz	3	8.2.4 Standing water bodies of extractive industrial sites	CODE_3_18	8240	vector
rpz	3	8.3.1 Lagoons	CODE_3_18	8310	vector
rpz	3	8.3.2 Estuaries	CODE_3_18	8320	vector
rpz	4	1.1.1.1 Continuous Urban Fabric (IM.D ≥ 80%)	CODE_4_18	1111	vector
rpz	4	1.1.1.2 Dense Urban Fabric (IM.D ≥ 30-80%)	CODE_4_18	1112	vector
rpz	4	1.1.1.3 Low Density Urban Fabric (IM.D < 30%)	CODE_4_18	1113	vector
rpz	4	7.1.2.1 Exploited peat bogs	CODE_4_18	7121	vector
rpz	4	7.1.2.2 Unexploited peat bogs	CODE_4_18	7122	vector
cz	1	1 Urban	CODE_1_18	10000	vector
cz	1	2 Cropland	CODE_1_18	20000	vector
cz	1	3 Woodland and forest	CODE_1_18	30000	vector

cz	1	4 Grassland	CODE_1_18	40000	vector
cz	1	5 Heathland and scrub	CODE_1_18	50000	vector
cz	1	6 Open spaces with little or no vegetation	CODE_1_18	60000	vector
cz	1	7 Wetland	CODE_1_18	70000	vector
cz	1	8 Water	CODE_1_18	80000	vector
cz	2	1.1 Urban fabric, industrial, commercial, public, military and private units	CODE_2_18	11000	vector
cz	2	1.2 Transport infrastructure	CODE_2_18	12000	vector
cz	2	1.3 Mineral extraction, dump and construction sites, land without current use	CODE_2_18	13000	vector
cz	2	1.4 Green urban, sports and leisure facilities	CODE_2_18	14000	vector
cz	2	2.1 Arable land	CODE_2_18	21000	vector
cz	2	2.2 Permanent crops	CODE_2_18	22000	vector
cz	2	2.3 Heterogeneous agricultural area	CODE_2_18	23000	vector
cz	2	3.1 Broadleaved forest	CODE_2_18	31000	vector
cz	2	3.2 Coniferous forest	CODE_2_18	32000	vector
cz	2	3.3 Mixed forest	CODE_2_18	33000	vector
cz	2	3.4 Transitional woodland and scrub	CODE_2_18	34000	vector
cz	2	3.5 Lines of trees and scrub	CODE_2_18	35000	vector
cz	2	3.6 Damaged forest	CODE_2_18	36000	vector
cz	2	4.1 Managed grassland	CODE_2_18	41000	vector
cz	2	4.2 Natural & semi-natural grassland	CODE_2_18	42000	vector
cz	2	5.1 Heathland and moorland	CODE_2_18	51000	vector
cz	2	5.2 Alpine scrub land	CODE_2_18	52000	vector
cz	2	5.3 Sclerophyllous scrubs	CODE_2_18	53000	vector
cz	2	6.1 Sparsely vegetated areas	CODE_2_18	61000	vector
cz	2	6.2 Beaches, dunes, river banks	CODE_2_18	62000	vector
cz	2	6.3 Bare rocks, burnt areas, glaciers and perpetual snow	CODE_2_18	63000	vector
cz	2	7.1 Inland wetlands	CODE_2_18	71000	vector
cz	2	7.2 Coastal wetlands	CODE_2_18	72000	vector
cz	2	8.1 Water courses	CODE_2_18	81000	vector

cz	2	8.2 Lakes and reservoirs	CODE_2_18	82000	vector
cz	2	8.3 Transitional waters	CODE_2_18	83000	vector
cz	2	8.4 Sea and ocean	CODE_2_18	84000	vector
cz	3	1.1.1 Urban fabric	CODE_3_18	11100	vector
cz	3	1.1.2 Industrial, commercial, public and military units	CODE_3_18	11200	vector
cz	3	1.2.1 Road networks and associated land	CODE_3_18	12100	vector
cz	3	1.2.2 Railways and associated land	CODE_3_18	12200	vector
cz	3	1.2.3 Port areas and associated land	CODE_3_18	12300	vector
cz	3	1.2.4 Airports and associated land	CODE_3_18	12400	vector
cz	3	1.3.1 Mineral extraction, dump and construction sites	CODE_3_18	13100	vector
cz	3	1.3.2 Land without current use	CODE_3_18	13200	vector
cz	3	2.1.1 Arable irrigated and non-irrigated land	CODE_3_18	21100	vector
cz	3	2.1.2 Greenhouses	CODE_3_18	21200	vector
cz	3	2.2.1 Vineyards, fruit trees and berry plantations	CODE_3_18	22100	vector
cz	3	2.2.2 Olive groves	CODE_3_18	22200	vector
cz	3	2.3.1 Annual crops associated with permanent crops	CODE_3_18	23100	vector
cz	3	2.3.2 Complex cultivation patterns	CODE_3_18	23200	vector
cz	3	2.3.3 Land principally occupied by agriculture with significant areas of natural vegetation	CODE_3_18	23300	vector
cz	3	2.3.4 Agro-forestry	CODE_3_18	23400	vector
cz	3	3.1.1 Natural & semi-natural broadleaved forest	CODE_3_18	31100	vector
cz	3	3.1.2 Highly artificial broadleaved plantations	CODE_3_18	31200	vector
cz	3	3.2.1 Natural & semi-natural coniferous forest	CODE_3_18	32100	vector
cz	3	3.2.2 Highly artificial coniferous plantations	CODE_3_18	32200	vector
cz	3	3.3.1 Natural & semi-natural mixed forest	CODE_3_18	33100	vector
cz	3	3.3.2 Highly artificial mixed plantations	CODE_3_18	33200	vector
cz	3	4.2.1 Semi-natural grassland	CODE_3_18	42100	vector
cz	3	4.2.2 Alpine and sub-alpine natural grassland	CODE_3_18	42200	vector
cz	3	6.1.1 Sparse vegetation on sands	CODE_3_18	61100	vector
cz	3	6.1.2 Sparse vegetation on rocks	CODE_3_18	61200	vector

cz	3	6.2.1 Beaches and dunes	CODE_3_18	62100	vector
cz	3	6.2.2 River banks	CODE_3_18	62200	vector
cz	3	6.3.1 Bare rocks, outcrops, cliffs	CODE_3_18	63100	vector
cz	3	6.3.2 Burnt areas (except burnt forest)	CODE_3_18	63200	vector
cz	3	6.3.3 Glaciers and perpetual snow	CODE_3_18	63300	vector
cz	3	7.1.1 Inland marshes	CODE_3_18	71100	vector
cz	3	7.1.2 Peat bogs	CODE_3_18	71200	vector
cz	3	7.2.1 Salt marshes	CODE_3_18	72100	vector
cz	3	7.2.2 Salines	CODE_3_18	72200	vector
cz	3	7.2.3 Intertidal flats	CODE_3_18	72300	vector
cz	3	8.1.1 Natural & semi-natural water courses	CODE_3_18	81100	vector
cz	3	8.1.2 Highly modified water courses and canals	CODE_3_18	81200	vector
cz	3	8.1.3 Seasonally connected water courses (oxbows)	CODE_3_18	81300	vector
cz	3	8.2.1 Natural lakes	CODE_3_18	82100	vector
cz	3	8.2.2 Reservoirs	CODE_3_18	82200	vector
cz	3	8.2.3 Aquaculture ponds	CODE_3_18	82300	vector
cz	3	8.2.4 Standing water bodies of extractive industrial sites	CODE_3_18	82400	vector
cz	3	8.3.1 Lagoons	CODE_3_18	83100	vector
cz	3	8.3.2 Estuaries	CODE_3_18	83200	vector
cz	3	8.3.3 Marine inlets and fjords	CODE_3_18	83300	vector
cz	3	8.4.1 Open sea	CODE_3_18	84100	vector
cz	3	8.4.2 Coastal waters	CODE_3_18	84200	vector
cz	4	1.1.1.1 Continuous urban fabric (IMD ≥80%)	CODE_4_18	11110	vector
cz	4	1.1.1.2 Dense urban fabric (IMD ≥30-80%)	CODE_4_18	11120	vector
cz	4	1.1.1.3 Low density fabric (IMD <30%)	CODE_4_18	11130	vector
cz	4	1.1.2.1 Industrial, commercial, public and military units (other)	CODE_4_18	11210	vector
cz	4	1.1.2.2 Nuclear energy plants and associated land	CODE_4_18	11220	vector
cz	4	1.2.3.1 Cargo port	CODE_4_18	12310	vector
cz	4	1.2.3.2 Passenger port	CODE_4_18	12320	vector

cz	4	1.2.3.3 Fishing port	CODE_4_18	12330	vector
cz	4	1.2.3.4 Naval port	CODE_4_18	12340	vector
cz	4	1.2.3.5 Marinas	CODE_4_18	12350	vector
cz	4	1.2.3.6 Local multi-functional harbours	CODE_4_18	12360	vector
cz	4	1.2.3.7 Shipyards	CODE_4_18	12370	vector
cz	4	1.3.1.1 Mineral extraction sites	CODE_4_18	13110	vector
cz	4	1.3.1.2 Dump sites	CODE_4_18	13120	vector
cz	4	1.3.1.3 Construction sites	CODE_4_18	13130	vector
cz	4	6.2.1.1 Beaches	CODE_4_18	62110	vector
cz	4	6.2.1.2 Dunes	CODE_4_18	62120	vector
cz	4	6.3.1.1 Bare rocks and outcrops	CODE_4_18	63110	vector
cz	4	6.3.1.2 Coastal cliffs	CODE_4_18	63120	vector
cz	4	7.1.2.1 Exploited peat bogs	CODE_4_18	71210	vector
cz	4	7.1.2.2 Unexploited peat bogs	CODE_4_18	71220	vector
cz	5	6.2.1.1.1 Sandy beaches	CODE_5_18	62111	vector
cz	5	6.2.1.1.2 Shingle beaches	CODE_5_18	62112	vector
clc	3	1.1.1 Continuous urban fabric		1	raster
clc	3	1.1.2 Discontinuous urban fabric		2	raster
clc	3	1.2.1 Industrial or commercial units		3	raster
clc	3	1.2.2 Road and rail networks and associated land		4	raster
clc	3	1.2.3 Port areas		5	raster
clc	3	1.2.4 Airports		6	raster
clc	3	1.3.1 Mineral extraction sites		7	raster
clc	3	1.3.2 Dump sites		8	raster
clc	3	1.3.3 Construction sites		9	raster
clc	3	1.4.1 Green urban areas		10	raster
clc	3	1.4.2 Sport and leisure facilities		11	raster
clc	3	2.1.1 Non-irrigated arable land		12	raster
clc	3	2.1.2 Permanently irrigated land		13	raster

clc	3	2.1.3 Rice fields		14	raster
clc	3	2.2.1 Vineyards		15	raster
clc	3	2.2.2 Fruit trees and berry plantations		16	raster
clc	3	2.2.3 Olive groves		17	raster
clc	3	2.3.1 Pastures		18	raster
clc	3	2.4.1 Annual crops associated with permanent crops		19	raster
clc	3	2.4.2 Complex cultivation patterns		20	raster
clc	3	2.4.3 Land principally occupied by agriculture, with significant areas of natural vegetation		21	raster
clc	3	2.4.4 Agro-forestry areas		22	raster
clc	3	3.1.1 Broad-leaved forest		23	raster
clc	3	3.1.2 Coniferous forest		24	raster
clc	3	3.1.3 Mixed forest		25	raster
clc	3	3.2.1 Natural grasslands		26	raster
clc	3	3.2.2 Moors and heathland		27	raster
clc	3	3.2.3 Sclerophyllous vegetation		28	raster
clc	3	3.2.4 Transitional woodland-shrub		29	raster
clc	3	3.3.1 Beaches, dunes, sands		30	raster
clc	3	3.3.2 Bare rocks		31	raster
clc	3	3.3.3 Sparsely vegetated areas		32	raster
clc	3	3.3.4 Burnt areas		33	raster
clc	3	3.3.5 Glaciers and perpetual snow		34	raster
clc	3	4.1.1 Inland marshes		35	raster
clc	3	4.1.2 Peat bogs		36	raster
clc	3	4.2.1 Salt marshes		37	raster
clc	3	4.2.2 Salines		38	raster
clc	3	4.2.3 Intertidal flats		39	raster
clc	3	5.1.1 Water courses		40	raster
clc	3	5.1.2 Water bodies		41	raster
clc	3	5.2.1 Coastal lagoons		42	raster

clc	3	5.2.2 Estuaries		43	raster
clc	3	5.2.3 Sea and ocean		44	raster
FTY_10m	-	1 Broadleaved forest		1	raster
FTY_10m	-	2 Coniferous forest		2	raster
FTY_100m	-	3 Mixed forest		3	raster
clcbb	-	1: Sealed		1	raster
clcbb	-	2: Woody – needle leaved trees		2	raster
clcbb	-	3: Woody – Broadleaved deciduous trees		3	raster
clcbb	-	4: Woody – Broadleaved evergreen trees		4	raster
clcbb	-	5: Low-growing woody plants (bushes, shrubs)		5	raster
clcbb	-	6: Permanent herbaceous		6	raster
clcbb	-	7: Periodically herbaceous		7	raster
clcbb	-	8: Lichens and mosses		8	raster
clcbb	-	9: Non- and sparsely-vegetated		9	raster
clcbb	-	10: Water		10	raster
clcbb	-	11: Snow and ice		11	raster
EUHydro	-	InlandWater		-	vector

13.1.1.3. European Ecosystem Typology

eta_code	class_level	class_name
100	1	1. Settlements and other artificial areas
110	2	1.1 Continuous settlement area
111	3	1.1.1 Continuous residential area
112	3	1.1.2 Continuous commercial and industrial area
120	2	1.2 Discontinuous settlement area
121	3	1.2.1 Discontinuous residential area
122	3	1.2.2 Discontinuous commercial and industrial area
130	2	1.3 Infrastructure
131	3	1.3.1 Road and rail networks and associated land
132	3	1.3.2 Port areas
133	3	1.3.3 Airports
134	3	1.3.4 Other infrastructure
135	3	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
136	3	1.3.6 Dump areas
137	3	1.3.7 Construction sites
140	2	1.4 Urban greenspace
141	3	1.4.1 Parks (including Zoos and botanical gardens)
142	3	1.4.2 Sports and recreation sites
143	3	1.4.3 Other urban green
150	2	1.5 Other artificial areas
151	3	1.5.1 Permanent Greenhouses
152	3	1.5.2 Cemeteries
153	3	1.5.3 Archaeological sites
154	3	1.5.4 Urban blue
200	1	2. Cropland

210	2	2.1 Annual cropland
211	3	2.1.1 Cereals excluding rice (C1000) excluding maize (C1500)
212	3	2.1.2 Maize (C1500 + G3000)
213	3	2.1.3 Dry pulses and protein crops (P0000)
214	3	2.1.4 Root crops, like sugar beet and potatoes (R0000)
215	3	2.1.5 Vegetables (including melons) and strawberries (V0000_S0000)
216	3	2.1.6 Industrial crops including annual bioenergy crops (I0000)
217	3	2.1.7 Flowers and ornamental plants (N0000)
218	3	2.1.8 Fallow land (Q0000)
219	3	2.1.9 Temporary grasses and grazing areas (G1000)
220	2	2.2 Rice fields
221	3	2.2.1 Rice fields (C2000)
230	2	2.3 Permanent crops
231	3	2.3.1 Olives (O1000)
232	3	2.3.2 Grapes (W1000)
233	3	2.3.3 Pome fruits (F1100)
234	3	2.3.4 Stone fruits (F1200)
235	3	2.3.5 Berries excluding strawberries (F3000)
236	3	2.3.6 Citrus fruits (T1000)
237	3	2.3.7 Nuts (F4000)
238	3	2.3.8 Hazelnut
239	3	2.3.9 Chestnut
240	2	2.4 Agro-forestry areas
241	3	2.4.1 Holm and cork oak forests
242	3	2.4.2 Other agro-forestry area
250	2	2.5 Mixed farmland
251	3	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
260	2	2.6 Other farmland
261	3	2.6.1 Nurseries

262	3	2.6.2 Christmas tree plantations
263	3	2.6.3 Perennial bioenergy crops
264	3	2.6.4 Field margins
300	1	3. Grassland (pastures, semi-natural and natural grasslands)
310	2	3.1 Sown pastures and fields (modified grassland)
311	3	3.1.1 Sown pastures used for grazing
312	3	3.1.2 Sown grassland mown frequently for fodder or silage
320	2	3.2 Natural and semi-natural grassland
321	3	3.2.1 Dry grassland
322	3	3.2.2 Seasonally wet and wet grassland
323	3	3.2.3 Alpine and subalpine grasslands
324	3	3.2.4 Woodland fringes and clearings and tall forb stands
325	3	3.2.5 Inland salt steppes
326	3	3.2.6 Sparsely wooded grasslands
327	3	3.2.7 Mesophilous extensive grassland
400	1	4. Forest and woodlands
410	2	4.1 Broadleaved deciduous forest
411	3	4.1.1 Riparian forest and woodland
412	3	4.1.2 Broadleaved swamp, woodland on non-acid and acid peat
413	3	4.1.3 Fagus dominated forest
414	3	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest
415	3	4.1.5 Acidophilous [Quercus]-dominated woodland
416	3	4.1.6 Temperate and boreal and Southern European Betula and Populus tremula forest on mineral soils
417	3	4.1.7 Other broadleaved deciduous forest, excluding highly modified plantations
418	3	4.1.8 Highly modified broadleaved deciduous forests, in particular plantations including stands of non-native trees species that have long been established in European ecosystems stands
420	2	4.2 Coniferous forests
421	3	4.2.1 Boreal and temperate fir and spruce forest
422	3	4.2.2 Mediterranean mountain fir and spruce forest

423	3	4.2.3 Temperate subalpine Larix, Pinus cembra and Pinus uncinata forest
424	3	4.2.4 Pine forest (excluding mires, non-thermophilous)
425	3	4.2.5 Mediterranean thermophilous lowland pine forest
426	3	4.2.6 Spruce, pine and larch mire forests
427	3	4.2.7 Taiga forests
428	3	4.2.8 Other coniferous forests, excluding plantations
429	3	4.2.9 Highly modified coniferous forests, in particular plantations
430	2	4.3 Broadleaved evergreen forest
431	3	4.3.1 Mediterranean evergreen Quercus forest
432	3	4.3.2 Mainland laurophyllous forest
433	3	4.3.3 Macaronesian laurophyllous forest
434	3	4.3.4 Olea europaea-Ceratonia siliqua forest
435	3	4.3.5 Palm groves
436	3	4.3.6 Other broadleaved evergeen forests
437	3	4.3.7 Highly modified broadleaved evergreen forests, in particular plantations (in particular Eucalyptus) including stands of non-native trees species that have long been established in European ecosystems stands
440	2	4.4 Mixed forests
441	3	4.4.1 Mixed forests dominated by coniferous species
442	3	4.4.2 Mixed forests dominated by broadleaved species
443	3	4.4.3 Other mixed forests including stands of non-native trees species that have long been established in European ecosystems stands
450	2	4.5 Transitional forest and woodland shrub
451	3	4.5.1 Transitional woodland/forest land
460	2	4.6 Plantations
461	3	4.6.1 Monoculture plantations
462	3	4.6.2 Mixed plantations
500	1	5. Heathlands and shrub
510	2	5.1 Tundra
511	3	5.1.1 Tundra
520	2	5.2 Heathland and (sub-) alpine shrub

521	3	5.2.1 Arctic alpine, subalpine and lowland shrub and heathland
522	3	5.2.2 Temperate and Mediterranean montane and hilly shrub and heathland
523	3	5.2.3 Temperate and Mediterranean lowland shrub and heathland
530	2	5.3 Sclerophyllous vegetation
531	3	5.3.1 Maquis, arborescent matorral and thermo-Mediterranean shrub
532	3	5.3.2 Garrigue
533	3	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)
534	3	5.3.4 Thermo-Atlantic xerophytic shrub (Madeira and Canary Islands)
600	1	6. Sparsely vegetated ecosystems
610	2	6.1 Bare rocks
611	3	6.1.1 Rocky pavements, outcrops, and screes
612	3	6.1.2 Lava flows
620	2	6.2 Semi-desert, desert and other sparsely vegetated areas
621	3	6.2.1 Semi-desert steppes
622	3	6.2.2 Cool deserts and semi-desert steppes
623	3	6.2.3 Other sparsely vegetated areas
630	2	6.3 Ice sheets, glaciers and perennial snowfields
631	3	6.3.1 Ice sheets, glaciers and perennial snowfields
700	1	7. Inland wetlands
710	2	7.1 Inland marshes on mineral soil
711	3	7.1.1 Reedbeds
712	3	7.1.2 Inland salt marshes
713	3	7.1.3 Other marshland and water-fringing ecosystems
720	2	7.2 Mires, bogs and fens
721	3	7.2.1 Raised bogs
722	3	7.2.2 Blanket bogs
723	3	7.2.3 Valley mires, poor fens and transition mires
724	3	7.2.4 Aapa, palsa and polygon mires
725	3	7.2.5 Base-rich fens and calcareous spring mires

726	3	7.2.6 Peat extraction sites
800	1	8. Rivers and canals
810	2	8.1 Rivers
811	3	8.1.1 Rivers
820	2	8.2 Canals, ditches and drains
821	3	8.2.1 Canals, ditches and drains
900	1	9. Lakes and reservoirs
910	2	9.1 Lakes
911	3	9.1.1 Lakes
920	2	9.2 Artificial reservoirs
921	3	9.2.1 Artificial reservoirs
930	2	9.3 Geothermal pools and wetlands (Iceland)
931	3	9.3.1 Geothermal pools and wetlands (Iceland)
1000	1	10 Marine inlets and transitional waters (lagoons, fjords)
1010	2	10.1 Coastal lagoons
1011	3	10.1.1 Coastal lagoons
1020	2	10.2 Estuaries and bays
1021	3	10.2.1 Estuaries and bays
1030	2	10.3 Intertidal flats
1031	3	10.3.1 Intertidal flats (e.g., Wadden Sea)
1040	2	10.4 Deepwater coastal inlets (fjords)
1041	3	10.4.1 Deepwater coastal inlets (fjords)
1100	1	11 Coastal beaches, dunes and wetlands
1110	2	11.1 Artificial shorelines
1111	3	11.1.1 Artificial shorelines
1120	2	11.2 Coastal dunes, beaches and sandy and muddy shores
1121	3	11.2.1 Coastal dunes
1122	3	11.2.2 Beaches and sandy shores
1123	3	11.2.3 Muddy shores

1130	2	11.3 Rocky shores
1131	3	11.3.1 Coastal shingle
1132	3	11.3.2 Rock cliffs, ledges and shores
1140	2	11.4 Coastal saltmarshes and salines
1141	3	11.4.1 Coastal saltmarshes
1142	3	11.4.2 Salines
1200	1	12 Marine ecosystems
1210	2	12.1 Marine macrophyte habitats
1211	3	12.1.1 Kelp forests
1212	3	12.1.2 Seagrass meadows
1220	2	12.2 Coral reefs
1221	3	12.2.1 Coral reefs
1230	2	12.3 Shellfish beds and reefs
1231	3	12.3.1 Shellfish beds and reefs
1241	3	12.4.1 Subtidal sand beds and mud plains
1250	2	12.5 Subtidal sand beds and mud plains
1251	3	12.5.1 Subtidal rocky substrates
1260	2	12.6 Continental and island slopes
1261	3	12.6.1 Continental and island slopes
1270	2	12.7 Deepwater benthic and pelagic ecosystems
1271	3	12.7.1 Deepwater benthic and pelagic ecosystems
1280	2	12.8 Sea ice
1281	3	12.8.1 Sea ice
2110	3	2.1.10 Other crops (further categories may be added by Member States, depending upon nationally important crop types).
238	3	2.3.8 Other perennial crops and orchards

13.1.2. Peloponnese National to ETA crosswalk

Overview of the dataset that was used to map the Peninsula of Peloponnese using only Copernicus Land Monitoring (CLMS) data crosswalked towards the European Ecosystem Typology (ETA).

Subsection	Description of Subsection	Columns	Description
National Crosswalk	This subsection provides an overview of the Natura 2000 (N2K) classes of Peloponnese crosswalked towards the ETA.	HABITAT4	This is the habitat code extracted from the N2K dataset provided by University of Patras.
		eta_code	Code used for European Typology.
		class_name	The names of each of the classes.
		raster_code	Code used for the European Typology.

13.1.2.1. National Crosswalk

HABITAT4	eta_code	crosswalked_eta_class	raster_code
1010	100	1. Settlements and other artificial areas	12
1011	100	1. Settlements and other artificial areas	12
1012	100	1. Settlements and other artificial areas	12
1013	100	1. Settlements and other artificial areas	12
1020	100	1. Settlements and other artificial areas	12
1021	151	1.5.1 Permanent Greenhouses	29
1022	131	1.3.1 Road and rail networks and associated land	20
1023	131	1.3.1 Road and rail networks and associated land	20
1024	131	1.3.1 Road and rail networks and associated land	20
1025	131	1.3.1 Road and rail networks and associated land	20
1028	131	1.3.1 Road and rail networks and associated land	20
1029	132	1.3.2 Port areas	21

1030	130	1.3 Infrastructure	19
1032	130	1.3 Infrastructure	19
1040	143	1.4.3 Other urban green	28
1041	142	1.4.2 Sports and recreation sites	28
1050	211	2.1.1 Cereals excluding rice (C1000) excluding maize (C1500)	30
1051	251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)	35
1056	210	2.1 Annual cropland	30
1057	251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)	35
1058	221	2.2.1 Rice fields (C2000)	31
1060	232	2.3.2 Grapes (W1000)	34
1061	251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)	35
1062	451	4.5.1 Transitional woodland/forest land	45
1063	238	8.1.1 Rivers	34
1065	429	4.2.9 Highly modified coniferous forests, in particular plantations	42
1066	230	2.3 Permanent crops	32
1067	251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)	35
1130	238	8.1.1 Rivers	34
1150	154	10.1.1 Coastal lagoons	29
1160	154	10.2.1 Estuaries and bays	29
1210	154	11.2.2 Beaches and sandy shores	29
2110	154	11.2.1 Coastal dunes	29
1310	325	3.2.5 Inland salt steppes	39
1410	154	11.4 Coastal saltmarshes and salines	29
1420	154	11.4 Coastal saltmarshes and salines	29
2120	154	11.2.1 Coastal dunes	29
2220	154	11.2.1 Coastal dunes	29
2230	154	11.2.1 Coastal dunes	29
21B0	154	11.2.2 Beaches and sandy shores	29
2240	154	11.2.1 Coastal dunes	29

2250	154	11.2.1 Coastal dunes	29
2260	154	11.2.1 Coastal dunes	29
2270	154	11.2.1 Coastal dunes	29
3150	238	9. Lakes and reservoirs	34
2190	154	11.2 Coastal dunes, beaches and sandy and muddy shores	29
2210	154	11.2 Coastal dunes, beaches and sandy and muddy shores	29
6420	154	11.2 Coastal dunes, beaches and sandy and muddy shores	29
3170	154	11.4 Coastal saltmarshes and salines	29
3190	238	9.1.1 Lakes	34
3280	623	6.2.3 Other sparsely vegetated areas	52
4090	533	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)	48
5150	451	4.5.1 Transitional woodland/forest land	45
5210	531	5.3.1 Maquis, arborescent matorral and thermo-Mediterranean shrub	48
5330	531	5.3.1 Maquis, arborescent matorral and thermo-Mediterranean shrub	48
5340	530	5.3 Sclerophyllous vegetation	48
5420	533	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)	48
5430	533	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)	48
6220	320	3.2 Natural and semi-natural grassland	38
6230	320	3.2 Natural and semi-natural grassland	38
6290	451	4.5.1 Transitional woodland/forest land	45
1240	154	11.3.2 Rock cliffs, ledges and shores	29
7210	238	7.2.5 Base-rich fens and calcareous spring mires	34
7230	238	7.2 Mires, bogs and fens	34
72A0	238	7. Inland wetlands	34
72B0	238	7. Inland wetlands	34
8140	611	6.1.1 Rocky pavements, outcrops, and screes	50
8210	610	6.1 Bare rocks	49
8250	610	6.1 Bare rocks	49
8310	600	6. Sparsely vegetated ecosystems	48

91M0	414	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest	41
9260	414	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest	41
92A0	411	4.1.1 Riparian forest and woodland	41
92C0	411	4.1.1 Riparian forest and woodland	41
92D0	500	5. Heathlands and shrub	46
9320	434	4.3.4 Olea europaea-Ceratonia siliqua forest	43
9340	431	4.3.1 Mediterranean evergreen Quercus forest	43
934A	431	4.3.1 Mediterranean evergreen Quercus forest	43
9350	414	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest	41
951B	422	4.2.2 Mediterranean mountain fir and spruce forest	42
9530	424	4.2.4 Pine Forest (excluding mires, non-thermophilous)	42
9540	425	4.2.5 Mediterranean thermophilous lowland pine forest	42
9560	428	4.2.8 Other coniferous forests, excluding plantations	42
9620	623	6.2.3 Other sparsely vegetated areas	52

````
# R Script – National only approach
# Excel files used as inputs for the scripts are too large to reasonably fit this report and are available upon request to the authors.

---
title: "National only approach"
output: html_notebook
---

```{r Convert n2k data into ETA crosswalked raster}
#Load libraries
library(terra)
library(dplyr)
library(here)
library(readr)
library(raster)
````
````

```
# Read boundary vector data
vector_data <- vect("path/coastline_buffered_10m.gpkg")

# Read the Natura 2000 (n2k) dataset received from Ioannes
n2k_nat <- vect("path/Natura_2000_Habitat_types_Peloponnese.shp")

# Project the n2k dataset to the coastline dataset
n2k_nat <- project(n2k_nat, crs(vector_data))

# what are the column names of n2k_nat:
n2k_nat @ptr[["layer"]]
n2k_nat@ptr[["names"]]


###############Join the new class codes to the existing local component vector layers#################

#Create a data frame of the n2k Peloponnese layer
n2k_nat_df <- as.data.frame(n2k_nat)

# Create a data frame of the reclassification table that needs to be imported
reclass_table <- as.data.frame(read.csv("path/n2k_national_unique_values.csv"))

# Perform the join to map old values to new values
n2k_nat_df_with_reclass <- left_join(n2k_nat_df, reclass_table, by = c("HABITAT4" = "original_values"))

# Update the SpatVector with the new class codes
n2k_nat$ETA <- n2k_nat_df_with_reclass$ETA
n2k_nat$ETA <- as.integer(n2k_nat$ETA)

writeVector(n2k_nat, "path/n2k_nat_reclassified.gpkg", overwrite=TRUE)

################Rasterize vector################

#Load reclassified vector layer
n2k_recl <- vect(here("path/n2k_nat_reclassified.gpkg"),extent=ext(vector_data), layer="n2k_nat_reclassified")

# Load reference raster
reference_raster <- rast("path/boundary_10m.tif")

# Rasterize the vector layer with 10m resolution and snap it to the reference raster grid
n2k_raster_10m <- rasterize(n2k_recl , rast(n2k_recl , resolution = res(reference_raster), crs = crs(reference_raster)), field = "ETA")
```

````
n2k_aligned_raster <- resample(n2k_raster_10m, reference_raster, method = "near")

n2k_aligned_raster <-mask(n2k_aligned_raster,reference_raster)

# Set values to 0 where ua_aligned_raster is NA and reference_raster is 1
n2k_aligned_raster[is.na(n2k_aligned_raster) & reference_raster == 1] <- 0

# Save the raster with 10m resolution
national_data <- writeRaster(n2k_aligned_raster, here("path/n2k_nat_ETA_10m_v2.tif"), filetype = "GTiff", overwrite = TRUE,
datatype='INT2U', gdal=c("COMPRESS=LZW"))

# Reclassify the national dataset to smaller numbers to be able to compare to CLMS only approach

# Define the reclassification rules
reclass_national_data <- c(
  `0` = 0, `100` = 12, `130` = 19, `131` = 20, `132` = 21, `142` = 28, `143` = 70,
  `151` = 29, `210` = 30, `211` = 71, `221` = 31, `230` = 32, `232` = 34,
  `251` = 35, `320` = 38, `325` = 73, `411` = 74, `414` = 75, `422` = 76, `424` = 77,
  `425` = 78, `428` = 79, `429` = 80, `431` = 81, `434` = 82, `451` = 45, `500` = 46,
  `530` = 48, `531` = 83, `533` = 84, `600` = 85, `610` = 49, `611` = 50, `623` = 52,
  `700` = 53, `720` = 86, `725` = 87, `811` = 56, `900` = 58, `911` = 60, `1011` = 88,
  `1021` = 63, `1120` = 65, `1121` = 89, `1122` = 90, `1132` = 67, `1140` = 91)

# Apply the reclassification
national_data_reclassified <- subst(national_data, from = as.numeric(names(reclass_national_data)), to = as.numeric(reclass_national_data),
others = 0)
national_data_reclassified_masked <-mask(national_data_reclassified,reference_raster)
national_data_reclassified_filled <- ifel(is.na(national_data_reclassified_masked) & reference_raster == 1, 0,
national_data_reclassified_masked)

# Save the National only reclassified map
national_data_reclassified_export<-writeRaster(national_data_reclassified_filled, "path/National_only.tif", overwrite = TRUE)

```
````
````

````
# R Script – National and CLMS approach
# Excel files used as inputs for the scripts are too large to reasonably fit this report and are available upon request to the authors.

---
title: "R National and CLMS Approach"
output: html_notebook
---


```{r Fill National classes with level 2 and 3 CLMS and local component}

# Load required library
pacman::p_load(terra, tidyverse, readxl)

# Load the crosswalked national raster layers
national_data <- rast("path/n2k_nat_ETA_10m_v2.tif")
# Load the crosswalked CLMS raster layers
clms_data <- rast("path/CLMS_only.tif")

#### STEP 1 ####

# Reclassify national data that only classes to be mapped at lower level can be considered (to level 2 and 3)

recl_national_data <- matrix(c(100  ,      100    ,
                    130      ,      200    ,
                    131      ,      0      ,
                    132      ,      0      ,
                    142      ,      0      ,
                    143      ,      0      ,
                    151      ,      0      ,
                    210      ,      0      ,
                    211      ,      0      ,
                    221      ,      0      ,
                    230      ,      300    ,
                    232      ,      0      ,
                    251      ,      0      ,
                    320      ,      400    ,
````
````

```
                    325     ,      0      ,
                    411     ,      0      ,
                    414     ,      0      ,
                    422     ,      0      ,
                    424     ,      0      ,
                    425     ,      0      ,
                    428     ,      0      ,
                    429     ,      0      ,
                    431     ,      0      ,
                    434     ,      0      ,
                    451     ,      0      ,
                    500     ,      500    ,
                    530     ,      0      ,
                    531     ,      0      ,
                    533     ,      0      ,
                    600     ,      600    ,
                    610     ,      700    ,
                    611     ,      0      ,
                    623     ,      0      ,
                    700     ,      800    ,
                    720     ,      0      ,
                    725     ,      0      ,
                    811     ,      0      ,
                    900     ,      900    ,
                    911     ,      0      ,
                    1011    ,      0      ,
                    1021    ,      0      ,
                    1120    ,      0      ,
                    1121    ,      0      ,
                    1122    ,      0      ,
                    1132    ,      0      ,
                    1140    ,      1000   ,
                    NA,NA), ncol=2, byrow=TRUE)

recl_national <- classify(national_data, recl_national_data)

unique(recl_national)


# Reclassify CLMS classes that should be considered to map the national classes to a lower level  (to level 2 and 3)

  # Define values to keep
```

```
values_to_keep <- c(
  13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29,
  33, 34, 39, 47, 48, 49, 50, 51, 52, 54, 59, 60, 61, 68, 69
)

# Reclassify raster values
recl_CLMS <- app(
  clms_data,
  function(x) ifelse(x %in% values_to_keep, x, 0)  # Vectorized condition
)


unique(clms_data)


# Sum both reclassified rasters
comb <- recl_national + recl_CLMS

#Check which class combinations to reclassify
unique(values(comb))


# Reclassify only the classes that are useful and can be mapped at level 2 and 3

reclass_rules <- c(
  `113` = 13, `114` = 14, `115` = 15, `116` = 16, `117` = 17, `118` = 18, `119` = 19,
  `120` = 20, `121` = 21, `122` = 22, `123` = 23, `124` = 24, `125` = 25, `126` = 26,
  `127` = 27, `128` = 28, `129` = 29, `220` = 20, `221` = 21, `222` = 22, `223` = 23,
  `224` = 24, `225` = 25, `333` = 33, `334` = 34, `439` = 39, `547` = 47, `548` = 48,
  `649` = 49, `650` = 50, `651` = 51, `652` = 52, `750` = 50, `751` = 51, `752` = 52,
  `854` = 54, `959` = 59, `960` = 60, `961` = 61, `1068` = 68, `1069` = 69
)

# Apply the reclassification. This is the raster dataset that has all the lower levels included (national & CLMS)
raster_reclassified <- subst(comb, from = as.numeric(names(reclass_rules)), to = as.numeric(reclass_rules), others = 0)

sort(unique(raster_reclassified))


#### Step 2 ####
# Fill the above produced layer with remaining national only crosswalked data
```

```
# National only crosswalked to ETA and reclassified dataset
national_data_reclassified_export <- rast("path/National_only.tif")

#Fill the area classified as 0 under raster_reclassified with national_data_reclassified_export
final_raster <- ifel(raster_reclassified == 0, national_data_reclassified_export, raster_reclassified)

#### Step 3 ####
#Final overlay between Step 2 and CLMS data
# As a final step, the above map which has been 1) Filled with CLMS data and 2) national data, now needs to be filled with "CLMS only" data
for the remaining area

#Load the CLMS only map
clms_data <-  rast("path/CLMS_only.tif")

#Fill the final_raster with the CLMS only data
national_and_CLMS <- ifel(final_raster == 0, clms_data, final_raster)

#Load the boundary mask
reference_raster <- rast("path/boundary_10m.tif")

# MAsk national_and_CLMS using the boundary mask
national_and_CLMS <-mask(national_and_CLMS,reference_raster)

# Export the CLMS and National ETA map
national_and_CLMS_map        <-        writeRaster(national_and_CLMS,"path/National_and_CLMS.tif",        overwrite=TRUE,datatype='INT1U',
gdal=c("COMPRESS=LZW"))

```
```
```

13.2. São Miguel Crosswalk

13.2.1. São Miguel CLMS to ETA Crosswalk

Overview of the datasets used to map São Miguel using only Copernicus Land Monitoring (CLMS) data crosswalked to European Ecosystem Typology (ETA). A supporting document with detailed descriptions of data sets can be provided upon request to the authors.

Subsections	Description of subsections	Columns	Description
CLMS Crosswalk	This sheet provides an overview of the combination of CLMS data sets to map São Miguel towards the ETA.	cz	See description in support document.
		clcbb	See description in support document.
		IMD	See description in support document.
		crosswalked_code	These are the codes used for the exported TIF files. The codes are linked to the crosswalked ETA class.
		crosswalked_eta_class	These are the ETA class names linked to the crosswalked_code.
CLMS Typologies	This sheet provides an overview of the typologies and their class names of each of the used CLMS data sets along with their class codes (at different levels)	dataset	Input data set.
		class_level	The thematic class level of each of the data sets (if they are split up into different levels).
		class_name	The names of each of the classes.
		column_name	The column names given to each class at different levels (only relevant for the vector data sets).
		class_code	The class code as numbered in the original dataset.
		original_data_format	The original format of the data that was downloaded from the CLMS portal.
European Ecosystem Typology	This sheet provides an overview of the European Ecosystem Typology at all levels of mapping	code	Code used for the European Typology.
		class_level	The thematic class level of each of the data sets (if they are split up into different levels).
		class_name	The names of each of the classes.

13.2.1.1. CLMS Crosswalk

cz	clcbb	IMD	crosswalked_code	crosswalked_eta_class
11110		76-100	111	1.1.1 Continuous residential area
11110		0-75	121	1.2.1 Discontinuous residential area
11120		76-100	111	1.1.1 Continuous residential area
11120		0-75	121	1.2.1 Discontinuous residential area
11130		76-100	111	1.1.1 Continuous residential area
11130		0-75	121	1.2.1 Discontinuous residential area
11210		76-100	112	1.1.2 Continuous commercial and industrial area
11210		0-75	122	1.2.2 Discontinuous commercial and industrial area
12100			131	1.3.1 Road and rail networks and associated land
12310			132	1.3.2 Port areas
12350			132	1.3.2 Port areas
12360			132	1.3.2 Port areas
12400			133	1.3.3 Airports
13110			135	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
13120			136	1.3.6 Dump areas
13130			137	1.3.7 Construction sites
13200			137	1.3.7 Construction sites
14000			140	1.4 Urban greenspace
21100	6		310	3.1 Sown pastures and fields (modified grassland)
21100	7		210	2.1 Annual cropland
21200			151	1.5.1 Permanent Greenhouses
22100			230	2.3 Permanent crops
23100			251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
23200			251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
23300			251	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)

23400			240	2.4 Agro-forestry areas
31100	2		420	4.2 Coniferous forests
31100	3		410	4.1 Broadleaved deciduous forest
31100	4		430	4.3 Broadleaved evergreen forest
32100	2		420	4.2 Coniferous forests
32100	3		410	4.1 Broadleaved deciduous forest
32100	4		430	4.3 Broadleaved evergreen forest
33100			440	4.4 Mixed forests
34000			451	4.5.1 Transitional woodland/forest land
41000	6		310	3.1 Sown pastures and fields (modified grassland)
41000	7		210	2.1 Annual cropland
42100			320	3.2 Natural and semi-natural grassland
51000			500	5. Heathlands and shrub
61200	9		600	6. Sparsely vegetated ecosystems
62111			1122	11.2.2 Beaches and sandy shores
62112			1131	11.3.1 Coastal shingle
63110			610	6.1 Bare rocks
63120			1132	11.3.2 Rock cliffs, ledges and shores
81100			810	8.1 Rivers
82100			910	9.1 Lakes
82200			921	9.2.1 Artificial reservoirs

R Script – São Miguel CLMS only approach

```
########## STEP 1: CLMS data rasterized ##########
# Three CLMS datasets were used to undertake the CLMS only mapping approach: Coastal Zones 2018 (CZ), Imperviousness 2018 (IMD) and CLCplus Backbone 2018 (CLC)
# The reprojection, clipping and rasterization of the data sets was undertaken in QGIS to the"reference_raster". For the CZ data set, CODE_3_18 = 111 and 112 had to be split up into continuous and discontinuous residential and commercial and industrial area to be able to be mapped according to the ETA.

#### Coastal Zones data set processing:####

#Firstly the CZ vector data set was reprojected and clipped to the "reference_raster". Then the median of IMD was calculated within the cz polygons to determine the sealing percentage, which is needed to discriminate continuous and discontinuous settlement. These classes in column "CODE_3_18"=: 111 and 112, were classified as 100 and then discriminated in QGIS with the following criteria:
#CASE WHEN "CODE_3_18" = 111 AND "_median" > 75 THEN 111 WHEN "CODE_3_18" = 111 AND "_median" <= 75 THEN 121 WHEN "CODE_3_18" = 112 AND "_median" > 75 THEN 112 WHEN "CODE_3_18" = 112 AND "_median" <= 75 THEN 122 ELSE "CODE_5_18" END

#The remaining classes were reclassified as follows:
# WHEN "ETA_IMD" = 12100 THEN 1
#WHEN "ETA_IMD" = 12310 THEN 2
#WHEN "ETA_IMD" = 12350 THEN 3
#WHEN "ETA_IMD" = 12360 THEN 4
#WHEN "ETA_IMD" = 12400 THEN 5
#WHEN "ETA_IMD" = 13110 THEN 6
#WHEN "ETA_IMD" = 13120 THEN 7
#WHEN "ETA_IMD" = 13130 THEN 8
#WHEN "ETA_IMD" = 13200 THEN 9
#WHEN "ETA_IMD" = 14000 THEN 10
#WHEN "ETA_IMD" = 21100 THEN 11
#WHEN "ETA_IMD" = 21200 THEN 12
#WHEN "ETA_IMD" = 22100 THEN 13
#WHEN "ETA_IMD" = 23100 THEN 14
#WHEN "ETA_IMD" = 23200 THEN 15
#WHEN "ETA_IMD" = 23300 THEN 16
#WHEN "ETA_IMD" = 23400 THEN 17
#WHEN "ETA_IMD" = 31100 THEN 18
#WHEN "ETA_IMD" = 32100 THEN 19
#WHEN "ETA_IMD" = 33100 THEN 20
#WHEN "ETA_IMD" = 34000 THEN 21
#WHEN "ETA_IMD" = 41000 THEN 22
```

```
#WHEN "ETA_IMD" = 42100 THEN 23
#WHEN "ETA_IMD" = 51000 THEN 24
#WHEN "ETA_IMD" = 61200 THEN 25
#WHEN "ETA_IMD" = 62111 THEN 26
#WHEN "ETA_IMD" = 62112 THEN 27
#WHEN "ETA_IMD" = 63110 THEN 28
#WHEN "ETA_IMD" = 63120 THEN 29
#WHEN "ETA_IMD" = 81100 THEN 30
#WHEN "ETA_IMD" = 82100 THEN 31
#WHEN "ETA_IMD" = 82200 THEN 32
#WHEN "ETA_IMD" = 84100 THEN 33
#WHEN "ETA_IMD" = 84200 THEN 34


# Thereafter the CZ dataset was rasterized and saved as "path/CLMS_only_data/cz_IMD_raster.tif"

########## STEP 2: Merge datasets to map towards the ETA ##########

# Load required library
pacman::p_load(terra, tidyverse, readxl,rast)

# Set working directory
setwd("path")

#Load datasets

CLC_BB <- rast("path/CLC_plus_BB_2018_5m_sao_miguel_v2.tif")
cz_IMD_raster <- rast("path/CLMS_only_data/cz_IMD_raster.tif")
extent_map <-  rast("path/reference_raster.tif")

# Reclassify CLC+BB to ETA so that it can be used for gap filling in the final step

recl_CLC_to_ETA <- matrix(c(1, 100,
                            2, 420,
                            3, 410,
                            4, 430,
                            5, 0,
                            6, 300,
                            7, 200,
                            9, 600,
                            10, 0,
                            254, NA,
```

```
                            NA,NA), ncol=2, byrow=TRUE)

reclassified_CLC_ETA <- classify(CLC_BB, recl_CLC_to_ETA)

# Crop the second raster using the extent of the first raster
cropped_raster2 <- mask(reclassified_CLC_ETA,extent_map)

writeRaster(cropped_raster2, filename = "CLC_BB_ETA_recla.tif", overwrite = TRUE,datatype='INT2U')

################################################################################################
#Reclassify coniferous and broadleaved forest from the coastal zones dataset to coniferous, deciduous and evergreen forest

#Reclassify cz_IMD_raster so that only broadleaved forest and coniferous forest are considered
recl <- matrix(c(1    ,      0      ,
                 2    ,      0      ,
                 3    ,      0      ,
                 4    ,      0      ,
                 5    ,      0      ,
                 6    ,      0      ,
                 7    ,      0      ,
                 8    ,      0      ,
                 9    ,      0      ,
                 10   ,      0      ,
                 11   ,      0      ,
                 12   ,      0      ,
                 13   ,      0      ,
                 14   ,      0      ,
                 15   ,      0      ,
                 16   ,      0      ,
                 17   ,      0      ,
                 18   ,      1      ,
                 19   ,      2      ,
                 20   ,      0      ,
                 21   ,      0      ,
                 22   ,      0      ,
                 23   ,      0      ,
                 24   ,      0      ,
                 25   ,      0      ,
                 26   ,      0      ,
                 27   ,      0      ,
                 28   ,      0      ,
                 29   ,      0      ,
```

```
                30   ,      0      ,
                31   ,      0      ,
                32   ,      0      ,
                34   ,      0      ,
                100  ,      0      ,
                111  ,      0      ,
                112  ,      0      ,
                121  ,      0      ,
                122  ,      0      ,
                NA,NA), ncol=2, byrow=TRUE)

reclassified_ETA_CZ_IMB <- classify(cz_IMD_raster, recl)


#Reclassify forest classes of CLCBB
recl_CLCBB <- matrix(c(1     ,      0      ,
                      2     ,      10     ,
                      3     ,      20     ,
                      4     ,      30     ,
                      5     ,      0      ,
                      6     ,      0      ,
                      7     ,      0      ,
                      9     ,      0      ,
                      10    ,      0      ,
                      254   ,      0      ,

                      NA,NA), ncol=2, byrow=TRUE)

reclassified_CLCBB <- classify(CLC_BB, recl_CLCBB)

#Sum up both layers to determine overlaps
combined <- reclassified_ETA_CZ_IMB + reclassified_CLCBB

unique(combined)

# Reclassify the combined classes of forest CLC+BB and the broadleaved (31100) and coniferous (32100) classes of the CZ which are reclassified based on CLC+BB classes

recl_comb <- matrix(c(0,0,
                     1,410,
                     2,420,
                     10,0,
```

```
                    20,0,
                    30,0,
                    11,420,
                    21,410,
                    31,430,
                    12,420,
                    22,410,
                    32,430,
                    NA,NA), ncol=2, byrow=TRUE)


#Use as input for final map#### Reclassify the raster to the ETA codes for forest classes
comb_to_eta_recl <- classify(combined, recl_comb)



setwd("N:/C2206_SELINA/Sao_Miguel/f02_secondary_data/CLMS_only_data")

writeRaster(comb_to_eta_recl, filename = "Forest_recl_with_original.tif", overwrite = TRUE,datatype='INT2U')



################################################################################################
#Reclassify grassland and arable land from the cz_IMD_raster

#Reclassify cz_IMD_raster so that only cropland and grassland are considered
recl_grass_arable <- matrix(c(1     ,      0      ,
                            2     ,      0      ,
                            3     ,      0      ,
                            4     ,      0      ,
                            5     ,      0      ,
                            6     ,      0      ,
                            7     ,      0      ,
                            8     ,      0      ,
                            9     ,      0      ,
                            10    ,      0      ,
                            11    ,      10     ,
                            12    ,      0      ,
                            13    ,      0      ,
                            14    ,      0      ,
                            15    ,      0      ,
                            16    ,      0      ,
```

```
                              17    ,      0      ,
                              18    ,      0      ,
                              19    ,      0      ,
                              20    ,      0      ,
                              21    ,      0      ,
                              22    ,      20     ,
                              23    ,      0      ,
                              24    ,      0      ,
                              25    ,      0      ,
                              26    ,      0      ,
                              27    ,      0      ,
                              28    ,      0      ,
                              29    ,      0      ,
                              30    ,      0      ,
                              31    ,      0      ,
                              32    ,      0      ,
                              34    ,      0      ,
                              100   ,      0      ,
                              111   ,      0      ,
                              112   ,      0      ,
                              121   ,      0      ,
                              122   ,      0      ,
                              NA,NA), ncol=2, byrow=TRUE)

reclassifiedcz_gr_ara <- classify(cz_IMD_raster, recl_grass_arable)

#Reclassify CLCBB to select only arable land and grassland
recl_CLCBB_ara_gra <- matrix(c(1    ,      0      ,
                               2    ,      0      ,
                               3    ,      0      ,
                               4    ,      0      ,
                               5    ,      0      ,
                               6    ,      1      ,
                               7    ,      2      ,
                               9    ,      0      ,
                               10   ,      0      ,
                               254  ,      0      ,

                               NA,NA), ncol=2, byrow=TRUE)


CLCBB_arable_grass <- classify(CLC_BB, recl_CLCBB_ara_gra)
```

```
#Sum up reclassified CLC+BB where arable land and grassland are reclassified with reclassified coastal zone layer where arable land and
grassland are reclassified

combined_gra_ara <- reclassifiedcz_gr_ara  + CLCBB_arable_grass

recl_combined_gra_ara <- matrix(c(0 , 0,
                                  1 ,      0      ,
                                  2 ,      0      ,
                                  10       ,      210    ,
                                  11       ,      310    ,
                                  12       ,      210    ,
                                  20       ,      310    ,
                                  21       ,      310    ,
                                  22       ,      210    ,

                                  NA,NA), ncol=2, byrow=TRUE)

###Use as input for final map

#Reclassify the reclassified cz based on CLC+BB to correct grassland and arable land and keep the cz layers where no CLC+BB exists - keep
the original codes
cz_arable_grass_CLC_BB <- classify(combined_gra_ara, recl_combined_gra_ara)

recl_combined_gra_ara_orig <- matrix(c(0 , 0,
                                       1   ,      0      ,
                                       2   ,      0      ,
                                       10  ,      10     ,
                                       11  ,      310    ,
                                       12  ,      210    ,
                                       20  ,      20     ,
                                       21  ,      310    ,
                                       22  ,      210    ,

                                       NA,NA), ncol=2, byrow=TRUE)

#Reclassify the reclassified cz based on CLC+BB to correct grassland and arable land and keep the cz layers where no CLC+BB exists
cz_arable_grass_CLC_BB_orig <- classify(combined_gra_ara, recl_combined_gra_ara_orig)

###############################################################################################
#Reclassify sparsely vegetated area
```

```
#Reclassify cz_IMD_raster so that only cropland and grassland are considered
recl_spars_veg <- matrix(c(1 ,      0       ,
                           2 ,      0       ,
                           3 ,      0       ,
                           4 ,      0       ,
                           5 ,      0       ,
                           6 ,      0       ,
                           7 ,      0       ,
                           8 ,      0       ,
                           9 ,      0       ,
                           10       ,       0       ,
                           11       ,       0       ,
                           12       ,       0       ,
                           13       ,       0       ,
                           14       ,       0       ,
                           15       ,       0       ,
                           16       ,       0       ,
                           17       ,       0       ,
                           18       ,       0       ,
                           19       ,       0       ,
                           20       ,       0       ,
                           21       ,       0       ,
                           22       ,       0       ,
                           23       ,       0       ,
                           24       ,       0       ,
                           25       ,       10      ,
                           26       ,       0       ,
                           27       ,       0       ,
                           28       ,       0       ,
                           29       ,       0       ,
                           30       ,       0       ,
                           31       ,       0       ,
                           32       ,       0       ,
                           34       ,       0       ,
                           100      ,       0       ,
                           111      ,       0       ,
                           112      ,       0       ,
                           121      ,       0       ,
                           122      ,       0       ,
                           NA,NA), ncol=2, byrow=TRUE)

reclassified_cz_spars_veg <- classify(cz_IMD_raster, recl_spars_veg)
```

```
#Reclassify CLCBB to select only sparsely vegetated
recl_CLCBB_spars_veg <- matrix(c(1  ,      0      ,
                                 2  ,      0      ,
                                 3  ,      0      ,
                                 4  ,      0      ,
                                 5  ,      0      ,
                                 6  ,      0      ,
                                 7  ,      0      ,
                                 9  ,      1      ,
                                 10 ,      0      ,
                                 254       ,      0      ,

                                 NA,NA), ncol=2, byrow=TRUE)


CLCBB_spars_veg <- classify(CLC_BB, recl_CLCBB_spars_veg)

#Sum up reclassified CLC+BB where sparsely vegetated are reclassified with reclassified coastal zone layer where sparsely vegetated are reclassified

combined_spars_veg <- reclassified_cz_spars_veg  + CLCBB_spars_veg

recl_combined_spars_veg <- matrix(c(0 , 0,
                                    1      ,      0      ,
                                    10     ,      61200,
                                    11     ,      600    ,
                                    NA,NA), ncol=2, byrow=TRUE)

###Use as input for final map
####Reclassify the reclassified cz based on CLC+BB to correct sparsely vegetated and keep the cz layers where no CLC+BB exists - keep the original codes
cz_spars_veg_CLC_BB <- classify(combined_spars_veg, recl_combined_spars_veg)

recl_combined_spars_veg_orig <- matrix(c(0 , 0,
                                         1 ,      0      ,
                                         10       ,      10     ,
                                         11       ,      600    ,
                                         NA,NA), ncol=2, byrow=TRUE)

#Reclassify the reclassified cz based on CLC+BB to correct sparsely vegetated and keep the cz layers where no CLC+BB exists
cz_spars_veg_CLC_BB_orig <- classify(combined_spars_veg, recl_combined_spars_veg_orig)
```

```
######################################################################################################
#Reclassify urban fabric


recl_urban_fab <- matrix(c(1 ,      0       ,
                           2 ,      0       ,
                           3 ,      0       ,
                           4 ,      0       ,
                           5 ,      0       ,
                           6 ,      0       ,
                           7 ,      0       ,
                           8 ,      0       ,
                           9 ,      0       ,
                           10       ,       0       ,
                           11       ,       0       ,
                           12       ,       0       ,
                           13       ,       0       ,
                           14       ,       0       ,
                           15       ,       0       ,
                           16       ,       0       ,
                           17       ,       0       ,
                           18       ,       0       ,
                           19       ,       0       ,
                           20       ,       0       ,
                           21       ,       0       ,
                           22       ,       0       ,
                           23       ,       0       ,
                           24       ,       0       ,
                           25       ,       0       ,
                           26       ,       0       ,
                           27       ,       0       ,
                           28       ,       0       ,
                           29       ,       0       ,
                           30       ,       0       ,
                           31       ,       0       ,
                           32       ,       0       ,
                           34       ,       0       ,
                           100      ,       100     ,
                           111      ,       111     ,
                           112      ,       112     ,
```

```
                           121      ,       121     ,
                           122      ,       122     ,
                           NA,NA), ncol=2, byrow=TRUE)

reclassified_cz_urban_fab <- classify(cz_IMD_raster, recl_urban_fab)

################################################################################################
#Reclassify remaining cz classes to ETA

recl_remain <- matrix(c(1    ,       131    ,
                        2    ,       132    ,
                        3    ,       132    ,
                        4    ,       132    ,
                        5    ,       133    ,
                        6    ,       135    ,
                        7    ,       136    ,
                        8    ,       137    ,
                        9    ,       100    ,
                        10   ,       140    ,
                        11   ,       0      ,
                        12   ,       151    ,
                        13   ,       230    ,
                        14   ,       251    ,
                        15   ,       251    ,
                        16   ,       251    ,
                        17   ,       240    ,
                        18   ,       0      ,
                        19   ,       0      ,
                        20   ,       440    ,
                        21   ,       451    ,
                        22   ,       0      ,
                        23   ,       320    ,
                        24   ,       500    ,
                        25   ,       0      ,
                        26   ,       1122   ,
                        27   ,       1131   ,
                        28   ,       611    ,
                        29   ,       1132   ,
                        30   ,       800    ,
                        31   ,       911    ,
                        32   ,       921    ,
                        33   ,       0      ,
```

```
                        34   ,      0      ,
                        100  ,      0      ,
                        111  ,      0      ,
                        112  ,      0      ,
                        121  ,      0      ,
                        122  ,      0      ,
                        NA,NA), ncol=2, byrow=TRUE)

reclassified_cz_remain <- classify(cz_IMD_raster, recl_remain)


################################################################################################
#Add all reclassified layers together


final_map <- comb_to_eta_recl + cz_arable_grass_CLC_BB + cz_spars_veg_CLC_BB + reclassified_cz_urban_fab + reclassified_cz_remain


################################################################################################
#Fill the area in the map that has not been mapped under the coastal zones layer using CLC+BB (translated to ETA) which has been cropped to
the reference_raster so that the CLMS & national map can be compared
x <-cover(final_map,cropped_raster2)

# replace all values that are 0 with the CLC+BB (translated to ETA)
x[x == 0] <- cropped_raster2[x == 0]

writeRaster(x, filename = "Final_CLMS_based_ETA_map_cz_filled_no_zero.tif", overwrite = TRUE,datatype='INT2U')
```

13.2.1.2. CLMS Typologies

dataset	class_level	class_name	column_name	class_code	original_data_format
cz	1	1 Urban	CODE_1_18	10000	vector
cz	1	2 Cropland	CODE_1_18	20000	vector
cz	1	3 Woodland and forest	CODE_1_18	30000	vector
cz	1	4 Grassland	CODE_1_18	40000	vector
cz	1	5 Heathland and scrub	CODE_1_18	50000	vector
cz	1	6 Open spaces with little or no vegetation	CODE_1_18	60000	vector
cz	1	7 Wetland	CODE_1_18	70000	vector
cz	1	8 Water	CODE_1_18	80000	vector
cz	2	1.1 Urban fabric, industrial, commercial, public, military and private units	CODE_2_18	11000	vector
cz	2	1.2 Transport infrastructure	CODE_2_18	12000	vector
cz	2	1.3 Mineral extraction, dump and construction sites, land without current use	CODE_2_18	13000	vector
cz	2	1.4 Green urban, sports and leisure facilities	CODE_2_18	14000	vector
cz	2	2.1 Arable land	CODE_2_18	21000	vector
cz	2	2.2 Permanent crops	CODE_2_18	22000	vector
cz	2	2.3 Heterogeneous agricultural area	CODE_2_18	23000	vector
cz	2	3.1 Broadleaved forest	CODE_2_18	31000	vector
cz	2	3.2 Coniferous forest	CODE_2_18	32000	vector
cz	2	3.3 Mixed forest	CODE_2_18	33000	vector
cz	2	3.4 Transitional woodland and scrub	CODE_2_18	34000	vector
cz	2	3.5 Lines of trees and scrub	CODE_2_18	35000	vector
cz	2	3.6 Damaged forest	CODE_2_18	36000	vector
cz	2	4.1 Managed grassland	CODE_2_18	41000	vector
cz	2	4.2 Natural & semi-natural grassland	CODE_2_18	42000	vector
cz	2	5.1 Heathland and moorland	CODE_2_18	51000	vector
cz	2	5.2 Alpine scrub land	CODE_2_18	52000	vector

cz	2	5.3 Sclerophyllous scrubs	CODE_2_18	53000	vector
cz	2	6.1 Sparsely vegetated areas	CODE_2_18	61000	vector
cz	2	6.2 Beaches, dunes, river banks	CODE_2_18	62000	vector
cz	2	6.3 Bare rocks, burnt areas, glaciers and perpetual snow	CODE_2_18	63000	vector
cz	2	7.1 Inland wetlands	CODE_2_18	71000	vector
cz	2	7.2 Coastal wetlands	CODE_2_18	72000	vector
cz	2	8.1 Water courses	CODE_2_18	81000	vector
cz	2	8.2 Lakes and reservoirs	CODE_2_18	82000	vector
cz	2	8.3 Transitional waters	CODE_2_18	83000	vector
cz	2	8.4 Sea and ocean	CODE_2_18	84000	vector
cz	3	1.1.1 Urban fabric	CODE_3_18	11100	vector
cz	3	1.1.2 Industrial, commercial, public and military units	CODE_3_18	11200	vector
cz	3	1.2.1 Road networks and associated land	CODE_3_18	12100	vector
cz	3	1.2.2 Railways and associated land	CODE_3_18	12200	vector
cz	3	1.2.3 Port areas and associated land	CODE_3_18	12300	vector
cz	3	1.2.4 Airports and associated land	CODE_3_18	12400	vector
cz	3	1.3.1 Mineral extraction, dump and construction sites	CODE_3_18	13100	vector
cz	3	1.3.2 Land without current use	CODE_3_18	13200	vector
cz	3	2.1.1 Arable irrigated and non-irrigated land	CODE_3_18	21100	vector
cz	3	2.1.2 Greenhouses	CODE_3_18	21200	vector
cz	3	2.2.1 Vineyards, fruit trees and berry plantations	CODE_3_18	22100	vector
cz	3	2.2.2 Olive groves	CODE_3_18	22200	vector
cz	3	2.3.1 Annual crops associated with permanent crops	CODE_3_18	23100	vector
cz	3	2.3.2 Complex cultivation patterns	CODE_3_18	23200	vector
cz	3	2.3.3 Land principally occupied by agriculture with significant areas of natural vegetation	CODE_3_18	23300	vector
cz	3	2.3.4 Agro-forestry	CODE_3_18	23400	vector
cz	3	3.1.1 Natural & semi-natural broadleaved forest	CODE_3_18	31100	vector
cz	3	3.1.2 Highly artificial broadleaved plantations	CODE_3_18	31200	vector
cz	3	3.2.1 Natural & semi-natural coniferous forest	CODE_3_18	32100	vector

cz	3	3.2.2 Highly artificial coniferous plantations	CODE_3_18	32200	vector
cz	3	3.3.1 Natural & semi-natural mixed forest	CODE_3_18	33100	vector
cz	3	3.3.2 Highly artificial mixed plantations	CODE_3_18	33200	vector
cz	3	4.2.1 Semi-natural grassland	CODE_3_18	42100	vector
cz	3	4.2.2 Alpine and sub-alpine natural grassland	CODE_3_18	42200	vector
cz	3	6.1.1 Sparse vegetation on sands	CODE_3_18	61100	vector
cz	3	6.1.2 Sparse vegetation on rocks	CODE_3_18	61200	vector
cz	3	6.2.1 Beaches and dunes	CODE_3_18	62100	vector
cz	3	6.2.2 River banks	CODE_3_18	62200	vector
cz	3	6.3.1 Bare rocks, outcrops, cliffs	CODE_3_18	63100	vector
cz	3	6.3.2 Burnt areas (except burnt forest)	CODE_3_18	63200	vector
cz	3	6.3.3 Glaciers and perpetual snow	CODE_3_18	63300	vector
cz	3	7.1.1 Inland marshes	CODE_3_18	71100	vector
cz	3	7.1.2 Peat bogs	CODE_3_18	71200	vector
cz	3	7.2.1 Salt marshes	CODE_3_18	72100	vector
cz	3	7.2.2 Salines	CODE_3_18	72200	vector
cz	3	7.2.3 Intertidal flats	CODE_3_18	72300	vector
cz	3	8.1.1 Natural & semi-natural water courses	CODE_3_18	81100	vector
cz	3	8.1.2 Highly modified water courses and canals	CODE_3_18	81200	vector
cz	3	8.1.3 Seasonally connected water courses (oxbows)	CODE_3_18	81300	vector
cz	3	8.2.1 Natural lakes	CODE_3_18	82100	vector
cz	3	8.2.2 Reservoirs	CODE_3_18	82200	vector
cz	3	8.2.3 Aquaculture ponds	CODE_3_18	82300	vector
cz	3	8.2.4 Standing water bodies of extractive industrial sites	CODE_3_18	82400	vector
cz	3	8.3.1 Lagoons	CODE_3_18	83100	vector
cz	3	8.3.2 Estuaries	CODE_3_18	83200	vector
cz	3	8.3.3 Marine inlets and fjords	CODE_3_18	83300	vector
cz	3	8.4.1 Open sea	CODE_3_18	84100	vector
cz	3	8.4.2 Coastal waters	CODE_3_18	84200	vector

cz	4	1.1.1.1 Continuous urban fabric (IMD ≥80%)	CODE_4_18	11110	vector
cz	4	1.1.1.2 Dense urban fabric (IMD ≥30-80%)	CODE_4_18	11120	vector
cz	4	1.1.1.3 Low density fabric (IMD <30%)	CODE_4_18	11130	vector
cz	4	1.1.2.1 Industrial, commercial, public and military units (other)	CODE_4_18	11210	vector
cz	4	1.1.2.2 Nuclear energy plants and associated land	CODE_4_18	11220	vector
cz	4	1.2.3.1 Cargo port	CODE_4_18	12310	vector
cz	4	1.2.3.2 Passenger port	CODE_4_18	12320	vector
cz	4	1.2.3.3 Fishing port	CODE_4_18	12330	vector
cz	4	1.2.3.4 Naval port	CODE_4_18	12340	vector
cz	4	1.2.3.5 Marinas	CODE_4_18	12350	vector
cz	4	1.2.3.6 Local multi-functional harbours	CODE_4_18	12360	vector
cz	4	1.2.3.7 Shipyards	CODE_4_18	12370	vector
cz	4	1.3.1.1 Mineral extraction sites	CODE_4_18	13110	vector
cz	4	1.3.1.2 Dump sites	CODE_4_18	13120	vector
cz	4	1.3.1.3 Construction sites	CODE_4_18	13130	vector
cz	4	6.2.1.1 Beaches	CODE_4_18	62110	vector
cz	4	6.2.1.2 Dunes	CODE_4_18	62120	vector
cz	4	6.3.1.1 Bare rocks and outcrops	CODE_4_18	63110	vector
cz	4	6.3.1.2 Coastal cliffs	CODE_4_18	63120	vector
cz	4	7.1.2.1 Exploited peat bogs	CODE_4_18	71210	vector
cz	4	7.1.2.2 Unexploited peat bogs	CODE_4_18	71220	vector
cz	5	6.2.1.1.1 Sandy beaches	CODE_5_18	62111	vector
cz	5	6.2.1.1.2 Shingle beaches	CODE_5_18	62112	vector
clcbb	-	1: Sealed		1	raster
clcbb	-	2: Woody – needle leaved trees		2	raster
clcbb	-	3: Woody – Broadleaved deciduous trees		3	raster
clcbb	-	4: Woody – Broadleaved evergreen trees		4	raster
clcbb	-	5: Low-growing woody plants (bushes, shrubs)		5	raster
clcbb	-	6: Permanent herbaceous		6	raster

clcbb	-	7: Periodically herbaceous		7	raster
clcbb	-	8: Lichens and mosses		8	raster
clcbb	-	9: Non- and sparsely-vegetated		9	raster
clcbb	-	10: Water		10	raster
clcbb	-	11: Snow and ice		11	raster

13.2.1.3. European Ecosystem Typology

Code	class_level	class_name
100	1	1. Settlements and other artificial areas
110	2	1.1 Continuous settlement area
111	3	1.1.1 Continuous residential area
112	3	1.1.2 Continuous commercial and industrial area
120	2	1.2 Discontinuous settlement area
121	3	1.2.1 Discontinuous residential area
122	3	1.2.2 Discontinuous commercial and industrial area
130	2	1.3 Infrastructure
131	3	1.3.1 Road and rail networks and associated land
132	3	1.3.2 Port areas
133	3	1.3.3 Airports
134	3	1.3.4 Other infrastructure
135	3	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
136	3	1.3.6 Dump areas
137	3	1.3.7 Construction sites
140	2	1.4 Urban greenspace
141	3	1.4.1 Parks (including Zoos and botanical gardens)
142	3	1.4.2 Sports and recreation sites
143	3	1.4.3 Other urban green
150	2	1.5 Other artificial areas

151	3	1.5.1 Permanent Greenhouses
152	3	1.5.2 Cemeteries
153	3	1.5.3 Archaeological sites
154	3	1.5.4 Urban blue
200	1	2. Cropland
210	2	2.1 Annual cropland
211	3	2.1.1 Cereals excluding rice (C1000) excluding maize (C1500)
212	3	2.1.2 Maize (C1500 + G3000)
213	3	2.1.3 Dry pulses and protein crops (P0000)
214	3	2.1.4 Root crops, like sugar beet and potatoes (R0000)
215	3	2.1.5 Vegetables (including melons) and strawberries (V0000_S0000)
216	3	2.1.6 Industrial crops including annual bioenergy crops (I0000)
217	3	2.1.7 Flowers and ornamental plants (N0000)
218	3	2.1.8 Fallow land (Q0000)
219	3	2.1.9 Temporary grasses and grazing areas (G1000)
220	2	2.2 Rice fields
221	3	2.2.1 Rice fields (C2000)
230	2	2.3 Permanent crops
231	3	2.3.1 Olives (O1000)
232	3	2.3.2 Grapes (W1000)
233	3	2.3.3 Pome fruits (F1100)
234	3	2.3.4 Stone fruits (F1200)
235	3	2.3.5 Berries excluding strawberries (F3000)
236	3	2.3.6 Citrus fruits (T1000)
237	3	2.3.7 Nuts (F4000)
238	3	2.3.8 Hazelnut
238	3	2.3.8 Other perennial crops and orchards
239	3	2.3.9 Chestnut
240	2	2.4 Agro-forestry areas
241	3	2.4.1 Holm and cork oak forests

242	3	2.4.2 Other agro-forestry area
250	2	2.5 Mixed farmland
251	3	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
260	2	2.6 Other farmland
261	3	2.6.1 Nurseries
262	3	2.6.2 Christmas tree plantations
263	3	2.6.3 Perennial bioenergy crops
264	3	2.6.4 Field margins
300	1	3. Grassland (pastures, semi-natural and natural grasslands)
310	2	3.1 Sown pastures and fields (modified grassland)
311	3	3.1.1 Sown pastures used for grazing
312	3	3.1.2 Sown grassland mown frequently for fodder or silage
320	2	3.2 Natural and semi-natural grassland
321	3	3.2.1 Dry grassland
322	3	3.2.2 Seasonally wet and wet grassland
323	3	3.2.3 Alpine and subalpine grasslands
324	3	3.2.4 Woodland fringes and clearings and tall forb stands
325	3	3.2.5 Inland salt steppes
326	3	3.2.6 Sparsely wooded grasslands
327	3	3.2.7 Mesophilous extensive grassland
400	1	4. Forest and woodlands
410	2	4.1 Broadleaved deciduous forest
411	3	4.1.1 Riparian forest and woodland
412	3	4.1.2 Broadleaved swamp, woodland on non-acid and acid peat
413	3	4.1.3 Fagus dominated forest
414	3	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest
415	3	4.1.5 Acidophilous [*Quercus*]-dominated woodland
416	3	4.1.6 Temperate and boreal and Southern European Betula and Populus tremula forest on mineral soils
417	3	4.1.7 Other broadleaved deciduous forest, excluding highly modified plantations
418	3	4.1.8 Highly modified broadleaved deciduous forests, in particular plantations including stands of non-native trees species that have long been established in European ecosystems stands

420	2	4.2 Coniferous forests
421	3	4.2.1 Boreal and temperate fir and spruce forest
422	3	4.2.2 Mediterranean mountain fir and spruce forest
423	3	4.2.3 Temperate subalpine Larix, Pinus cembra and Pinus uncinata forest
424	3	4.2.4 Pine forest (excluding mires, non-thermophilous)
425	3	4.2.5 Mediterranean thermophilous lowland pine forest
426	3	4.2.6 Spruce, pine and larch mire forests
427	3	4.2.7 Taiga forests
428	3	4.2.8 Other coniferous forests, excluding plantations
429	3	4.2.9 Highly modified coniferous forests, in particular plantations
430	2	4.3 Broadleaved evergreen forest
431	3	4.3.1 Mediterranean evergreen *Quercus* Forest
432	3	4.3.2 Mainland laurophyllous forest
433	3	4.3.3 Macaronesian laurophyllous forest
434	3	4.3.4 Olea europaea-Ceratonia siliqua forest
435	3	4.3.5 Palm groves
436	3	4.3.6 Other broadleaved evergeen forests
437	3	4.3.7 Highly modified broadleaved evergreen forests, in particular plantations (in particular *Eucalyptus*) including stands of non-native trees species that have long been established in European ecosystems stands
440	2	4.4 Mixed forests
441	3	4.4.1 Mixed forests dominated by coniferous species
442	3	4.4.2 Mixed forests dominated by broadleaved species
443	3	4.4.3 Other mixed forests including stands of non-native trees species that have long been established in European ecosystems stands
450	2	4.5 Transitional forest and woodland shrub
451	3	4.5.1 Transitional woodland/forest land
460	2	4.6 Plantations
461	3	4.6.1 Monoculture plantations
462	3	4.6.2 Mixed plantations
500	1	5. Heathlands and shrub
510	2	5.1 Tundra
511	3	5.1.1 Tundra

520	2	5.2 Heathland and (sub-) alpine shrub
521	3	5.2.1 Arctic alpine, subalpine and lowland shrub and heathland
522	3	5.2.2 Temperate and Mediterranean montane and hilly shrub and heathland
523	3	5.2.3 Temperate and Mediterranean lowland shrub and heathland
530	2	5.3 Sclerophyllous vegetation
531	3	5.3.1 Maquis, arborescent matorral and thermo-Mediterranean shrub
532	3	5.3.2 Garrigue
533	3	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)
534	3	5.3.4 Thermo-Atlantic xerophytic shrub (Madeira and Canary Islands)
600	1	6. Sparsely vegetated ecosystems
610	2	6.1 Bare rocks
611	3	6.1.1 Rocky pavements, outcrops, and screes
612	3	6.1.2 Lava flows
620	2	6.2 Semi-desert, desert and other sparsely vegetated areas
621	3	6.2.1 Semi-desert steppes
622	3	6.2.2 Cool deserts and semi-desert steppes
623	3	6.2.3 Other sparsely vegetated areas
630	2	6.3 Ice sheets, glaciers and perennial snowfields
631	3	6.3.1 Ice sheets, glaciers and perennial snowfields
700	1	7. Inland wetlands
710	2	7.1 Inland marshes on mineral soil
711	3	7.1.1 Reedbeds
712	3	7.1.2 Inland salt marshes
713	3	7.1.3 Other marshland and water-fringing ecosystems
720	2	7.2 Mires, bogs and fens
721	3	7.2.1 Raised bogs
722	3	7.2.2 Blanket bogs
723	3	7.2.3 Valley mires, poor fens and transition mires
724	3	7.2.4 Aapa, palsa and polygon mires
725	3	7.2.5 Base-rich fens and calcareous spring mires

726	3	7.2.6 Peat extraction sites
800	1	8. Rivers and canals
810	2	8.1 Rivers
811	3	8.1.1 Rivers
820	2	8.2 Canals, ditches and drains
821	3	8.2.1 Canals, ditches and drains
900	1	9. Lakes and reservoirs
910	2	9.1 Lakes
911	3	9.1.1 Lakes
920	2	9.2 Artificial reservoirs
921	3	9.2.1 Artificial reservoirs
930	2	9.3 Geothermal pools and wetlands (Iceland)
931	3	9.3.1 Geothermal pools and wetlands (Iceland)
1000	1	10 Marine inlets and transitional waters (lagoons, fjords)
1010	2	10.0 Coastal lagoons
1011	3	10.1.1 Coastal lagoons
1020	2	10.2 Estuaries and bays
1021	3	10.2.1 Estuaries and bays
1030	2	10.3 Intertidal flats
1031	3	10.3.1 Intertidal flats (e.g., Wadden Sea)
1040	2	10.4 Deepwater coastal inlets (fjords)
1041	3	10.4.1 Deepwater coastal inlets (fjords)
1100	1	11 Coastal beaches, dunes and wetlands
1110	2	11.1 Artificial shorelines
1111	3	11.1.1 Artificial shorelines
1120	2	11.2 Coastal dunes, beaches and sandy and muddy shores
1121	3	11.2.1 Coastal dunes
1122	3	11.2.2 Beaches and sandy shores
1123	3	11.2.3 Muddy shores
1130	2	11.3 Rocky shores

1131	3	11.3.1 Coastal shingle
1132	3	11.3.2 Rock cliffs, ledges and shores
1140	2	11.4 Coastal saltmarshes and salines
1141	3	11.4.1 Coastal saltmarshes
1142	3	11.4.2 Salines
1200	1	12 Marine ecosystems
1210	2	12.1 Marine macrophyte habitats
1211	3	12.1.1 Kelp forests
1212	3	12.1.2 Seagrass meadows
1220	2	12.2 Coral reefs
1221	3	12.2.1 Coral reefs
1230	2	12.3 Shellfish beds and reefs
1231	3	12.3.1 Shellfish beds and reefs
1241	3	12.4.1 Subtidal sand beds and mud plains
1250	2	12.5 Subtidal sand beds and mud plains
1251	3	12.5.1 Subtidal rocky substrates
1260	2	12.6 Continental and island slopes
1261	3	12.6.1 Continental and island slopes
1270	2	12.7 Deepwater benthic and pelagic ecosystems
1271	3	12.7.1 Deepwater benthic and pelagic ecosystems
1280	2	12.8 Sea ice
1281	3	12.8.1 Sea ice
2110	3	2.1.10 Other crops (further categories may be added by Member States, depending upon nationally important crop types).

13.2.2. São Miguel National to ETA Crosswalk

Overview of the datasets used to map São Miguel using only Land Use 2018 (LU) data crosswalked to the European Ecosystem Typology (ETA).

13.2.2.1. Land Use to ETA

Subsections	Description of Subsections	Columns	Description
Land Use Crosswalk	This section provides an overview of the Land Use 2018 (LU) classes of São Miguel crosswalked towards the ETA.	codigo	Code of the LU data set
		designacao_translation	Translation of the class name of the LU data set
		eta_code	This is the crosswalked ETA code
		crosswalked_eta_class	This is the crosswalked ETA name
Land Use Translation	The National Land Use 2018 (LU) data set was translated from Portuguese to English	codigo	Code of the LU data set
		designacao	Class name of the LU data set
		designacao_translation	Translation of the class name of the LU data set
European Ecosystem Typology	This section provides an overview of the European Ecosystem Typology at all levels of mapping	Code	Code used for the European Typology
		class_level	The thematic class level of each of the data sets (if they are split up into different levels)
		class_name	The names of each of the classes

13.2.2.2. Land Use Crosswalk

codigo	designacao_translation	eta_code	crosswalked_eta_class
111	Continuous urban fabric	111	1.1.1 Continuous residential area
112	Discontinuous urban fabric	121	1.2.1 Discontinuous residential area
121	Industry, commerce, general equipment and infrastructure	100	1. Settlements and other artificial areas
122	Road networks and associated spaces	131	1.3.1 Road and rail networks and associated land
123	Port areas	132	1.3.2 Port areas
124	Airports and aerodromes	133	1.3.3 Airports
131	Areas of Mineral Extraction	135	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
132	Waste management areas	136	1.3.6 Dump areas
133	Construction areas	137	1.3.7 Construction sites
141	Urban green spaces	141	1.4.1 Parks (including Zoos and botanical gardens)
142	Sports, cultural, tourist and leisure facilities	142	1.4.2 Sports and recreation sites
211	Arable land	210	2.1 Annual cropland
212	Permanent crops	230	2.3 Permanent crops
213	Meadows/pastures	310	3.1 Sown pastures and fields (modified grassland)
214	Heterogeneous agricultural areas	250	2.5 Mixed farmland
311	Deciduous Forests	410	4.1 Broadleaved deciduous forest
312	Coniferous forests	420	4.2 Coniferous forests
313	Natural forests	433	4.3.3 Macaronesian laurophyllous forest
314	Riparian galleries	411	4.1.1 Riparian forest and woodland
315	Natural herbaceous vegetation	320	3.2 Natural and semi-natural grassland
316	Scrublands	523	5.2.3 Temperate and Mediterranean lowland shrub and heathland
321	Sparse vegetation	600	6. Sparsely vegetated ecosystems
322	Beaches	1120	11.2 Coastal dunes, beaches and sandy and muddy shores
324	Bare rock	610	6.1 Bare rocks
411	Flooded areas	700	7. Inland wetlands
511	Watercourses	800	8. Rivers and canals
512	Lakes	900	9. Lakes and reservoirs

13.2.2.3. Land Use Translation

codigo	designacao	designacao_translation
111	Tecido urbano continuo	Continuous urban fabric
112	Tecido urbano descontinuo	Discontinuous urban fabric
121	Indústria, comércio, equipamentos gerais e infraestruturas	Industry, commerce, general equipment and infrastructure
122	Redes viárias e espaços associados	Road networks and associated spaces
123	Áreas portuárias	Port areas
124	Aeroportos e aeródromos	Airports and aerodromes
131	Áreas de extração de massas minerais	Areas of Mineral Extraction
132	Áreas de gestão de resíduos	Waste management areas
133	Áreas em construção	Construction areas
141	Espaços verdes urbanos	Urban green spaces
142	Equipamentos desportivos, culturais, turisticos e de lazer	Sports, cultural, tourist and leisure facilities
211	Terras aráveis	Arable land
212	Culturas permanentes	Permanent crops
213	Prados/pastagens	Meadows/pastures
214	Áreas agrícolas heterogéneas	Heterogeneous agricultural areas
311	Florestas de folhosas	Deciduous Forests
312	Florestas de resinosas	Coniferous forests
313	Florestas naturais	Natural forests
314	Galerias ripícolas	Riparian galleries
315	Vegetação herbácea natural	Natural herbaceous vegetation
316	Matos	Scrublands
321	Vegetação esparsa	Sparse vegetation
322	Praias	Beaches
324	Rocha nua	Bare rock
411	Zonas apauladas	Flooded areas
511	Cursos de água	Watercourses
512	Lagoas	Lakes

13.2.2.4. European Ecosystem Typology

Code	class_level	class_name
100	1	1. Settlements and other artificial areas
110	2	1.1 Continuous settlement area
111	3	1.1.1 Continuous residential area
112	3	1.1.2 Continuous commercial and industrial area
120	2	1.2 Discontinuous settlement area
121	3	1.2.1 Discontinuous residential area
122	3	1.2.2 Discontinuous commercial and industrial area
130	2	1.3 Infrastructure
131	3	1.3.1 Road and rail networks and associated land
132	3	1.3.2 Port areas
133	3	1.3.3 Airports
134	3	1.3.4 Other infrastructure
135	3	1.3.5 Mineral extraction sites (excluding peat extraction sites, see 7.3.1)
136	3	1.3.6 Dump areas
137	3	1.3.7 Construction sites
140	2	1.4 Urban greenspace
141	3	1.4.1 Parks (including Zoos and botanical gardens)
142	3	1.4.2 Sports and recreation sites
143	3	1.4.3 Other urban green
150	2	1.5 Other artificial areas
151	3	1.5.1 Permanent Greenhouses
152	3	1.5.2 Cemeteries
153	3	1.5.3 Archaeological sites
154	3	1.5.4 Urban blue
200	1	2. Cropland
210	2	2.1 Annual cropland
211	3	2.1.1 Cereals excluding rice (C1000) excluding maize (C1500)

212	3	2.1.2 Maize (C1500 + G3000)
213	3	2.1.3 Dry pulses and protein crops (P0000)
214	3	2.1.4 Root crops, like sugar beet and potatoes (R0000)
215	3	2.1.5 Vegetables (including melons) and strawberries (V0000_S0000)
216	3	2.1.6 Industrial crops including annual bioenergy crops (I0000)
217	3	2.1.7 Flowers and ornamental plants (N0000)
218	3	2.1.8 Fallow land (Q0000)
219	3	2.1.9 Temporary grasses and grazing areas (G1000)
220	2	2.2 Rice fields
221	3	2.2.1 Rice fields (C2000)
230	2	2.3 Permanent crops
231	3	2.3.1 Olives (O1000)
232	3	2.3.2 Grapes (W1000)
233	3	2.3.3 Pome fruits (F1100)
234	3	2.3.4 Stone fruits (F1200)
235	3	2.3.5 Berries excluding strawberries (F3000)
236	3	2.3.6 Citrus fruits (T1000)
237	3	2.3.7 Nuts (F4000)
238	3	2.3.8 Hazelnut
239	3	2.3.9 Chestnut
240	2	2.4 Agro-forestry areas
241	3	2.4.1 Holm and cork oak forests
242	3	2.4.2 Other agro-forestry area
250	2	2.5 Mixed farmland
251	3	2.5.1 Mosaic farmland (comprising cropland, grassland and (semi-)natural components)
260	2	2.6 Other farmland
261	3	2.6.1 Nurseries
262	3	2.6.2 Christmas tree plantations
263	3	2.6.3 Perennial bioenergy crops
264	3	2.6.4 Field margins

300	1	3. Grassland (pastures, semi-natural and natural grasslands)
310	2	3.1 Sown pastures and fields (modified grassland)
311	3	3.1.1 Sown pastures used for grazing
312	3	3.1.2 Sown grassland mown frequently for fodder or silage
320	2	3.2 Natural and semi-natural grassland
321	3	3.2.1 Dry grassland
322	3	3.2.2 Seasonally wet and wet grassland
323	3	3.2.3 Alpine and subalpine grasslands
324	3	3.2.4 Woodland fringes and clearings and tall forb stands
325	3	3.2.5 Inland salt steppes
326	3	3.2.6 Sparsely wooded grasslands
327	3	3.2.7 Mesophilous extensive grassland
400	1	4. Forest and woodlands
410	2	4.1 Broadleaved deciduous forest
411	3	4.1.1 Riparian Forest and woodland
412	3	4.1.2 Broadleaved swamp, woodland on non-acid and acid peat
413	3	4.1.3 Fagus dominated forest
414	3	4.1.4 Temperate, Submediterranean and Mediterranean thermophilous deciduous forest
415	3	4.1.5 Acidophilous [Quercus]-dominated woodland
416	3	4.1.6 Temperate and boreal and Southern European Betula and Populus tremula forest on mineral soils
417	3	4.1.7 Other broadleaved deciduous forest, excluding highly modified plantations
418	3	4.1.8 Highly modified broadleaved deciduous forests, in particular plantations including stands of non-native trees species that have long been established in European ecosystems stands
420	2	4.2 Coniferous forests
421	3	4.2.1 Boreal and temperate fir and spruce forest
422	3	4.2.2 Mediterranean mountain fir and spruce forest
423	3	4.2.3 Temperate subalpine Larix, Pinus cembra and Pinus uncinata forest
424	3	4.2.4 Pine Forest (excluding mires, non-thermophilous)
425	3	4.2.5 Mediterranean thermophilous lowland pine forest
426	3	4.2.6 Spruce, pine and larch mire forests
427	3	4.2.7 Taiga forests

428	3	4.2.8 Other coniferous forests, excluding plantations
429	3	4.2.9 Highly modified coniferous forests, in particular plantations
430	2	4.3 Broadleaved evergreen forest
431	3	4.3.1 Mediterranean evergreen Quercus Forest
432	3	4.3.2 Mainland laurophyllous forest
433	3	4.3.3 Macaronesian laurophyllous forest
434	3	4.3.4 Olea europaea-Ceratonia siliqua forest
435	3	4.3.5 Palm groves
436	3	4.3.6 Other broadleaved evergeen forests
437	3	4.3.7 Highly modified broadleaved evergreen forests, in particular plantations (in particular *Eucalyptus*) including stands of non-native trees species that have long been established in European ecosystems stands
440	2	4.4 Mixed forests
441	3	4.4.1 Mixed forests dominated by coniferous species
442	3	4.4.2 Mixed forests dominated by broadleaved species
443	3	4.4.3 Other mixed forests including stands of non-native trees species that have long been established in European ecosystems stands
450	2	4.5 Transitional Forest and woodland shrub
451	3	4.5.1 Transitional woodland/forest land
460	2	4.6 Plantations
461	3	4.6.1 Monoculture plantations
462	3	4.6.2 Mixed plantations
500	1	5. Heathlands and shrub
510	2	5.1 Tundra
511	3	5.1.1 Tundra
520	2	5.2 Heathland and (sub-) alpine shrub
521	3	5.2.1 Arctic alpine, subalpine and lowland shrub and heathland
522	3	5.2.2 Temperate and Mediterranean montane and hilly shrub and heathland
523	3	5.2.3 Temperate and Mediterranean lowland shrub and heathland
530	2	5.3 Sclerophyllous vegetation
531	3	5.3.1 Maquis, arborescent matorral and thermo-Mediterranean shrub
532	3	5.3.2 Garrigue
533	3	5.3.3 Spiny Mediterranean heaths (phrygana, hedgehog-heaths & coastal cliff vegetation)

534	3	5.3.4 Thermo-Atlantic xerophytic shrub (Madeira and Canary Islands)
600	1	6. Sparsely vegetated ecosystems
610	2	6.1 Bare rocks
611	3	6.1.1 Rocky pavements, outcrops, and screes
612	3	6.1.2 Lava flows
620	2	6.2 Semi-desert, desert and other sparsely vegetated areas
621	3	6.2.1 Semi-desert steppes
622	3	6.2.2 Cool deserts and semi-desert steppes
623	3	6.2.3 Other sparsely vegetated areas
630	2	6.3 Ice sheets, glaciers and perennial snowfields
631	3	6.3.1 Ice sheets, glaciers and perennial snowfields
700	1	7. Inland wetlands
710	2	7.1 Inland marshes on mineral soil
711	3	7.1.1 Reedbeds
712	3	7.1.2 Inland salt marshes
713	3	7.1.3 Other marshland and water-fringing ecosystems
720	2	7.2 Mires, bogs and fens
721	3	7.2.1 Raised bogs
722	3	7.2.2 Blanket bogs
723	3	7.2.3 Valley mires, poor fens and transition mires
724	3	7.2.4 Aapa, palsa and polygon mires
725	3	7.2.5 Base-rich fens and calcareous spring mires
726	3	7.2.6 Peat extraction sites
800	1	8. Rivers and canals
810	2	8.1 Rivers
811	3	8.1.1 Rivers
820	2	8.2 Canals, ditches and drains
821	3	8.2.1 Canals, ditches and drains
900	1	9. Lakes and reservoirs
910	2	9.1 Lakes

911	3	9.1.1 Lakes
920	2	9.2 Artificial reservoirs
921	3	9.2.1 Artificial reservoirs
930	2	9.3 Geothermal pools and wetlands (Iceland)
931	3	9.3.1 Geothermal pools and wetlands (Iceland)
1000	1	10 Marine inlets and transitional waters (lagoons, fjords)
1010	2	10.1 Coastal lagoons
1011	3	10.1.1 Coastal lagoons
1020	2	10.2 Estuaries and bays
1021	3	10.2.1 Estuaries and bays
1030	2	10.3 Intertidal flats
1031	3	10.3.1 Intertidal flats (e.g., Wadden Sea)
1040	2	10.4 Deepwater coastal inlets (fjords)
1041	3	10.4.1 Deepwater coastal inlets (fjords)
1100	1	11 Coastal beaches, dunes and wetlands
1110	2	11.1 Artificial shorelines
1111	3	11.1.1 Artificial shorelines
1120	2	11.2 Coastal dunes, beaches and sandy and muddy shores
1121	3	11.2.1 Coastal dunes
1122	3	11.2.2 Beaches and sandy shores
1123	3	11.2.3 Muddy shores
1130	2	11.3 Rocky shores
1131	3	11.3.1 Coastal shingle
1132	3	11.3.2 Rock cliffs, ledges and shores
1140	2	11.4 Coastal saltmarshes and salines
1141	3	11.4.1 Coastal saltmarshes
1142	3	11.4.2 Salines
1200	1	12 Marine ecosystems
1210	2	12.1 Marine macrophyte habitats
1211	3	12.1.1 Kelp forests

1212	3	12.1.2 Seagrass meadows
1220	2	12.2 Coral reefs
1221	3	12.2.1 Coral reefs
1230	2	12.3 Shellfish beds and reefs
1231	3	12.3.1 Shellfish beds and reefs
1241	3	12.4.1 Subtidal sand beds and mud plains
1250	2	12.5 Subtidal sand beds and mud plains
1251	3	12.5.1 Subtidal rocky substrates
1260	2	12.6 Continental and island slopes
1261	3	12.6.1 Continental and island slopes
1270	2	12.7 Deepwater benthic and pelagic ecosystems
1271	3	12.7.1 Deepwater benthic and pelagic ecosystems
1280	2	12.8 Sea ice
1281	3	12.8.1 Sea ice
2110	3	2.1.10 Other crops (further categories may be added by Member States, depending upon nationally important crop types).
238	3	2.3.8 Other perennial crops and orchards

R Script – São Miguel National only approach

```
########## STEP 1: Rasterize Land Use 2018 Map ##########

# As a first step the National (Land Use 2018) map was rasterized to 5 meters (based on the "codigo" column) and reprojected (to EPSG:3035)
in QGIS.
# This dataset was used as an input for STEP 1

# Load relevant library
library(terra)

setwd("path")

# Load the rasterized LU 2018 file
LU_2018 <- rast("path/LU_2018_5m_EPSG3035_Sao_miguel.tif")

#Create a reclassification matrix from Land Use to ETA
rcl_matrix  <-  value_matrix <- matrix(c(
  111, 111,
  112, 121,
  121, 100,
  122, 131,
  123, 132,
  124, 133,
  131, 135,
  132, 136,
  133, 137,
  141, 141,
  142, 142,
  211, 210,
  212, 230,
  213, 310,
  214, 251,
  311, 410,
  312, 420,
  313, 433,
  314, 411,
  315, 320,
  316, 523,
```

```
  321, 600,
  322, 1120,
  324, 610,
  411, 700,
  511, 800,
  512, 900
), ncol = 2, byrow = TRUE, dimnames = list(NULL, c("Input", "Output")))


# Reclassify the raster based on the reclassification ruleset
tif_reclassified <- classify(LU_2018, rcl_matrix)

#Export the reclassified raster
LU_ETA <- writeRaster(tif_reclassified, filename = "LU_2018_5m_EPSG3035_Sao_miguel_ETA.tif",  overwrite = TRUE,datatype='INT2U')


##########STEP 2: Create boundary raster from LU_ETA ##########

#Create a raster from LU_ETA (which will act as the boundary of Sao Miguel)
unique_values <- na.omit(unique(values(LU_ETA)))  # Get unique non-null values
reclass_matrix <- cbind(unique_values, 1)  # Map all to 1

# Apply reclassification
r_reclassified <- classify(LU_ETA, reclass_matrix)

# Save the rasterized output
reference_raster <- writeRaster(r_reclassified, here("path/reference_raster.tif"), filetype = "GTiff", overwrite = TRUE,  datatype='INT2U',
gdal=c("COMPRESS=LZW"))


##########STEP 3: Rasterize Forest Inventory Map ##########

#The Forest Inventory (FI) map was crosswalked to ETA and this table joined in QGIS

#Load Forest Inventory which has been reclassified to ETA in QGIS using a join. The column was named "ETA_code_E"
FI <- vect("path/IF_graf_Z26_Perimetro_Florestal_ETA_joined.shp")

#Make the ETA_code_E column numeric
FI$ETA_code_E <- as.numeric(as.character(FI$ETA_code_E))

#Rasterize FI to the same resolution & coordinate system as LU_ETA and where ETA is not null, take those values. Remainder is null.
raster_data_5m <- rasterize(FI, rast(FI, resolution = res(LU_ETA), crs = crs(LU_ETA)), field = "ETA_code_E")
```

```
values(raster_data_5m)[is.nan(values(raster_data_5m))] <- NA
aligned_raster <- resample(raster_data_5m, LU_ETA, method = "near")

# Save the rasterized output
FI_ETA <- writeRaster(aligned_raster, here("path/forest_inventory_ETA_5m.tif"), filetype = "GTiff", overwrite = TRUE,  datatype='INT2U', gdal=c("COMPRESS=LZW"))


##########STEP 4: Combine LU_ETA and FI_ETA ##########

# Set values to 0 where FI_ETA is NA and reference_raster is 1
FI_ETA[is.na(FI_ETA) & reference_raster == 1] <- 0

# Fill the remaining area that is still classified as 1 from the blank canvas with LU_ETA
FI_ETA[FI_ETA == 0] <- LU_ETA[FI_ETA == 0]


# Save the final national map
writeRaster(FI_ETA, here("path/National_ETA.tif"), filetype = "GTiff", overwrite = TRUE,  datatype='INT2U', gdal=c("COMPRESS=LZW"))
```

```
# R Script – São Miguel National and CLMS approach

# Load required library
pacman::p_load(terra, tidyverse, readxl)
# Set working directory
setwd("path")

########## STEP 1: Rasterize Land Use 2018 Map ##########
# The "codigo" classes 1111 and 1112 from the National Land use 2018 map could be discriminated into continuous and discontinuous settlement in QGIS in the same manner how the coastal zones data was discriminated using IMD (based on average percentage of imperviousness inside the polygon). Here is the SQL query:
#CASE
#WHEN "_mean" >75 THEN 1111
#WHEN "_mean" <= 75 and "_mean" >=0 THEN 1112
#ELSE "codigo"
#END

# This dataset was then crosswalked to ETA, rasterized, clipped and reprojected to the reference_raster
# Thereafter this dataset was combined with crosswalked forest inventory layer

# Step 1: Load the raster files
reference_raster <- rast("path/reference_raster.tif")  # The file you want to resample and clip
FI_ETA <- rast("path/forest_inventory_eta_5m.tif")
LU_data <- rast("path/LU_2018_5m_EPSG3035_Sao_miguel_IMD_to_ETA_v2.tif")


# Add these values to the blank canvas

# Set values to 0 where aligned_raster is NA and reference_raster is 1
FI_ETA[is.na(FI_ETA) & reference_raster == 1] <- 0

# Fill the remaining area that is still classified as 1 from the blank canvas with the land use map created for the first mapping

FI_ETA[FI_ETA == 0] <- LU_data[FI_ETA == 0]

# Save the the national map which has been crosswalked towards the European Typology, rasterized and settlement areas split into continuous and discontinuous using IMD
LU_to_ETA <-writeRaster(aligned_raster, here("path/LU_forest_inventory_ETA_final_including_IMD.tif"), filetype = "GTiff", overwrite = TRUE, datatype='INT2U', gdal=c("COMPRESS=LZW"))
```

```
# Load the coastal zones dataset which has been rasterized to 5m
CZ <- rast("CZ_2018_5m_EPSG3035_Sao_miguel.tif")


# Create a matrix to reclassify the LU_to_ETA dataset so that only the classes that can be mapped towards a lower level or need to be improved can be reclassified
recl_lu <- matrix(c(111,0,
                    112,0,
                    121,0,
                    122,0,
                    131,0,
                    132,0,
                    133,0,
                    135,0,
                    136,0,
                    137,0,
                    141,0,
                    142,0,
                    210,1,
                    230,2,
                    250,0,
                    310,3,
                    320,0,
                    410,4,
                    411,5,
                    420,6,
                    428,0,
                    433,7,
                    436,0,
                    441,0,
                    442,0,
                    461,0,
                    462,0,
                    523,0,
                    600,0,
                    610,8,
                    700,0,
                    800,0,
                    900,9,
                    1120,10,
                    NA,NA), ncol=2, byrow=TRUE
)
```

```
# Reclassify LU_to_ETA
lu_to_eta_recl <- classify(LU_to_ETA, recl_lu)

# Create a matrix to reclassify relevant coastal zones classes to improve LU_to_ETA using CLMS data

recl_cz <- matrix(c(11110,0,
                    11120,0,
                    11130,0,
                    11210,0,
                    12100,0,
                    12310,0,
                    12350,0,
                    12360,0,
                    12400,0,
                    13110,0,
                    13120,0,
                    13130,0,
                    13200,0,
                    14000,0,
                    21100,0,
                    21200,100,
                    22100,0,
                    23100,200,
                    23200,300,
                    23300,400,
                    23400,0,
                    31100,0,
                    32100,0,
                    33100,0,
                    34000,500,
                    41000,0,
                    42100,0,
                    51000,0,
                    61200,0,
                    62111,600,
                    62112,700,
                    63110,0,
                    63120,800,
                    81100,0,
                    82100,0,
                    82200,900,
```

```
                    84100,0,
                    84200,0,
                    NA,NA), ncol=2, byrow=TRUE)

# Reclassify coastal zones
cz_to_eta_recl <- classify(CZ, recl_cz)

# Sum both reclassified rasters (LU_to_ETA and CZ)
comb <- lu_to_eta_recl + cz_to_eta_recl

# Look at combinations of the two rasters
sort(unique(values(comb)))

# Create a reclassification matrix to recode the "comb" dataset so that the improved classes (lower level or improved) can be used for further mapping
recl_comb <- matrix(c(0     ,      0      ,
                      1     ,      0      ,
                      2     ,      0      ,
                      3     ,      0      ,
                      4     ,      0      ,
                      5     ,      0      ,
                      6     ,      0      ,
                      7     ,      0      ,
                      8     ,      0      ,
                      9     ,      0      ,
                      10    ,      0      ,
                      100   ,      0      ,
                      101   ,      0      ,
                      102   ,      151    ,
                      103   ,      0      ,
                      104   ,      0      ,
                      200   ,      0      ,
                      201   ,      251    ,
                      204   ,      0      ,
                      300   ,      0      ,
                      301   ,      251    ,
                      302   ,      0      ,
                      303   ,      251    ,
                      304   ,      0      ,
                      305   ,      0      ,
                      306   ,      0      ,
                      308   ,      0      ,
```

```
310    ,      0      ,
400    ,      0      ,
401    ,      251    ,
402    ,      0      ,
403    ,      251    ,
404    ,      0      ,
405    ,      0      ,
408    ,      0      ,
500    ,      0      ,
501    ,      0      ,
502    ,      0      ,
503    ,      0      ,
504    ,      451    ,
505    ,      451    ,
506    ,      451    ,
507    ,      451    ,
508    ,      0      ,
509    ,      0      ,
600    ,      0      ,
601    ,      0      ,
602    ,      0      ,
603    ,      0      ,
604    ,      0      ,
605    ,      0      ,
608    ,      0      ,
610    ,      1122   ,
700    ,      0      ,
708    ,      1131   ,
800    ,      0      ,
801    ,      0      ,
802    ,      0      ,
803    ,      0      ,
804    ,      0      ,
805    ,      0      ,
808    ,      1132   ,
810    ,      1132   ,
900    ,      0      ,
903    ,      0      ,
904    ,      0      ,
909    ,      921    ,
NA,NA), ncol=2, byrow=TRUE)
```

```
# Recode all combinations to the correct ETA code based on Coastal Zones coverage
comb_to_eta_recl <- classify(comb, recl_comb)

#Check whether the classes were reclassified correctly
unique(values(comb_to_eta_recl))

#### STEP 2 ###
#The above created dataset (combination of LU_to_ETA and CZ) should be mapped as a first priority and thereafter the LU_to_ETA should be
used to fill the rest of the area
comb_to_eta_recl[comb_to_eta_recl == 0] <- LU_to_ETA[comb_to_eta_recl == 0]

#The coastal zones layer contained a gap (NA) which led to removal of area. Here the empty area from the coastal zones layer is replaced
with the Land Use layer (which has been crosswalked to the ETA)
x<-cover(comb_to_eta_recl,LU_to_ETA)

### STEP 3 ###
# As a third step the CLC Plus Backbone dataset was used to improve and distinguish between cropland and sown pastures and fields which have
been falsely mapped under the national typology
CLCBB <- rast("path/CLC_plus_BB_2018_5m_sao_miguel_v2.tif")

# Create a matrix which reclassifies intermediate product
recl_x <- matrix(c(111      ,      0      ,
                   112      ,      0      ,
                   121      ,      0      ,
                   122      ,      0      ,
                   131      ,      0      ,
                   132      ,      0      ,
                   133      ,      0      ,
                   135      ,      0      ,
                   136      ,      0      ,
                   137      ,      0      ,
                   141      ,      0      ,
                   142      ,      0      ,
                   151      ,      0      ,
                   210      ,      1      ,
                   230      ,      0      ,
                   250      ,      0      ,
                   251      ,      0      ,
                   310      ,      2      ,
                   320      ,      0      ,
                   410      ,      0      ,
                   411      ,      0      ,
```

```
                   420      ,      0      ,
                   428      ,      0      ,
                   433      ,      0      ,
                   436      ,      0      ,
                   441      ,      0      ,
                   442      ,      0      ,
                   451      ,      0      ,
                   461      ,      0      ,
                   462      ,      0      ,
                   523      ,      0      ,
                   600      ,      0      ,
                   610      ,      0      ,
                   700      ,      0      ,
                   800      ,      0      ,
                   900      ,      0      ,
                   921      ,      0      ,
                   1120     ,      0      ,
                   1122     ,      0      ,
                   1131     ,      0      ,
                   1132     ,      0      ,
                   NA,NA), ncol=2, byrow=TRUE)

#Reclassify x (cropland and sown pastures and fields)
reclassified_x <- classify(x, recl_x)


# Create a matrix which reclassifies CLCBB
recl_CLCBB <- matrix(c(1,0,
                       2,0,
                       3,0,
                       4,0,
                       5,0,
                       6,10,
                       7,20,
                       9,0,
                       10,0,
                       254,0,
                       NA,NA), ncol=2, byrow=TRUE)

#Reclassify CLCBB
reclassified_CLCBB <- classify(CLCBB, recl_CLCBB)
```

```
#Sum up both reclassified layers to determine overlaps
combined <- reclassified_x + reclassified_CLCBB
#Check which combinations exist and need to be reclassified
unique(combined)


# Final reclassification of class combinations

# Create a reclassification matrix to recode the "combined" dataset so that the improved classes (lower level or improved) can be used for further mapping
#Combinations 11 and 12 become grassland and combinations 21 and 22 become cropland
recl_comb <- matrix(c(0,0,
                      1,0,
                      2,0,
                      10,0,
                      11,310,
                      12,310,
                      20,0,
                      21,210,
                      22,210,
                      NA,NA), ncol=2, byrow=TRUE)


# Reclassify the raster to the ETA codes where 210 and 310 have been changed
comb_to_eta_recl <- classify(combined, recl_comb)

# Combine the above reclassified raster (comb_to_eta_recl) with the original raster (x) which includes the coastal zones improvement and the land use data. Where the above raster is 0 there the values will be replaced
comb_to_eta_recl[comb_to_eta_recl == 0] <- x[comb_to_eta_recl == 0]

## Reclassify 250 to 251. This was an error during the crosswalk
recl_mosaic <- matrix(c(250,251,
                        NA,NA), ncol=2, byrow=TRUE)

reclassified_final <- classify(comb_to_eta_recl, recl_mosaic)

unique(reclassified_final)

writeRaster(reclassified_final, filename = "Final_National_and_CLMS_v3.tif", overwrite = TRUE,datatype='INT2U')
```

13.2.3. São Miguel Forest Inventory (2024) to ETA

Overview of the crosswalk between the São Miguel Forest Inventory (FI) updated as of 2024 towards the European Ecosystem Typology (ETA).

Subsections	Description of Subsections	Columns	Description
Forest Inventory to ETA Crosswalk	A crosswalk was undertaken between the Forest Inventory (FI) and the European Ecosystem Typology (ETA) using the unique combinations of plant species occurring under the broader forest class (IFUsolo_ID). Under the broader Urban (IFUsolo_ID) class all the roads and related areas were also crosswalked to the relevant ETA. All columns taken from the FI data set will not be explained in further detail as they are described in the Forest Inventory document.	ETA l1	This is the level 1 ETA class
		ETA l2	This is the level 2 ETA class
		ETA l3	This is the level 3 ETA class
		COMMENT	This is a short comment reasoning why the class combination was crosswalked to a certain ETA
		Definition from FI doc	This is the definition relevant to the species taken from the FI document
Forest Inventory Translation	The FI dataset was delivered in Portuguese and was internally translated to English. The translation is shown under this subsection.	Column	This is the column under FI dataset translated from Portuguese to English
		Portuguese	Portuguese name
		English	English name
Forest Inventory ID Crosswalked	Each unique ID of the FI dataset that was mapped as either forest (IFUsolo_ID) or roads and related areas (IFOcSo_ID) was crosswalked to the ETA (as seen under the FI_Crosswalk sheet) and given a code to the level it could be crosswalked to. For example, 441 would represent the level 3 ETA class 4.4.1 Mixed forests dominated by coniferous species.	ID	This is the ID extracted from the ID column in the FI dataset
		ETA_code	This is the ETA code (without full stops) that each polygon has been crosswalked to linked to its respective ID

13.2.3.1. Forest Inventory to ETA Crosswalk

This table is too large to reasonably fit within this report. It is available upon request to the authors.

13.2.3.2. Forest Inventory Translation

Column	Portuguese	English
IFUsolo_ID	Floresta	Forest
IFUsolo_ID	Urbano	Urban
IFOcSo_ID	Vegetação exótica	Exotic vegetation
IFOcSo_ID	Vegetação autóctone	Native vegetation
IFOcSo_ID	Vias de comunicação e áreas conexas	Road and related areas
IFEvert_ID	Multi-estratificada	multi-stratified
IFEvert_ID	Uni-estratificada	Uni-stratified
IFNcob_ID	Espontânea	Spontaneous
IFNcob_ID	Cultivada	Cultivated
IFcomp_ID	Misto	mixed
IFcomp_ID	Puro	pure

13.2.3.3. Forest Inventory ID Crosswalked

This table is too large to reasonably fit within this report. It is available upon request to the authors.

13.3. Feature list for habitat mapping

Abbreviation	Name	Input layer spatial resolution	Feature category	Original dataset
scd	Climate snow covered days	100m	Biophysical	CHELSA Climate, 1981-2010
gst	Climate mean temperature of the growing season			CHELSA Climate, 1981-2010
gdd5	Climate number of growing days with T>5°C			CHELSA Climate, 1981-2010
gsp	Climate precipitation in growing season			CHELSA Climate, 1981-2010
bio12	Climate annual precipitation			CHELSA Climate, 1981-2010
bdod	Soil bulk density in topsoil (0-30 cm)			ISRIC Soilgrids
cec	Cation exchange capacity in topsoil			ISRIC Soilgrids
cfvo	Soil coarse fragments in topsoil			ISRIC Soilgrids
soc	Soil organic carbon concentration in topsoil			ISRIC Soilgrids
sand	Soil texture sand fraction in topsoil			ISRIC Soilgrids
clay	Soil texture clay fraction in topsoil			ISRIC Soilgrids
phh20	Soil pH in topsoil measured in water			ISRIC Soilgrids
dist	Distance to inland water		Additional features	Copernicus EU-Hydro River Network Database 2006-2012 (vector), Europe
pop18	Population density in 2018		Biophysical features	EUROSTAT
occur	Inundation occurrence			EUROSTAT
vppamp	HRVPP: Season amplitude	10m		Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
vpplen	HRVPP: Length of season			Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
vppgup	HRVPP: Slope of the greenup period			Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
vppmax	HRVPP: Vegetation index value at MAXD			Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
vpptpr	HRVPP: Total productivity			Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
vppmin	HRVPP: Average vegetation index value of minima on left and right sides of each season			Copernicus: Vegetation – Vegetation Phenology and Productivity Parameters 2017-present (raster 10m), Europe, yearly
lai01	Leaf area index January	100m		Copernicus Leaf Area Index 2014-present (raster 300m), global, 10-daily – version 1 (non-clouded image)
lai04	Leaf area index April			Copernicus Leaf Area Index 2014-present (raster 300m), global, 10-daily – version 1 (non-clouded image)

lai07	Leaf area index July			Copernicus Leaf Area Index 2014-present (raster 300m), global, 10-daily – version 1 (non-clouded image)
lai10	Leaf area index October			Copernicus Leaf Area Index 2014-present (raster 300m), global, 10-daily – version 1 (non-clouded image)
veg-height	Vegetation height	10m	Additional features	ETH Zürich: Research Collection – Global Canopy Height, 10m, 2020, version 1
B02-p10-10m	Reflection of 10th percentile in Sentinel-2 Band 2 (Blue)		Worldcover features	ESA WorldCover (Terrascope)
B02-p50-10m	Reflection of 50th percentile in Sentinel-2 Band 2 (Blue)			
B02-p90-10m	Reflection of 90th percentile in Sentinel-2 Band 2 (Blue)			
B02-iqr-10m	Reflection of interquartile range (IQR) in Sentinel-2 Band 2 (Blue)			
B03-p10-10m	Reflection of 10th percentile in Sentinel-2 Band 3 (Green)			
B03-p50-10m	Reflection of 50th percentile in Sentinel-2 Band 3 (Green)			
B03-p90-10m	Reflection of 90th percentile in Sentinel-2 Band 3 (Green)			
B03-iqr-10m	Reflection of interquartile range (IQR) in Sentinel-2 Band 3 (Green)			
B04-p10-10m	Reflection of 10th percentile in Sentinel-2 Band 4 (Red)			
B04-p50-10m	Reflection of 50th percentile in Sentinel-2 Band 4 (Red)			
B04-p90-10m	Reflection of 90th percentile in Sentinel-2 Band 4 (Red)			
B04-iqr-10m	Reflection of interquartile range (IQR) in Sentinel-2 Band 4 (Red)			
B08-p10-10m	Reflection of 10th percentile in Sentinel-2 Band 8 (NIR)			
B08-p50-10m	Reflection of 50th percentile in Sentinel-2 Band 8 (NIR)			
B08-p90-10m	Reflection of 90th percentile in Sentinel-2 Band 8 (NIR)			
B08-iqr-10m	Reflection of interquartile range (IQR) in Sentinel-2 Band 8 (NIR)			
ndvi-p10-10m	10th percentile in Sentinel-2 NDVI (Normalized Difference Vegetation Index)			
ndvi-p50-10m	50th percentile in Sentinel-2 NDVI			
ndvi-p90-10m	90th percentile in Sentinel-2 NDVI			
ndvi-iqr-10m	Interquartile range in Sentinel-2 NDVI			
ndvi-ts0-10m	NDVI time series 0			
ndvi-ts1-10m	NDVI time series 1			
ndvi-ts2-10m	NDVI time series 2			
ndvi-ts3-10m	NDVI time series 3			
ndvi-ts4-10m	NDVI time series 4			
ndvi-ts5-10m	NDVI time series 5			
evi-p10-10m	10th percentile in Sentinel-2 EVI (Enhanced Vegetation Index)			
evi-p50-10m	50th percentile in Sentinel-2 EVI			

evi-p90-10m	90th percentile in Sentinel-2 EVI			
evi-iqr-10m	Interquartile range in Sentinel-2 EVI			
nirv-p10-10m	Reflection of 10th percentile in Sentinel-2 Near-Infrared			
nirv-p50-10m	Reflection of 50th percentile in Sentinel-2 Near-Infrared			
nirv-p90-10m	Reflection of 90th percentile in Sentinel-2 Near-Infrared			
nirv-iqr-10m	Reflectance of Interquartile range in Sentinel-2 Near-Infrared			
ndwi-p10-10m	10th percentile in Sentinel-2 NDWI (Normalized Difference Water Index)			
ndwi-p50-10m	50th percentile in Sentinel-2 NDWI			
ndwi-p90-10m	90th percentile in Sentinel-2 NDWI			
ndwi-iqr-10m	Interquartile range in Sentinel-2 NDWI			
ndgi-p10-10m	10th percentile in Sentinel-2 NDGI (Normalized Difference Greenness Index)			
ndgi-p50-10m	50th percentile in Sentinel-2 NDGI			
ndgi-p90-10m	90th percentile in Sentinel-2 NDGI			
ndgi-iqr-10m	Interquartile range in Sentinel-2 NDGI			
ndmi-p10-20m	10th percentile in Sentinel-2 NDMI (Normalized Difference Moisture Index)	20m		
ndmi-p50-20m	50th percentile in Sentinel-2 NDMI			
ndmi-p90-20m	90th percentile in Sentinel-2 NDMI			
ndmi-iqr-20m	Interquartile range in Sentinel-2 NDMI			
nbr-p10-20m	10th percentile in Sentinel-2 NBR (Normalized Burn Ratio)			
nbr-p50-20m	50th percentile in Sentinel-2 NBR			
nbr-p90-20m	90th percentile in Sentinel-2 NBR			
nbr-iqr-20m	Interquartile range in Sentinel-2 NBR			
nbr2-p10-20m	10th percentile in Sentinel-2 NBR (with water sensitivity in vegetation)			
nbr2-p50-20m	50th percentile in Sentinel-2 NBR (with water sensitivity in vegetation)			
nbr2-p90-20m	90th percentile in Sentinel-2 NBR (with water sensitivity in vegetation)			
nbr2-iqr-20m	Interquartile range in Sentinel-2 NBR (with water sensitivity in vegetation)			
rep-p10-20m	Sentinel-2 Red-Edge Position at 10th percentile			
rep-p50-20m	Sentinel-2 Red-Edge Position at 50th percentile			
rep-p90-20m	Sentinel-2 Red-Edge Position at 90th percentile			
rep-iqr-20m	Sentinel-2 Red-Edge Position at Interquartile range			
anir-p10-20m	Sentinel-2 angle at NIR at 10th percentile			
anir-p50-20m	Sentinel-2 angle at NIR at 50th percentile			

anir-p90-20m	Sentinel-2 angle at NIR at 90th percentile			
anir-iqr-20m	Sentinel-2 angle at NIR at interquartile range			
ndre2-p10-20m	Normalized Difference Red Edge Index v2 at 10th percentile			
ndre2-p50-20m	Normalized Difference Red Edge Index v2 at 50th percentile			
ndre2-p90-20m	Normalized Difference Red Edge Index v2 at 90th percentile			
ndre2-iqr-20m	Normalized Difference Red Edge Index v2 at interquartile range			
ndre3-p10-20m	Normalized Difference Red Edge Index v3 at 10th percentile			
ndre3-p50-20m	Normalized Difference Red Edge Index v3 at 50th percentile			
ndre3-p90-20m	Normalized Difference Red Edge Index v3 at 90th percentile			
ndre3-iqr-20m	Normalized Difference Red Edge Index v3 at interquartile range			
B05-p10-10m	Reflection of 10th percentile in Sentinel-2 Band 5 (Red edge: 0,705 µm)			
B05-p50-20m	Reflection of 50th percentile in Sentinel-2 Band 5 (Red edge: 0,705 µm)			
B05-p90-20m	Reflection of 90th percentile in Sentinel-2 Band 5 (Red edge: 0,705 µm)			
B05-iqr-20m	Reflection of interquartile range (IQR) in Sentinel-2 Band 5 (Red edge: 0,705 µm)			
B06-p10-20m	Reflection of 10th percentile in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B06-p50-20m	Reflection of 50th percentile in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B06-p90-20m	Reflection of 90th percentile in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B06-iqr-20m	Reflection of interquartile range (IQR) in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B07-p10-20m	Reflection of 10th percentile in Sentinel-2 Band 7 (Red edge: 0,783 µm)			
B07-p50-20m	Reflection of 50th percentile in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B07-p90-20m	Reflection of 90th percentile in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B07-iqr-20m	Reflection of interquartile range (IQR) in Sentinel-2 Band 6 (Red edge: 0,740 µm)			
B11-p10-20m	Reflection of 10th percentile in Sentinel-2 Band 11 (SWIR: 1,610 µm)			

B11-p50-20m	Reflection of 50^{th} percentile in Sentinel-2 Band 11 (SWIR: 1,610 µm)			
B11-p90-20m	Reflection of 90^{th} percentile in Sentinel-2 Band 11 (SWIR: 1,610 µm)			
B11-iqr-20m	Reflection of interquartile range (IQR) in Sentinel-2 Band 11 (SWIR: 1,610 µm)			
B12-p10-20m	Reflection of 10^{th} percentile in Sentinel-2 Band 12 (SWIR: 2,190 µm)			
B12-p50-20m	Reflection of 50^{th} percentile in Sentinel-2 Band 12 (SWIR: 2,190 µm)			
B12-p90-20m	Reflection of 90^{th} percentile in Sentinel-2 Band 12 (SWIR: 2,190 µm)			
B12-iqr-20m	Reflection of interquartile range (IQR) in Sentinel-2 Band 12 (SWIR: 2,190 µm)			
VH-p10-20m	Sentinel 1 vertical-horizontal backscatter at 10^{th} percentile			
VH-p50-20m	Sentinel 1 vertical-horizontal backscatter at 50^{th} percentile			
VH-p90-20m	Sentinel 1 vertical-horizontal backscatter at 90^{th} percentile			
VH-iqr-20m	Sentinel 1 vertical-horizontal backscatter at interquartile range			
VV-p10-20m	Sentinel 1 vertical-vertical backscatter at 10^{th} percentile			
VV-p50-20m	Sentinel 1 vertical-vertical backscatter at 50^{th} percentile			
VV-p90-20m	Sentinel 1 vertical-vertical backscatter at 90^{th} percentile			
VV-iqr-20m	Sentinel 1 vertical-vertical backscatter at interquartile range			
vh_vv-p10-20m	Sentinel 1 VH to VV backscatter at 10^{th} percentile			
vh_vv-p50-20m	Sentinel 1 VH to VV backscatter at 50^{th} percentile			
vh_vv-p90-20m	Sentinel 1 VH to VV backscatter at 90^{th} percentile			
vh_vv-iqr-20m	Sentinel 1 VH to VV backscatter at interquartile range			
rvi-p10-20m	Sentinel 1 Radar Vegetation Index at 10^{th} percentile			
rvi-p50-20m	Sentinel 1 Radar Vegetation Index at 50^{th} percentile			
rvi-p90-20m	Sentinel 1 Radar Vegetation Index at 90^{th} percentile			
rvi-iqr-20m	Sentinel 1 Radar Vegetation Index at interquartile range			
DEM-alt-20m	Digital Elevation Model – Altitude		Additional features (but extracted as Worldcover features)	Copernicus DEM - Global and European Digital Elevation Model / ESA WorldCover (Terrascope)
DEM-slo-20m	Digital Elevation Model – Slope			Copernicus DEM - Global and European Digital Elevation Model / ESA Worldcover (Terrascope)
contextft_1	Deep Contextual Feature #1	10m	Contextual features	Derived from RGB median images of the year
contextft_2	Deep Contextual Feature #2			

contextft_3	Deep Contextual Feature #3			
contextft_4	Deep Contextual Feature #4			
contextft_5	Deep Contextual Feature #5			
contextft_6	Deep Contextual Feature #6			
contextft_7	Deep Contextual Feature #7			
contextft_8	Deep Contextual Feature #8			
contextft_9	Deep Contextual Feature #9			
contextft_10	Deep Contextual Feature #10			
contextft_11	Deep Contextual Feature #11			
contextft_12	Deep Contextual Feature #12			
contextft_13	Deep Contextual Feature #13			
contextft_14	Deep Contextual Feature #14			
contextft_15	Deep Contextual Feature #15			
contextft_16	Deep Contextual Feature #16			

13.4. EUNIS habitat classes used in this report

Note that the latest EUNIS typology ‘EUNIS2021’ is currently under review. Therefore, it is still incomplete, missing classes like ‘inland surface waters’ or ‘urban area’. Some classes were added from the previous EUNIS typology ‘EUNIS2012’ (e.g. C and J). Besides, in interaction with local experts from São Miguel TS partner, some personalized habitat classes were defined (e.g. S4x, T2x, V6a, V6b) to improve the habitat map detail.

Level	EUNIS code	EUNIS name
1	C	Inland surface waters
2	C1	Surface standing waters
3	C13	Permanent eutrophic lakes, ponds and pools
3	C16	Temporary lakes, ponds and pools
2	C2	Surface running waters
3	C23	Permanent non-tidal, smooth-flowing watercourses
1	J	Constructed, industrial and other artificial habitats
2	J1	Buildings of cities, towns and villages
3	J11	Residential buildings of city and town centers
3	J12	Residential buildings of villages and urban peripheries
3	J14	Urban and suburban industrial and commercial sites still in active use
3	J16	Urban and suburban construction and demolition sites
2	J2	Low density buildings
3	J24	Agricultural constructions
2	J3	Extractive industrial sites
3	J32	Active opencast mineral extraction sites, including quarries
2	J4	Transport networks and other constructed hard-surfaced areas
3	J42	Road networks
3	J44	Airport runways and aprons
3	J45	Hard-surfaced areas of ports
2	J5	Highly artificial man-made waters and associated structures
3	J53	Highly artificial non-saline standing waters
2	J6	Waste deposits
3	J62	Household waste and landfill sites
1	N	Coastal habitats
2	N1	Coastal dunes and sandy shores
3	N12	Mediterranean and Black Sea sand beach
3	N14	Mediterranean, Macaronesian and Black Sea shifting coastal dun
3	N16	Mediterranean and Macaronesian coastal dune grassland (grey dune)

3	N1B	Mediterranean and Black Sea coastal dune scrub
3	N1G	Mediterranean coniferous coastal dune forest
3	N1J	Mediterranean and Black Sea moist and wet dune slack
2	N2	Coastal shingle
3	N21	Atlantic, Baltic and Arctic coastal shingle beach
2	N3	Rock cliffs, ledges and shores, including the supralittoral
3	N33	Macaronesian rocky sea cliff and shore
3	N35	Mediterranean and Black Sea soft sea cliff
1	Q	Wetlands
2	Q1	Raised and blanket bogs
3	Q11	Raised bog
3	Q12	Blanket bog
2	Q2	Valley mires, poor fens and transition mires
3	Q21	Oceanic valley mire
3	Q23	Relict mire of Mediterranean mountains
2	Q5	Helophyte beds
3	Q53	Tall-sedge bed
2	Q6	Periodically exposed shores
3	Q62	Periodically exposed shore with stable, mesotrophic sediments with pioneer or ephemeral vegetation
3	Q63	Periodically exposed saline shore with pioneer or ephemeral vegetation
1	R	Grasslands and lands dominated by forbs, mosses or lichens
2	R1	Dry grasslands
3	R14	Perennial rocky grassland of the Italian Peninsula
3	R1D	Mediterranean closely grazed dry grassland
3	R1K	Balkan and Anatolian or mediterranean dry grassland
3	R1T	Azorean open, dry, acid to neutral grassland
2	R2	Mesic grasslands
3	R21	Mesic permanent pasture of lowlands and mountains
3	R22	Low and medium altitude hay meadow
2	R3	Seasonally wet and wet grasslands
3	R35	Moist or wet mesotrophic to eutrophic hay meadow
3	R36	Moist or wet mesotrophic to eutrophic pasture
3	R37	Temperate and boreal moist or wet oligotrophic grassland
2	R4	Alpine and subalpine grasslands
3	R43	Temperate acidophilous alpine grassland
3	R44	Arctic-alpine calcareous grassland
2	R6	Inland salt steppes and salt marshes
3	R61	Mediterranean inland salt steppe
1	S	Heathland, scrub and tundra

2	S2	Arctic, alpine and subalpine scrub
3	S24	Subalpine genistoid scrub of the Amphi-Adriatic region
2	S4	Temperate shrub heathland
3	S43	Macaronesian heath
3	S4x	All temperate shrub heathland not further classified to level 3
2	S5	Maquis, arborescent matorral and thermo-Mediterranean scrub
3	S51	Mediterranean maquis and arborescent matorral
3	S52	Submediterranean pseudomaquis
3	S54	Thermomediterranean arid scrub
2	S6	Garrigue
3	S62	Western acidophilous garrigue
2	S7	Spiny Mediterranean heaths (phrygana, hedgehog-heaths and related coastal cliff vegetation)
3	S72	Eastern Mediterranean spiny heath (phrygana)
3	S73	Western Mediterranean mountain hedgehog-heath
3	S75	Eastern Mediterranean mountain hedgehog-heath
2	S9	Riverine and fen scrubs
3	S93	Mediterranean riparian scrub
1	T	Forest and other wooded land
2	T1	Deciduous broadleaved forest
3	T14	Mediterranean and Macaronesian riparian forest
3	T16	Broadleaved mire forest on acid peat
3	T19	Temperate and submediterranean thermophilous deciduous forest
3	T1A	Mediterranean thermophilous deciduous forest
3	T1H	Broadleaved deciduous plantation of non site-native trees
3	T1J	Deciduous self sown forest of non site-native trees
2	T2	Broadleaved evergreen forest
3	T21	Mediterranean evergreen *Quercus* forest
3	T23	Macaronesian laurophyllous forest
3	T24	*Olea europaea-Ceratonia siliqua* forest
3	T29	Broadleaved evergreen plantation of non site-native trees
3	T2A	Broadleaved evergreen plantation of site-native trees
3	T2B	Mediterranean evergreen *Quercus forest* (Greece)
3	T2x	All broadleaved evergreen forest not further classified to level 3
2	T3	Coniferous forest
3	T36	Temperate and submediterranean montane *Pinus sylvestris-Pinus nigra* forest
3	T3A	Mediterranean lowland to submontane *Pinus* forest
3	T3D	Mediterranean *Cupressaceae* forest
3	T3L	Coniferous self sown forest of non site-native trees
3	T3M	Coniferous plantation of non site-native trees

3	T3N	Coniferous plantation of site-native trees
3	T3P	Mediterranean mountain *Abies* forest (Greece)
1	U	Inland habitats with no or little soil and mostly with sparse vegetation
2	U2	Screes
2	U3	Inland cliffs, rock pavements and outcrops
3	U3C	Macaronesian inland cliff
2	U5	Miscellaneous inland habitats usually with very sparse or no vegetation
2	U6	Recent volcanic features
3	U62	Mediterranean, Macaronesian and temperate volcanic field
2	U7	Unvegetated or sparsely vegetated gravel bars
1	V	Vegetated man-made habitats
2	V1	Arable land and market gardens
2	V2	Cultivated areas of gardens and parks
3	V21	Large-scale ornamental garden areas
2	V3	Artificial grasslands and herb dominated habitats
3	V31	Agriculturally-improved, re-seeded and heavily fertilised grassland, including sports fields and grass lawns
2	V5	Shrub plantations
2	V6	Tree dominated man-made habitats
3	V6a	Cultivated broadleaved (deciduous and evergreen) tree orchards.
3	V6b	Planted and intensively managed woodlands (Forest Reserves of São Miguel).

13.5. Overview of training data for habitat mapping

13.5.1. Overview of training data for habitat mapping of Peloponnese TS

The number of training points per habitat class and the indication of this point were used for training a model at EUNIS levels 1, 2, 3.

EUNIS Habitat Class	Number of Sampled Points	Used in Level 1	Used in Level 2	Used in Level 3
J	551	True	False	False
C	96	True	False	False
N12	20	True	True	True
N14	50	True	True	True
N16	20	True	True	True
N1B	20	True	True	True
N1G	15	True	True	True
N1J	20	True	True	True
N35	20	True	True	False
Q23	15	True	True	False
Q53	43	True	True	False
Q63	12	True	True	False
R1D	85	True	True	True
R1E	15	True	True	True
R1K	73	True	True	True
R61	28	True	True	False
S24	13	True	True	False
S51	564	True	True	True
S52	10	True	True	True
S54	45	True	True	True
S62	3378	True	True	False
S72	816	True	True	True
S75	1234	True	True	True
S93	92	True	True	False

T14	202	True	True	True
T19	231	True	True	True
T1A	23	True	True	True
T21	259	True	True	True
T24	285	True	True	True
T2B	254	True	True	True
T36	317	True	True	True
T3A	273	True	True	True
T3D	83	True	True	True
T3N	27	True	True	True
T3P	1320	True	True	True
U29	31	True	True	False
U38	169	True	True	False
U5	19	True	True	False
U72	11	True	True	False
V	5063	True	False	False
V11	41	True	True	False
V13	221	True	True	False
V16	11	True	True	False
V17	344	True	True	False
V31	8	True	True	False
V54	61	True	True	False
V55	19	True	True	False
V61	50	True	True	False
V62	345	True	True	False
V67	196	True	True	False
V68	302	True	True	False

13.5.2. Overview of training data for habitat mapping of São Miguel TS

The number of training points per habitat class and the indication of this point was used for training a model at EUNIS level 1, 2 or 3.

EUNIS Habitat Class	Total Number Points	Priority source dataset	Sampled Points	Used in Level 1	Used in Level 2	Used in Level 3
C13	274	COS2018A	242	True	True	True
C16	32	IF2024	15	True	True	True
C23	362	IF2024	319	True	True	False
J11	3488	IF2024	1184	True	True	True
J12	3607	COS2018A	1224	True	True	True
J14	744	COS2018A	252	True	True	True
J16	47	COS2018A	20	True	True	True
J24	363	COS2018A	123	True	True	False
J32	431	COS2018A	146	True	True	False
J42	3245	IF2024	1101	True	True	True
J44	121	COS2018A	41	True	True	True
J45	51	COS2018A	20	True	True	True
J53	36	IF2024	20	True	True	False
J62	44	COS2018A	20	True	True	False
N11	164	LifeIPNaturaAzores	33	True	True	False
N21	246	LifeIPNaturaAzores	49	True	True	False
N33	1540	LifeIPNaturaAzores	310	True	True	False
Q11	511	LifeIPNaturaAzores	35	True	True	True
Q12	4359	LifeIPNaturaAzores	297	True	True	True
Q21	53	LifeIPNaturaAzores	20	True	True	False
Q62	827	LifeIPNaturaAzores	100	True	True	False
R1T	1548	LifeIPNaturaAzores	792	True	True	False
R21	9801	IF2024 or TerceiraLandUse	5011	True	True	False
S4x	8607	IF2024	2915	True	False	True
S43	8527	LifeIPNaturaAzores	2888	True	False	True
T16	1463	LifeIPNaturaAzores	733	True	True	True
T1H	166	IF2024	83	True	True	True
T1J	30	IF2024	20	True	True	True
T2x	4663	IF2024	2336	True	True	True
T23	5510	LifeIPNaturaAzores	2760	True	True	True
T29	336	IF2024	168	True	True	True

T2A	176	IF2024	72	True	True	True
T3L	65	IF2024	33	True	True	True
T3M	14602	IF2024	7314	True	True	True
T3N	153	IF2024	77	True	True	True
U3C	47	LifeIPNaturaAzores	20	True	True	False
U62	707	LifeIPNaturaAzores	298	True	True	False
V1	918	COS2018A	221	True	True	False
V21	123	COS2018A	29	True	True	False
V31	87938	IF2024	21078	True	True	False
V52	43	GoogleEarth	20	True	True	False
V6a	1379	TerceiraLandUse	331	True	True	True
V6b	196	IF2024	47	True	True	True

13.6. Raw Forest Condition variable accounts for São Miguel 2017-2023

- T1: Broadleaved deciduous forest
- T2: Broadleaved evergreen forest
- T3: Coniferous forest
- S4: Temperate shrub heathland

Net Primary Productivity (NPP)

NPP [kgC/ha]	T1	T2	T3	S4
2017	14268.00	14454.81	13815.87	13643.52
2018	13642.32	14188.29	13587.44	13311.28
2019	13785.35	14080.09	13430.30	13285.35
2020	14407.34	14389.76	13803.37	13570.07
2021	12750.17	13743.99	13146.08	12911.65
2022	12964.84	13454.44	12767.02	12614.50
2023	11923.42	13251.60	12467.37	12255.47

Aboveground biomass (AGB)

AGB [ton/ha]	T1	T2	T3	S4
2017	81.34	71.60	104.87	42.98
2018	80.77	73.92	104.87	43.48
2019	87.43	75.60	110.83	45.46
2020	91.62	79.99	117.02	48.16
2021	90.25	79.02	116.54	48.63
2022	91.65	78.87	116.87	49.29
2023	91.65	78.87	116.87	49.29

Normalized Difference Water Index (NDWI)

NDWI	T1	T2	T3	S4
2017	-0.71	-0.78	-0.77	-0.74
2018	-0.72	-0.78	-0.78	-0.75
2019	-0.71	-0.78	-0.77	-0.75
2020	-0.71	-0.78	-0.77	-0.75
2021	-0.72	-0.78	-0.77	-0.75
2022	-0.72	-0.78	-0.77	-0.75
2023	-0.72	-0.78	-0.77	-0.75

Forest Connectivity (FC)

FC [%]	T1	T2	T3	S4
2017	24.38	22.07	30.46	9.73
2018	24.38	22.07	30.46	9.73
2019	24.38	22.07	30.46	9.73
2020	24.38	22.07	30.46	9.73
2021	24.38	22.07	30.46	9.73
2022	24.38	22.07	30.46	9.73
2023	24.38	22.07	30.46	9.73

Threatened Forest Bird Species Diversity (TFBSD)

TFBSD	T1	T2	T3	S4
2017	0.11	0.10	0.11	0.10
2018	0.11	0.10	0.11	0.10
2019	0.11	0.10	0.11	0.10
2020	0.11	0.10	0.11	0.10
2021	0.11	0.10	0.11	0.10
2022	0.11	0.10	0.11	0.10
2023	0.11	0.10	0.11	0.10

Soil Organic Carbon (SOC)

SOC [tonC/ha]	T1	T2	T3	S4
2017	405.25	368.01	394.66	325.43
2018	405.25	368.01	394.66	325.43
2019	405.25	368.01	394.66	325.43
2020	405.25	368.01	394.66	325.43
2021	405.25	368.01	394.66	325.43
2022	405.25	368.01	394.66	325.43
2023	405.25	368.01	394.66	325.43