

**Horizon 2020**  
Space Call - Earth Observation: EO-3-2016: Evolution of Copernicus services  
Grant Agreement No. 730008

# ECoLaSS

## Evolution of Copernicus Land Services based on Sentinel data



## D12.2

### “D42.1b - Prototype Report: Consistent HR Layer Time Series/Incremental Updates (Issue 2)”

Issue/Rev.: 2.0

Date Issued: 03.12.2019

submitted by:



an e-GEOS (ASI / Telespazio) Company

in collaboration with the consortium partners:



submitted to:



European Commission – Research Executive Agency

This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme, under Grant Agreement No. 730008.

## CONSORTIUM PARTNERS

| No. | PARTICIPANT ORGANISATION NAME  | SHORT NAME | CITY, COUNTRY                 |
|-----|--|------------|-------------------------------|
| 1   | GAF AG   | GAF        | Munich,<br>Germany            |
| 2   | Systèmes d'Information à Référence Spatiale SAS                                    | SIRS       | Villeneuve<br>d'Ascq, France  |
| 3   | JOANNEUM RESEARCH Forschungsgesellschaft mbH                                       | JR         | Graz,<br>Austria              |
| 4   | Université catholique de Louvain, Earth and Life Institute (ELI)                   | UCL        | Louvain-la-<br>Neuve, Belgium |
| 5   | German Aerospace Center (DLR), German Remote Sensing<br>Data Center (DFD), Weßling | DLR        | Weßling,<br>Germany           |

## CONTACT:

GAF AG  
Arnulfstr. 199 – D-80634 München – Germany  
Phone: ++49 (0)89 121528 0 – FAX: ++49 (0)89 121528 79  
E-mail: copernicus@gaf.de – Internet: www.gaf.de

## DISCLAIMER:

The contents of this document are the copyright of GAF AG and Partners. It is released by GAF AG on the condition that it will not be copied in whole, in section or otherwise reproduced (whether by photographic, reprographic or any other method) and that the contents thereof shall not be divulged to any other person other than of the addressed (save to the other authorised officers of their organisation having a need to know such contents, for the purpose of which disclosure is made by GAF AG) without prior consent of GAF AG. Nevertheless single or common copyrights of the contributing partners remain unconditionally unaffected.

### DOCUMENT RELEASE SHEET

|               | NAME, FUNCTION  | DATE       | SIGNATURE   |
|---------------|---|------------|---|
| Author(s):    | David Herrmann (GAF),<br>Roger Almengor González (GAF),<br>Sophie Villerot (SIRS),<br>Christophe Sannier (SIRS),<br>Alexandre Pennec (SIRS) | 20.09.2019 |   |
| Review:       | Eva Sevillano Marco (GAF)<br>Sophie Villerot (SIRS)   | 21.10.2019 |  |
| Approval:     | Eva Sevillano Marco (GAF)   | 03.12.2019 |  |
| Acceptance:   | Massimo Ciscato (REA)   |            |   |
| Distribution: | Public  |            |   |

### DISSEMINATION LEVEL

| DISSEMINATION LEVEL |  |   |
|---------------------|--|---|
| PU                  | Public   | X |
| CO                  | Confidential: only for members of the consortium (including the Commission Services) |   |

### DOCUMENT STATUS SHEET

| ISSUE/REV | DATE       | PAGE(S) | DESCRIPTION / CHANGES  |
|-----------|------------|---------|--|
| 1.0       | 23.07.2018 | 76      | First Issue of the Prototype Report on Consistent HR Layer Time Series and Incremental Updates   |
| 2.0       | 03.12.2019 | 110     | Second Issue of the Prototype Report on Consistent HR Layer Time Series and Incremental Updates. Minor changes in section 3, and major updates from sections 4-6, with final results from phase 2 implementation |

## APPLICABLE DOCUMENTS

| ID   | DOCUMENT NAME / ISSUE DATE  |
|------|---|
| AD01 | Horizon 2020 Work Programme 2016 – 2017, 5 iii. Leadership in Enabling and Industrial Technologies – Space. Call: EO-3-2016: Evolution of Copernicus services. Issued: 13.10.2015 |
| AD02 | Guidance Document: Research Needs Of Copernicus Operational Services. Final Version issued: 30.10.2015  |
| AD03 | Proposal: Evolution of Copernicus Land Services based on Sentinel data. Proposal acronym: ECoLaSS, Proposal number: 730008. Submitted: 03.03.2016                                 |
| AD04 | Grant Agreement – ECoLaSS. Grant Agreement number: 730008 – ECoLaSS – H2020-EO-2016/H2020-EO-2016, Issued: 18.10.2016   |
| AD05 | D3.2 : D21.1b – Service Evolution Requirements Report (Issue 1), Issued: December 2019  |
| AD06 | D7.2 : D32.1b - Methods Compendium: Time Series Preparation (Issue 1), Issued: 15.05.2019   |
| AD07 | D8.2 : D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification (Issue 2), Issued: December 2019   |
| AD08 | D9.2 : D34.1b - Methods Compendium: Time Series Analysis for Change Detection (Issue 2), Issued: December 2019  |
| AD09 | D10.3 : D35.1b- Methods Compendium: HRL Time Series Consistency for HRL Product Updates (Issue 2), Issued: December 2019  |
| AD10 | D1.5: D11.3b – Interim Progress Report (Issue 2), Issued: 09.07.201   |

## EXECUTIVE SUMMARY

The Horizon 2020 (H2020) project, “Evolution of Copernicus Land Services based on Sentinel data” (ECoLaSS) addresses the H2020 Work Programme 5 iii. Leadership in Enabling and Industrial technologies - Space, specifically the Topic EO-3-2016: Evolution of Copernicus services. ECoLaSS is being conducted from 2017–2019 and aims at developing and prototypically demonstrating selected innovative products and methods as candidates for future next-generation operational Copernicus Land Monitoring Service (CLMS) products of the pan-European and Global Components. ECoLaSS assesses the operational readiness of such candidate products and eventually suggests some of these for implementation. This shall enable the key CLMS stakeholders (i.e. mainly the Entrusted European Entities (EEE) EEA and JRC) to take informed decisions on potential procurement as (part of) the next generation of Copernicus Land services from 2020 onwards.

To achieve this goal, ECoLaSS makes full use of dense time series of High-Resolution (HR) Sentinel-2 optical and Sentinel-1 Synthetic Aperture Radar (SAR) data, complemented by Medium-Resolution (MR) Sentinel-3 optical data if needed and feasible. Rapidly evolving scientific developments as well as user requirements are continuously analysed in a close stakeholder interaction process, targeting a future pan-European roll-out of new/improved CLMS products, and assessing the potential transferability to global applications.

This Deliverable **D12.2: “D42.1b - Prototype Report: Consistent HR Layer Time Series/Incremental Updates”** constitutes the final report of the Work Package (WP) 42 which is working on developing improved status layers and proposed incremental update prototypes for the High Resolution Layer (HRL) Imperviousness and HRL Forest. The applied methodologies build directly on the methods and processing lines developed in Task 3, especially WP 33 “Time Series Analysis for Thematic Classification” [AD07] and WP 34 “Time Series Analysis for Change Detection” [AD08]. The demonstration of incremental updates is oriented towards the update frequency proposed in WP 35 “HRL Time Series Consistency for HRL Product (incremental) Updates” [AD09].

This document is structured as follows: after a short introduction on the purpose and scope of this WP, the HRL incremental update feasibility is analysed in chapter 2. Besides layer-specific background information, user requirements are taken into account which are relevant for the proposed prototypes [AD05]. Chapter 3 provides an overview of the relevant demonstration sites in which the prototypes are finally implemented. In chapter 4, an overview of the consistently applied methods fully described in Task 3 reports is given. These have been developed and assessed in Task 3 and have been transferred to and applied in specific demonstration sites to implement the respective prototypes. The main results and outcomes are presented in chapter 5. First, input data, general pre-processing steps and the layer-specific prototype setup are presented, followed by the classification and validation results of the improved status layers for Imperviousness and Forest, which are analysed and discussed in detail. Subsequently, the selected change detection approaches and the incremental updates resulting thereof are presented, together with dedicated assessments, conclusions and proposed specifications. Finally, chapter 6 concludes the findings and provides an outlook on the activities and research aspects from the second project phase.

The ECoLaSS project follows a two-phased approach of two times 18 months duration. The first deliverable constitutes was issued at month 18, presenting preliminary results. In the second 18-month project cycle, a second issue of this deliverable is published, containing all relevant updates and final results.

## Table of Contents

|                        |  |            |
|------------------------|--|------------|
| <b>1</b>               | <b>INTRODUCTION .....</b>  | <b>1</b>   |
| <b>2</b>               | <b>HRL INCREMENTAL UPDATE NEEDS AND FEASIBILITY.....</b>             | <b>2</b>   |
| <b>3</b>               | <b>DEMONSTRATION SITES.....</b>                                      | <b>5</b>   |
| 3.1                    | DEMONSTRATION SITE SOUTH-WEST .....                                  | 6          |
| 3.2                    | DEMONSTRATION SITE NORTH .....                                       | 7          |
| 3.3                    | DEMONSTRATION SITE SOUTH-EAST .....                                  | 9          |
| 3.4                    | DEMONSTRATION SITE CENTRAL.....                                      | 10         |
| <b>4</b>               | <b>OVERVIEW OF APPLIED METHODS.....</b>                              | <b>12</b>  |
| 4.1                    | METHODS FOR IMPLEMENTING A PROTOTYPE ON IMPERVIOUSNESS .....         | 12         |
| 4.1.1                  | Derivation of biophysical variables .....                            | 12         |
| 4.1.2                  | Active learning Classification approach based on DAP profiles.....   | 13         |
| 4.2                    | METHODS FOR IMPLEMENTING A PROTOTYPE ON FOREST.....                  | 15         |
| <b>5</b>               | <b>PROTOTYPE IMPLEMENTATION.....</b>                                 | <b>17</b>  |
| 5.1                    | PROTOTYPE OF A POTENTIAL FUTURE HRL IMPERVIOUSNESS .....             | 17         |
| 5.1.1                  | Data and Processing Setup .....                                      | 17         |
| 5.1.2                  | Classification Results and Validation .....                          | 38         |
| 5.1.3                  | Change Detection and Incremental Update Results and Validation ..... | 48         |
| 5.2                    | PROTOTYPE OF A POTENTIAL FUTURE HRL FOREST .....                     | 52         |
| 5.2.1                  | Data and Processing Setup .....                                      | 52         |
| 5.2.2                  | Classification Results and Validation .....                          | 63         |
| 5.2.3                  | Change Detection and Incremental Update Results and Validation ..... | 85         |
| 5.3                    | PROTOTYPE SPECIFICATIONS .....                                       | 93         |
| <b>6</b>               | <b>CONCLUSIONS AND OUTLOOK.....</b>                                  | <b>107</b> |
| <b>REFERENCES.....</b> |  | <b>109</b> |

## List of Figures

|   |    |
|---|----|
| Figure 3-1: Biogeographic Regions of Europe (2015) and European ECoLaSS Demonstration Sites .....   | 5  |
| Figure 3-2: Demonstration site South-West, with CLC 2018 background layer.....  | 7  |
| Figure 3-3: Demonstration site North with CLC 2018 background layer .....   | 8  |
| Figure 3-4: Overview of DEMO Site South-East. ....  | 9  |
| Figure 3-5: Overview of DEMO Site Central .....   | 11 |
| Figure 4-1: Example of SSUs organised in a 5x5 10m grid.....  | 14 |
| Figure 5-1: Theia Level-2A data production extent. (21.07.2018) .....   | 18 |
| Figure 5-2: Reference calibration samples overlaid on the change mask 2015-2017 (SW).....   | 22 |
| Figure 5-3: - Reference calibration samples overlaid on the change mask 2015/17-2018.....   | 23 |
| Figure 5-4: NDVI Sentinel-2 based maximum feature for the year 2017 .....   | 24 |
| Figure 5-5: 2018 NDVI Sentinel-2 based maximum feature for the year 2018.....   | 25 |
| Figure 5-6: SAR statistical features .....  | 26 |
| Figure 5-7: Sentinel-2 based initial built-up and non-built-up mask for 2017 for the South-West site .....  | 27 |
| Figure 5-8: Sentinel-2 based initial sealed and non-sealed mask for 2018 for the demonstration sites....  | 28 |
| Figure 5-9: Post-classification processing.....   | 30 |
| Figure 5-10: Final HRL Imperviousness 2017 prototype for the South-West demonstration site.....   | 31 |
| Figure 5-11: Final HRL Imperviousness 2018 prototype for the demonstration sites .....  | 32 |
| Figure 5-12: Final HRL Imperviousness Change 2015-2018 prototypes for the demonstration sites .....   | 33 |
| Figure 5-13: Final HRL Imperviousness Change 2015-2017 prototype South-West .....   | 34 |
| Figure 5-14: Final HRL Imperviousness Classified Change 2015-2018 or 2017-2018 prototypes .....   | 35 |
| Figure 5-15: 2018 PanTex Sentinel-2 based feature for the Central demonstration site .....  | 36 |
| Figure 5-16: Sentinel-2 based initial built-up and non-built-up masks for 2018 .....  | 37 |
| Figure 5-17: Final HRL Built-up 2018 prototypes for the demonstration sites .....   | 38 |
| Figure 5-18: Scatterplots for the continuous IMD layers 2018 validation .....   | 40 |
| Figure 5-19: Representation of the behaviour of the 95% confidence interval for a 5x5 SSU grid over the whole range of imperviousness degree values .....               | 44 |
| Figure 5-20: Case of omission (yellow dots) in isolated built-up features in the IMD 2017 prototype.....  | 46 |
| Figure 5-21: Examples of commission errors in the IMD 2017 prototype .....  | 47 |
| Figure 5-22: Example of calibration points for the newly detected built-up in 2017 .....  | 49 |
| Figure 5-23: Example of omission errors in 2015 (undetected built-up in 2015) .....   | 49 |
| Figure 5-24: Example of commission errors in 2017, false built-up in 2017 .....   | 50 |
| Figure 5-25: Sentinel-2 Data Score Layer (S2DSL): Number of cloud-free S-2 observations per pixel .....   | 54 |
| Figure 5-26: Forest sample layer derived from the integrated Copernicus High-Resolution Layers 2015 ..  | 55 |
| Figure 5-27: Forest prototypes workflow.....  | 58 |
| Figure 5-28: Median Time Feature stacks of S-2 bands for generation of Tree Cover Density products ...  | 59 |
| Figure 5-29: Tree Cover Density workflow .....  | 60 |
| Figure 5-30: Examples of 100m x 100m polygon samples for validation of Tree Cover Density products, collected on the ESA DWH ADDITIONAL D2_MG2b_ECOL_011a dataset. .... | 61 |
| Figure 5-31: Distribution of 100m x 100m samples (blue) for validation of Tree Cover Density products.  | 62 |
| Figure 5-32: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demosite North.....  | 65 |
| Figure 5-33: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demosite North.....  | 66 |
| Figure 5-34: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demosite Central .....   | 67 |
| Figure 5-35: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demosite Central .....   | 68 |
| Figure 5-36: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demosite South-East....  | 69 |
| Figure 5-37: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demosite South-East....  | 70 |
| Figure 5-38: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demosite North .....  | 71 |
| Figure 5-39: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demosite Central....  | 72 |
| Figure 5-40: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demosite South-East   | 73 |
| Figure 5-41: Tree Cover Density 2018 in 10m spatial resolution for the demosite North.....  | 74 |
| Figure 5-42: Tree Cover Density 2018 in 10m spatial resolution for the demosite Central .....   | 75 |

|   |    |
|---|----|
| Figure 5-43: Tree Cover Density 2018 in 10m spatial resolution for the demosite South-East.....                                   | 76 |
| Figure 5-44: Distribution of LUCAS 2018 validation samples within the demosite South-East .....                                   | 78 |
| Figure 5-45: Cloud masking issues in the South-East demonstration site.. .....  | 81 |
| Figure 5-46: Proportion of DLT 2018 land cover classes per demonstration site .....   | 83 |
| Figure 5-47: Scatter plots for validation of the TCD 2018 prototypes .....  | 84 |
| Figure 5-48: Modified Change detection workflow applied to the forest prototypes.....   | 86 |
| Figure 5-49: Intermediate Products of the modified change detection workflow .....  | 86 |
| Figure 5-50: Incremental Update Layer Tree Cover Change 2017-2018 at 10m spatial resolution for the demonstration site North..... | 87 |
| Figure 5-51: Distribution of LUCAS 2018 points and additional random points for validation of the TCC 1718 layer.....             | 88 |
| Figure 5-52: Examples of forest loss provided by the TCC 1718 layer for the demosite North.....                                   | 89 |
| Figure 5-53: Examples of forest loss provided by the TCC 1718 layer for the demosite Central .....                                | 90 |
| Figure 5-54: Examples of forest loss provided by the TCC 1718 layer for the demosite South-East .....                             | 91 |
| Figure 5-55: Final statistics of the Incremental Update layers for all FOR demonstration sites.....                               | 92 |

## List of Tables

|   |     |
|---|-----|
| Table 3-1: Description of the ECoLaSS demonstration sites .....   | 6   |
| Table 5-1: Used Sentinel-2 reflectance bands (adapted from Suhet, 2015). ....   | 18  |
| Table 5-2: SAR annual statistical features. ....  | 25  |
| Table 5-3: Specifications of the ‘classified change’ layer .....  | 34  |
| Table 5-4: Confusion matrix of the internal validation of the IMD 2017 in demo site South-West .....                              | 41  |
| Table 5-5: Confusion matrix of the internal validation of the IMD 2018 in demo site South-West .....                              | 41  |
| Table 5-6: Confusion matrix of the internal validation of the IMD 2018 in demo site Central .....                                 | 41  |
| Table 5-7: Internal validation of the IMD 2018 in demo site South-East .....  | 42  |
| Table 5-8: Confusion matrix of the internal validation of the IMCC 1718 in demo site South-West .....                             | 42  |
| Table 5-9: Confusion matrix of the internal validation of the IMCC 1518 in demo site Central.....                                 | 43  |
| Table 5-10: Confusion matrix of the internal validation of the IMCC 1518 in demo site South-East .....                            | 43  |
| Table 5-11: Confusion matrix of the internal validation of the IBU 2018 in demo site South-West .....                             | 45  |
| Table 5-12: Confusion matrix of the internal validation of the IBU 2018 in demo site Central.....                                 | 45  |
| Table 5-13: Confusion matrix of the internal validation of the IBU 2018 in demo site South-East .....                             | 45  |
| Table 5-14: Proportional distribution of detected changes within the IMC layer (SW, phase 1) .....                                | 48  |
| Table 5-15: Proportional distribution of detected changes within the IMC layers (phase 2).....                                    | 50  |
| Table 5-16: Confusion matrix of the internal validation of the IMD 2018 in demo site South-West .....                             | 51  |
| Table 5-17: Confusion matrix of the internal validation of the IMD 2018 in demo site Central.....                                 | 51  |
| Table 5-18: Internal validation of the IMD 2018 in demo site South-West, Central and South-East .....                             | 52  |
| Table 5-19: Summary of Sentinel-2 and Sentinel-1 scenes pre-processed per Demo Site and Tile with a max. cloud cover of 60% ..... | 53  |
| Table 5-20: Land Cover statistics of the derived HRL 2015 sample layer.....   | 56  |
| Table 5-21: Parameter settings for TCM and DLT products .....   | 57  |
| Table 5-22: Sample distribution of the initial training dataset for TTCTM and DLT products per site. ....                         | 57  |
| Table 5-23: Accuracy figures of the internal validation & model training (demosite CENTRAL) .....                                 | 63  |
| Table 5-24: Accuracy figures of the internal validation & model training (demosite NORTH) .....                                   | 63  |
| Table 5-25: Accuracy figures of the internal validation & model training (demosite SOUTH-EAST) .....                              | 64  |
| Table 5-26: Model parameters for Tree Cover Density products .....  | 64  |
| Table 5-27: Error matrix for the improved TCM 2017 of the demosite NORTH .....  | 79  |
| Table 5-28: Error matrix for the improved TCM 2018 of the demosite NORTH .....  | 79  |
| Table 5-29: Error matrix for the improved TCM 2017 of the demosite CENTRAL .....  | 79  |
| Table 5-30: Error matrix for the improved TCM 2018 of the demonsite CENTRAL .....   | 80  |
| Table 5-31: Error matrix for the improved TCM 2017 of the demosite SOUTH-EAST .....   | 80  |
| Table 5-32: Error matrix for the improved TCM 2018 of the demosite SOUTH-EAST .....   | 80  |
| Table 5-33: Error matrix for the improved DLT 2018 status layer of the demosite NORTH .....                                       | 82  |
| Table 5-34: Error matrix for the improved DLT 2018 status layer of the demosite CENTRAL .....                                     | 82  |
| Table 5-35: Error matrix for the improved DLT 2018 status layer of the demosite SOUTH-EAST .....                                  | 83  |
| Table 5-36: Distribution of leaf type classes per demonstration site .....  | 83  |
| Table 5-37: Error matrix for the Incremental Update Layer TCC 1718 of the demosite NORTH. ....                                    | 89  |
| Table 5-38: Error matrix for the Incremental Update Layer TCC 1718 of the demosite CENTRAL .....                                  | 90  |
| Table 5-39: Error matrix for the Incremental Update Layer TCC 1718 of the demosite SOUTH-EAST .....                               | 91  |
| Table 5-40: Detailed specifications for the improved primary status layer Imperviousness 2017/2018 ...                            | 95  |
| Table 5-41: Detailed specifications for the built-up status layer 2018 .....  | 97  |
| Table 5-42: Detailed specifications for the Incremental Update Layer Imperviousness Change Classified                             | 98  |
| Table 5-43: Detailed specifications for the Incremental Update Layer Imperviousness Change .....                                  | 99  |
| Table 5-44: Detailed specifications for the improved Tree Cover Mask.....   | 101 |
| Table 5-45: Detailed specifications for the improved primary status layer Dominant Leaf Type.....                                 | 102 |
| Table 5-46 : Detailed specifications for the improved primary status layer Tree Cover Density .....                               | 103 |
| Table 5-47: Detailed specifications for the Incremental Update Layer Tree Cover Change .....                                      | 104 |
| Table 5-48: Colour palette for the probability layers.....  | 106 |

## Abbreviations

|         |   |
|---------|---|
| AD      | Applicable Document   |
| AL      | Active Learning   |
| ATCOR   | Atmospheric and Topographic Correction                                |
| BU      | Built Up Layer  |
| CI      | Confidence Interval   |
| CLC     | CORINE Land Cover   |
| CLMS    | Copernicus Land Monitoring Service                                    |
| CNES    | Centre national d'études spatiales                                    |
| CRM     | Crop Mask   |
| CRT     | Crop Type   |
| DAP     | Differential Attribute Profile  |
| DEM     | Digital Elevation Model   |
| DLT     | Dominant Leaf Type  |
| DMP     | Differential Morphological Profile                                    |
| DWH     | Data Warehouse  |
| ECoLaSS | Evolution of Copernicus Land Services based on Sentinel data          |
| EEA     | European Environment Agency   |
| EEA-39  | The 33 EEA member states plus the 6 cooperating Wets Balkan countries |
| EEE     | Entrusted European Entities   |
| EO      | Earth Observation   |
| EPSG    | European Petroleum Survey Group                                       |
| ESA     | European Space Agency   |
| ESM     | European Settlement Map   |
| EU      | European Union  |
| ETRS89  | European Terrestrial Reference System 1989                            |
| FADSL   | Forest Additional Support Layer                                       |
| FOR     | Forest  |
| FP7     | 7 <sup>th</sup> Framework Programme for research of the EU            |
| FRE     | Flat Reflectance  |
| FTS     | Fast Track Service  |
| GeoTIFF | Georeferenced TIFF (file format)                                      |
| GHSL    | Global Human Settlement Layer   |
| GIO     | GMES Initial Operations   |
| GIS     | Geographic Information System   |
| GMES    | Global Monitoring for Environment and Security                        |
| GRA     | Grassland   |
| H2020   | Horizon 2020  |
| HR      | High Resolution   |
| HRL     | High Resolution Layer   |
| IMC     | Imperviousness Change Classified                                      |
| IMD     | Imperviousness Density  |
| IMP     | Imperviousness  |
| INSPIRE | Infrastructure for Spatial Information in Europe                      |
| IRECI   | Inverted Red-edge Chlorophyll Index                                   |
| IRS     | Indian Remote Sensing satellite                                       |
| ISO     | International Organisation for Standardisation                        |
| ITT     | Invitation to Tender  |
| JRC     | Joint Research Centre   |
| LC      | Land Cover  |
| LM      | Land Monitoring   |

|         |  |
|---------|--|
| LU      | Land Use   |
| LUCAS   | Land Use/Cover Area frame statistical Survey                         |
| LZW     | Lempel-Ziv-Welch (data compression algorithm)                        |
| MACCS   | Multi-sensor Atmospheric Correction and Cloud Screening              |
| MAJA    | MACCS-ATCOR Joint Algorithm  |
| MAP     | Mapped value from product  |
| MAX     | Maximum  |
| MEAN    | Arithmetic mean value  |
| MGRS    | Military Grid Reference System                                       |
| MIN     | Minimum  |
| MMU     | Minimum Mapping Unit   |
| MR      | Medium Resolution  |
| MSGI    | Metadata Standard for Geographic Information                         |
| MUSCATE | Multi-satellite, multi-sensor ground segment for multi-temporal data |
| NDVI    | Normalized Difference Vegetation Index                               |
| NDVImax | Normalized Difference Vegetation Index maximal                       |
| NDWI    | Normalized Difference Water Index                                    |
| OA      | Overall Accuracy   |
| OSM     | Open Street Map  |
| PA      | Producer Accuracy  |
| PSU     | Primary Sampling Units   |
| RF      | Random Forest  |
| RSG     | Remote Sensing Software Graz   |
| S-1     | Sentinel-1   |
| S-2     | Sentinel-2   |
| S2DSL   | Sentinel-2 Data Score Layer  |
| SAR     | Synthetic Aperture Radar   |
| SCL     | Scene Classification Layer   |
| SPOT    | Satellite Pour l'Observation de la Terre                             |
| SRE     | Surface Reflectance  |
| SRTM    | Shuttle Radar Topography Mission                                     |
| SSU     | Secondary Sampling Units   |
| STD     | Standard Deviation   |
| SVM     | Support Vector Machines  |
| SWF     | Small Woody Features   |
| TCC     | Tree Cover Change  |
| TCD     | Tree Cover Density   |
| TCM     | Tree Cover Mask  |
| TIFF    | Tagged Image File Format   |
| UA      | User Accuracy  |
| USGS    | United States Geological Survey                                      |
| UTM     | Universal Transverse Mercator  |
| VH      | Vertical transmit/Horizontal receive (polarisation)                  |
| VHR     | Very High Resolution   |
| VHR1    | Very High Resolution 1 where resolution <=1m                         |
| VHSR    | Very High Spatial Resolution   |
| VHR2    | Very High Resolution 2 where 1m < resolution <=4m                    |
| VV      | Vertical transmit/Vertical receive (polarisation)                    |
| WaW     | Water and Wetness  |
| WGS84   | World Geodetic System 1984   |
| WP      | Work Package   |
| XML     | Extensible Markup Language   |

# 1 Introduction

The Horizon 2020 (H2020) project, “Evolution of Copernicus Land Services based on Sentinel data” (ECoLaSS) addresses the H2020 Work Programme 5 iii. Leadership in Enabling and Industrial technologies - Space, specifically the Topic EO-3-2016: Evolution of Copernicus services. ECoLaSS is being conducted from 2017–2019 and aims at developing and prototypically demonstrating selected innovative products and methods as candidates for future next-generation operational Copernicus Land Monitoring Service (CLMS) products of the pan-European and Global Components. ECoLaSS assesses the operational readiness of such candidate products and eventually suggests some of these for implementation. This shall enable the key CLMS stakeholders (i.e. mainly the Entrusted European Entities (EEE) EEA and JRC) to take informed decisions on potential procurement as (part of) the next generation of Copernicus Land services from 2020 onwards.

To achieve this goal, ECoLaSS makes full use of dense time series of High-Resolution (HR) Sentinel-2 optical and Sentinel-1 Synthetic Aperture Radar (SAR) data, complemented by Medium-Resolution (MR) Sentinel-3 optical data if needed and feasible. Rapidly evolving scientific developments as well as user requirements are continuously analysed in a close stakeholder interaction process, targeting a future pan-European roll-out of new/improved CLMS products, and assessing the potential transferability to global applications.

This Deliverable **D12.2: “D42.1b - Prototype Report: Consistent HR Layer Time Series/Incremental Updates”** aims to demonstrate the feasibility of future consistent time series as well as incremental update products with a focus on Imperviousness and Forest which have the longest time series history. These improved HRL prototypes are implemented in specific demonstration sites in Europe. This work is part of **WP 42** of Task 4: “Thematic Proof-of-Concept/Prototype on Continental/Global Scale”. This report is accompanied by the Deliverable **D12.4: “P.42.2b - Data Sets of HR Layer Incremental Updates”**. This report serves as documentation of the prototype datasets.

In the ECoLaSS project, a prototype is defined as a thematic proof-of-concept implementation of an improved or newly defined potential future Copernicus Land layer, building on the methods and processing lines developed in Task 3. The consortium has selected representative demonstration sites both in Europe and Africa, covering various bio-geographic regions and biomes. All prototype products and services are prototypically implemented in a selection of these sites in the frame of the Task 4 WPs. In ECoLaSS, proofs-of-concept / prototype demonstrations are carried out with respect to five topics of relevance: (i) Time series derived indicators and variables (WP 41), (ii) Incremental Updates of HR Layers (WP 42), (iii) Improved permanent grassland identification (WP 43), (iv) Crop area and crop status / parameters monitoring (WP 44), and (v) New LC/LU products (WP 45).

After methods have been tested by the Task 3 WPs (cf. Deliverables D31.1b, D32.1b, D33.1b, D34.1b, D35.1b) in all test sites and algorithms are described, the demonstration activities of Task 4 have set up the developed processing lines in the demonstration sites in order to derive the final prototypes. This deliverable focusses on the two time series consistency/incremental update prototypes on **Imperviousness and Forest** as part of **WP 42**.

This report comprises a Chapter on the background to an Imperviousness and Forest Prototypes and associated requirements (Chapter 2); a description of the Demonstration Sites where the prototypes are implemented (Chapter 3); an overview of the methodologies carried over from the testing and benchmarking in Task 3 (Chapter 4); followed by a Chapter on the prototype implementation itself, the results and validation, including a description of the dataset (Chapter 5); and a summary and outlook (Chapter 6).

## 2 HRL Incremental Update Needs and Feasibility

This section discusses the feasibility of incremental updates of the HR Layers, with a focus on preserving a consistent product time series. At present, the investigations focus on the two most mature HRLs with the longest time series, i.e. Imperviousness and Forest. During the project timeline, exploring the potential for incremental updates has become relevant specifically also for the HRL Grassland, for which in phase 1 only the 2015 implementation (with the current product specifications) had been published. However, at present, the direction to be taken appears largely dependent on the HRL 2018 operational ITT already launched by the EEA and currently under production. In ECoLaSS, Grasslands improved status layers, change production and the incremental updates are addressed in WP 43 (and the associated Task 3 WPs 33, 34 and 35). Likewise, the HR Layer Water & Wetness 2015 was just recently been published by the EEA in phase 1, and this significantly differed in its definition from the other HRLs, making use of continuous radar based observations over seven consecutive years. Conceptualising a change product for this HRL will be a challenge of its own, and also largely depends on the direction which EEA has taken in the operational HRL 2018 ITT. The HRL on Small Woody Features (SWF) is anyway out of scope for incremental updates due to much higher spatial (VHR) resolution and associated longer EO coverage repetition cycles. For these reasons, the report currently addresses the HRLs Imperviousness and Forest.

The IMD layer was the first HRL to be produced for the reference year 2006. However, at that time it was described as a “soil sealing database for Europe” and delivered as part of the GMES (Global Monitoring for Environment and Security, former name of Copernicus) Fast Track Service on Land Monitoring (Land FTS LM). It was produced during 2006-2008 from multi-sensor and bi-temporal, orthorectified satellite imagery. The production of IMD2006 covered 38 European countries (32 EEA Member States and 6 West-Balkan countries) and was implemented in two phases:

- an initial soil sealing product based on the EEA specification, and
- a soil sealing enhancement product based on evaluation of the initial product by some Member States.

The production approach used an automatic algorithm based on calibrated NDVI. The main deliverable was a raster dataset of continuous degree of soil sealing ranging from 0 – 100 % in full spatial resolution (20 m x 20 m) with no MMU (Minimum Mapping Unit, i.e. the minimum number of pixels to form a patch) and with the associated metadata. A derived product, a raster dataset of continuous degree of soil sealing ranging from 0 - 100% in aggregated spatial resolution (100 m x 100 m) in European projection was generated.

Since the production of the HRL IMD 2006, a time series of imperviousness has been delivered for the reference years 2009, under the FP7 geoland2 project, and 2012, under the GMES Initial Operations (GIO). The lastest update for 2015 has been produced and the current one for 2018 is in production at the moment of the writing of thie report. For each of these iterations, the results contain two products: a status layer for each reference year (e.g. IMD2012), as well as an imperviousness density change layer between reference years (e.g. IMC2009-2012), and based on the already existing imperviousness product for that previous reference year. It is worth noting that there have been revisions to the previous years' IMD products during updating and the IMC time series has been reworked. The update for 2012 and 2015 was produced as part of the operational Copernicus Land Monitoring Service (CLMS), on the same procedure and for the same product types. A re-analysis of all existing Imperviousness Degree (IMD 2006, 2009 and 2012) and Imperviousness Change products (IMC 2006-2009, 2009-2012, 2006-2012) was also conducted during the production of IMD2015. A new layer regarding the presence of buildings in the impervious surfaces, called Built-Up layer (IBU2018), has been requested in the tender for the 2018 production and its feasibility is demonstrated in this project.

The most recent HRL IMD 2015 production was conducted through semi-automatic classification of calibrated biophysical parameters, derived from multi-temporal high-resolution optical satellite images from the reference period 2015 +/- 1 year. The result for the pan-European IMD2015 product shows a producer's accuracy of 84.58% and a user's accuracy of 90.14%, as determined by visual checks of open source VHR images in combination with the multi-temporal, multi-seasonal EO database (e.g. Sentinel-2, Landsat). The production of the IMD2015 was achieved in a very short time (15-18 months) thanks to multi-sensors and multi-temporal time series data available through the European Space Agency's (ESA) Data Warehouse (DWH) (mainly SPOT, IRS-P6, Sentinel-2A and Landsat). The availability of time-series along with Landsat-8 for gap filling permitted to shorten the production phase compared to 2012. But the user's point of view (as expressed by various stakeholders, cf. WP 21 results) typically is to rather have the products available as close as possible to the reference year of image acquisition and thus in an even more shortened time. The reduction of the production time is tied to the question of the input data and their availability for all areas in sufficient quantity and reliable quality at the beginning of the project. The higher spatial resolution of the time series used for IMD2015 leads to generate products in the actual resolution of 20m (compared to former used data from the ESA DWH such as IRS with a nominal resolution of 23.5m or Landsat 8 with a nominal resolution of 30m). It is noticeable that the use of full S-2 time series will also lead to an improvement of the currently applied spatial resolution from 20m to 10m – which has already been taken into account in the 2018 ITT for the IMP production by the EEA. As a proof-of-concept, in ECoLaSS the IMP prototypes are actually produced at 10 m spatial resolution.

The HRL Forest at the status of the reference year 2015 consisted of two pixel-based primary status layers at 20m spatial resolution: the Dominant Leaf Type (DLT), providing information on the corresponding leaf type (broadleaved or coniferous) per pixel, and the Tree Cover Density (TCD), providing information on the proportional tree (crown) coverage per pixel. These two status layers are fully identical in their spatial extents and are currently produced for the EEA-39 area in a 3-year update cycle for certain reference years (2012, 2015, 2018, etc.), whereas the reference year is defined as comprising +/- 1 year. Both products represent the main input data source for any derived layers of the HRL Forest product portfolio and enable users to apply any (national) forest definition (when being further processed to a desired TCD and MMU), whichever fits best to the specific needs.

However, the data situation of the reference year 2012 and 2015 differed significantly in terms of temporal data availability, available sensors and, connected to both: data quality. Whereas in 2012, a mono-temporal optical EO data coverage from 2011 to 2013 had been used in form of the European Space Agency's (ESA) Data Warehouse (DWH) dataset MG2b\_CORE\_01 (also referred to as HR\_IMAGE\_2012, consisting of SPOT-4, SPOT-5, IRS-P6, IRS-ResourceSat-2 and RapidEye data), multi-temporal time series data became available for the 2015 production with the launch of Sentinel-2A and the utilization of the Landsat archive of the United States Geological Survey (USGS), strongly focussing on the acquisition year 2016 (i.e. reference year 2015 +1). Consequently, the production workflow of the HRLs 2015 was already adopted towards the processing of mass data and thematic analysis of multi-temporal EO data, resulting in high quality products, which are partially not directly comparable with the derived 2012 products due to their specific lineage (e.g. mono-temporal data basis, radiometric resolution, data quality issues -haze, cloud cover, phenology-, distribution in geographical production units –lots-, etc.).

Even though significant improvements in the production speed were achieved in the 2015 production (15-18 months) compared to 2012 (40 months), users argue that the time span between production start and final provision to the end-users should be shortened, and that product specifications could be improved in terms of spatial and temporal resolution (cf. WP 21). It has to be noted that the HRL Forest 2015 has already strongly benefitted from a generally good data situation, with more than 80% of the image acquisitions having taken place within one year. ECoLaSS is able to explore taking the full advantage of the complete Sentinel-2A+B and Sentinel-1A+B time series, which provides a significantly improved temporal resolution, supporting an improved tree cover and leaf type mapping as well as change detection approaches.

Even though the HRLs 2015 on Imperviousness and Forest are of high quality, requirements for improved products have been voiced by users and stakeholders in WP 21 and WP 51 as well as in terms of ongoing Copernicus developments. Moreover, ECoLaSS aims to demonstrate the feasibility of HRL incremental updates (WP 35 [AD 09]). The recognised potential for improvement of the methodological approaches for both Layers can be summarised as follows:

- an improved level of automation to allow a faster production and
- shorter monitoring intervals (e.g. for future yearly incremental updates);
- improving the thematic classification accuracy
- fully exploiting optical Sentinel-2 and SAR Sentinel-1 time series instead of using limited temporal EO data coverages or pre-selected, best-suited EO data scenes
- applying an integrated SAR/optic time series data analysis to benefit from the multi-sensor characteristics and ability for gap filling of clouds
- improving the status layer's spatial detail from 20m spatial resolution to 10m. Product definitions consequently might have to be adapted, such as e.g. the Minimum Mapping Unit
- refine the change detection approach to detect both increase and decrease of Imperviousness or Forest areas

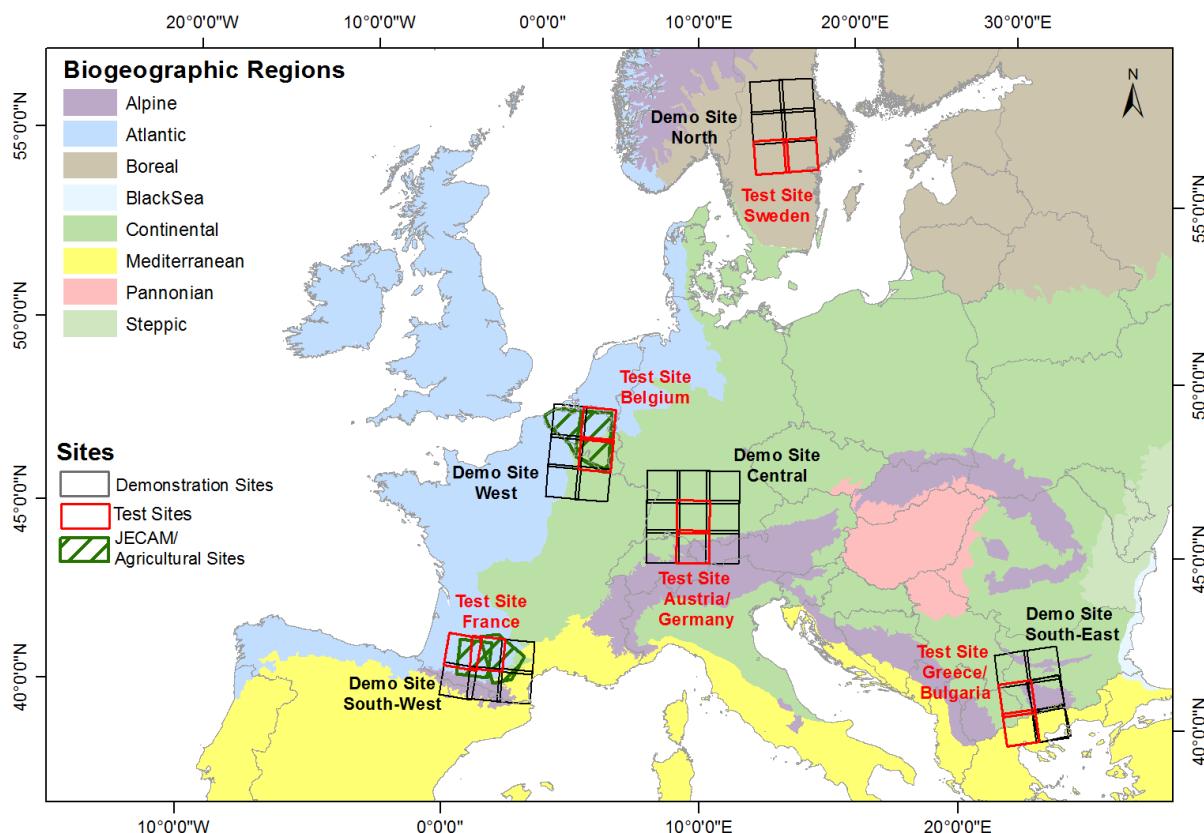
These points have been considered and implemented in the technical developments within ECoLaSS and the specifications of the recent HRL2018 production are benefiting from these testbeds experiences. Both, the HRL Imperviousness 2018 and the HRL Forest 2018 are now consistently produced at an enhanced 10 m spatial resolution, providing an unprecedented level of detail and accuracy.

The tested methods for the Imperviousness and Forest prototypes, addressing the abovementioned user demands, are presented in Task 3 deliverables (WP 33 [AD 07], WP 34 [AD 08], WP 35 [AD D09]) and the implementation in the larger area prototypes is subsequently presented in the next chapters, alongside with investigations on the feasibility for future operational implementation of such prototypes at pan-European and global scales.

### 3 Demonstration Sites

The incremental update prototypes investigated by WP 42 (i.e. Imperviousness and Forest) are implemented in selected representative demonstration sites which cover various bio-geographic regions and biomes, as shown in the following.

The selected demonstration sites (60,000/90,000km<sup>2</sup> per demonstration site) spatially contain the 5 test sites used by Task 3. The demonstration sites are relevant in Task 4 for demonstrating the proposed candidates for a Copernicus Land Service Evolution roll-out on a larger scale. As shown in Figure 3-1, the pre-selected demonstration sites are covering the Atlantic and Continental zones (Source EEA: <http://www.eea.europa.eu/data-and-maps/data/biogeographical-regions-europe-3#tab-gis-data>) of the member and associated states of the EEA-39, and are determined by the position of Sentinel-2 tiles of the official tiling grid. These selected ECoLaSS demonstration sites are located in the North of Europe, in the Alpine/Central region, in the West, in the South-West and in the South-East of Europe.



**Figure 3-1: Biogeographic Regions of Europe (2015) and European ECoLaSS Demonstration Sites**  
*(Map: © European Environment Agency; © EuroGeographics for the administrative boundaries)*

A short description of these prototype sites is given in Table 3-1 below. The land cover figures are derived from CORINE Land Cover (CLC) 2018 data with a Minimum Mapping Unit (MMU) of 25 ha. Therefore, strong generalisation effects have to be assumed.

**Table 3-1: Description of the ECoLaSS demonstration sites**

| Location                | Biogeographical region(s)          | Countries   | Distribution of CORINE land cover classes 2018 (Level 1) per demonstration site  |
|-------------------------|------------------------------------|---|--|
| Northern Europe         | Boreal                             | Sweden, Norway  | Artificial areas: 1.90%, Agricultural areas: 11.87%, Forest and semi-natural areas: 69.01%, Wetlands: 3.25%, Waterbodies: 13.94% |
| Alpine / Central Europe | Continental, Alpine                | Germany, Austria, Switzerland, Italy and Czech Republic | Artificial areas: 9.03%, Agricultural areas: 44.55%, Forest and semi-natural areas: 44.65%, Wetlands: 0.23%, Waterbodies: 1.55%  |
| West Europe             | Atlantic, Continental              | Belgium, France, Luxembourg                             | Artificial areas: 13.47%, Agricultural areas: 63.08%, Forest and semi-natural areas: 21.43%, Wetlands: 0.39%, Waterbodies: 1.61% |
| South-East Europe       | Mediterranean, Continental, Alpine | Serbia, Macedonia, Greece, Bulgaria and Kosovo          | Artificial areas: 3.34%, Agricultural areas: 34.87%, Forest and semi-natural areas: 56.67%, Wetlands: 0.17%, Waterbodies: 4.93%  |
| South-West Europe       | Atlantic, Mediterranean, Alpine    | France, Spain   | Artificial areas: 3.26%, Agricultural areas: 46.73%, Forest and semi-natural areas: 49.02%, Wetlands: 0.01%, Waterbodies: 0.40%  |

\* figures for “Agricultural areas” comprise both, arable land and grassland

All prototype products and services were prototypically implemented in one or more demonstration sites in the first Reporting Period, and are now complemented to a total of three demonstration sites in Reporting Period 2. In the first Reporting Period, the Imperviousness prototype was implemented in the South-West demonstration site (which covers the Mediterranean, Atlantic and Alpine biogeographic regions), and the Forest prototype in the North demonstration site (which covers the Boreal zone). In phase 2, the demonstration sites where the Imperviousness prototypes are developed are Central, South-West and South-East, and in turn the Forest prototypes are located in the Central, South-East and North demo sites. Since these two demonstration sites constitute the more relevant ones for WP 42, they are shortly described in the following sections.

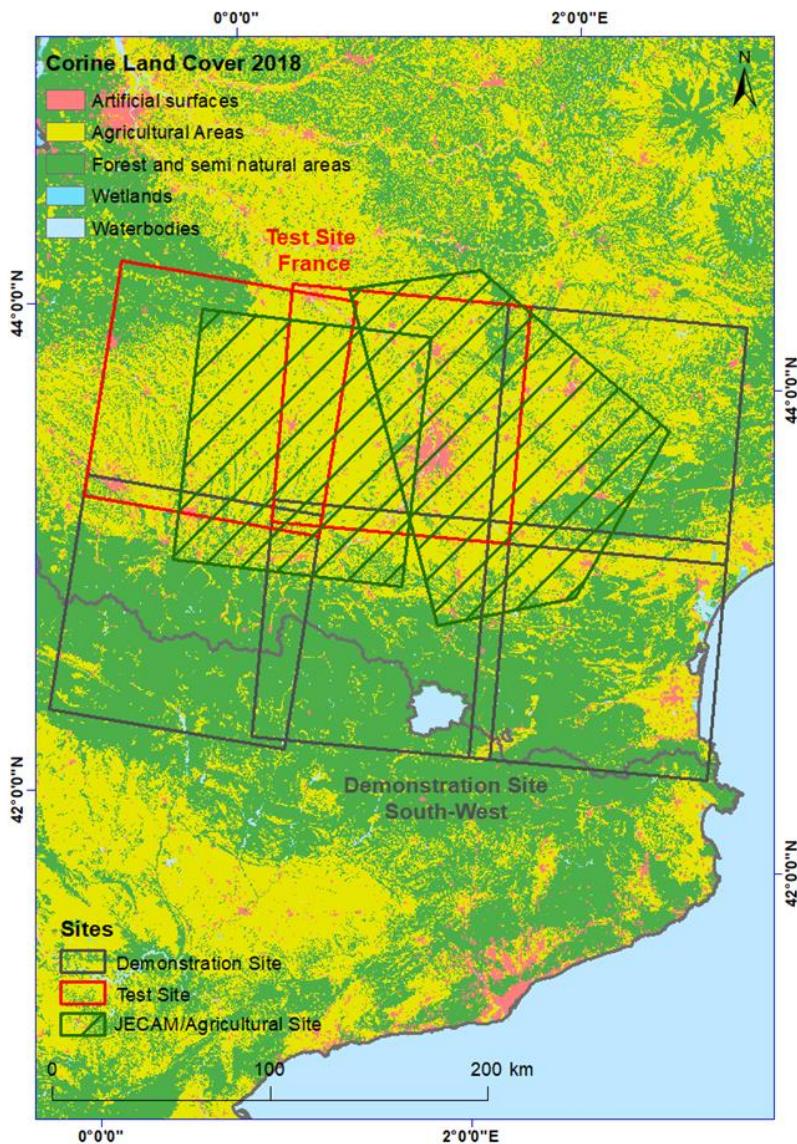
### 3.1 Demonstration Site South-West

The demonstration site South-West (approx. 65,000 km<sup>2</sup>) is covering southern France and some parts of northern Spain, and includes the primary test site for the method developments in Task 3 related to the improvement of the HRL Imperviousness (IMP). It serves to demonstrate the implementation of the prototype of a potential future HRL Imperviousness, as part of WP 42, and the New Land Cover products, as part of WP 45, in the first Reporting Period. In the second project phase, the demonstration sites Central and South-East have been added for the Imperviousness prototyping.

The landscape in the demonstration site South-West is composed of different biogeographic regions such as Mediterranean, Alpine and Atlantic. Three Sentinel-2 tiles are dominated by mountain landscapes, a mix of bare soils and natural grasslands, due to the presence of the Pyrenees. The 31TCJ tile is dominated by a strong proportion of impervious surfaces, because of Toulouse, a major French city. Toulouse is the 4<sup>th</sup> city in France in terms of urban and demographic expansions and is the most dynamic city of the South-West region. The presence of the city leads the region to be the second most attractive and dynamic region in France. The plains surrounding the city are mainly rurally dominated areas composed of croplands mixed with grassland and an increasing amount of forest with the proximity of the coastal region. But rurally dominated areas also show a dynamic increase of population and a dynamic situation of settlements due to the proximity of Toulouse. So the region shows a real tendency towards urban expansion in the surroundings of Toulouse and in the rural parts of the region with small cities like Montauban, Auch,

Carcassonne, Tarbes, Castres or Albi. In general, the Mediterranean area in the East of the demonstration site is a patchwork of cropland, dry grassland and vineyards. There is also a small portion of the Landes forest in the North-West of the demonstration site. The unique situation of the region in terms of urban dynamic is the reason for the selection of this site for the Imperviousness prototype.

A more detailed map of the characteristics of the selected demonstration site South-West as used for the Imperviousness prototype is provided in Figure 3-2 below.



**Figure 3-2: Demonstration site South-West, with CLC 2018 background layer**

(©European Environment Agency, © EuroGeographics for the administrative boundaries)

### 3.2 Demonstration Site North

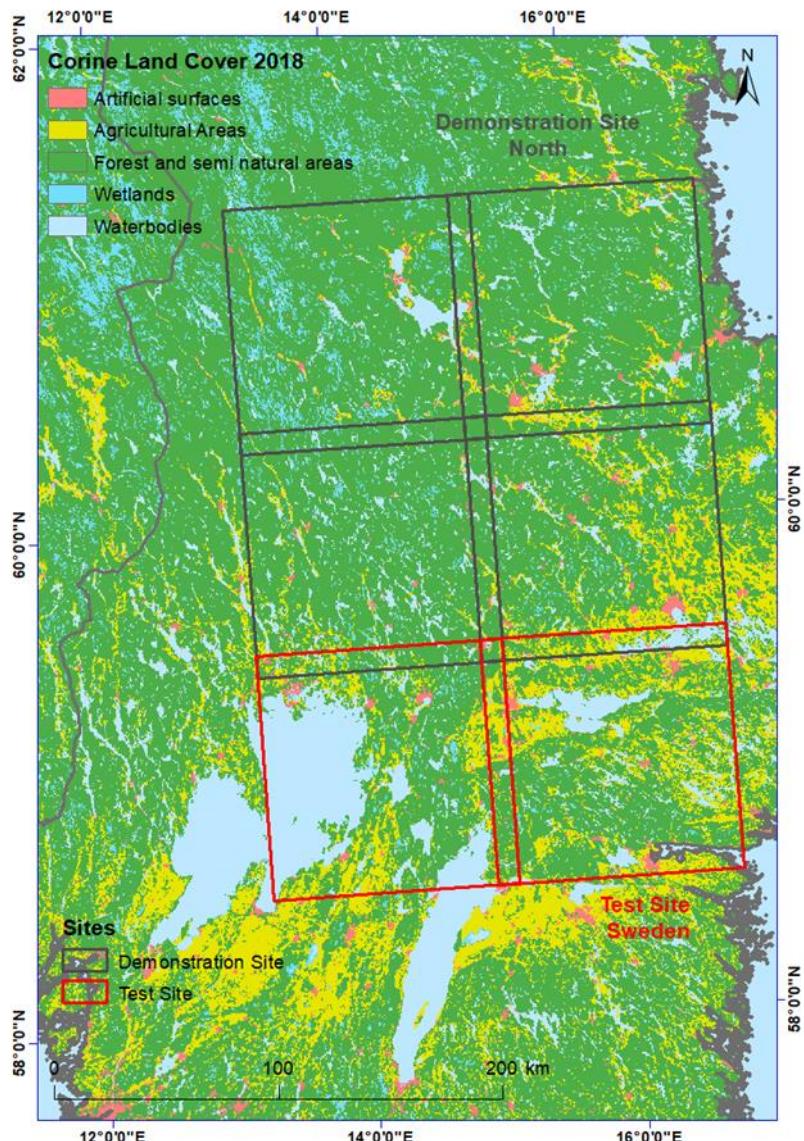
The demonstration site North in southern Sweden includes the primary test site for the method developments in Task 3 related to the improvement of the HRL Forest (FOR) in phase 1. It served to demonstrate the implementation of a prototype on a potential future HRL Forest status layer, and the incremental update thereof, as part of WP 42, in the first Reporting Period. In the second project phase the demonstration sites Central and South-East have been added as forest prototypes sites.

The site has a spatial extent of approx. 65,000 km<sup>2</sup> and is dominated by forests and water bodies. In some central parts and in the South, larger agricultural areas are included. The forested land is intersected by lakes and smaller water bodies, as well as by peat bogs and grassland.

The particular challenges of the demonstration site North are related to:

- a typically medium to high cloud cover throughout the year;
- the existence of very large tree and forest stands within the area;
- extensive forest management practices with clear cuts frequently > 0.5 ha, resulting in a fragmented pattern of forest stand ages and densities due to frequent tree harvesting;
- a generally difficult differentiation between the two dominant leaf types (broadleaved / coniferous) due to:
  - the influence of soil moisture and water content (small water bodies and peat bogs);
  - over-radiation effects of vital vegetative undergrounds (e.g. grassland and bushes) in less dense coniferous tree/forest stands.

A map of the selected demonstration site North is provided in Figure 3-3 below.



**Figure 3-3: Demonstration site North with CLC 2018 background layer**  
 (@European Environment Agency, © EuroGeographics for the administrative boundaries)

### 3.3 Demonstration Site South-East

The demonstration site “South-East” incorporates the areas of the test sites, and serve for demonstrating the proposed candidates for Copernicus Land Service evolution in terms of roll-out to a larger scale (Task 4) in ECoLaSS for the following thematic topics addressed: Grasslands, Forest and Imperviousness. The selected prototype sites cover the most important environmental zones (Source: EEA) of Europe and the member states of the EEA-39. In particular, South-East site is located in the Mediterranean and Steppic European region (Greece/Bulgaria), including continental and alpine areas.

Within the demonstration site South-East, with an area of approximately 65545.32 km<sup>2</sup>, 34.87% are agriculturally used while the rest is covered by water bodies (4.93%), forests and semi-natural areas (34.87%), artificial areas (3.34%), as well as wetlands (0.17%). In this statistic the various grassland types appear under different land cover classes. Natural grasslands (CLC code 321) with 5.73% are generalised under forests and semi-natural areas, whereas pastures (CLC code 231) with 2.56 % are generalised under agricultural areas. Further, green urban areas (CLC code 141) with 0.04% and sport and leisure facilities (CLC code 142) with 0.08% are generalised under artificial areas. A map of the selected demonstration site South-East for the HRL Grassland (GRA) prototype is provided in Figure 3-4 below.

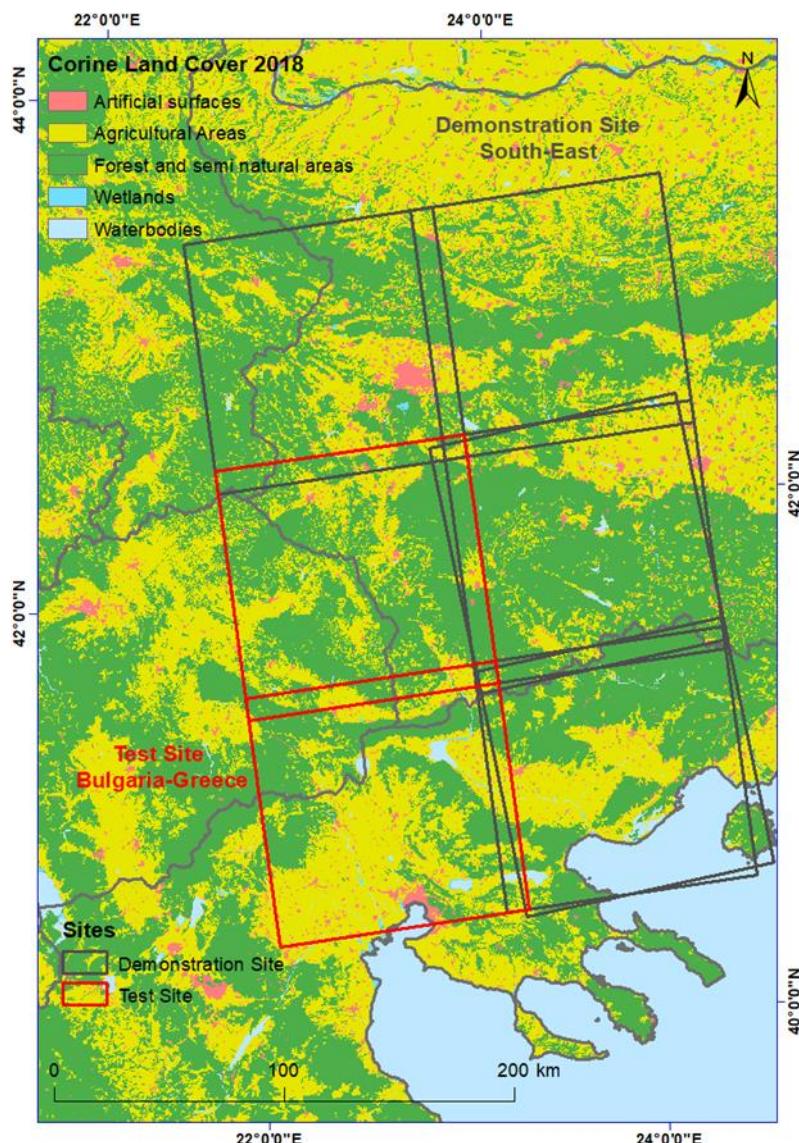


Figure 3-4: Overview of DEMO Site South-East.

### 3.4 Demonstration Site Central

The demonstration site “Central” incorporates the areas of the test sites, and serve for demonstrating the proposed candidates for Copernicus Land Service evolution in terms of roll-out to a larger scale (Task 4) in ECoLaSS for all thematic topics addressed: Grasslands, Forest, Imperviousness, Indicators/Variables, Agriculture and New Land Cover products (including the CLC approach and the consistency of the combined HRL layers approach). The selected prototype sites cover the most important environmental zones (Source: EEA) of Europe and the member states of the EEA-39. In particular, Central site is located in the Continental and Alpine Central European region (Germany, Austria, Switzerland, Italy). During phase 1, the sites were reviewed as was explained in the Periodic Technical Report and second issue of the Interim Progress Report [AD10]. Accordingly, the Central site remains a large site with 9 Sentinel-2 granules (approx. 90,000 km<sup>2</sup>).

The test site in **Germany/Austria** is dominated in the North by cropland areas, mixed with grassland (pastures). The Southern part, covering the Bavarian Alpine Foreland, is dominated by forest cover and grassland, including extensively used grassland and wetland areas. The test site covers the “Wetterstein mountain range” as part of the Alps with mountain-specific vegetation zones and stretches South down to the Inn valley. The whole demo is very challenging regarding grasslands and agriculture discrimination as it is heterogeneous in terms of climate/microclimate conditions as well as in terms of altitude leading to heterogeneous farming management systems (time windows for sowing and harvesting) as well as to shifted vegetation periods. On the North/South gradient, there is a high variety of croptypes especially in the more favourable areas of north and central tiles), with more fodder crops/grasslands and pastures towards the southern areas closer to the Alps. In the alpine region agricultural areas are restricted mainly to valleys whereas grassland and pastures are dominant in higher altitudes. In the western part, towards the north, the River Rhine valley, with very mild climate even in wintertime, favours the presence of vineyards and agriculture uses, whereas towards the south the black forest domains and the agriculture land use is present up to moderate altitudes. Croplands are also present along the Danube riverside and lake Constanza. According to CLC, within the demonstration site Central, with an area of approximately 96013.29 km<sup>2</sup>, 44.55% are agriculturally used while the rest is covered by water bodies (1.55%), forests and semi-natural areas (44.64%), artificial areas (9.03%), as well as wetlands (0.22%). In this statistic the various grassland types appear under different land cover classes. Natural grasslands (CLC code 321) with 3.76% are generalised under forests and semi-natural areas, whereas pastures (CLC code 231) with 16.12 % are generalised under agricultural areas. Further, green urban areas (CLC code 141) with 0.06% and sport and leisure facilities (CLC code 142) with 0.42% are generalised under artificial areas. This shows that the demonstration site Central is mainly comprised by forests and seminatural grasslands and agricultural grasslands. A map of the selected demonstration site Central for the HRL Grassland (GRA) prototype is provided in Figure 3-5 below.

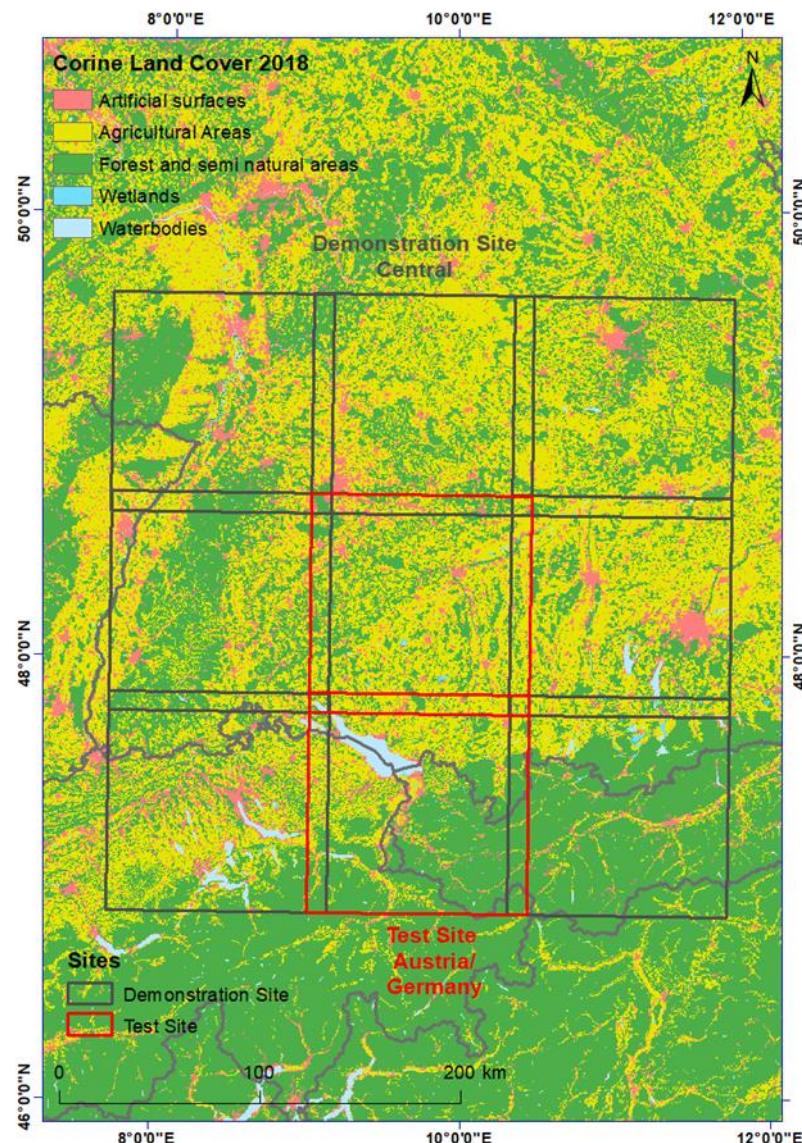


Figure 3-5: Overview of DEMO Site Central

## 4 Overview of Applied Methods

In the following, a brief overview on the methods applied to derive the prototypes on Imperviousness and Forest is given. Further details and background information are provided in [AD07], [AD08] and [AD09] as outcomes of Task 3.

### 4.1 Methods for Implementing a Prototype on Imperviousness

This section shows an overview of the methods implemented for the Imperviousness prototypes Imperviousness Degree (IMD) and Imperviousness Change Classified (IMC) following the outcomes of the Task 3. Further details and background information are provided in the three Deliverables on “D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification” [AD07] and “D34.1b - Methods Compendium: Time Series Analysis for Change Detection” [AD08], as well as “D35.1b- Methods Compendium: HRL Time Series Consistency for HRL Product Updates” [AD09], as outcomes of Task 3.

#### 4.1.1 Derivation of biophysical variables

Many spectral indices and biophysical variables have been defined in the past three decades. Some have been and are still widely used, such as the Normalized Difference Vegetation Index (NDVI), while others have been proposed as alternatives only in the recent years. Thus, the PanTex textural index is now known to bring interesting results in order to study urban (especially in the frame of built-up features). In WP 31, further details on spectral indices are listed, focusing on vegetation and built-up discrimination in order to explore their potential contribution to the evolution of CLMS products.

In the frame of the Task 4, it has been chosen to focus on the NDVI, adapted for the Sentinel-2 sensor as defined by Henrich et al. (2012) and the PanTex to focus on the built-up areas as shown by (Pesaresi et al., 2013). Further details on the NDVI and the PanTex are provided in WP 31. Conceptual details on the NDVI are provided in WP 31. However, although multispectral information such as from the NDVI is essential to discriminate landscape elements, it has limitations in the detection of built-up area and calculation of imperviousness density. An effective detection could require advanced feature computing able to discriminate features:

- **Texture / Structure:** Texture and structure analysis consists in extracting information on the spatial arrangement of pixels. Amongst numerous existing techniques, the Sobel filter is particularly interesting. The **Sobel operator** is used in image processing particularly within edge detection algorithms where it creates an image emphasizing edges.
- **Granulometry by Mathematical morphology:** Mathematical morphology is the analysis of the image constructions and their distribution at different scale. It consists in simplifying the image progressively through the preservation of bright elements (with closing operators) or dark elements (opening operators). Amongst numerous existing techniques, the following one is particularly interesting and is implementing for the HRL IMD 2017 Prototype:

**Differential Attribute Profiles (DAP)/Haar attributes:** Multiscale features often appear as a relevant alternative, with Gabor filters and Differential Morphological Profile (DMP) having achieved great classification performances. However, even such features come with a significant cost. DMP is relying on a series of morphological filters by reconstruction and it has shown for more than a decade its ability to deal with VHSR images (Pesaresi & Benediktsson, 2001). Recently, an alternative multiscale feature, called Differential Attribute Profile (Dalla Mura et al., 2010) has been built upon DMP to achieve more discriminative power, a higher flexibility, for a lower computational cost. DAP is very appealing since it is computed from a tree-based image representation that can be built with very efficient algorithms (see Carlinet et al. (2014) for a review). Once the tree is built, the description of each pixel (or object, node) is straightforward and relies on the analysis of all its ancestors up to the root. As such, it has been embedded in large-

scale analysis performed by the Joint Research Center such as the Global Human Settlement Layer (GHSL Release 2019) (Pesaresi et al., 2013) and European Settlement Map (ESM Release 2019) (Florczyk et al., 2015).

#### 4.1.2 Active learning Classification approach based on DAP profiles

Multiple algorithms could be used to map artificial lands. Classification methods range from unsupervised algorithms such as K-means to parametric supervised algorithms to machine learning algorithms such as artificial neural networks (Mas & Flores, 2008), decision trees (Breiman, 1984), Support Vector Machines (SVM) (Mountrakis et al., 2011) and ensembles of classifiers such as Random Forest (Breiman, 2001). A selection of these best algorithms for classification has been tested and detailed in the frame of the WP 33 [AD 07].

The automated supervised classification used to derive the built-up mask follows the outcomes of the Task 3 of the project. So the Imperviousness built-up mask layer for 2017 was performed using supervised machine learning methods to create the updated built-up mask for 2017. The production of the built-up mask is achieved with a selection of reference (or training) data. Following the results of the WP 31 (separability of the information for thematic classifications) and WP 33 (Time Series Analysis for Thematic Classification), the input data selected rely on multispectral information and granulometry by mathematical morphology (Differential Attribute Profiles). Indeed, the active learning algorithm shows great classification performances whilst being very computer efficient, thus substantially reducing processing time overall and dealing with large dataset, as it has been tested for the IMD layers over South-West in phase 1 and then applied to the BU layer over the South-West, Central and South-East test sites in phase 2.

##### Active Learning (AL) and Differential Attribute Profiles (DAP)

Production of Land Cover maps is usually achieved with a selection of reference (or training) data, supervised classification, and manual map refinement/correction. The classification accuracy is directly related to the quality of the training samples, i.e. their ability to represent the data to be classified. Collecting training samples is done through a costly operation consisting of manually labelling the pixels. Furthermore, such pixels may not be representative of the land cover classes, thus requiring important corrections in the post processing step. To alleviate these issues, active learning has been introduced a couple of decades ago, and used in remote sensing since more than 5 years (Tuia et al., 2009). It works in both interactive and batch mode. In the former case, the user is given some specific pixels to label (e.g. by photo-interpretation), while in the latter case only relevant samples from the training sets will be used (leading to a better modelling of land cover classes as well as a more efficient classification process). It has been a very active field of research (see Tuia et al. (2011) for a review) reaching similar accuracies than supervised classifiers but with only 5 to 10% of the training samples. It is now considered as a well-established framework (Crawford et al., 2013). Recent developments are related to large-scale analysis and domain adaptation (Alajlan et al., 2013) or multiscale classification (Zhang et al., 2016).

##### Set-up of reference databases for validation

A stratified random sampling approach is used adapting the number of sample units to each stratum. For the HRL IMD 2017 and 2018 Prototypes, a stratification has been applied on a series of omission/commission strata. The number of sample units called Primary Sampling Units (PSUs) per stratum based on LUCAS and densified LUCAS grid ensured a sufficient level of precision at reporting level. Different sampling intensity has been applied and the stratification was defined as follows:

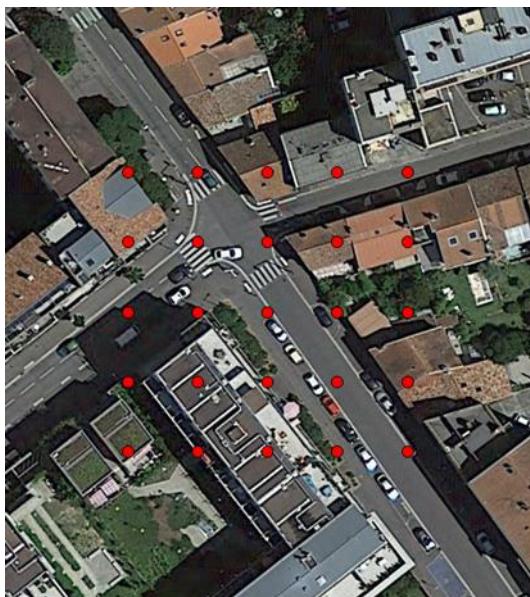
- Commission: impervious class (1-100%);
- Omission high probability: no IMP areas and CLC/OSM IMP areas;
- Omission low probability: Remaining areas.

In the phase 2 of the project, an additional stratum has been introduced to better estimate the accuracy of the change features:

- Change 2017-2018: IMP change areas between 2015/2017 and 2018.

The number of primary sample units (PSUs) per stratum should be such to ensure a sufficient level of precision at reporting level. The minimum number of PSUs per stratum is set at 20. The validation exercise covers the whole study area to be valid (e.g. use of low and high probability omission strata for HRL with low sampling intensity in low probability stratum). 1,000 sample units for the prototype product were selected.

Each PSU corresponds to one 0.25ha square. Each PSU is then associated to secondary sampling units (SSUs) corresponding to a 5x5 grid with 10m between each SSU (see Figure 4-1). The idea is that each SSU can then be associated with the corresponding HRL 10m layer pixel.



**Figure 4-1: Example of SSUs organised in a 5x5 10m grid**

In a second stage, the sample units were provided to the interpretation team as separate shapefiles in which the information about the product class was included to perform a plausibility analysis. Following the introduction of the Built-up layer in the phase 1, the interpretation of the SSUs also allows the differentiation of the sealed areas and the built-up features to derive statistics on both IMD and BU layers.

The approach set up since the phase 1 is fully consistent with the procedures developed and applied as part of HRL 2015 and 2018 production and as part of the Copernicus land lot 1 external validation contract. Thus, it is guaranteed that internal validation to be consistent with the internal validation performed during the reference year production.

### **Estimation and analyses procedures**

Thematic accuracy is presented in the form of an error matrix. Unequal sampling intensity resulting from the stratified systematic sampling approach is accounted for by applying a weight factor to each sample unit. According to the accuracy guidelines integrated as part of the D33 “Time Series Analysis for Thematic Classification” [AD07], thematic accuracy is usually assessed based on the construction of confusion or error matrix and associated metrics such as: Overall Accuracy (OA), Producer’s Accuracy (PA) (related to the omission error), User’s Accuracy (UA) (related to the commission error), the associated 95% confidence intervals (CI), the kappa coefficient and the F1-Score.

Regarding the density values, a scatterplot extracted from the sample units for both the reference and prototype is made with a view to assess the correlation between reference and map values and identify any systematic bias (slope and intercept of the regression line significantly different for 1 and 0 respectively). A scatterplot is a way of displaying data against Cartesian coordinates to show and compare values for two variables within a dataset. The data is displayed as a series of points, where the x and y locations relate two variables assigned to a particular recording instance, in this case a PSU. The available measurements for each PSU are the reference data and the mapped value from the product. To quantitatively summarise the results displayed in the scatterplots above a linear regression analysis is performed to estimate the relationships between the reference and mapped product information. The analysis produces a coefficient of determination ( $R^2$ ) which is gives information about the goodness of fit of the estimated regression model. In this case as the reference and map information are meant to represent the same information then it is useful to also consider the slope and intercept of the estimated regression model. The slope should therefore approach 1 and the intercept should be close to 0 for the required relationships. Deviations from the expected values will give an indication of the correspondence of the reference and mapped imperviousness data.

## 4.2 Methods for Implementing a Prototype on Forest

This section provides an overview of the methods implemented for the Forest prototypes Dominant Leaf Type (DLT), Tree Cover Density (TCD) and Tree Cover Change (TCC) following the outcomes of Task 3. A pixel-based Tree Cover Mask (TCM) forms the basis for all three products and represents the first step in the thematic workflow. Further details and background information are provided in the three Deliverables on “D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification” [AD07] and “D34.1b - Methods Compendium: Time Series Analysis for Change Detection” [AD08], as well as “D35.1b- Methods Compendium: HRL Time Series Consistency for HRL Product Updates” [AD09], as outcomes of Task 3.

According to the outcomes of the Testing and Benchmarking exercise as part of the methods compendium of WP 33 “Time Series Analysis for Thematic Classification” [AD07], the Random Forest (RF) classifier has been selected as the best rated classification algorithm in terms of processing time and achieved accuracy for creation of the Tree Cover Mask and the improved primary status layer Dominant Leaf Type (DLT). Spatio-temporal input features that capture important time series properties and patterns are used. An automated reference sampling approach has been applied to derive the necessary sample basis for the classifier. According to the tests carried out in WP 33, the combined use of Sentinel-1 and Sentinel-2 for the spring and summer periods (vegetation period) provided the best results in terms of classification accuracy, but show the highest processing cost. In order to provide the best possible results, this data scenario has been applied for all three Forest demonstration sites. First, a 10 m Tree Cover Mask is calculated from time series features derived from Sentinel-2 and Sentinel-1. This mask is subsequently intersected with a seamless and independently derived leaf type layer to create the improved DLT status layer.

The continuous-scale Tree Cover Density (TCD) layer is representing the second improved status layer of the Forest product portfolio within ECoLaSS and has been calculated on the median time features based on the spectral bands of Sentinel-2. A multiple linear regression estimator is used to calculate density values for each pixel within the median time feature stack. The resulting raster is subsequently masked with the TCM to generate the improved TCD status layer at 10 m spatial resolution. Compared to its 2012 and 2015 precursors at 20m spatial resolution, the improved 10 m TCD 2018 status layer provides much more level of detail and benefits from the improved tree cover detection and overall accuracy of the TCM. The method for generation of the TCD layer is described in the final WP 33 report “D33.1b - Methods Compendium: Time Series Analysis for Thematic Classification” [AD07].

In view of a potential future HRL Forest Incremental Update layer, the delineation of forest change/loss is based on the comparison of a pre- and post-change tree cover mask as described in the methods

compendium of WP 34 “Time Series Analysis for Change Detection” [AD08]. The method has been firstly applied in project phase 1 in order to simulate an Incremental Update at 20 m spatial resolution. Thereby, the TCM 2015 at 20 m spatial resolution (derived from the HRL Forest Dominant Leaf Type 2015 product) represents the pre-change mask whereas the newly classified TCM 2017 (derived from optical Sentinel-2 data and resampled to 20m) represents the post-change mask with reference year 2017. The Incremental Update layer resulting thereof, hereinafter explicitly named as Tree Cover Change (TCC), compares the pre- and post-change mask to detect areas of forest loss under consideration of a specific Minimum Mapping Unit (MMU). Due to the very short time interval of mostly < 1 year between the two masks (the HRL 2015 data are mainly from spring and summer 2016 and the ECoLaSS data from spring and summer 2017), this layer concentrates on negative changes (forest loss) only.

This map-to-map change detection method is fully detached from the input data used for the tree cover/forest mask generation and therefore completely independent from the input data applied for production. Issues in view of the fusion of images from different optical sensors or even the fusion of optical and SAR data can be solved in complete isolation from the change detection. On the other hand, it has to be noted, that the quality of the input masks is key for the overall quality of the outputs.

The presented methodology can incorporate both, Sentinel-1 and Sentinel-2 data without significant adjustments, which provides more flexibility in areas of frequent cloud cover. However, in project phase 1, only optical Sentinel-2 data has been utilised. Based on the experiences and lessons learned from project phase 1, an extension to SAR data has been implemented in project phase 2. In addition to a more deep exploration of the multisensor integration approach in the TCM and DLT status layer, a more consistent and reliable change detection at 10 m spatial resolution considering a NDVI plausibility approach has been applied to derive the incremental update layers for the consecutive years 2017 and 2018.

Even though this short time period is not suitable in terms of the forest growth pace, hindering forest increase detection as described in Task 3 reports, it has been proved that yearly incremental updates are feasible and the operational roll-out to larger scales would be possible by focussing on forest loss.

Furthermore, differences between the biogeographic regions are further explored, as the prototypes are produced in representative regions in Northern Europe, Central Europe and Mediterranean environments.

## 5 Prototype Implementation

This chapter presents the implementation of the prototypes of the improved HRL Imperviousness (section 5.1) and HRL Forest (section 5.2) within the demonstration sites South-West, Central, South-East and North. For each prototype, the following aspects are examined: Data and Processing Setup (sections 5.1.1 & 5.2.1), the Classification Results and Validation (sections 5.1.2 and 5.2.2), and the Change Detection and Incremental Update Results and Validation (sections 5.1.3 & 5.2.3). Finally, the description of the dataset properties and the associated metadata are provided in detailed Prototype Specifications (section 5.3).

### 5.1 Prototype of a potential Future HRL Imperviousness

This section shows the prototypical implementation of the Imperviousness prototypes, both the improved status layer IMD and the incremental update layer IMC. Firstly, the integrated EO and ancillary data are described, followed by explaining the pre-processing steps (section 5.1.1), the demonstration of the classification results of the actual prototype in the demonstration site (section 5.1.2), and the demonstration of the prototypic results of change and incremental updates including the accuracy assessment (section 5.1.3).

#### 5.1.1 Data and Processing Setup

Firstly, the integrated EO and ancillary data are described, followed by explaining the pre-processing steps for optical and SAR data, as well as the experimental setup for the classification and incremental update approach.

##### 5.1.1.1 Input Data and Data Integration

Based on the outcomes of the phase 1 and 2 (respectively Task 3 and Task 4), a multi-sensor approach combining Sentinel-1 and Sentinel-2 was adopted in phase 1 and 2 to perform the classification that finally leads to the impervious prototypes.

##### SAR DATA - SENTINEL-1

The Sentinel-1 sensor system has an overall number of 2 bands (both polarisation signals VV and VH) at 10m pixel spacing. Pre-processing has been performed following the processing chain as detailed in WP 32 [AD 06]. Selected scenes cover the time frame from 01-January to 15-November 2018 and represent a total of 1 836 Sentinel-1 images which were used to produce the impervious prototype.

##### OPTICAL DATA - SENTINEL-2

The South-West demonstration site comprises six adjacent Sentinel-2 tiles (30TYN, 30TYP, 31TCH, 31TCJ, 31TDH, 31TDJ).

The Central demonstration site is composed of nine adjacent Sentinel-2 tiles (32UMV, 32UNV, 32UPV, 32UMU, 32UNU, 32UPU, 32TMT, 32TNT, 32TPT).

The South-East demonstration site is made of eight adjacent Sentinel-2 (34TFN, 34TGN, 34TFM, 34TGM, 34TFL, 34TGL, 35TKG, 35TKF).

All Sentinel-2A+B data in 10m resolution have been pre-processed.

The Sentinel-2 sensor system has an overall number of 12 bands from 10m to 60m spatial resolution. For the ECoLaSS processing, only the 10m bands are used, which are in total 4 bands. The list of the used bands with their central wavelengths and abbreviations is shown in Table 5-1.

**Table 5-1: Used Sentinel-2 reflectance bands (adapted from Suhet, 2015).**

| Sentinel-2 Bands | Description | Central Wavelength ( $\mu\text{m}$ ) | Stack number |
|------------------|-------------|--------------------------------------|--------------|
| <b>Band 2</b>    | Blue        | 0.490                                | 1            |
| <b>Band 3</b>    | Green       | 0.560                                | 2            |
| <b>Band 4</b>    | Red         | 0.665                                | 3            |
| <b>Band 8</b>    | NIR         | 0.842                                | 4            |

The CNES MUSCATE production centre produces the Sentinel-2 Level-2A data in near real time, which are corrected for atmospheric effects using the MACCS-ATCOR Joint Algorithm (MAJA) software. The products are available for download at <https://theia.cnes.fr/atdistrib/rocket/#/home>. The data is acquired in large areas shown in Figure 5-1.



**Figure 5-1: Theia Level-2A data production extent. (21.07.2018) Source: [http://www.cesbio.ups-tlse.fr/multitemp/?page\\_id=7501](http://www.cesbio.ups-tlse.fr/multitemp/?page_id=7501)**

The Sentinel-2 images are provided as a GeoTiff per spectral band, for the 10m bands (B2, B3, B4, B8) and the 20m bands (B5, B6, B7, B8A, B11, B12) of Sentinel-2. For the South-West site, the data comes along with two types of surface reflectances:

- Surface Reflectance (SRE) which is corrected for atmospheric effects, including adjacency effects
- Flat Reflectance (FRE) which is corrected for atmospheric effects, including adjacency effects and also corrected for terrain effects, which consists in suppressing the apparent reflectance variations. The corrected images look like if the land was flat.

The dataset used for the Imperviousness prototypes for this demonstration site is based on the Flat Reflectance to take into account the topographic effects as described and tested in the deliverables of the WP 32 [AD 06].

Selected scenes cover the time frame from 01-January to 14-November 2018 and represent a total of 2,607 Sentinel-2A+B images which were used to produce the impervious prototypes for all three demonstrations sites.

### 5.1.1.2 Pre-processing

Sentinel-2 and Sentinel-1 data were pre-processed according to the recommendations of the Task 3 and 4 – phase 1 and Task 3 – phase 2, as summarised in the subsequent sub-sections.

### 5.1.1.2.1 Pre-Processing methods for optical time series

As mentioned in the WP 32, the processing methods for optical time images include the generation of spatio-temporally consistent optical images with top of atmosphere reflectance values. Therefore, the following pre-processing steps are applied:

- Atmospheric correction,
- Topographic normalisation,
- Cloud, cloud shadow and snow masking.

#### ATMOSPHERIC CORRECTION

The Sentinel-2 data produced by CNES' Theia Land Data Centre and available for download are corrected for atmospheric effects, including adjacency effects. These atmospheric corrections include compensating the light absorption by air molecules and the light scattering by molecules and aerosols.

Several models may be used to perform atmospheric corrections. In the case of the MAJA software, the MACCS processor is the model used. It pre-computes "Look-up Tables" using an accurate radiative transfer code (Successive Orders of Scattering), that simulates the light propagation through the atmosphere. The MACCS/MAJA method combines different approaches to obtain robust estimates of aerosol optical thickness.

#### TOPOGRAPHIC NORMALISATION

A topographic correction is necessary if the test sites are characterized by mountainous terrain as it is the case for the South-West Demonstration site. The topography can significantly influence the radiometric properties of the signal received from the satellite (see Wulder and Franklin, 2012). This effect is caused by the different lighting angles resulting from the topography (cf. Gallaun et al., 2007). The aim of a topographical correction is to compensate for the differences in reflectance intensity between the areas with varying slope, exposure and inclination and to obtain the radiation values that the sensor would measure in the case of a flat surface.

The Sentinel-2 data using the MAJA software and available for download are corrected from the topographic effects.

#### CLOUD, CLOUD SHADOW AND SNOW MASKING

The MAJA cloud detection method is based on a number of threshold tests using the cirrus band (B10). Additionally, multi-temporal tests are carried out to detect clouds by measuring a steep increase of the blue surface reflectance. Finally, the correlation of the pixel neighbourhood with previous images is calculated to avoid over detections based on the assumption that two different clouds at the same location on successive dates will not have the same shape. If a large correlation is observed, the pixel is excluded from the cloud mask as it is likely to be a bright land surface.

### 5.1.1.2.2 Pre-Processing methods for SAR time series

Pre-processing has been performed with the Remote Sensing Software Graz (RSG) module "Space Suite". It comprises the following processing steps:

- Image ingestion: bulk import of original images to RSG \*.rsx files, orbit update (precise orbits), automated combination of adjacent scenes
- Image pre-processing: definition of image frame extent (based on selected granules), full image resolution, no speckle filtering, no multitemporal filtering, radiometric terrain correction to gamma naught based on SRTM 4.1 model (Central demonstration site: also tests with sigma naught), combine polarizations in one image stack (band1: VH; band2: VV)

- Orthorectification: based on an interpolated Digital Elevation Model (DEM) (SRTM 4.1), output image resolution is 10m, output image resampling method (nearest neighbour), coordinate system: UTM WGS84
- Calculation of incidence angle map

### 5.1.1.3 Experimental Setup

The developed processing chain is able to process a large amount of input data within a reasonable amount of time to provide the classification results. The achieved level of automation ensures the effective application of the process to map impervious areas of almost the entirety of Europe.

The workflow/methodological steps for the production of the Imperviousness Prototype is listed hereafter:

1. Set-up of reference databases for calibration
2. Production of the Imperviousness 2017 (for phase 1) and 2018 (for phase 2)
  - a. Data preparation (Sentinel-1, Sentinel-2)
  - b. Biophysical variables and additional image parameters (NDVI, textural metrics for S-2, time features for S-1)
  - c. Derivation of classification training samples from additional reference data (HR layers)
  - d. Production of initial built-up masks for 2017 by automated supervised classification (Active learning)
  - e. Fusion of S-1/S-2 built-up masks
  - f. Absolute calibration of IMD2017 and IMD2018
  - g. Post-processing (filtering, contextual analysis based on change probability)
  - h. Validation
3. Change detection:
  - a. 2015-2018 for the demonstration sites Central and South-East
  - b. 2015-2017 and 2017-2018 for the demonstration site South-West
4. Production of the Built-up 2018
  - a. Biophysical variables and additional image parameters (NDVI, Pantex, textural metrics for S-2)
  - b. Derivation of classification training samples from additional reference data: Open Street Map (OSM) and European Settlement Map (ESM)
  - c. Production of initial built-up masks for 2018 by automated supervised classification (Active learning)
  - d. Post-processing (filtering, contextual analysis)
  - e. Validation

#### 5.1.1.3.1 Set-up of reference databases for calibration

The development of a dataset for calibration of the IMD and IMC prototypes 2017 and 2018 is needed for two main reasons:

- To provide a reference dataset for the absolute calibration of the HRL2017 and HRL2018 10m status layer Imperviousness degree (1-100%).
- To provide a reference dataset for the statistical calibration of the changes for the time intervals 2015-2017 and 2017-2018 as a basis for the re-processing of the 2015 layers thus ensuring the temporal consistency of the products.

The stratification and sampling approach required to create these two different reference datasets is quite different and therefore they are performed separately.

### **REFERENCE DATASET FOR ABSOLUTE CALIBRATION OF IMPERVIOUSNESS DEGREE LEVEL (1-100%)**

The reference imperviousness density values are collected for selected sample cells (PSU of 1ha) within the Sentinel-2 tiles. Imperviousness degree levels from 1-100% are obtained for each PSU. The sealing Information, sealed surfaces vs. non sealed surfaces, is collected through Secondary Sampling Units (SSUs – 5x5 grid) within each PSU.

In order to be a representative methodology, the approach chosen combines random and stratified approaches and benefits. The stratification is based on the previous 2015 Imperviousness layer (IMD density value [1-100%]).

### **REFERENCE DATASET FOR THE STATISTICAL CALIBRATION OF THE CHANGES**

One of the key requirement is to ensure the temporal consistency and comparability between the different time intervals and that there should be no spatial inconsistencies between the layers of the different epochs. Due to the semi-automated nature of the HRL production workflow, it is not possible to guarantee that all errors can be removed from the change layer. However, the relative magnitude of actual change versus the errors contained in the change layer for each time interval should be known in order to provide a basis for improving the temporal consistency between each layer. Therefore, there is a need to develop a reference dataset that will be used to determine the relative proportion of actual change versus all the error components described above. To be valid, this calibration dataset should be selected based on a probability sampling approach. Sampling design refers to the protocol whereby the samples are selected. A probability sampling design is preferred for its objectivity. “Simple random, stratified random, clustered random and systematics designs are all examples of probability sampling designs” (Stehman et al., 1998). Even though a simple random design is easy to implement, its main drawback is that some portions of the population may not be adequately sampled. Cluster sampling is often used to reduce the costs of the collection of reference data but does not resolve geographic distribution problems. A systematic approach would solve this problem, yet it is not appropriate if the map contains cyclic patterns. A stratified approach consists in allocating a pre-defined number of samples per land-cover class. As explained in (Stehman et al., 1998), stratification ensures that each class is represented. Since the focus is on change, the approach will help assess:

- The new built-up for the year 2017 or for the year 2018;
- The omission errors from 2015 or from 2017 – the undetected built-up pixels of 2015 or from 2017 respectively;
- And the commission errors from 2017 or from 2018 – the pixels falsely flagged as built-up in 2017 or from 2018 respectively.

Therefore, there is a strong need to develop a reference calibration dataset that will be used to determine the relative proportion of actual change versus all the error components described in the previous paragraph. This calibration dataset has been selected based on a probability sampling approach similar, but independent, to that of the validation dataset as implemented in the reports of the Task 3.

The stratification and the sampling design primarily consist of selecting an appropriate sampling frame and a sampling unit. In case of changes, a sampling design based on points is considered most appropriate. Figure 5-2 displays the datasets for the years 2017-2018 and Figure 5-3 for the phase 2 prototypes.

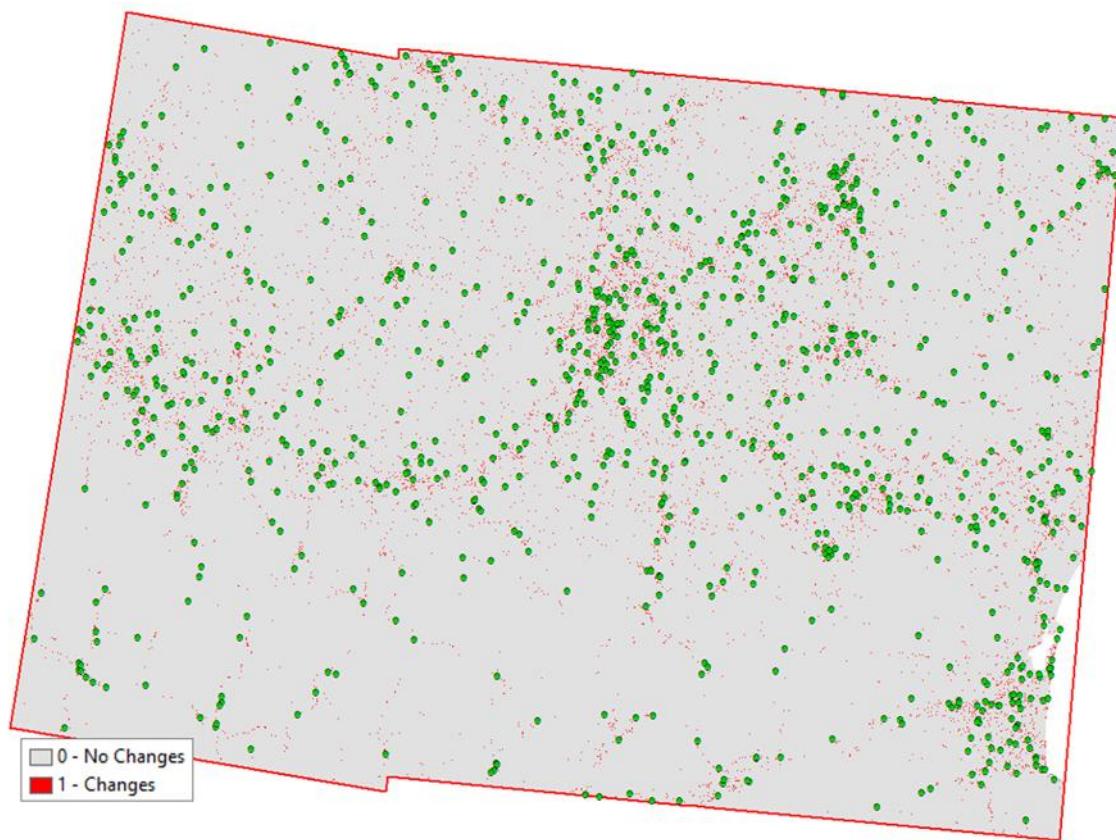
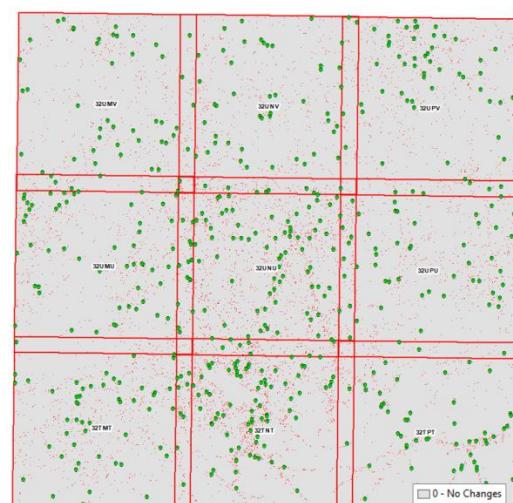


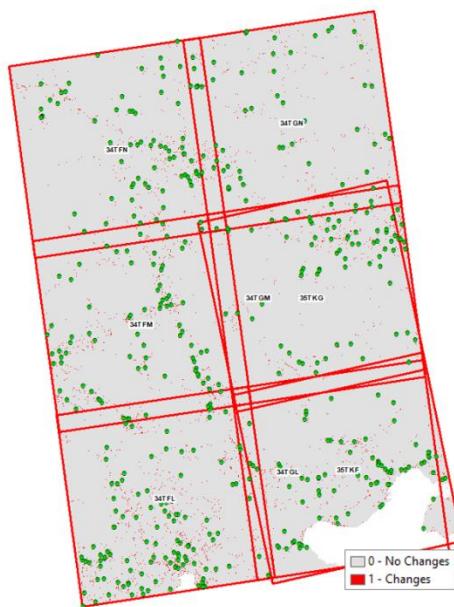
Figure 5-2: Reference calibration samples overlaid on the change mask 2015-2017 (SW demonstration site)



South-West demonstration site for 2017-2018



Central demonstration site for 2015-2018



*South-East demonstration site for 2015-2018*

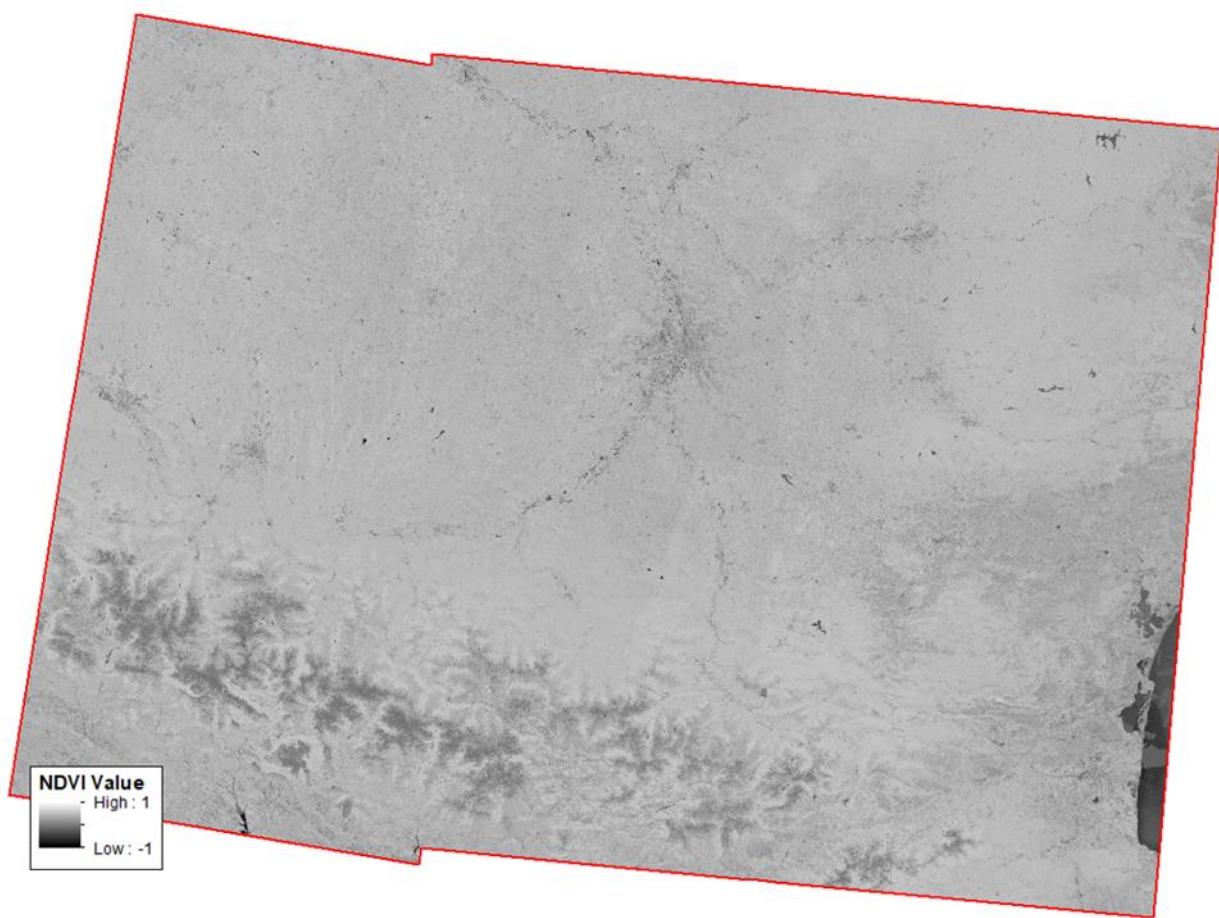
**Figure 5-3: - Reference calibration samples overlaid on the change mask 2015/17-2018**

#### 5.1.1.3.2 Production of the Imperviousness 2017 & 2018

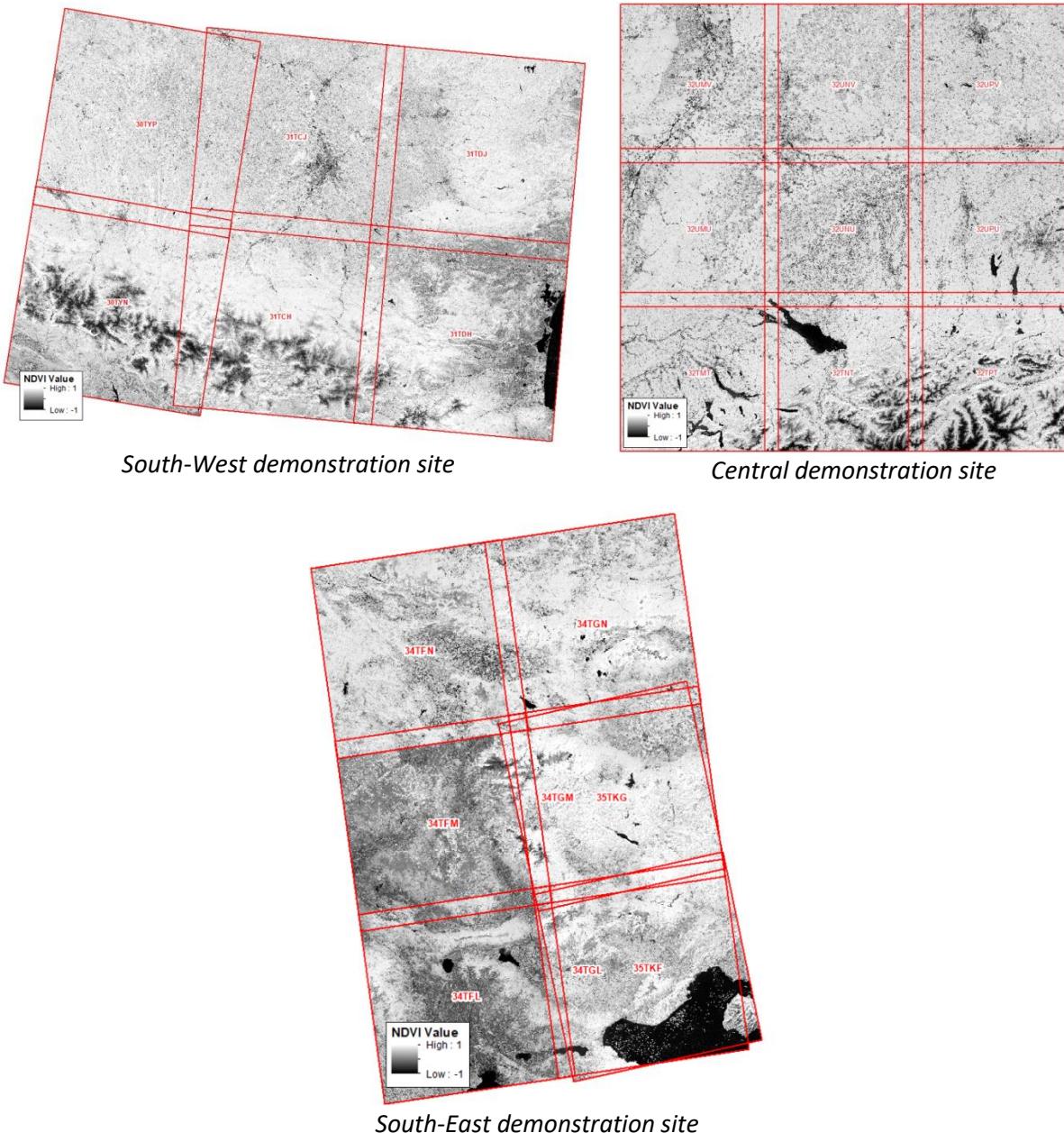
##### **DATA PREPARATION (SENTINEL-1, SENTINEL-2)**

This step includes all the pre-processing required to prepare the data which can be listed as: downloading, data and metadata extraction, best-scene selection (based on cloud coverage), layer stacking, preprocessing, for S-2 or S-1 – and finally cloud masking for S-2.

For the purpose of the calibration task, the NDVI is derived per single Sentinel-2 image, then mosaicked to a maximum NDVI as shown in Figure 5-4.



**Figure 5-4: NDVI Sentinel-2 based maximum feature for the year 2017**



**Figure 5-5: 2018 NDVI Sentinel-2 based maximum feature for the year 2018**

Following annual SAR features are generated using S-1 data and both polarisation signals (VV, VH) including 1,826 images from 01.01.2017 to 15.11.2018, covering the South-West demonstration site. Examples for such statistical features are presented in Figure 5-6.

**Table 5-2: SAR annual statistical features.**

| feature     | description        |
|-------------|--------------------|
| <b>MIN</b>  | Minimum            |
| <b>MAX</b>  | Maximum            |
| <b>MEAN</b> | Mean               |
| <b>STD</b>  | Standard deviation |

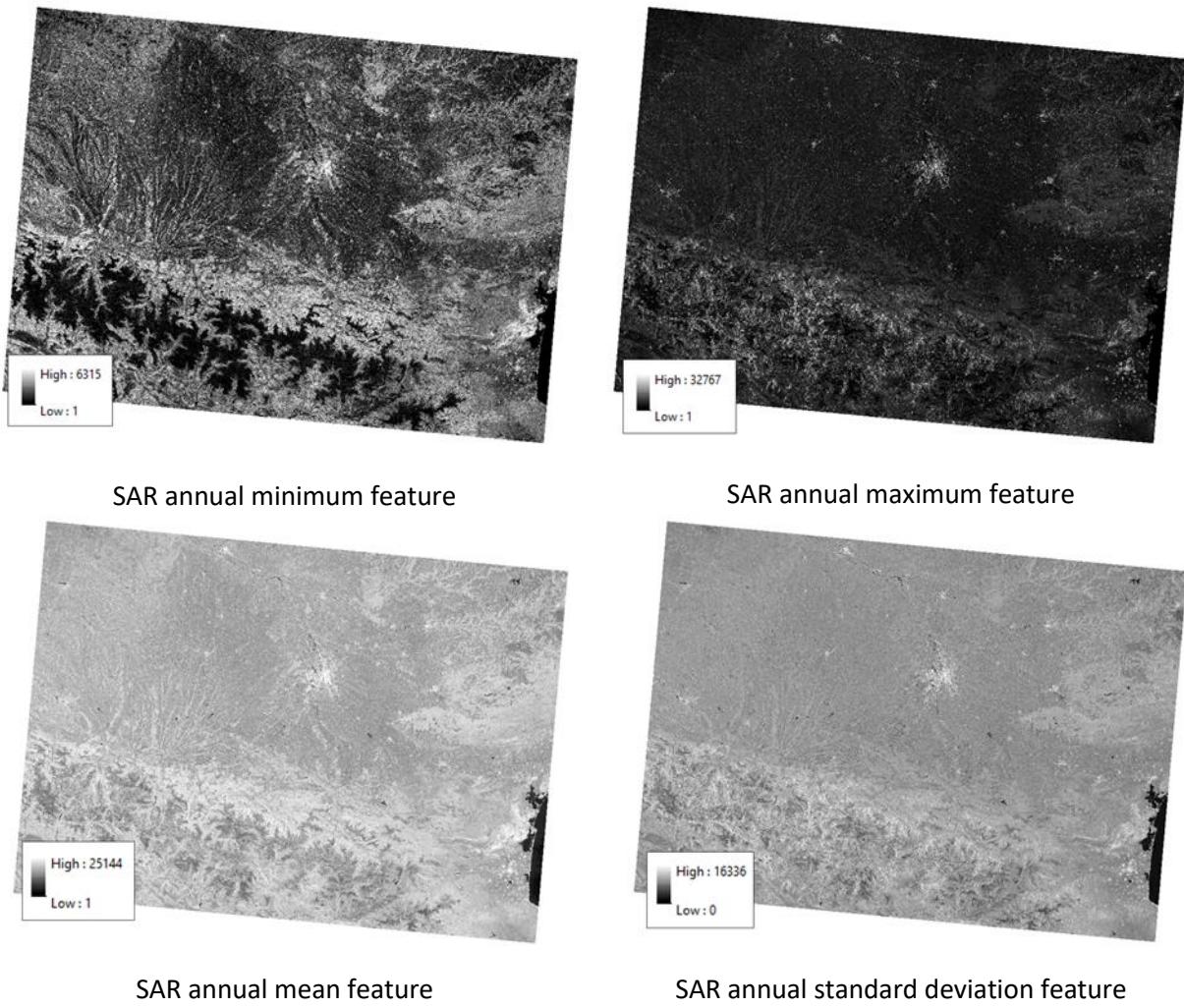


Figure 5-6: SAR statistical features (NB: the 4 features have different value ranges and scaling)

#### AUTOMATED DERIVATION OF CLASSIFICATION TRAINING SAMPLES

As input for these machine learning algorithms, a set of training data is required. The training data chosen must therefore be representative of the whole study area in order to cover all the reflectance variations of the classes, as well as to go further and take into account the local variability of the environmental classes due to the soil type, moisture, etc. The training sites must be exempt from anomalies and must be a suitable statistical representation of the area. There must be a substantial number of them. That is why, the historical High Resolution Layers have been used as training data:

Reliable training samples have been derived from relevant in-situ sources: historical HRL 2015 Imperviousness, Forest, Grassland, Water and Small Woody Features. In order to best reflect the different imperviousness classes, an automated random point sampling within buffered IMD 2015 has been applied. Samples in non-built-up areas have been selected in different land cover classes such as grassland, bare soil, vegetation and water in order to obtain a representative distribution of non-imperviousness samples.

Based on the spectral information, biophysical indicators and texture parameters at the training sample points, the algorithm ‘learns’ how to classify the features (Tan et al. 2006, Camp-Valls, 2009) and identifies the most significant combinations of input parameters to differentiate built-up areas from other land cover.

For the purpose of the automated derivation of the training sample, a stratified random approach, based on the HRLs, has been preferred as described in the accuracy guidelines section in WP 33 “Time Series Analysis for Thematic Classification” [AD07].

### **PRODUCTION OF INITIAL SEALED AND NON-SEALED MASKS FOR 2017 AND 2018**

The results of the initial Sentinel-2 based sealed and non-sealed mask are shown in Figure 5-7 and Figure 5-8.

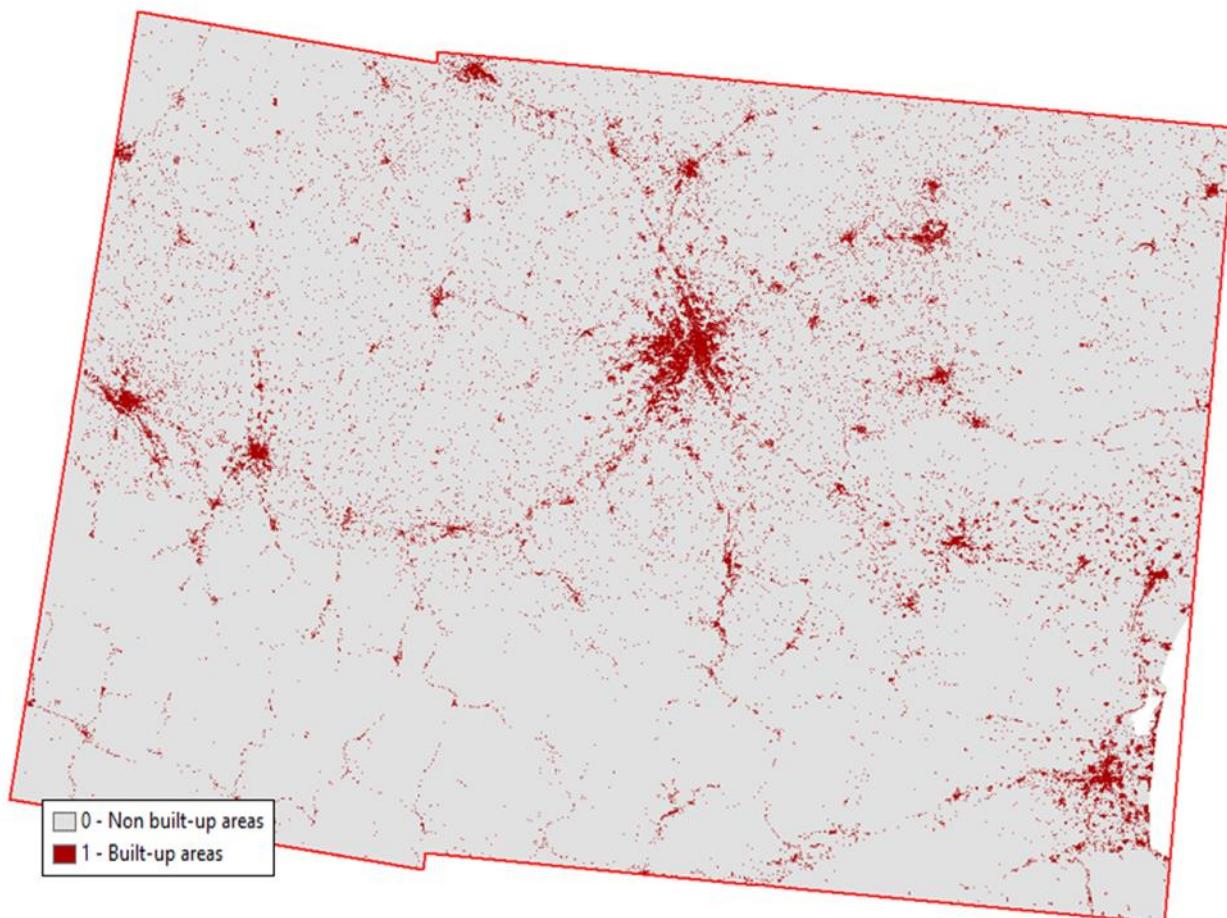
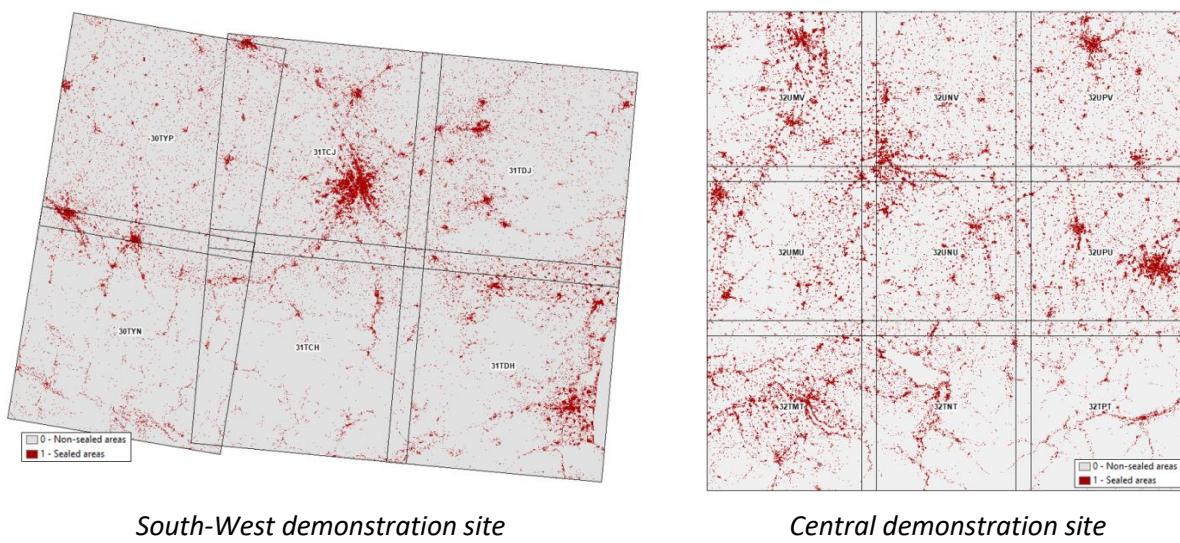
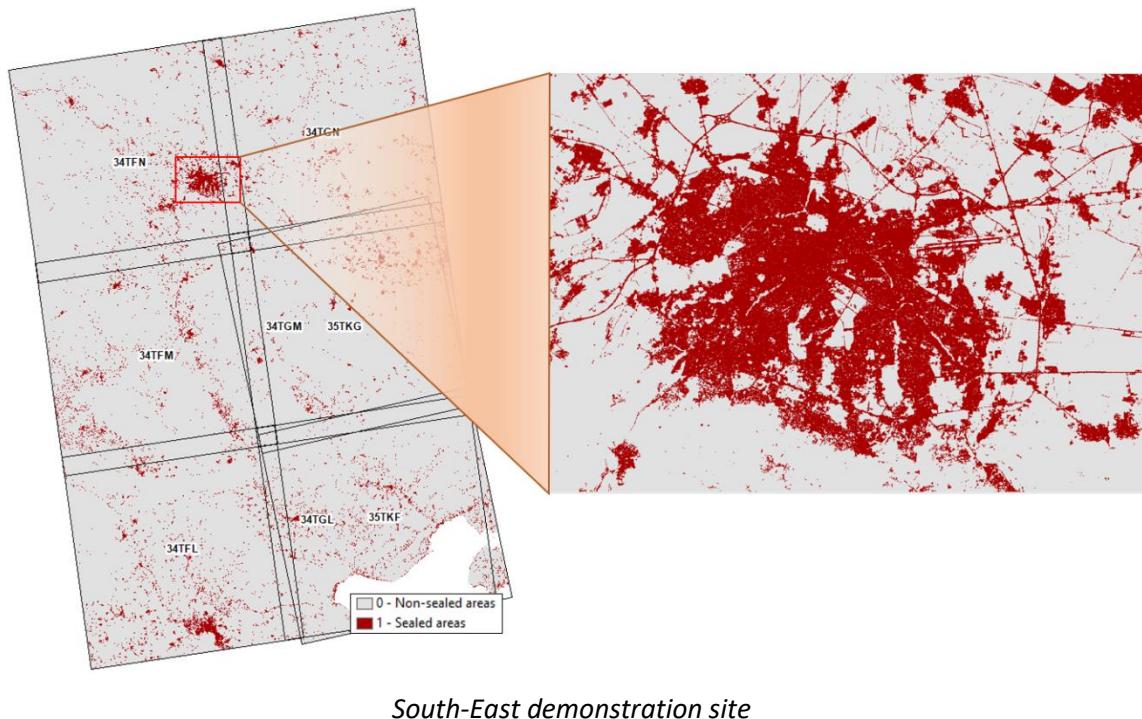


Figure 5-7: Sentinel-2 based initial built-up and non-built-up mask for 2017 for the South-West site





*South-East demonstration site*

**Figure 5-8: Sentinel-2 based initial sealed and non-sealed mask for 2018 for the demonstration sites**

#### **Absolute and relative calibration of IMD2017 and IMD2018**

Besides the production of initial sealed and non-sealed masks for 2017 and 2018 (previous steps described in section 4), one key step in the HRL Imperviousness production is the estimation of the degree of imperviousness and linking these IMD measurements over time. Each single pixel in the built-up mask will be assigned an imperviousness density value of 1 to 100%. The linkage between the biophysical variables and the IMD measurements will be done through an absolute (linking the biophysical variables to IMD) and relative calibration procedure. This combination improves the accuracy of imperviousness density estimates, correct any over-/underestimation of values and assure comparability and consistency over time.

The reference calibration database serves as calibration input for an absolute calibration of the 2017 and 2018 IMD measurements. For the prediction of the imperviousness degree, a linear regression method is used to model the relationship between the collected reference samples and meaningful metrics from the biophysical variables (e.g. NDVImax) derived from the seasonal image composites. The established linear equation is applied to transform the input data into imperviousness degree values between 1 – 100%. This results in absolutely calibrated IMD measurements derived from the 2017 and 2018 imagery.

Then, the calibrated 20m IMD 2015 status layer will be used as input to adjust the imperviousness density values of 2017 and 2018 by relative calibration. Indeed, despite the absolute calibration based on a well-established procedure (with the use of a reference calibration dataset), there will always remain some obvious and local issues in the imperviousness density derivation which will lead to wrongly detection changes in the change Layers. The relative approach is so needed to correct these local artefacts

The IMD 2017 and 2018 values, limited to the newly created 2017 and 2018 sealed mask, are re-analysed by an automatic cross-calibration approach: the IMD 2017 and 2018 values are compared to the IMD 2015 values resampled to 10 meters spatial resolution and further corrected using a rule-based approach. Indeed, a filtering approach is needed to adequately map sealing changes. In 20m spatial resolution, changes of imperviousness density within sealed areas are not that frequent compared to changes of the

sealed area. Despite all the calibration efforts in image pre-processing and subsequent adaptation procedures, there will always remain a certain error budget for sealing change detection mainly caused by:

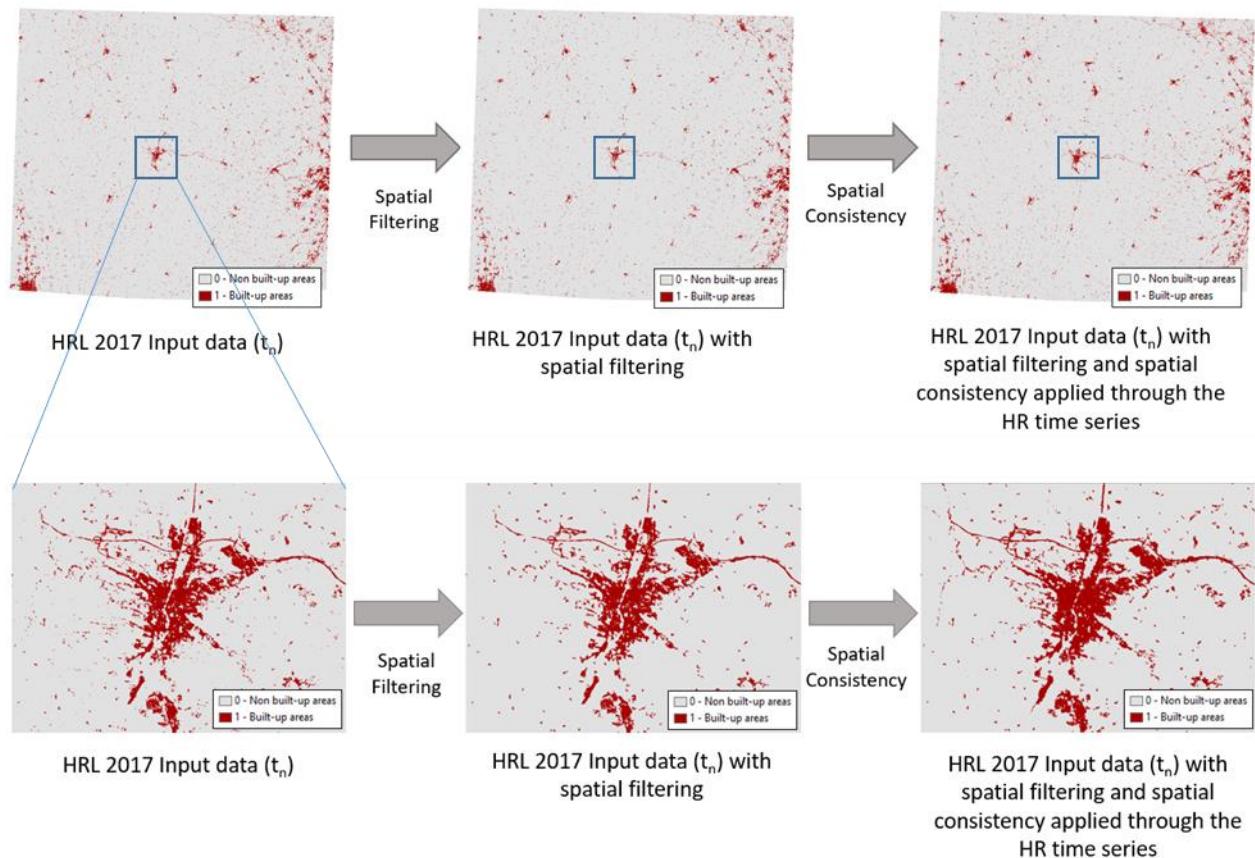
- Persisting spectral differences due to even subtle deviations in illumination, shadow effects, atmospheric conditions and vegetation status, often in conjunction with:
- Geometric misalignments of the IMAGE databases. This occurs quite frequently and often exceeds a range of 1 pixel (>20m).

Hence, in order to derive a reliable and realistic picture of sealing changes (within existing built-up areas), thresholds are applied. Differences of >20% of sealing increase will be considered acceptable if a contiguous area of at least 16 (10m x 10m) pixels is concerned. The threshold of 16 contiguous pixels permits to overcome the scaling issue (10 vs 20m spatial resolution). Differences  $\leq$  20% sealing increase will be considered as stable. The special case of imperviousness decrease is rare and, if occurring, it will rather be due to a re-greening (full de-sealing) of an impervious surface than an actual decrease. With regard to this assumption sealing decrease within built-up areas will only be accepted as valid if a remarkable change of 80% decrease takes place. Differences  $\leq$  80% sealing decrease will be considered as stable. In phase 2, for the South-West prototype, the same rule-based approach was applied to correct the 2018 values with comparison to the 2017 IMD values obtained in phase 1.

### **POST-PROCESSING**

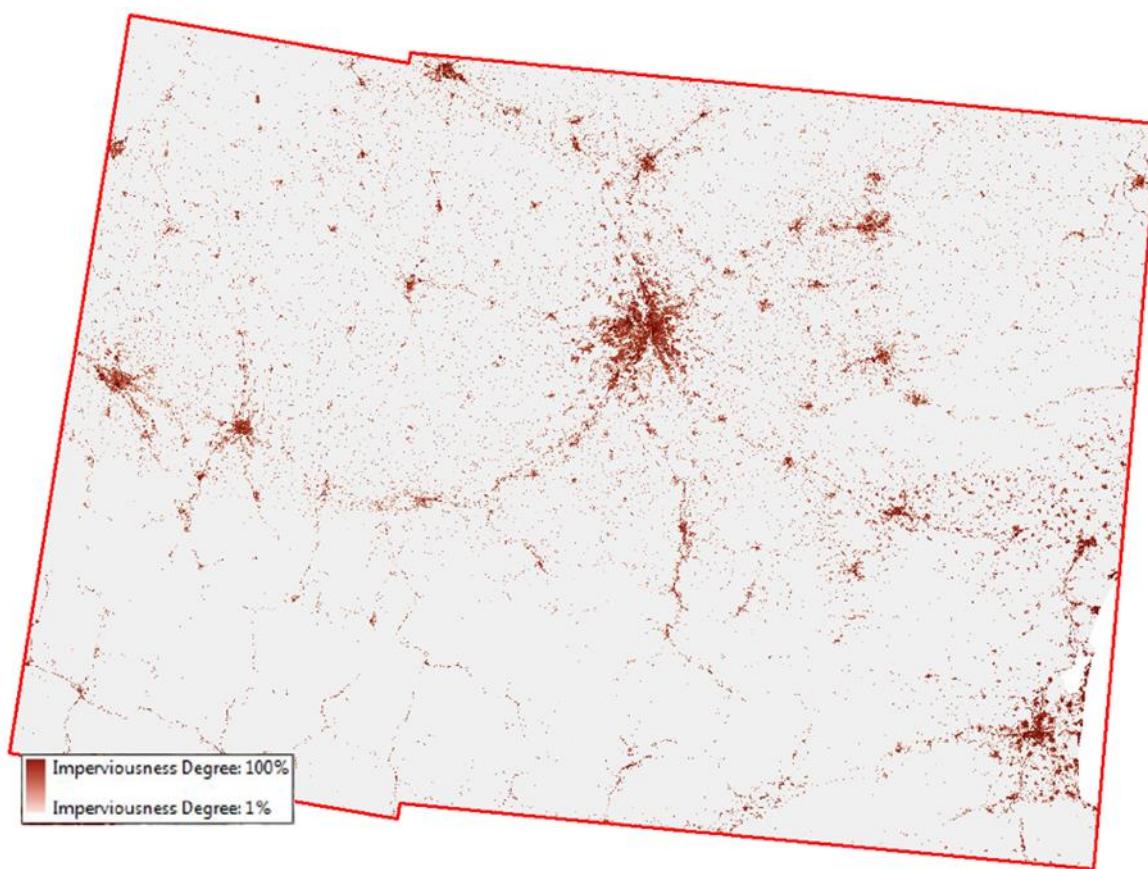
The post-classification (see Figure 5-9) implies post-processing of the layers in order to be spatially consistent, including:

- Post-processing filtering usually used in global urban maps. Indeed, there is a significant portion of noise due to single pixels or isolated pixels (small aggregated group of pixels), which are most likely misclassifications. Such noises should be reduced/removed with post-classification filtering approaches.
- Contextual analysis based on change probability (Lefebvre et al., 2016). The aim is to take into account the built-up pixels in the 2015 sealed mask in order to establish a probability map of changes. The analysis describes each cell's relationship/membership to a source or a set of sources based on probabilities. The assumption made using the contextual analysis is that urbanized areas spread more than they appear randomly in the landscape (Lefebvre et al., 2016).



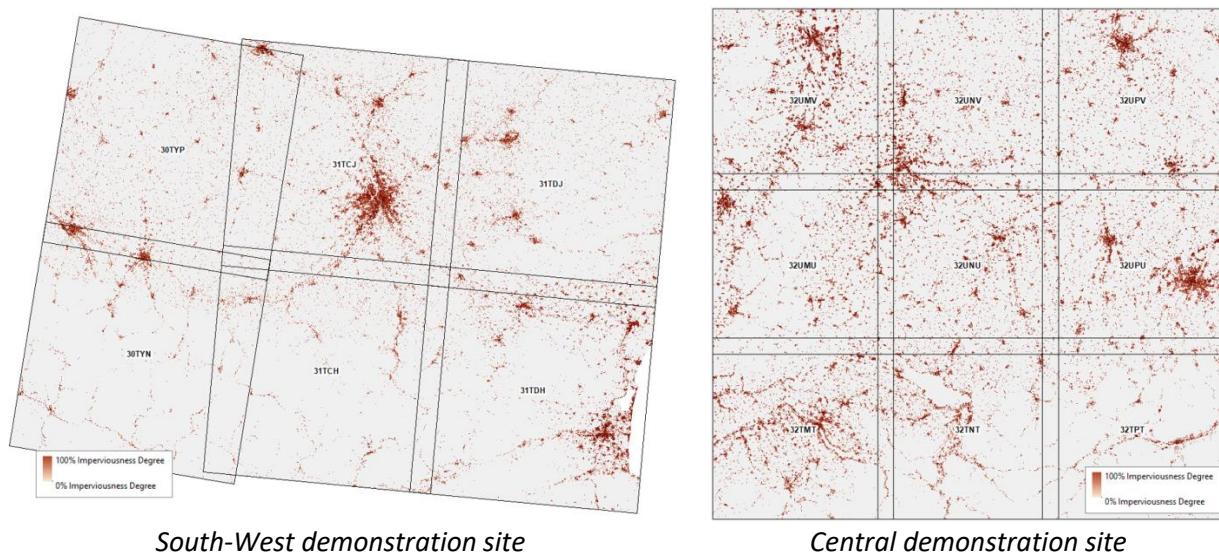
**Figure 5-9: Post-classification processing**

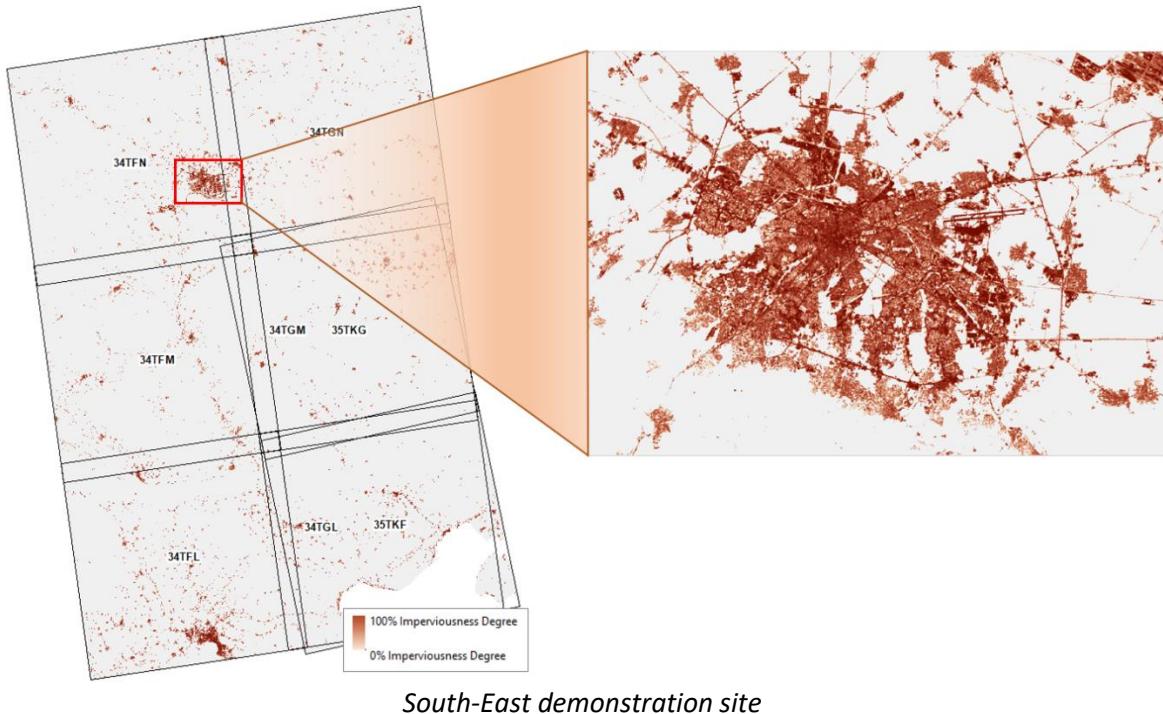
The final result of the implemented Imperviousness prototype 2017 is shown in Figure 5-10.



**Figure 5-10: Final HRL Imperviousness 2017 prototype for the South-West demonstration site**

The final results of the implemented Imperviousness prototypes 2018 is shown in Figure 5-11.





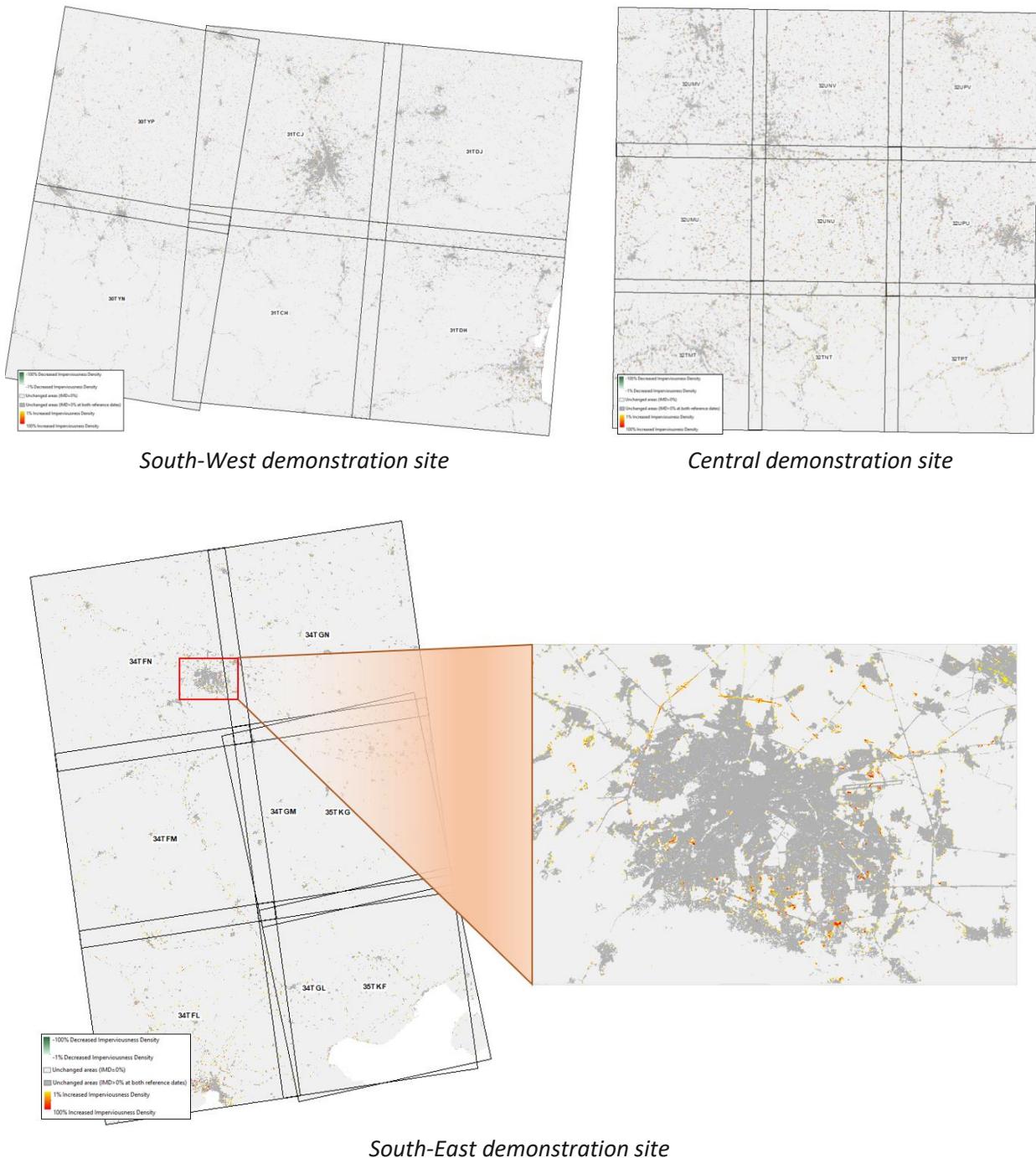
**Figure 5-11: Final HRL Imperviousness 2018 prototype for the demonstration sites**

#### 5.1.1.3.3 Production of the Imperviousness Change 2015-2017, 2017-2018 and 2015-2018

The change detection 2015-2017 (SW demonstration site, phase 1), 2015-2018 (CE and SE demonstration sites, phase 2) and 2017-2018 (SW demonstration site, phase 2), is applied based on the HRL IMD 2017 and 2018 Prototypes resampled to 20m and the IMD 2015 from the HRL Imperviousness 2015 produced during the operational HRL production outside this project. The production of the imperviousness change 2017-2018 for the SW demonstration site is done at 10m based on both IMD products from phase 1 and 2. It is important to note that this step not only reveals 2015-2017 (respectively 2015-2018 and 2017-2018) sealed changes, but, as stated before, it also detects potential omission errors of the sealed initial mask (2015 or 2017) as well as potential commission errors of the new period (2017 or 2018) sealed mask.

At first, a 20m change layer is calculated by direct subtraction of the 20m imperviousness values but without any further filtering, thereby guaranteeing full consistency of all products. The result displays the total imperviousness degree change values from -100% to +100%, according to the thresholds set at the relative calibration and without any thematic classification applied. In other words, the first-step change layer only consists of a continuous layer with change values from -100% decrease to +100% increase and not a categorisation of changes.

Then, a spatial filtering is applied in order to take into account the different specifications. Indeed, the HRL IMD 2015 data presents a 20-meters spatial resolution whereas the HRL 2017 and 2018 Prototype, based on S-1 and S-2, show a 10-meters resolution. A minimum mapping unit has been applied on the HRL 2017 Prototype based on a 4 pixels contiguous rule. The results of the HRL Imperviousness change prototypes are shown in Figure 5-12.



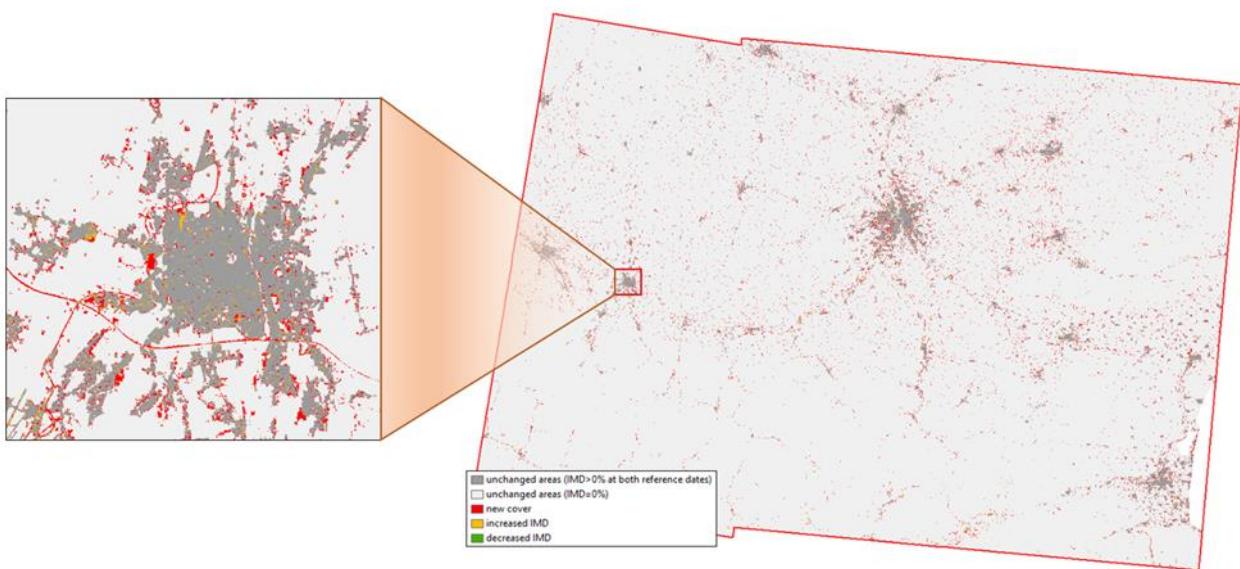
**Figure 5-12: Final HRL Imperviousness Change 2015-2018 prototypes for the demonstration sites**

To be fully compliant with the actual specification of the products, the derived change layer is then converted into a ‘classified change’ layer. For this purpose, the continuous change values will be thematically aggregated into the following categorical classes according to the rule base defined in Table 5-3 below.

**Table 5-3: Specifications of the ‘classified change’ layer**

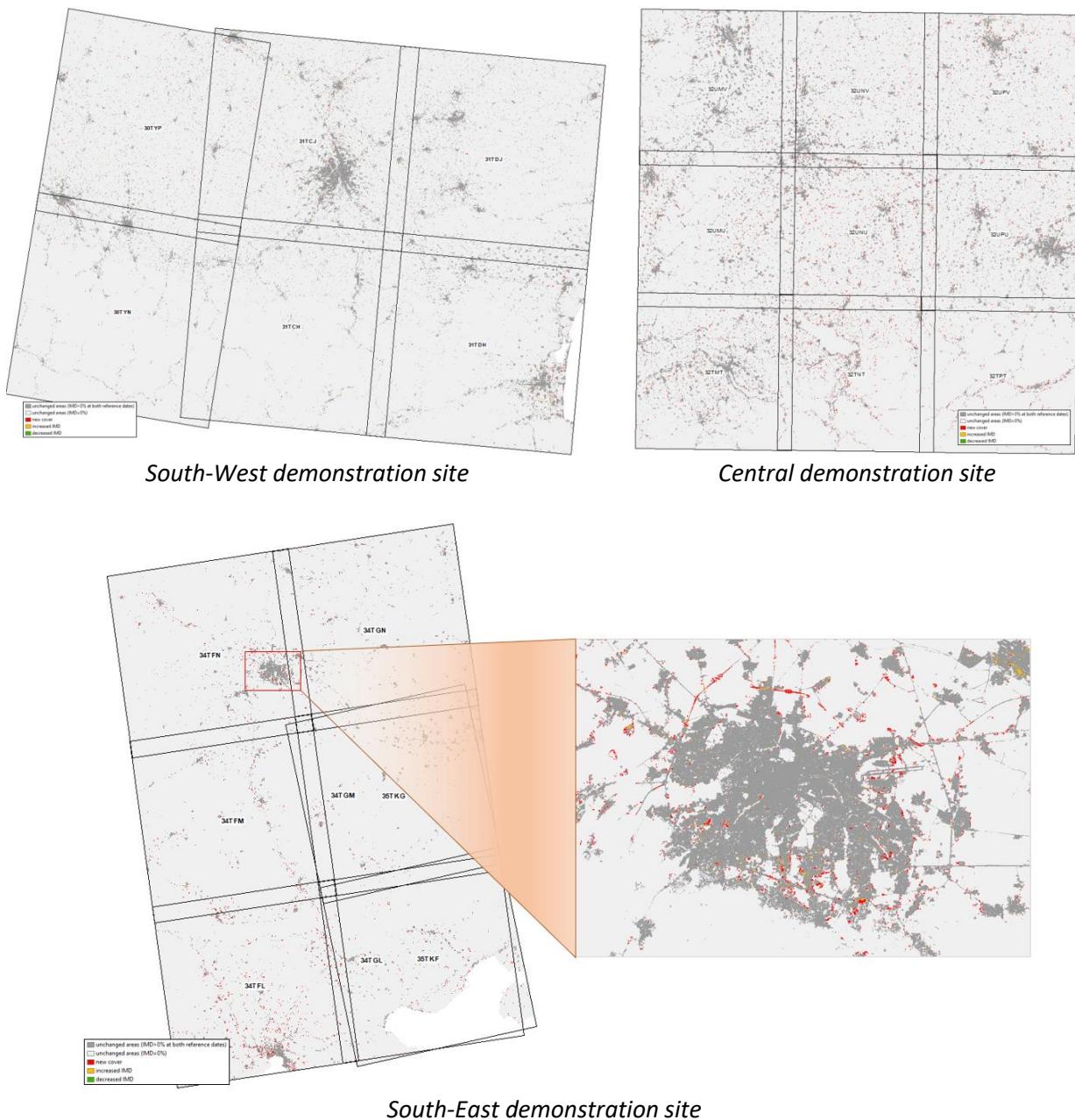
| Class Code | Classified change<br>Class name   | IMD t1 | IMD t2 |
|------------|---|--------|--------|
| <b>0</b>   | unchanged areas with IMD=0%   | 0%     | 0%     |
| <b>1</b>   | new cover (increased imperviousness density, 0% IMD at first reference date)                          | 0%     | >0%    |
| <b>2</b>   | loss of cover (decreasing imperviousness density, 0% IMD at second reference date)                    | >0%    | 0      |
| <b>10</b>  | unchanged areas (IMD>0% at both reference dates)  | >0%    | >0%    |
| <b>11</b>  | increased IMD (IMD>0% at both reference dates)<br>subject to a 20% threshold for 4px contiguous areas | >0%    | >>0%   |
| <b>12</b>  | decreased IMD (IMD>0 at both reference dates)<br>subject to a 80% threshold for 4px contiguous areas  | >>0%   | >0     |
| <b>254</b> | unclassifiable in any of parent status layers   | 254    | 254    |
| <b>255</b> | outside area  | 255    | 255    |

The final result of the implementation of the production of the Imperviousness Change Prototype 2015-2017 for the South-West demonstration site (as part of the Phase 1) is presented in the Figure 5-13 :



**Figure 5-13: Final HRL Imperviousness Change 2015-2017 prototype for the South-West demonstration site**

The final results of the implementation of the production of the Imperviousness Change Prototypes 2015-2018 (as part of the phase 2 for the Central and South-East demo sites) and the Prototype 2017-2018 over the South-West demo site are presented in the Figure 5-14:



**Figure 5-14: Final HRL Imperviousness Classified Change 2015-2018 or 2017-2018 prototypes for the demonstration sites**

#### 5.1.1.3.4 Production of the Built-up Layer 2018

##### DATA PREPARATION (SENTINEL-1, SENTINEL-2)

This step includes all the pre-processing mentioned in section 5.1.1.2.1 for Sentinel-2.

For the purpose of the BU layer classification, the PanTex is derived per single S-2 image, as shown in Figure 5-15. The PanTex is used as input data along with the 10 m spectral bands for the classification.

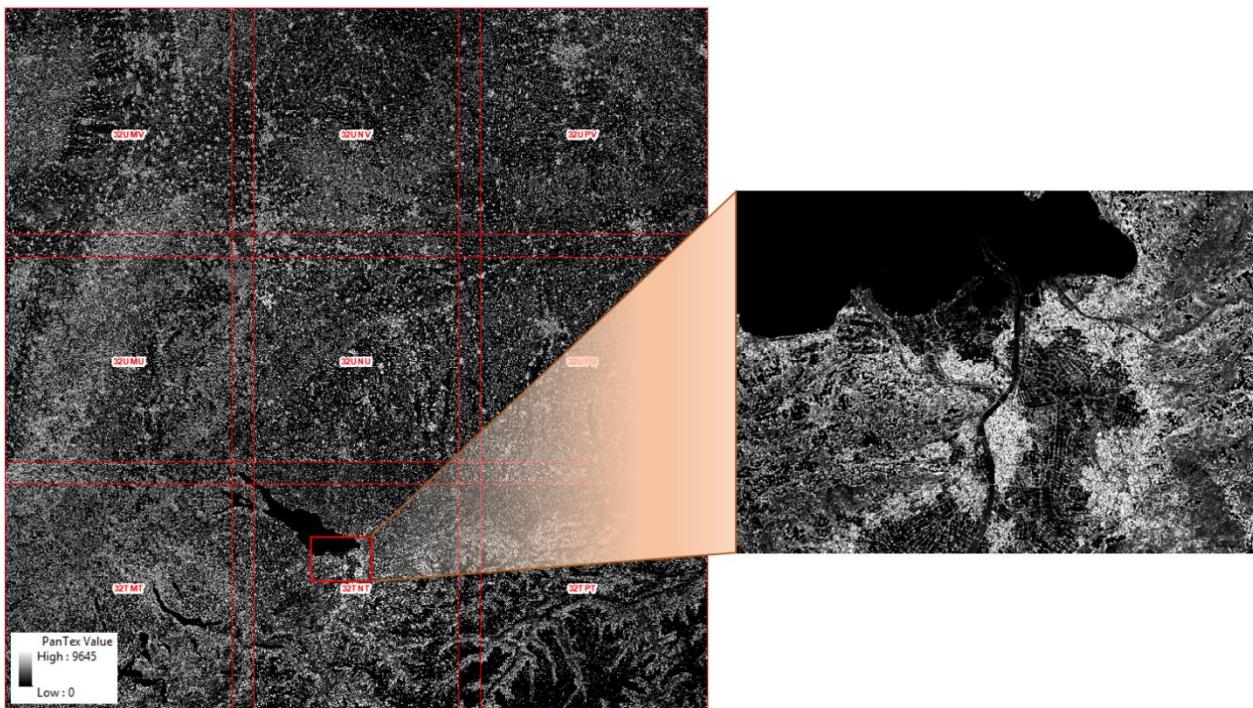


Figure 5-15: 2018 PanTex Sentinel-2 based feature for the Central demonstration site

#### AUTOMATED DERIVATION OF CLASSIFICATION TRAINING SAMPLES

As input for the machine learning algorithms, a specific set of training data, different from the one implemented for the sealed surface classification, is also required for the Built-up layer. The Open Street Map (OSM) and European Settlement Map (ESM) have been used as training data.

In order to best reflect the different built-up features, an automated random point sampling has been applied. Samples in non-built-up areas have been selected in different features/classes such as roads, railways or parking lots to obtain a representative distribution of non-built-up samples.

Based on the spectral information, biophysical indicators and texture parameters at the training sample points, the algorithm ‘learns’ how to classify the features (Tan et al. 2006, Camp-Valls, 2009) and identifies the most significant combinations of input parameters to differentiate sealed areas from other land cover.

For the purpose of the automated derivation of the training sample, a stratified random approach, based on the HRLs, has been preferred as described in the accuracy guidelines section in WP 33 “Time Series Analysis for Thematic Classification” [AD 07].

#### PRODUCTION OF INITIAL BUILT-UP AND NON-BUILT-UP MASK FOR 2018

The results of the initial Sentinel-2 based built-up and non-built-up mask are shown in Figure 5-16.



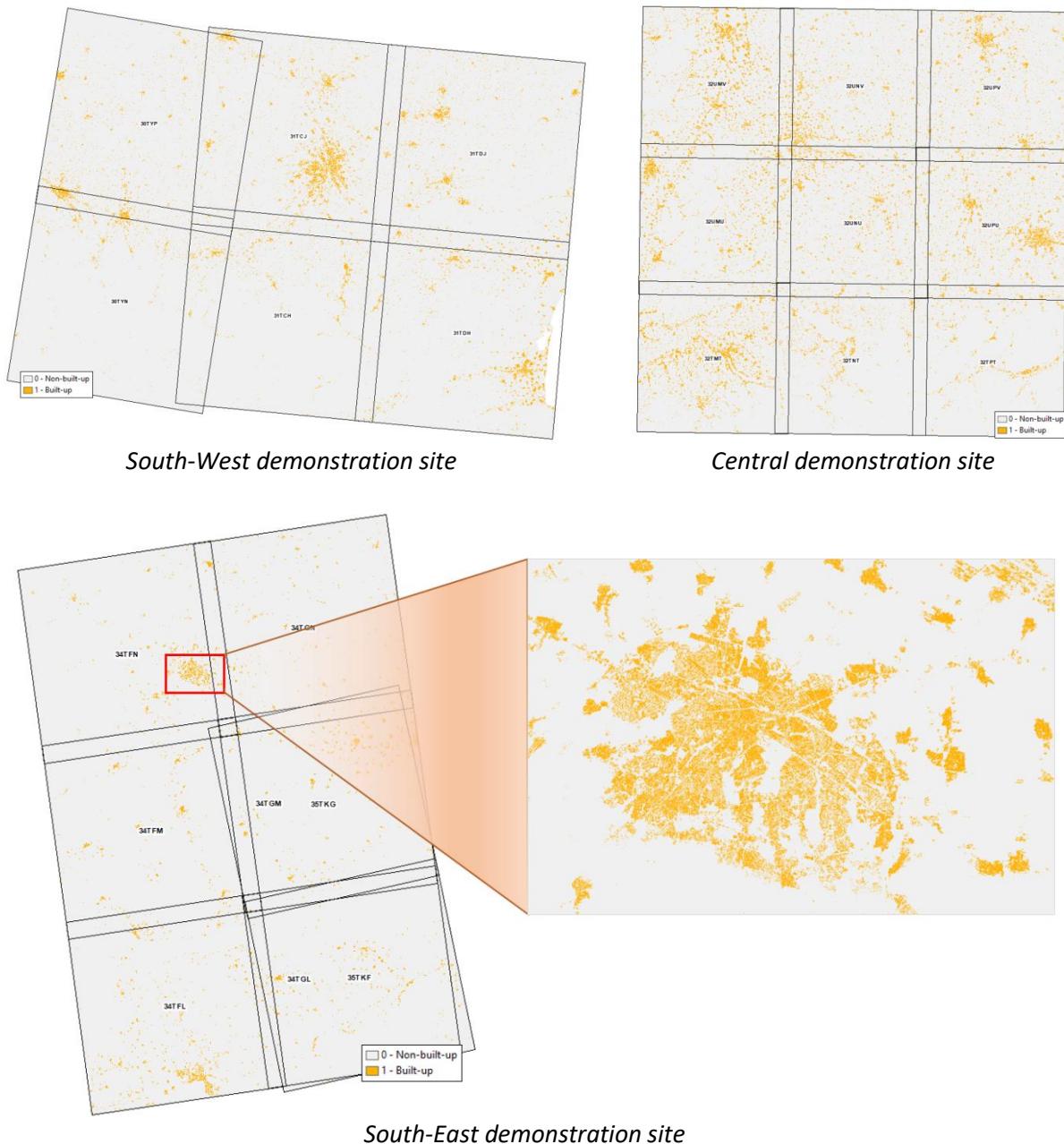
**Figure 5-16: Sentinel-2 based initial built-up and non-built-up masks for 2018 for the demonstration sites**

#### POST-PROCESSING

The post-classification implies post-processing of the layers in order to be spatially consistent with the HRL Imperviousness (IMD) layers including:

- Post-processing filtering using the sealed mask. Indeed, built-up pixels which are not sealed in the IMD classification should be removed in a post-classification step to ensure the spatial consistency between products.

The final results of the implemented Built-up prototypes 2018 is shown in Figure 5-17.



**Figure 5-17: Final HRL Built-up 2018 prototypes for the demonstration sites**

### 5.1.2 Classification Results and Validation

This chapter depicts the results of the classification as well as their validation. Firstly, the thematic accuracies are summarized (see section 5.1.2.2). The thematic accuracies are followed by a discussion of the validation results (see section 5.1.2.3).

#### 5.1.2.1 IMD 2018 scatterplots & regression analysis

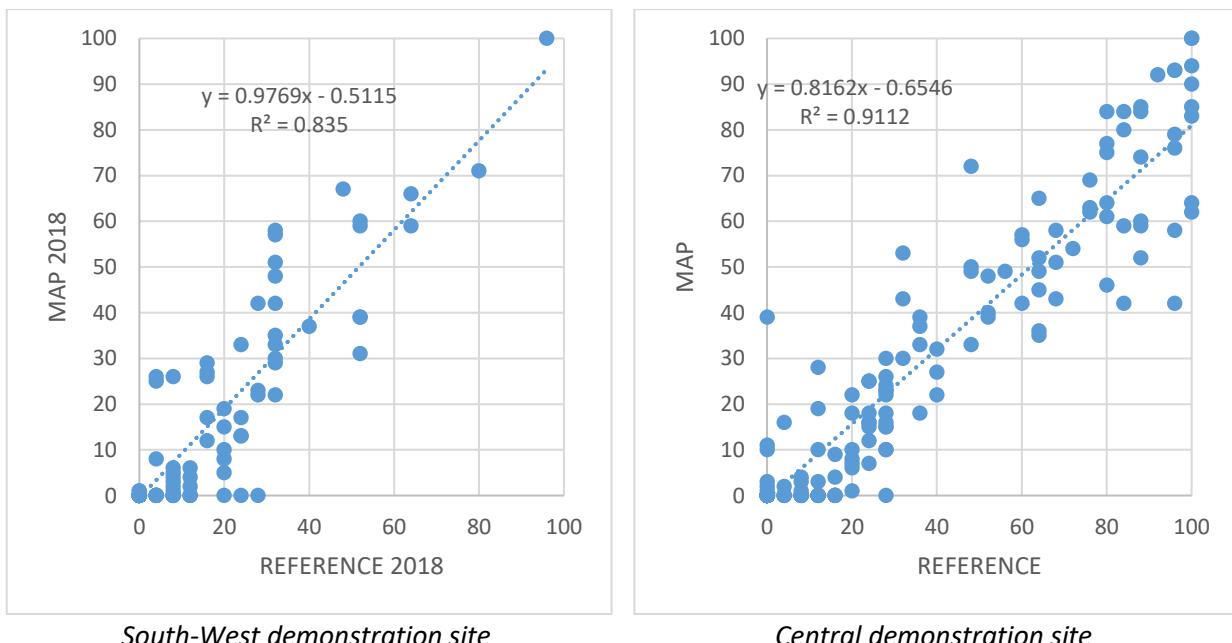
As mentioned in section 4.1.2, a scatterplot is a way of displaying data against Cartesian coordinates to show and compare values for two variables within a dataset. The data is displayed as a series of points, where the x and y locations relate two variables assigned to a particular recording instance, in this case a PSU. The available measurements for each PSU are the original reference data (called REFERENCE in each figure) and the mapped value from the product (called MAP on the figures). For this validation exercise the position / value on the horizontal axis represented the reference information and the position / value

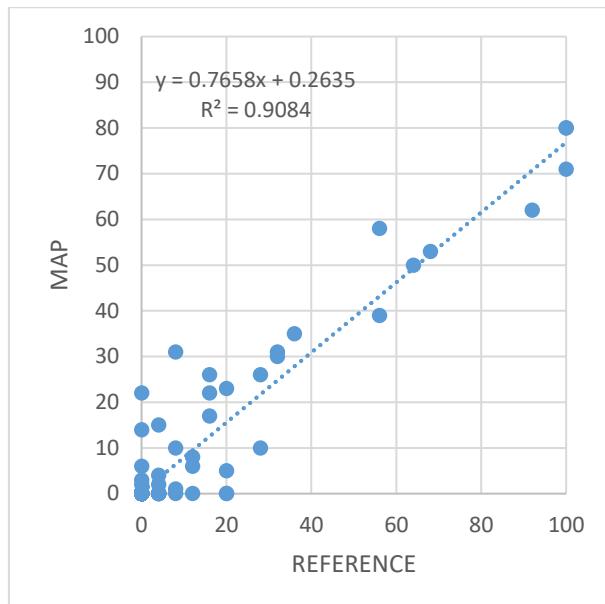
on the vertical axis represents product (MAP) information. In this way the relation of the reference and product information for a point can be compared to a 1:1 line which runs diagonally across the scatter plot. The closeness of a point to the point to the 1:1 line is an indication of the similarity between the reference and mapped results. The points that lie exactly on the x and y axes are related to omission and commission rather than the calibration of the IMD values themselves.

The scatterplots presented in Figure 5-18 show very limited scatter of up to 20 % each side of the best fit line. There are a few numbers of point on the x and y axes showing commission, where sealing is mapped that is not present in reality, and omission, where sealed areas are missed which seems to indicate that the commission and omission errors are also limited. It can also be seen that the distribution is almost centred on the 1:1 line.

The scatterplots and the resulting best fit lines are controlled by the actual geography of the demonstration sites. The results for the South-West and South-East show that there are limited number of sealed areas, the areas present tend to have low imperviousness values. There are also a few omissions and commissions which is highly likely for Mediterranean regions where vegetation is limited and there may be extensive areas of bare soils.

The Central demonstration site is more representative of Europe as a whole and contains significant sealed areas of varying imperviousness. There is a group of points which are marked as completely sealed (100 % imperviousness) in the reference data but range from around 60 to 100 % in the mapped data, which may indicate that certain sealed surfaces are more difficult to detect or calibrated for in the EO data.





*South-East demonstration site*

**Figure 5-18: Scatterplots for the continuous IMD layers 2018 validation**

To quantitatively summarise the results displayed in the scatterplots above a linear regression analysis is performed to estimate the relationships between the reference and mapped product information. The analysis produces a coefficient of determination ( $R^2$ ) which is gives information about the goodness of fit of the estimated regression model. Coefficients of determination closer to 1 represent a better fit. In this case as the reference and map information are meant to represent the same information then it is useful to also consider the slope and intercept of the estimated regression model. The slope should therefore approach 1 and the intercept should be close to 0 for the required relationships. Deviations from the expected values give an indication of the correspondence of the reference and mapped imperviousness data.

All regression coefficients are greater than 0.80 which indicate a very good relationship between the reference and map density values. The lowest value is related to the South-West demonstration site with complex landscapes with a  $R^2$  close to 0.8. Regression coefficients are greater than 0.90 for the Central and South-East demonstration sites.

Regression slopes are consistently close to 1 with very few variability except for the South-East demonstration site (close to 0.75). All slope values are smaller than 1,0 which is mainly a result of commission errors (IMD detected in non-sealed areas). The intercept parameter shows good values close to 0 but slightly over 0 which is a result of the commission errors (positive intercept).

### 5.1.2.2 Thematic accuracy

#### 5.1.2.2.1 Imperviousness 2017 & 2018

The below confusion matrices give a summary of the internal accuracy assessment of the improved HRL Imperviousness 2017 and 2018 for the demonstration site South-West.

**Table 5-4: Confusion matrix of the internal validation of the IMD 2017 in demo site South-West (area-weighted)**

| IMD_2017_10m_SW_03035 |            | REFERENCE  |                     |         | User Accuracy | Confidence Interval |
|-----------------------|------------|------------|---------------------|---------|---------------|---------------------|
|                       |            | Non-Sealed | Sealed              | Total   |               |                     |
| PRODUCT               | Non-Sealed | 890.28     | 7.21                | 897.49  | 99.20%        | 0.57%               |
|                       | Sealed     | 15.30      | 87.21               | 102.51  | 85.07%        | 1.67%               |
|                       | Total      | 905.58     | 94.42               | 1,000   |               |                     |
|                       |            |            | Producer Accuracy   | 98.31 % | 92.37 %       | 97.75 %             |
|                       |            |            | Confidence Interval | 0.55 %  | 1.61 %        | 0.75                |
|                       |            |            |                     |         |               | Overall Accuracy    |
|                       |            |            |                     |         |               | Confidence Interval |
|                       |            |            |                     |         |               | 0.99                |
|                       |            |            |                     |         |               | F-Score Non Sealed  |
|                       |            |            |                     |         |               | 0.89                |
|                       |            |            |                     |         |               | F-Score Sealed      |
|                       |            |            |                     |         |               | 0.87                |
|                       |            |            |                     |         |               | Kappa               |

The below confusion matrices give a summary of the internal accuracy assessment of the HRL Imperviousness 2018 for the demonstration sites, see Table 5-5, Table 5-6, and Table 5-7, respectively for the South-West, the Central and then the South-East demonstration sites.

**Table 5-5: Confusion matrix of the internal validation of the IMD 2018 in demo site South-West (area-weighted)**

| IMD_2018_10m_SW_03035 |            | REFERENCE  |                     |         | User Accuracy | Confidence Interval |
|-----------------------|------------|------------|---------------------|---------|---------------|---------------------|
|                       |            | Non-Sealed | Sealed              | Total   |               |                     |
| PRODUCT               | Non-Sealed | 179.49     | 1.61                | 181.10  | 99.11 %       | 0.51 %              |
|                       | Sealed     | 4.42       | 17.93               | 22.36   | 80.21 %       | 1.69 %              |
|                       | Total      | 183.91     | 19.55               | 203.46  |               |                     |
|                       |            |            | Producer Accuracy   | 97.59 % | 91.74 %       | 97.03 %             |
|                       |            |            | Confidence Interval | 0.59 %  | 1.63 %        | 0.77                |
|                       |            |            |                     |         |               | Overall Accuracy    |
|                       |            |            |                     |         |               | Confidence Interval |
|                       |            |            |                     |         |               | 0.98                |
|                       |            |            |                     |         |               | F-Score Non IMD     |
|                       |            |            |                     |         |               | 0.86                |
|                       |            |            |                     |         |               | F-Score IMD         |
|                       |            |            |                     |         |               | 0.84                |
|                       |            |            |                     |         |               | Kappa               |

**Table 5-6: Confusion matrix of the internal validation of the IMD 2018 in demo site Central (area-weighted)**

| IMD_2018_10m_CE_03035 |            | REFERENCE  |                     |         | User Accuracy | Confidence Interval |
|-----------------------|------------|------------|---------------------|---------|---------------|---------------------|
|                       |            | Non-Sealed | Sealed              | Total   |               |                     |
| PRODUCT               | Non-Sealed | 181.10     | 181.10              | 181.10  | 99.46 %       | 0.09 %              |
|                       | Sealed     | 22.36      | 22.36               | 22.36   | 95.70 %       | 1.43 %              |
|                       | Total      | 203.46     | 203.46              | 203.46  |               |                     |
|                       |            |            | Producer Accuracy   | 99.56 % | 94.76%        | 99.11 %             |
|                       |            |            | Confidence Interval | 0.08 %  | 1.37 %        | 0.12%               |
|                       |            |            |                     |         |               | Overall Accuracy    |
|                       |            |            |                     |         |               | Confidence Interval |
|                       |            |            |                     |         |               | 1.00                |
|                       |            |            |                     |         |               | F-Score Non IMD     |
|                       |            |            |                     |         |               | 0.95                |
|                       |            |            |                     |         |               | F-Score IMD         |
|                       |            |            |                     |         |               | 0.95                |
|                       |            |            |                     |         |               | Kappa               |

**Table 5-7: Internal validation of the IMD 2018 in demo site South-East (area-weighted)**

| IMD_2018_10m_SE_03035 |            | REFERENCE           |         |        | User Accuracy | Confidence Interval |
|-----------------------|------------|---------------------|---------|--------|---------------|---------------------|
|                       |            | Non-Sealed          | Sealed  | Total  |               |                     |
| PRODUCT               | Non-Sealed | 214.45              | 0.42    | 214.87 | 99.81 %       | 0.07 %              |
|                       | Sealed     | 0.70                | 4.46    | 5.16   | 86.37%        | 2.12 %              |
|                       | Total      | 215.16              | 4.88    | 220.03 |               |                     |
|                       |            | Producer Accuracy   | 99.67 % | 91.42% | 99.41 %       | Overall Accuracy    |
|                       |            | Confidence Interval | 0.08 %  | 1.98 % | 0.12 %        | Confidence Interval |
|                       |            |                     |         |        | 1.00          | F-Score Non IMD     |
|                       |            |                     |         |        | 0.89          | F-Score IMD         |
|                       |            |                     |         |        | 0.89          | Kappa               |

### 5.1.2.2.2 Imperviousness Classified Change 2015/17-2018

The below confusion matrices give a summary of the internal accuracy assessment of the HRL Imperviousness Classified Change 2015/17-2018 for the demonstration sites, see Table 5-8, Table 5-9 and Table 5-10 respectively for the South-West, the Central and then the South-East demonstration sites.

**Table 5-8: Confusion matrix of the internal validation of the IMCC 1718 in demo site South-West (area-weighted)**

| IMCC_1718_10m_SW_03035 |                     | REFERENCE          |              |                  |           |        | User Accuracy | Confidence Interval        |
|------------------------|---------------------|--------------------|--------------|------------------|-----------|--------|---------------|----------------------------|
|                        |                     | Unchanged unsealed | New Built-up | Unchanged sealed | Increased | Total  |               |                            |
| PRODUCT                | Unchanged unsealed  | 213.13             | 0.01         | 0.01             |           | 213.14 | 99.99%        | 0.01%                      |
|                        | New Built-up        |                    | 0.35         | 0.01             |           | 0.35   | 97.29%        | 0.04%                      |
|                        | Unchanged sealed    | 0.70               |              | 8.98             |           | 9.68   | 92.80%        | 0.07%                      |
|                        | Increased           |                    |              |                  | 0.01      | 0.01   | 100.00%       | 0.00%                      |
|                        | Total               | 213.82             | 0.35         | 9.00             | 0.01      | 223.18 |               |                            |
|                        | Producer Accuracy   | 99.67%             | 98.24%       | 99.81%           | 100.00%   |        | 99.68%        | Overall Accuracy           |
|                        | Confidence Interval | 0.02%              | 0.09%        | 0.01%            | 0.00%     |        | 0.03%         | Confidence Interval        |
|                        |                     |                    |              |                  |           |        | 1.00          | F-Score unchanged unsealed |
|                        |                     |                    |              |                  |           |        | 0.98          | F-Score new built-up       |
|                        |                     |                    |              |                  |           |        | 0.96          | F-Score unchanged sealed   |
|                        |                     |                    |              |                  |           |        | 1.00          | F-Score increased          |
|                        |                     |                    |              |                  |           |        | 0.96          | Kappa                      |

**Table 5-9: Confusion matrix of the internal validation of the IMCC 1518 in demo site Central (area-weighted)**

| IMCC_1518_20m_CE_03035 |                     | REFERENCE          |              |                  |        | User Accuracy | Confidence Interval        |
|------------------------|---------------------|--------------------|--------------|------------------|--------|---------------|----------------------------|
|                        |                     | Unchanged unsealed | New Built-up | Unchanged sealed | Total  |               |                            |
| PRODUCT                | Unchanged unsealed  | 327.29             | 0.00         | 3.85             | 331.15 | 98.83%        | 0.27%                      |
|                        | New Built-up        |                    | 0.34         |                  | 0.34   | 100.00%       | 0.00%                      |
|                        | Unchanged sealed    | 1.90               | 0.04         | 21.62            | 23.56  | 91.77%        | 1.06%                      |
|                        | Total               | 329.18             | 0.39         | 25.48            | 355.06 |               |                            |
|                        | Producer Accuracy   | 99.42%             | 87.55%       | 84.87%           |        | 98.37%        | Overall Accuracy           |
|                        | Confidence Interval | 0.13%              | 1.69%        | 2.24%            |        | 0.37%         | Confidence Interval        |
|                        |                     |                    |              |                  |        | 0.99          | F-Score unchanged unsealed |
|                        |                     |                    |              |                  |        | 0.93          | F-Score new built-up       |
|                        |                     |                    |              |                  |        | 0.88          | F-Score unchanged sealed   |
|                        |                     |                    |              |                  |        | 0.87          | Kappa                      |

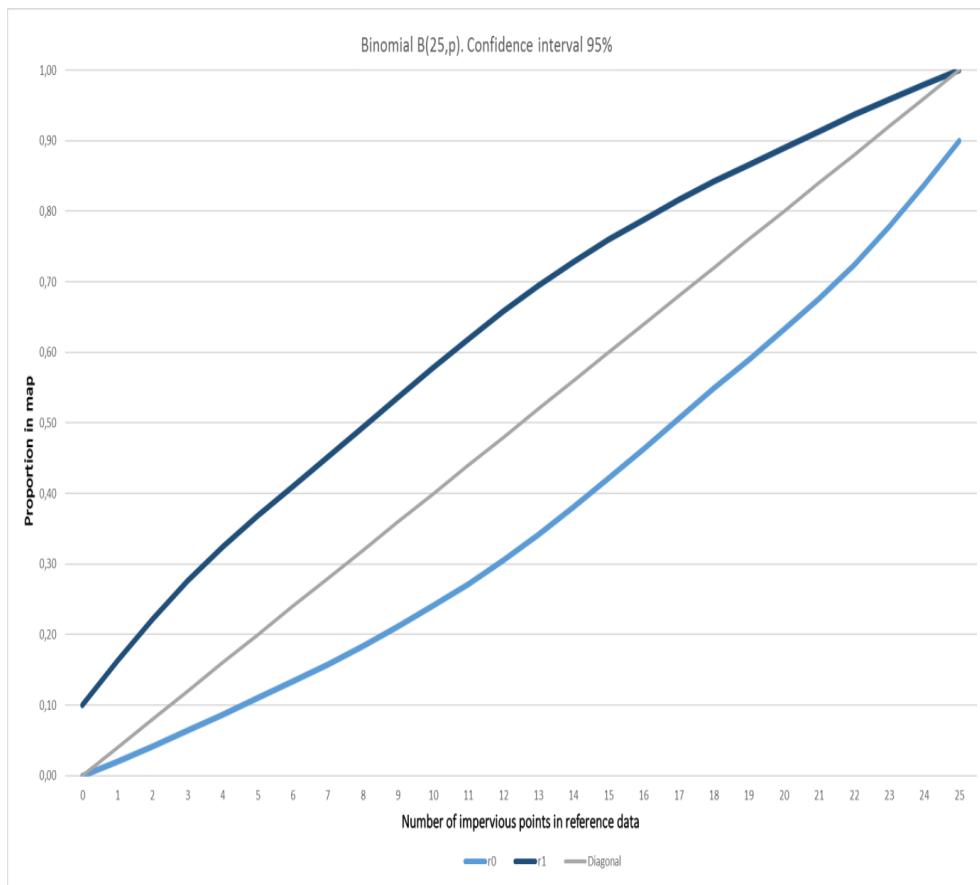
**Table 5-10: Confusion matrix of the internal validation of the IMCC 1518 in demo site South-East (area-weighted)**

| IMCC_1518_20m_SE_03035 |                     | REFERENCE          |              |                  |        | User Accuracy | Confidence Interval        |
|------------------------|---------------------|--------------------|--------------|------------------|--------|---------------|----------------------------|
|                        |                     | Unchanged unsealed | New Built-up | Unchanged sealed | Total  |               |                            |
| PRODUCT                | Unchanged unsealed  | 248.01             |              | 0,01             | 248,02 | 100,00%       | 0.00%                      |
|                        | New Built-up        |                    | 0.01         |                  | 0,01   | 100,00%       | 0.00%                      |
|                        | Unchanged sealed    | 1.05               | 0.01         | 4,61             | 5,67   | 81,38%        | 1.75%                      |
|                        | Total               | 249.06             | 0.02         | 4,62             | 253,69 |               |                            |
|                        | Producer Accuracy   | 99.58%             | 60.82%       | 99,83%           |        | 99.58%        | Overall Accuracy           |
|                        | Confidence Interval | 0.05%              | 3.87%        | 0.04%            |        | 0.18%         | Confidence Interval        |
|                        |                     |                    |              |                  |        | 1.00          | F-Score unchanged unsealed |
|                        |                     |                    |              |                  |        | 0.76          | F-Score new built-up       |
|                        |                     |                    |              |                  |        | 0.90          | F-Score unchanged sealed   |
|                        |                     |                    |              |                  |        | 0.89          | Kappa                      |

Based on the confusion matrix, the internal validation results show, as expected, very good area-weighted user's and producer's accuracies for the stable unsealed and sealed classes (above the 90% threshold required). Moreover, the user's and producer's accuracies for the new built-up class also show very good results but have to be put into perspective since the change layer strata is very small and few sample units are concerned. The section **Error! Reference source not found.** will show that a lot of 2015 omission errors still exist and should impact the change validation. It should be notice that the accuracy assessment of all change classes is not feasible. Indeed, in the phase 2 of the project, an additional stratum has been introduced to better estimate the accuracy of the change features; but includes all type of changes without any distinction. Based on the relative small areas of the change classes, such approach leads to a non-

representation of certain classes. Nevertheless, a sampling design and a stratification of each change classes would not be relevant because it would lead to an over-sampling of the change classes that would impact the accuracy assessment. The accuracy figures way above the required 90% for the South-West demo site shows the relative good level of reliability of the reclassification procedure.

Regarding the Imperviousness change layers (IMC products), an internal validation was not performed since the IMC layer is a continuous change layer with very few exact matches between validation and map layer data. Indeed, based on the IMD internal validation (see Figure 5-19), the classification accuracies that would be obtained for each change degree class will be rather low which would not be representative of the quality of the products. In fact, when performing a simple linear regression analysis between reference and map data for the IMD layers, a high coefficient of regression is obtained ( $r^2=0.83$ ,  $r^2=0.91$  and  $r^2=0.91$  respectively for the demo site SW, CE and SE), but there is still a considerable amount of scatter as shown in Figure 5-18 with most map values lying within a 20-30% range of the reference value. It has to be highlighted that this is not only due to noisy map data, but there can also be substantial errors in the reference data due to the sampling procedure. This error is maximised for an IMD value of 50% which in fact could be within 30 and 70%. This can be modelled as illustrated in Figure 5-19. One solution would be to perform an internal validation by grouping IMC values into classes of urban density changes (3 to 5 classes respectively for the increase degree and decrease) and applying a “one-class tolerance” (+/- IMC class) which reflect the inherent variability in the IMD dataset. Such approach would require a specific validation dataset with a good representativity of each IMC classes; which do not exist right now.



**Figure 5-19: Representation of the behaviour of the 95% confidence interval for a 5x5 SSU grid over the whole range of imperviousness degree values**

### 5.1.2.2.3 Built-up 2018

The below confusion matrices give a summary of the internal accuracy assessment of the HRL Imperviousness 2018 for the demonstration sites, see Table 5-11, Table 5-12 and Table 5-13 respectively for the South-West, the Central and then the South-East demonstration sites.

**Table 5-11: Confusion matrix of the internal validation of the IBU 2018 in demo site South-West (area-weighted)**

| IBU_2018_10m_SW_03035 |                   | REFERENCE    |          |        | User Accuracy | Confidence Interval  |
|-----------------------|-------------------|--------------|----------|--------|---------------|----------------------|
|                       |                   | Non-built-up | Built-up | Total  |               |                      |
| PRODUCT               | Non-built-up      | 177.93       | 0.29     | 178.22 | 99.84%        | 0.37%                |
|                       | Built-up          | 1.17         | 5.26     | 6.43   | 81.82%        | 1.93%                |
|                       | Total             | 179.10       | 5.55     | 184.65 |               |                      |
|                       | Producer Accuracy | 99.35%       | 94.74%   |        | 99.21%        | Overall Accuracy     |
|                       |                   | 0.48%        | 0.33%    | 1.11%  | 0.48%         | Confidence Interval  |
|                       |                   |              |          |        | 1.00          | F-Score Non-built-up |
|                       |                   |              |          |        | 0.88          | F-Score Built-up     |
|                       |                   |              |          |        | 0.87          | Kappa                |

**Table 5-12: Confusion matrix of the internal validation of the IBU 2018 in demo site Central (area-weighted)**

| IBU_2018_10m_CE_03035 |                   | REFERENCE    |          |        | User Accuracy | Confidence Interval  |
|-----------------------|-------------------|--------------|----------|--------|---------------|----------------------|
|                       |                   | Non-built-up | Built-up | Total  |               |                      |
| PRODUCT               | Non-built-up      | 278.16       | 1.46     | 279.62 | 99.48%        | 0.08%                |
|                       | Built-up          | 2.63         | 19.00    | 21.63  | 87.84%        | 1.82%                |
|                       | Total             | 280.79       | 20.46    | 301.25 |               |                      |
|                       | Producer Accuracy | 99.06%       | 92.86%   |        | 98.64%        | Overall Accuracy     |
|                       |                   | 0.08%        | 1.23 %   |        | 0.27%         | Confidence Interval  |
|                       |                   |              |          |        | 0.99          | F-Score Non-built-up |
|                       |                   |              |          |        | 0.90          | F-Score Built-up     |
|                       |                   |              |          |        | 0.90          | Kappa                |

**Table 5-13: Confusion matrix of the internal validation of the IBU 2018 in demo site South-East (area-weighted)**

| IBU_2018_10m_SE_03035 |                   | REFERENCE    |          |        | User Accuracy | Confidence Interval  |
|-----------------------|-------------------|--------------|----------|--------|---------------|----------------------|
|                       |                   | Non-built-up | Built-up | Total  |               |                      |
| PRODUCT               | Non-built-up      | 207.74       | 0.29     | 208.04 | 99.86%        | 0.09%                |
|                       | Built-up          | 0.88         | 3.80     | 4.68   | 81.25%        | 2.47%                |
|                       | Total             | 208.62       | 4.09     | 212.71 |               |                      |
|                       | Producer Accuracy | 99.58%       | 92.86%   |        | 99.45%        | Overall Accuracy     |
|                       |                   | 0.09%        | 2.03%    |        | 0.13%         | Confidence Interval  |
|                       |                   |              |          |        | 1.00          | F-Score Non-built-up |
|                       |                   |              |          |        | 0.87          | F-Score Built-up     |
|                       |                   |              |          |        | 0.86          | Kappa                |

### 5.1.2.3 Discussion of the validation results

Regarding the HRL IMD 2017 and 2018 prototypes, the look and feel assessment shows very satisfying results. Based on the confusion matrices, the internal validation results show very good area-weighted user's and producer's accuracies. The IMD 2017 and 2018 products show a very high producer's accuracies above the requirements threshold (91.74% +/- 1.63% for the SW site; 94.76% +/- 1.37% for the CE site and 91.41% +/- 1.98% for the SE site). In other words, the validation shows a very limited amount of omission errors. The user's accuracy is slightly under the required 90% except for the Central demo site but shows good level of reliability (above 80.00%).

It is important to note that these results are based on the IMD product fully obtained automatically without manual enhancement, post-processing and without reclassification based on the change detection approach. So, user's and producer's accuracies could easily be largely improved.

Most of the omission errors concern small and isolated built-up features as exemplarily shown in Figure 5-20 below.



Figure 5-20: Case of omission (yellow dots) in isolated built-up features in the IMD 2017 prototype

Regarding the commission errors, as known from previous HRL IMP productions, the sources of errors are related to the following land use features: arable lands without crops, construction sites, quarries, dry riverbeds (e.g. bare soils). Examples are provided in Figure 5-21 .

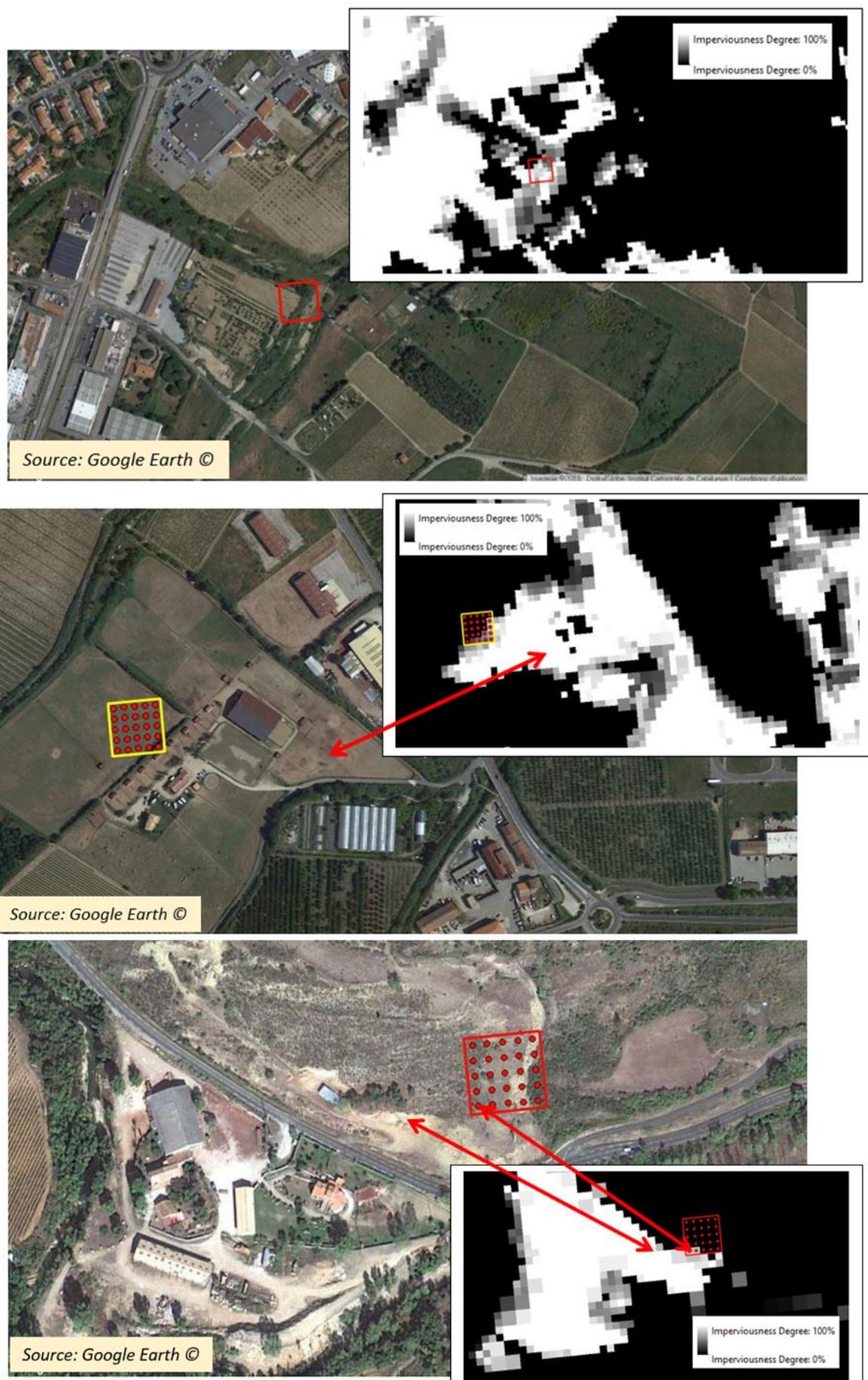


Figure 5-21: Examples of commission errors in the IMD 2017 prototype

Regarding the HRL BU 2018 prototypes, the look and feel assessment shows also satisfying results. Based on the confusion matrices, the internal validation results show very good area-weighted user's and producer's accuracies. It should be notice the high adding value of the Pantex index for the detection of the built-up features. Nevertheless, in comparison with the HRL IMD results, the accuracy assessment shows lower accuracy figures; especially regarding the user accuracy. Indeed, BU products show a producer accuracies above the requirements threshold (94.74% +/- 1.23% for the SW site; 92.86% +/- 1.23% for the CE site and 92.86% +/- 2.03% for the SE site). In other words, the validation shows a very limited amount of omission errors but a greater rate of commission errors. The spatial resolution of the S-2 data (e.g. 10 m) explains mostly the commission errors. Indeed, based on HR images, it is difficult to separate built-up (buildings) from other sealed areas (roads, parking lots) in urban areas.

It is important to note that these results are based on the IMD product fully obtained automatically without manual enhancement and post-processing. So, the accuracy assessment figures could easily be improved.

### 5.1.3 Change Detection and Incremental Update Results and Validation

As described before, in phase 1, the benchmarking of the best change detection approach is only done on the basis of the best status layer classification for 2017 based on Sentinel-2 cloud-free images (high resolution spectral and spatial). The implementation has been done on the demonstration site South-West. The change results obtained from the reference calibration dataset are presented in Table 5-14 .

**Table 5-14: Proportional distribution of detected changes within the IMC layer (SW demonstration site, phase 1)**

|  | % of total change areas |
|--|-------------------------|
| New built-up 2017                                | 5.80%                   |
| Omission errors 2015 (undetected built-up 2015): | 78.90%                  |
| Commission errors 2017 (false built-up 2017)     | 15.30%                  |

Based on the calibration dataset, the relative magnitude of actual change is estimated to 5.8% of the total change areas. Thus, the errors concern the remaining 94% of the change areas detected, which is mostly due to the different spatial resolutions, i.e. 10m vs 20m. There is a high amount of omission errors coming from the reference data (2015), impacting the change areas detected.

Although the majority of errors is detected in the reference layer, new errors can also appear in the detection of change between two-time periods. The change layers' errors can be, as mentioned before, due to the following factors:

- Omission of change (quantifiable only in the validation, but not as a delineation from the product);
- Commission errors added for the new period (cf. Figure 5-24 );
- Omission errors detected for the previous period (cf. Figure 5-23 ).

Regarding the latter, it should be noted that the specifications of the HRL 2015 layer are different from the specifications and the input data situation for the HRL 2017, hence a comparison is not fully "fair". Indeed, for 2015 the product was obtained through a 20-meter spatial resolution production mostly based on Landsat data, whereas the HRL 2017 could draw upon Sentinel-2 and a 10-meter resolution, which explains that most of the omission errors concern small and isolated built-up features (cf Figure 5-22, Figure 5-23, Figure 5-24). It would therefore be a wrong conclusion to consider the reference layer of 2015 being of poor quality.

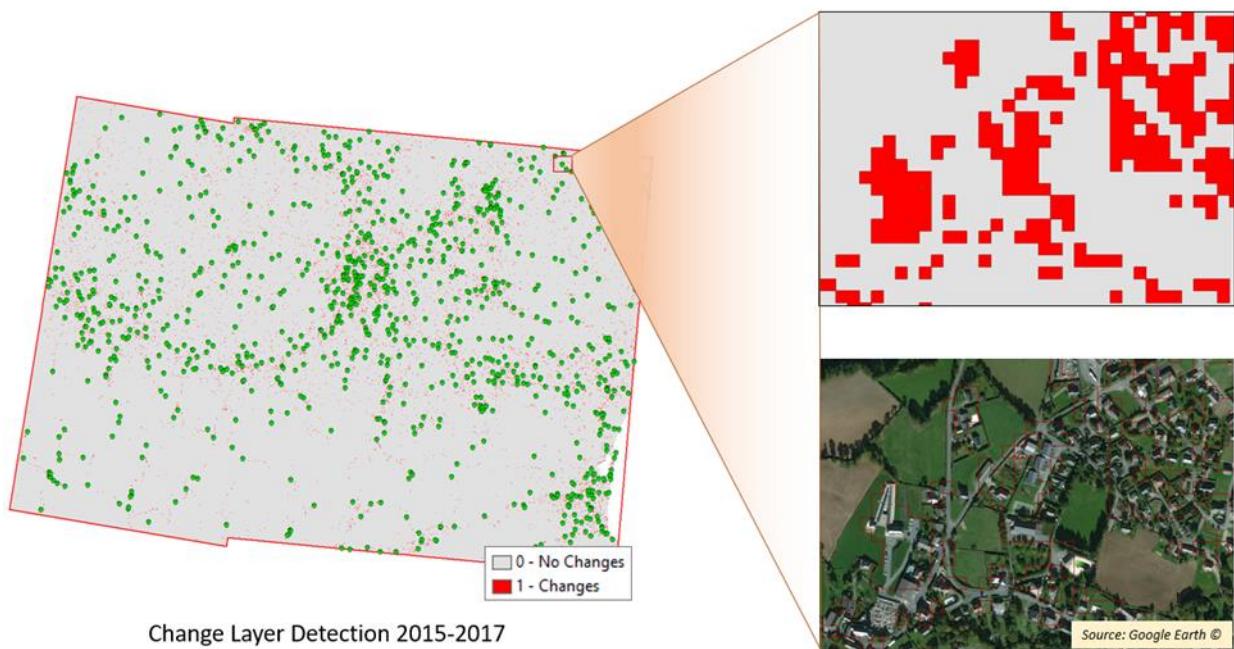


Figure 5-22: Example of calibration points for the newly detected built-up in 2017

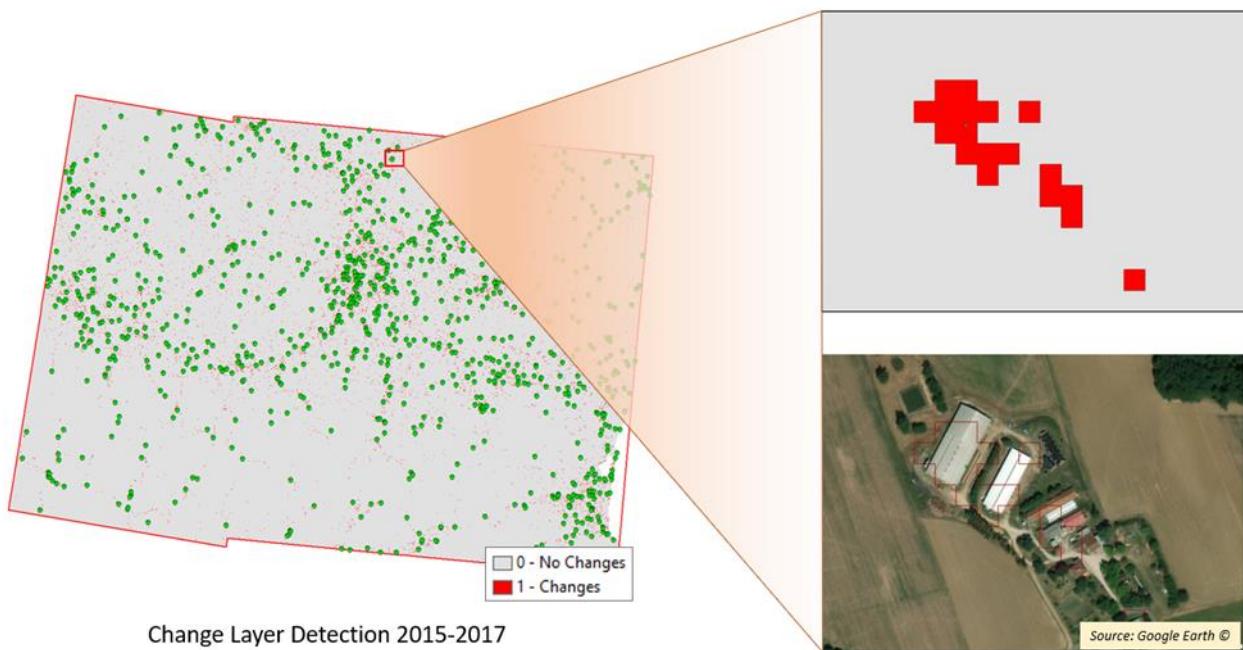
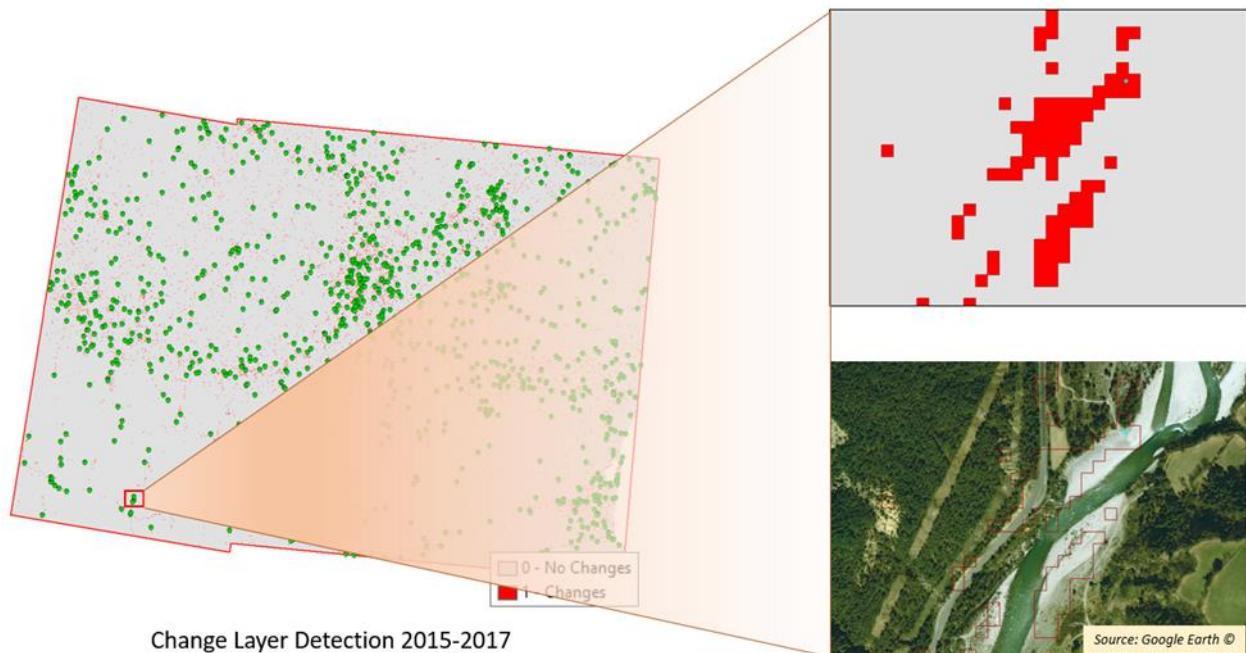


Figure 5-23: Example of omission errors in 2015 (undetected built-up in 2015)



**Figure 5-24: Example of commission errors in 2017, false built-up in 2017**

In phase 2, the change results and tendency obtained are similar as shown in Table 5-15.

**Table 5-15: Proportional distribution of detected changes within the IMC layers (phase 2)**

| % of total change areas                           |            |   |          |            |
|---|------------|---|----------|------------|
|   | South-West |   | Central  | South-East |
| New built-up 2018                                 | 7.60%      | New built-up 2018                                 | 15.14%   | 9.09%      |
| Omission errors 2017* (undetected built-up 2017): | 35.05%     | Omission errors 2015* (undetected built-up 2015): | 71.35%   | 66.38%     |
| Commission errors 2018* (false built-up 2018)     | 57.35%     | Commission errors 2015* (false built-up 2015)     | 13.51%   | 24.52%     |
| Total Change Area (km <sup>2</sup> )              | 198.57     | Total Change Area (km <sup>2</sup> )              | 1,092.32 | 554.82     |

\* For the South-West demonstration site: change layer 2017-2018. The HRL IMD 2017 has been created as part of Phase 1 (explaining the reduced amount of change areas and the relative proportionality between omission and commission errors)

Based on the calibration dataset, the relative magnitude of actual change is estimated between 7.6% (SW) and 15.1% (Central) of the total change areas. Thus, the errors concern the remaining 92% to 85% of the change areas detected; results which are in line with the ones obtained in phase 2. For the Central and South-East demonstration sites for which the change detection is based on the HRL 2015, most of the error is due to the different spatial resolutions, i.e. 10m vs 20m. As shown in phase 1, there is a high amount of omission errors coming from the reference data (2015), impacting the change areas detected. For the South-West demonstration site, most of the errors comes from the new 2018 classification (57.4% vs 35.1% for the omission errors 2017) but must be counterbalance with the low proportion of change areas (198.6 km<sup>2</sup> representing 0.3% of the total demonstration area).

Again, it should be stressed out that the specifications of the HRL 2015 layer are different from the specifications and the input data situation for the HRL 2017 (10m vs 20m) which explains most of the omission errors.

Those statistics are then used in a reduction-of-bias algorithm to ensure temporal and spatial consistency of changes or the mask. This approach consists in a re-classification step linking the three categories (real change, omission, commission) with suitable training data from 2018 imagery. The reclassification procedure over the initially detected change areas is done:

- either by adjusting the classification confidence thresholds
- or by re-classifying, following the same procedure used for the generation of the initial 2018 status layer, but focusing only on the gain strata and using the Reference Database for Change Calibration and specific training data collected in each stratum.

Once the re-classification procedure is finalized, actual change areas can be isolated from errors and the updated sealed surface masks can be generated.

The below confusion matrices give a summary of the internal accuracy assessment of the improved/reclassified HRL Imperviousness 2018 for the demonstration sites (Table 5-16, Table 5-17, Table 5-18).

**Table 5-16: Confusion matrix of the internal validation of the IMD 2018 in demo site South-West (area-weighted)**

| IMD_2018_10m_SW_03035 |                     | REFERENCE  |         |        | User Accuracy | Confidence Interval |
|-----------------------|---------------------|------------|---------|--------|---------------|---------------------|
|                       |                     | Non-Sealed | Sealed  | Total  |               |                     |
| PRODUCT               | Non-Sealed          | 173.10     | 1.62    | 174.72 | 99.07 %       | 0.51 %              |
|                       | Sealed              | 3.91       | 24.83   | 28.74  | 86.40 %       | 1.57 %              |
|                       | Total               | 177.01     | 26.45   | 203.46 |               |                     |
|                       | Producer Accuracy   | 97.79 %    | 93.89 % |        | 97.28 %       | Overall Accuracy    |
|                       | Confidence Interval | 0.57 %     | 1.41 %  |        | 0.70 %        | Confidence Interval |
|                       |                     |            |         |        | 0.98          | F-Score Non IMD     |
|                       |                     |            |         |        | 0.90          | F-Score IMD         |
|                       |                     |            |         |        | 0.88          | Kappa               |
|                       |                     |            |         |        |               |                     |

**Table 5-17: Confusion matrix of the internal validation of the IMD 2018 in demo site Central (area-weighted)**

| IMD_2018_10m_CE_03035 |                     | REFERENCE  |        |        | User Accuracy | Confidence Interval |
|-----------------------|---------------------|------------|--------|--------|---------------|---------------------|
|                       |                     | Non-Sealed | Sealed | Total  |               |                     |
| PRODUCT               | Non-Sealed          | 285.52     | 1.55   | 287.07 | 99.46 %       | 0.06 %              |
|                       | Sealed              | 0.68       | 28.09  | 28.77  | 97.64 %       | 1.21 %              |
|                       | Total               | 286.20     | 29.64  | 315.84 |               |                     |
|                       | Producer Accuracy   | 99.76 %    | 94.76% |        | 99.29 %       | Overall Accuracy    |
|                       | Confidence Interval | 0.05 %     | 1.37 % |        | 0.06 %        | Confidence Interval |
|                       |                     |            |        |        | 1.00          | F-Score Non IMD     |
|                       |                     |            |        |        | 0.96          | F-Score IMD         |
|                       |                     |            |        |        | 0.96          | Kappa               |
|                       |                     |            |        |        |               |                     |
|                       |                     |            |        |        |               |                     |

**Table 5-18: Internal validation of the IMD 2018 in demo site South-West, Central and South-East (area-weighted)**

| IMD_2018_10m_SE_03035 |            | REFERENCE  |         |        | User Accuracy | Confidence Interval |
|-----------------------|------------|------------|---------|--------|---------------|---------------------|
|                       |            | Non-Sealed | Sealed  | Total  |               |                     |
| PRODUCT               | Non-Sealed | 214.83     | 0.44    | 215.26 | 99.80 %       | 0.06 %              |
|                       | Sealed     | 0.33       | 4.44    | 4.77   | 93.09 %       | 2.12 %              |
|                       | Total      | 215.16     | 4.88    | 220.03 |               |                     |
| Producer Accuracy     |            | 99.85 %    | 91.04 % |        | 99.65 %       | Overall Accuracy    |
| Confidence Interval   |            | 0.05 %     | 1.98 %  |        | 0.08 %        | Confidence Interval |
|                       |            |            |         |        | 1.00          | F-Score Non IMD     |
|                       |            |            |         |        | 0.92          | F-Score IMD         |
|                       |            |            |         |        | 0.92          | Kappa               |

The internal validation results show improved area-weighted user's and producer's accuracies due to the reclassification procedure. The technical errors from the reference year 2018 are reduced which lead to greater accuracies for the IMD 2018 products.

Again, it is important to note that these results are based on the IMD product fully obtained automatically without manual enhancement and post-processing. So, user's and producer's accuracies could easily be improved.

## 5.2 Prototype of a potential Future HRL Forest

This section shows the prototypical implementation of the Forest Prototypes (both the improved status layers DLT and TCD and the incremental update layer TCC). Firstly, the data and processing setup is described (section 5.2.1), followed by presenting the classification results and validation (section 5.2.2), and lastly the demonstration of the change detection and incremental update results and validation (section 5.2.3).

### 5.2.1 Data and Processing Setup

Besides Sentinel-2 and Sentinel-1, representing the main data sources, also other data have been integrated into the workflow and specifically prepared (e.g. the Copernicus High-Resolution Layers 2015, LUCAS 2018). Thereafter, the conducted pre-processing steps are detailed and the experimental setup for the classification and incremental update approach is explained.

#### 5.2.1.1 Input Data and Data Integration

The ECoLaSS demonstration site North in Sweden comprises six adjacent Sentinel-2 tiles (33VVF, 33VVH, 33VVG, 33VWF, 33VWH, 33VWG (Figure 5-15). The ECoLaSS demonstration site Central in Central Europe consists of 9 adjacent Sentinel-2 tiles (32TPT, 32TNT, 32TMT, 32UNU, 32UPU, 32UMU, 32UPV, 32UNV, 32UMV) (Figure 5-16). The third demonstration site (southeast), consists as in the case of North, of 6 grids Sentinel-2 (34TFL, 34TFN, 34TFM, 34TGL, 34TGN, 34TGM) (Figure 5-26). For all demo sites Sentinel-2A+B data in 10m resolution have been processed, with the addition of Sentinel-1A+B in phase 2 (see [AD06]).

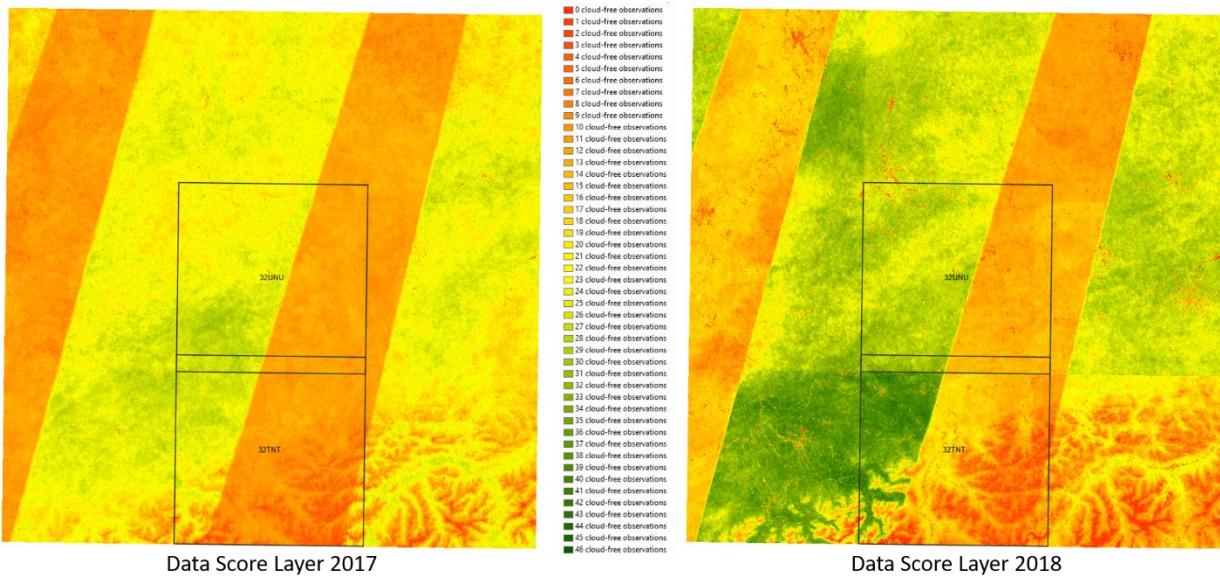
In phase 2, the time frame was extended from 15-March to 15-September, to increase the number of observations per pixel. This time frame was finally applied for all demo sites. Table 5-19 provides an overview on the number of scenes processed for implementation of the FOR prototypes. A total of 946 Sentinel-2A+B scenes were pre-processed in phase 2 for the years 2017 and 2018 and the maximum cloud cover threshold has been set to 60% for all demo sites. The number of Sentinel-1 scenes processed in

phase 2 in VV and VH polarisations were 3,946 for the years 2017 and 2018. In total, 4,892 scenes have been pre-processed using the Sen2Cor and MACCS processors.

**Table 5-19: Summary of Sentinel-2 and Sentinel-1 scenes pre-processed per Demo Site and Tile with a max. cloud cover of 60%**

| Demo site  | Tile  | No. of Sentinel-2 scenes |      | No. of Sentinel-1 scenes |       |
|------------|-------|--------------------------|------|--------------------------|-------|
|            |       | 2017                     | 2018 | 2017                     | 2018  |
| NORTH      | 33VVF | 19                       | 31   |                          |       |
|            | 33VVG | 15                       | 34   |                          |       |
|            | 33VVH | 11                       | 30   |                          |       |
|            | 33VWF | 17                       | 32   |                          |       |
|            | 33VWG | 15                       | 35   |                          |       |
|            | 33VWH | 19                       | 47   |                          |       |
|            | Total | 96                       | 210  | 708                      | 654   |
| CENTRAL    | 32TPT | 13                       | 14   |                          |       |
|            | 32TNT | 17                       | 25   |                          |       |
|            | 32TMT | 17                       | 20   |                          |       |
|            | 32UPU | 15                       | 19   |                          |       |
|            | 32UNU | 16                       | 15   |                          |       |
|            | 32UMU | 17                       | 17   |                          |       |
|            | 32UPV | 29                       | 30   |                          |       |
|            | 32UNV | 17                       | 19   |                          |       |
|            | 32UMV | 19                       | 28   |                          |       |
| Total      |       | 160                      | 187  | 644                      | 946   |
|            |       |                          |      |                          |       |
| SOUTH-EAST | 34TFL | 39                       | 32   |                          |       |
|            | 34TFM | 18                       | 26   |                          |       |
|            | 34TGL | 25                       | 25   |                          |       |
|            | 34TGM | 20                       | 23   |                          |       |
|            | 34TFN | 17                       | 22   |                          |       |
|            | Total | 143                      | 150  | 488                      | 506   |
| SUM        | 21    | 399                      | 547  | 1,840                    | 2,106 |

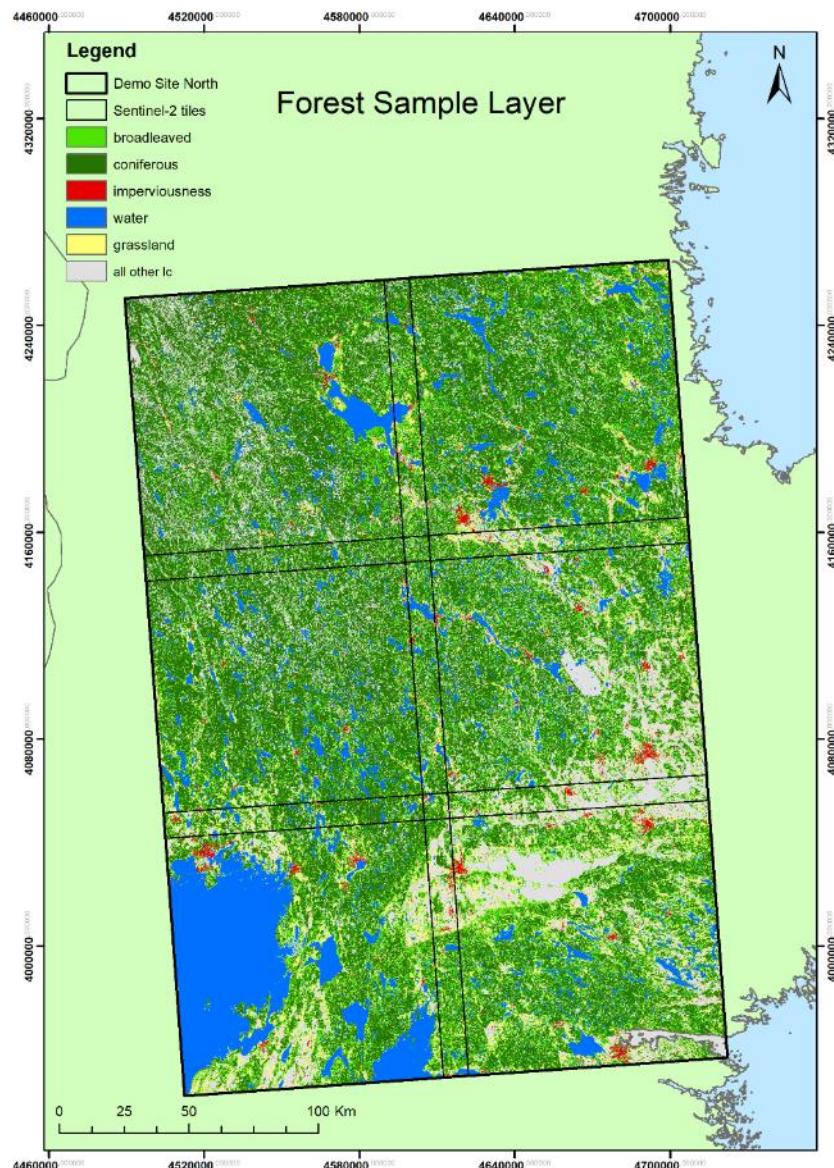
From the Scene Classification Layer (SCL) obtained through Sen2Cor processed Level-2A data, pixel-based cloud masks have been calculated. Figure 5-25 shows exemplarily the Sentinel-2 Data Score Layer (S2DSL) for the demonstration site Central (overlaid by the test site Austria/Germany). It provides the respective data score (inverted cloud score) for each pixel and thus information on the number of available cloud-free Sentinel-2 observations for the selected time period 15-March to 15-September in 2017 and 2018. The significantly higher number of cloud-free observations in 2018 is due to the full availability of Sentinel-2B data from the second half of the year 2017 onwards.



**Figure 5-25: Sentinel-2 Data Score Layer (S2DSL): Number of cloud-free Sentinel-2 observations per pixel**

The created S2DSL is provided together with the products and provides information on the number of cloud-free Sentinel-2 observations per pixel. This information can be additionally used to assess the accuracy of the DLT and TCC products in areas with a poor data situation, since areas with fewer observations might tend to have a higher uncertainty in classification accuracy. Furthermore, the Data Score Layer can be also used to relocate samples from areas of low data density to regions with a higher observation density.

Additionally, the existing Copernicus High-Resolution Layers 2015 (Imperviousness Degree, Dominant Leaf Type, Grassland, Water and Wetness –WaW-) in 20m spatial resolution have been combined to generate a forest sample layer, which was used for the implemented automatic sampling approach. The layer has been derived by integration of the HRLs 2015 in a hierarchical order. An exemplary demonstration of the Sample Layer is presented in Figure 5-26. It consists of six thematic classes: broadleaved, coniferous, imperviousness, water, grassland, and all other land cover classes. Table 5-20 provides relevant statistics on the derived sample layer per demonstration site. Forests (broadleaved/coniferous) represent the dominant land cover type in all sites, followed by the all other land cover class.



**Figure 5-26: Forest sample layer derived from the integrated Copernicus High-Resolution Layers 2015**  
 (© EuroGeographics for the administrative boundaries)

A systematic random sampling approach has been applied to collect 250 samples per tile from the generated Forest Sample Layer. The aim was to automatically collect a sufficient number of samples within broadleaved/coniferous forests and in areas with no tree cover (represented by specific thematic LC classes). This pool comprises the basis for the sample selection for the TCM and DLT classifications on tile-basis. More details on the selected sampling approach are given in [AD07].

**Table 5-20: Land Cover statistics of the derived HRL 2015 sample layer**

| Demo site  | Class Code | Class Name     | Area km <sup>2</sup> | Area % |
|------------|------------|----------------|----------------------|--------|
| CENTRAL    | 1          | Broadleaved    | 20,941.01            | 21.25  |
|            | 2          | Coniferous     | 19,678.75            | 19.97  |
|            | 3          | Imperviousness | 5,470.47             | 5.55   |
|            | 4          | Water          | 1,371.31             | 1.39   |
|            | 5          | Grassland      | 15,736.33            | 15.97  |
|            | 6          | All other LC   | 35,335.32            | 35.86  |
| NORTH      | 1          | Broadleaved    | 12,457.50            | 19.15  |
|            | 2          | Coniferous     | 26,369.55            | 40.54  |
|            | 3          | Imperviousness | 794.21               | 1.22   |
|            | 4          | Water          | 7,576.86             | 11.64  |
|            | 5          | Grassland      | 3,715.23             | 5.71   |
|            | 6          | All other LC   | 14,130.01            | 21.72  |
| SOUTH-EAST | 1          | Broadleaved    | 26,099.16            | 41.71  |
|            | 2          | Coniferous     | 8,084.86             | 12.92  |
|            | 3          | Imperviousness | 1,180.02             | 1.89   |
|            | 4          | Water          | 333.55               | 0.53   |
|            | 5          | Grassland      | 5,514.55             | 8.81   |
|            | 6          | All other LC   | 21,366.04            | 34.14  |

### 5.2.1.2 Pre-processing

The EO pre-processing steps include an atmospheric correction (including cirrus removal), cloud mask processing, topographic normalization with Minnaert correction and image export in TIFF format. For the atmospheric correction and topographic normalization, the ESA Sen2Cor software was utilized. Details on the specific pre-processing steps of the optical Sentinel-2 data are provided in [AD06].

On the other hand, the preprocessing of radar images (Sentinel-1) includes:

- automatic preparation of the digital elevation model
- automatic update of orbit files
- border noise removal
- radiometric calibration
- multi-looking and image filtering (speckle, adaptive)
- spatio-temporal speckle filtering
- radiometric terrain corrections and geo-referencing.

Further details about the geometric pre-processing and radiometric pre-processing steps for SAR Data research undertaken are provided in the final issue of WP 32 [AD06].

### 5.2.1.3 Experimental Setup

Methods and experiences from the previous testing and benchmarking exercise described in [AD07] have been considered to set up dedicated classification models for the forest prototypes in all three demonstration sites.

For the implementation of the TCM and DLT prototypes, a specific workflow has been set up within a python environment for machine learning. The workflow aims for an automated tree cover and dominant leaf type classification using time features derived from the Sentinel time series within a pre-defined observation period (15-March to 15-September). The maximum cloud cover threshold was set to 60% and automatically collected samples were subsequently checked, through the implementation of an automatic outlier detection using scatter plots. Table 5-21 shows the different parameters settings for the TCM and DLT products in the three FOR demonstration sites.

**Table 5-21: Parameter settings for TCM and DLT products**

| Product              | Demo Site  | Cloud Cover | Feature Selection                | No.of features | Processing time |
|----------------------|------------|-------------|----------------------------------|----------------|-----------------|
| <i>TCM 2017 10 m</i> | CENTRAL    | 60%         | Analysis of Variance Mean values | 49             | 1.5 h / Tile    |
| <i>TCM 2018 10 m</i> |            | 60%         | Analysis of Variance Mean values | 64             | 1.5 h / Tile    |
| <i>DLT 2018 10 m</i> |            | 60%         | Kfold Cross Validation           | 66             | 1.5 h / Tile    |
| <i>TCM 2017 10 m</i> | NORTH      | 60%         | Kfold Cross Validation           | 50             | 1.5 h / Tile    |
| <i>TCM 2018 10 m</i> |            | 60%         | K-fold Cross-Validation          | 68             | 1.5 h / Tile    |
| <i>DLT 2018 10 m</i> |            | 60%         | Kfold Cross- Validation          | 154            | 18 h / Tile     |
| <i>TCM 2017 10 m</i> | SOUTH-EAST | 60%         | K-fold Cross-Validation + visual | 52             | 1.5h / Tile     |
| <i>TCM 2018 10 m</i> |            | 60%         | K-fold Cross-Validation + visual | 73             | 1.5h / Tile     |
| <i>DLT 2018 10 m</i> |            | 60%         | K-fold Cross-Validation + visual | 48             | 1.5 h / Tile    |

As explained in the final WP 33 report [AD 07], the decision on the final settings regarding time window, number of features, sensors (i.e., Sentinel-1 only, Sentinel-2 only, Sentinel-1/-2 combined) was made considering the results of the testing and benchmarking phase, and the lessons learned from the first Forest prototype implementation in project phase 1. The final model parameter settings from the test sites have been finally transferred to the full demonstration site extent.

Two independent sample sets were used for classification and validation. The training sets for the classification of tree cover and dominant leaf type were extracted from the HRL 2015 sample layer as described above. For validation purposes, the LUCAS 2018 points were used, after performing a recoding of classes according to the ECoLaSS product specifications. These points served as ground-truth in the internal validation procedure (see section 5.2.2.1). Table 5-22 shows the distribution of sample points of the initial training datasets and its corresponding raster products for all three demonstration sites. Samples were weighted according to the estimated area proportion as derived from the Forest Sample Layer.

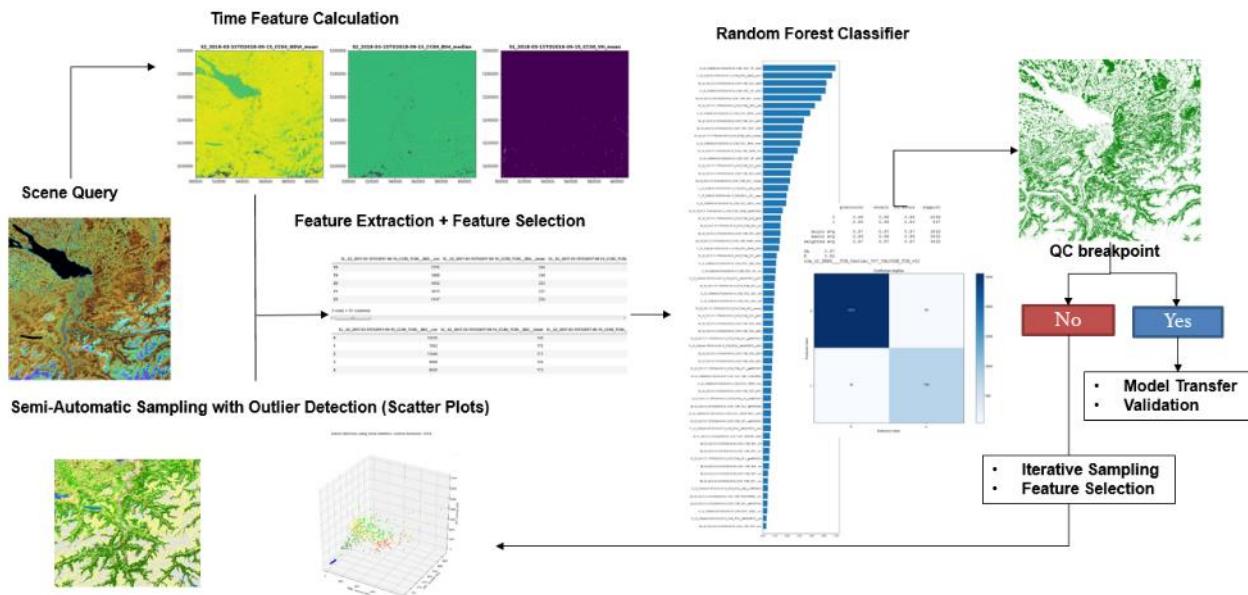
**Table 5-22: Sample distribution of the initial training dataset for TCM and DLT products per site.**

| Product              | Demo Site  | Class Name        | Class Code | Number |
|----------------------|------------|-------------------|------------|--------|
| <i>TCM 2017 10 m</i> | CENTRAL    | No Tree Cover     | 0          | 600    |
|                      |            | Tree Cover        | 1          | 300    |
|                      |            | No Tree Cover     | 0          | 600    |
|                      | NORTH      | Tree Cover        | 1          | 300    |
|                      |            | Broadleaved trees | 1          | 150    |
|                      |            | Coniferous trees  | 2          | 150    |
| <i>TCM 2018 10 m</i> | NORTH      | No Tree Cover     | 0          | 400    |
|                      |            | Tree Cover        | 1          | 400    |
|                      |            | No Tree Cover     | 0          | 433    |
|                      | SOUTH-EAST | Tree Cover        | 1          | 400    |
|                      |            | Broadleaved Trees | 1          | 214    |
|                      |            | Coniferous Trees  | 2          | 219    |
| <i>DLT 2018 10 m</i> | SOUTH-EAST | No Tree Cover     | 0          | 420    |
|                      |            | Tree Cover        | 1          | 220    |
|                      |            | No Tree Cover     | 0          | 420    |
|                      | CENTRAL    | Tree Cover        | 1          | 220    |
|                      |            | Broadleaved Trees | 1          | 115    |
|                      |            | Coniferous Trees  | 2          | 115    |

The classification workflow for both, TCM and DLT products is shown in Figure 5-27 and comprises the following steps:

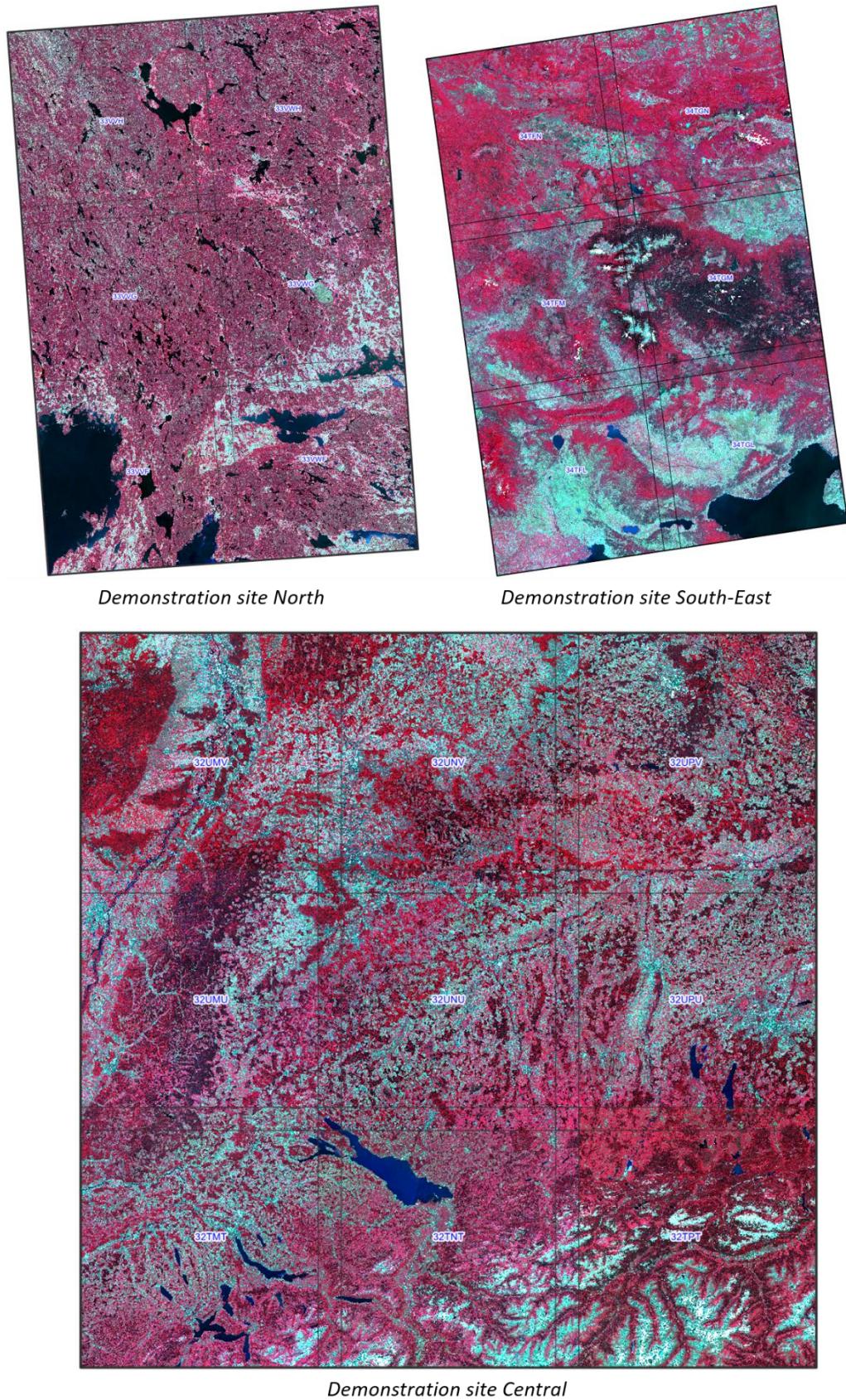
- Derivation of time features (complex and simple) from a time series of Sentinel bands/indices
- Creation of reference dataset, consisting of:
  - a set of training samples for the complete demo site
  - a set of independently interpreted validation samples

- Extraction of raster values at reference data locations (based on a set of samples) for the selected time features
- Training the classifier based on the training set resulting in a classification model
- Generation of an accuracy report based on the independent validation data
- Prediction/mapping with the calculated raster features via the final model. This step yields the predictions (classes), class-probabilities (one layer per class), and their reliability layers (max, probability, breaking ties, entropy).
- Potential iterative sampling and feature selection



**Figure 5-27: Forest prototypes workflow**

Independently from the TCM and DLT classifications, an improved Tree Cover Density classification at 10m spatial resolution has been performed for each demonstration site. Density values have been calculated on median time feature stacks which have been created from the Sentinel-2 bands (Figure 5-28). Thereby, the time period from 01-June to 31-August has been rated as the most promising one in order to obtain stable spectral characteristics. Based on the HRL2015 Tree Cover Density product, more than 1,000 samples (30m x 30m polygons) have been randomly collected in each demonstration site in order to cover the full range of density values. Subsequently, spectral mean values have been extracted and a multiple linear regression analysis has been performed. Potential outliers can be removed by visual inspection of the scatter plots or directly corrected using VHR reference data, if applicable. Previews on small scale allow a first assessment (look & feel) of the model accuracy which can be additionally checked by performing a cross-validation. Subsequently, the model is transferred to the image data stack in order to produce a seamless TCD “raw” raster. The final Tree Cover Density product is then derived by masking the seamless density raster with the generated TCM.



**Figure 5-28: Median Time Feature stacks of Sentinel-2 bands for generation of Tree Cover Density products.**  
Produced using Copernicus Sentinel data [2018]

Figure 5-29 provides the workflow for the TCD classification which comprises of the following steps:

- Derivation of time features from a time series of Sentinel-2 bands (all bands except the 60m bands)
- Mosaicking and stacking of tile-based time features to a seamless mosaic
- Automatic random sampling (30m x 30m polygons) with subsequent extraction of spectral mean values at reference data locations
- Iterative regression analysis (multiple linear regression) and removal of outliers using scatter plots
- Small-scale assessment of the model accuracy using previews across the region of interest
- Cross-validation of the sample dataset in order to assess the model accuracy
- Transferring of the model to the image data stack and subsequent classification of density values on pixel level, resulting in a seamless TCD “raw” raster
- Masking of the seamless TCD “raw” raster with its associated Tree Cover Mask

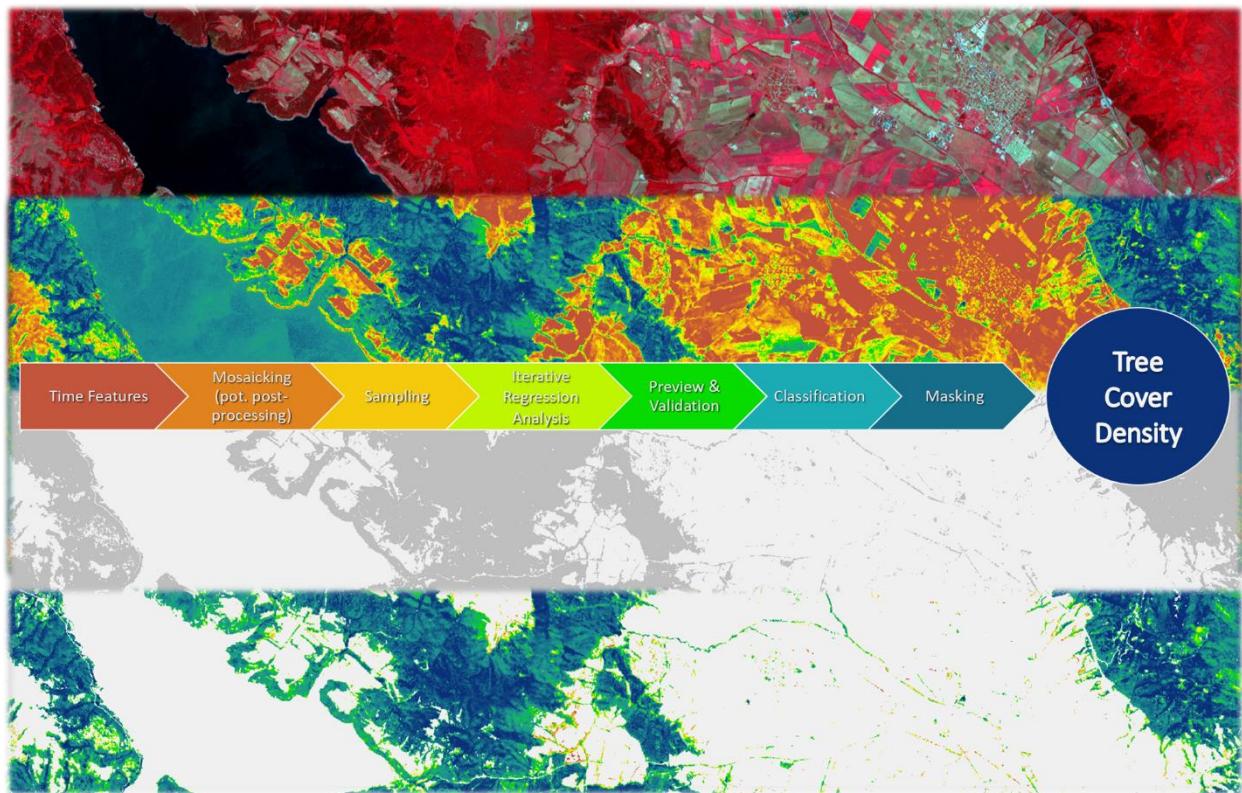
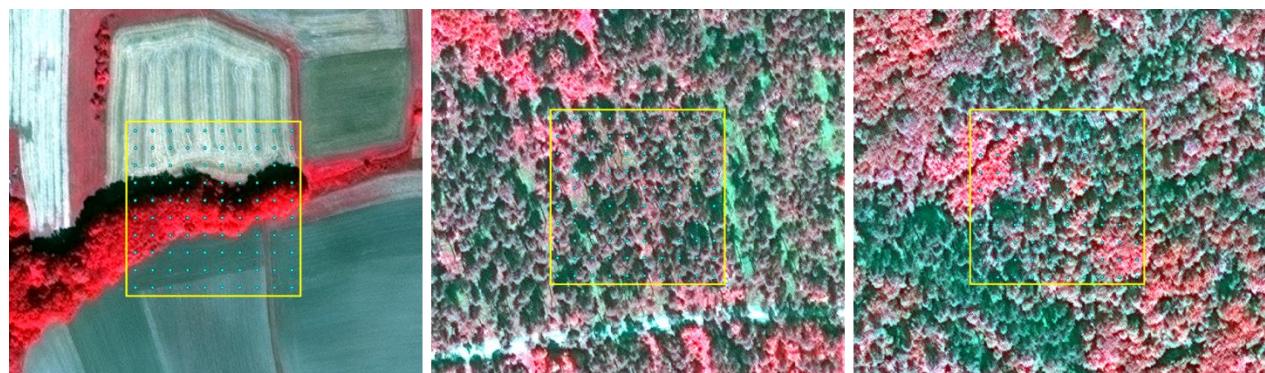
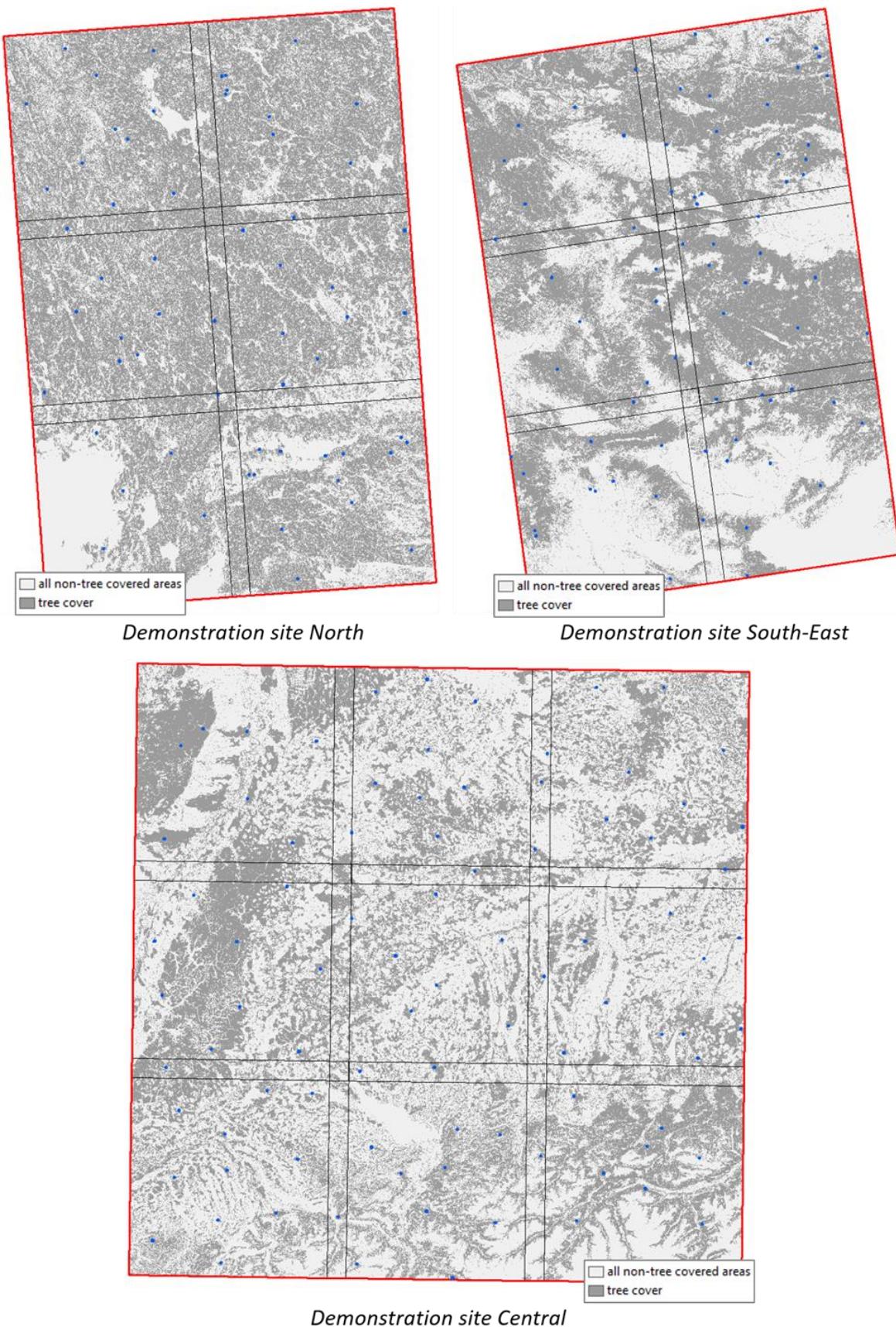


Figure 5-29: Tree Cover Density workflow

TCD products have been validated with an independent reference dataset for validation comprising of 100m x 100m polygons (PSUs) with a regular 10 x 10 point grid (SSUs) in order to assess the tree cover density (see also section 5.2.2.1). For this purpose, a reference dataset comprising of 223 polygon samples has been set up. Samples have been interpreted on the best available VHR data sources. Figure 5-30 shows some examples for the reference samples and Figure 5-31 presents the distribution of the samples within the three demonstration sites.



**Figure 5-30: Examples of 100m x 100m polygon samples for validation of Tree Cover Density products, collected on the ESA DWH ADDITIONAL D2\_MG2b\_ECOL\_011a dataset. © DigitalGlobe Inc (2018), all rights reserved.**



**Figure 5-31: Distribution of 100m x 100m samples (blue) for validation of Tree Cover Density products.**

## 5.2.2 Classification Results and Validation

This chapter depicts the results of the classification as well as their validation. Firstly, the thematic accuracies are summarized (see section 5.2.2.1). The thematic accuracies are followed by a discussion of the validation results (see section 5.2.2.2).

For all demo sites, the classification of the improved TCM and DLT at 10 m spatial resolution for the years 2017 and 2018 has been performed independently with a Random Forest classifier using time features calculated for the period 15<sup>th</sup> March to 15<sup>th</sup> September. In case of the TCM, the result is a binary raster with values of 0 and 1 for no tree-covered areas and tree-covered areas respectively. The DLT is a seamless raster layer with values of 1 for broadleaved trees and 2 for coniferous trees. The final Dominant Leaf Type product is generated by masking the seamless raster layer with the TCM. The model training shows for all products an Overall Accuracy above 90 %. The following tables summarize the values obtained for each Tree Cover Mask and DLT layer of the three FOR demonstration sites.

**Table 5-23: Accuracy figures of the internal validation & model training (demonstration site CENTRAL)**

| Demo site | Product                | Model Training   |          |             |
|-----------|------------------------|------------------|----------|-------------|
|           |                        | Class Code       | 0        | 1           |
| CENTRAL   | Tree Cover Mask 2017   | Precision        | 0.97     | 0.86        |
|           |                        | Recall           | 0.96     | 0.89        |
|           |                        | F1-Score         | 0.96     | 0.88        |
|           |                        | Overall Accuracy |          | <b>0.94</b> |
|           |                        | Kappa Statistic  |          | 0.84        |
|           | Tree Cover Mask 2018   | Class Code       | <b>0</b> | <b>1</b>    |
|           |                        | Precision        | 0.98     | 0.95        |
|           |                        | Recall           | 0.98     | 0.93        |
|           |                        | F1-score         | 0.98     | 0.94        |
|           |                        | Overall Accuracy |          | <b>0.97</b> |
|           | Dominan Leaf Type 2018 | Kappa Statistic  |          | 0.92        |
|           |                        | Class Code       | <b>1</b> | <b>2</b>    |
|           |                        | Precision        | 0.97     | 0.99        |
|           |                        | Recall           | 0.99     | 0.97        |
|           |                        | F1-Score         | 0.98     | 0.98        |
|           |                        | Overall Accuracy |          | <b>0.98</b> |
|           |                        | Kappa Statistic  |          | 0.96        |

**Table 5-24: Accuracy figures of the internal validation & model training (demonstration site NORTH)**

| Demo site | Product                | Model Training   |          |             |
|-----------|------------------------|------------------|----------|-------------|
|           |                        | Class Code       | 0        | 1           |
| NORTH     | Tree Cover Mask 2017   | Precision        | 1.00     | 0.88        |
|           |                        | Recall           | 0.93     | 0.99        |
|           |                        | F1-Score         | 0.96     | 0.93        |
|           |                        | Overall Accuracy |          | <b>0.95</b> |
|           |                        | Kappa Statistic  |          | 0.90        |
|           | Tree Cover Mask 2018   | Class Code       | <b>0</b> | <b>1</b>    |
|           |                        | Precision        | 0.97     | 0.97        |
|           |                        | Recall           | 0.97     | 0.97        |
|           |                        | F1-Score         | 0.97     | 0.97        |
|           |                        | Overall Accuracy |          | <b>0.97</b> |
|           | Dominan Leaf Type 2018 | Kappa Statistic  |          | 0.94        |
|           |                        | Class Code       | <b>1</b> | <b>2</b>    |
|           |                        | Precision        | 0.97     | 0.94        |
|           |                        | Recall           | 0.94     | 0.97        |
|           |                        | F1-Score         | 0.95     | 0.96        |
|           |                        | Overall Accuracy |          | <b>0.95</b> |
|           |                        | Kappa Statistic  |          | 0.91        |

**Table 5-25: Accuracy figures of the internal validation & model training (demonstration site SOUTH-EAST)**

| Demo site  | Product                       | Model Training   |          |             |
|------------|-------------------------------|------------------|----------|-------------|
|            |                               | Class Code       | 0        | 1           |
| SOUTH-EAST | <b>Tree Cover Mask 2017</b>   | Precision        | 1.00     | 0.99        |
|            |                               | Recall           | 0.99     | 0.99        |
|            |                               | F1-Score         | 0.99     | 0.99        |
|            |                               | Overall Accuracy |          | <b>0.99</b> |
|            |                               | Kappa Statistic  |          | 0.98        |
|            |                               | Class Code       | <b>0</b> | <b>1</b>    |
| SOUTH-EAST | <b>Tree Cover Mask 2018</b>   | Precision        | 0.99     | 0.99        |
|            |                               | Recall           | 1.00     | 0.98        |
|            |                               | F1-Score         | 0.99     | 0.99        |
|            |                               | Overall Accuracy |          | <b>0.99</b> |
|            |                               | Kappa Statistic  |          | 0.98        |
|            |                               | Class Code       | <b>1</b> | <b>2</b>    |
| SOUTH-EAST | <b>Dominan Leaf Type 2018</b> | Precision        | 1.00     | 0.96        |
|            |                               | Recall           | 0.96     | 1.00        |
|            |                               | F1-Score         | 0.98     | 0.98        |
|            |                               | Overall Accuracy |          | <b>0.98</b> |
|            |                               | Kappa Statistic  |          | 0.96        |

For all demo sites, the classification of the improved TCD at 10 m spatial resolution for the reference year 2018 has been performed using a multiple linear regression estimator. For each pixel in the median time feature stack, a density value is being calculated. The thereof resulting raw density raster represents a seamless raster layer with values ranging from 0 to 100. The final Tree Cover Density product is generated by masking the seamless TCD raster layer with the TCM.

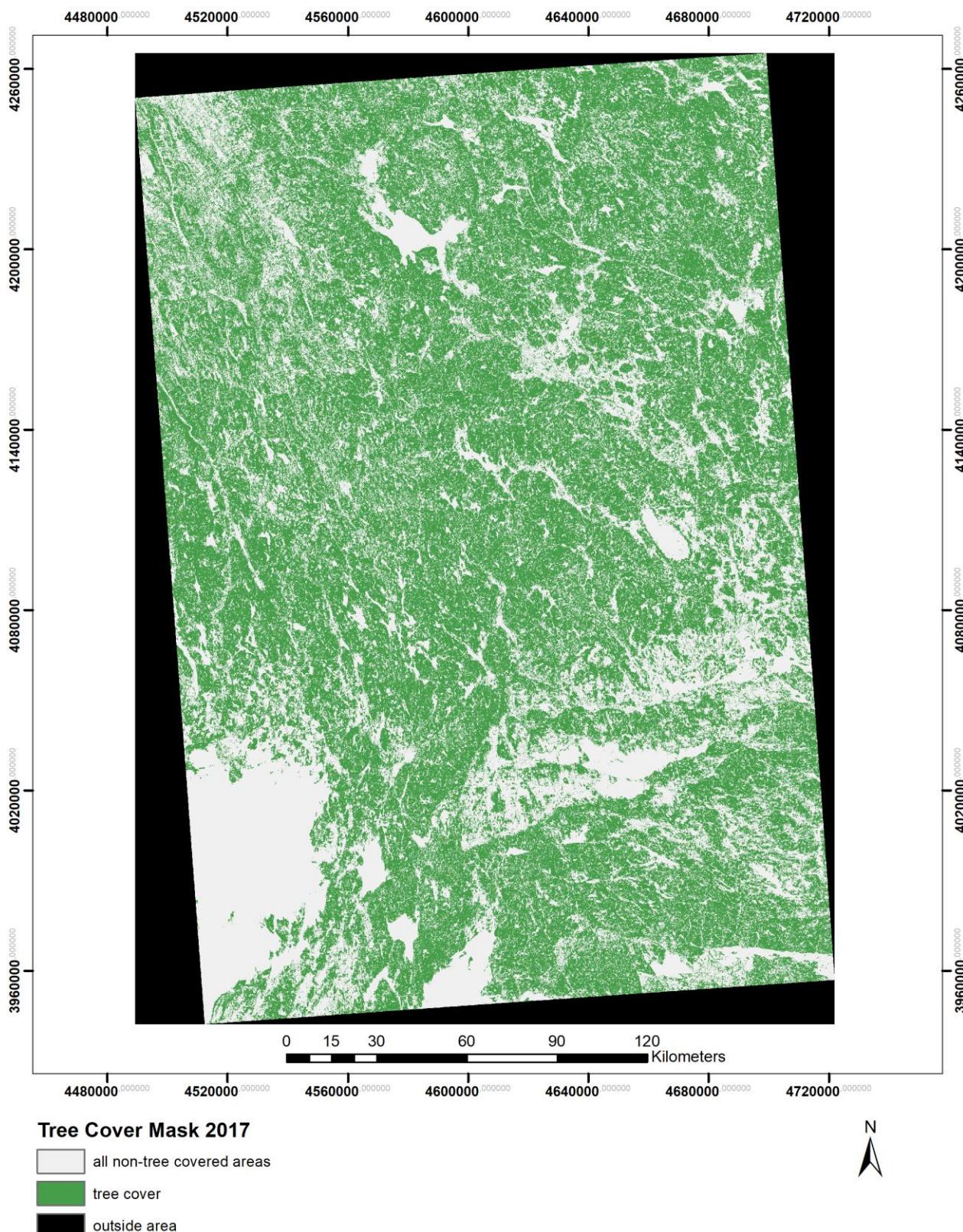
Table 5-26 provides the model parameters for each TCD 2018 product within the three FOR demonstration sites. The model training shows for all demonstration sites a coefficient of determination  $R^2$  greater 0.9, which is exceptionally high.

**Table 5-26: Model parameters for Tree Cover Density products**

| Demo site  | No. of samples | Mean Absolute Error | RMS  | $R^2$         |
|------------|----------------|---------------------|------|---------------|
| CENTRAL    | 1,818          | 6.15                | 8.17 | <b>0.9407</b> |
| NORTH      | 1,387          | 3.45                | 4.23 | <b>0.9573</b> |
| SOUTH-EAST | 1,293          | 2.52                | 3.27 | <b>0.9737</b> |

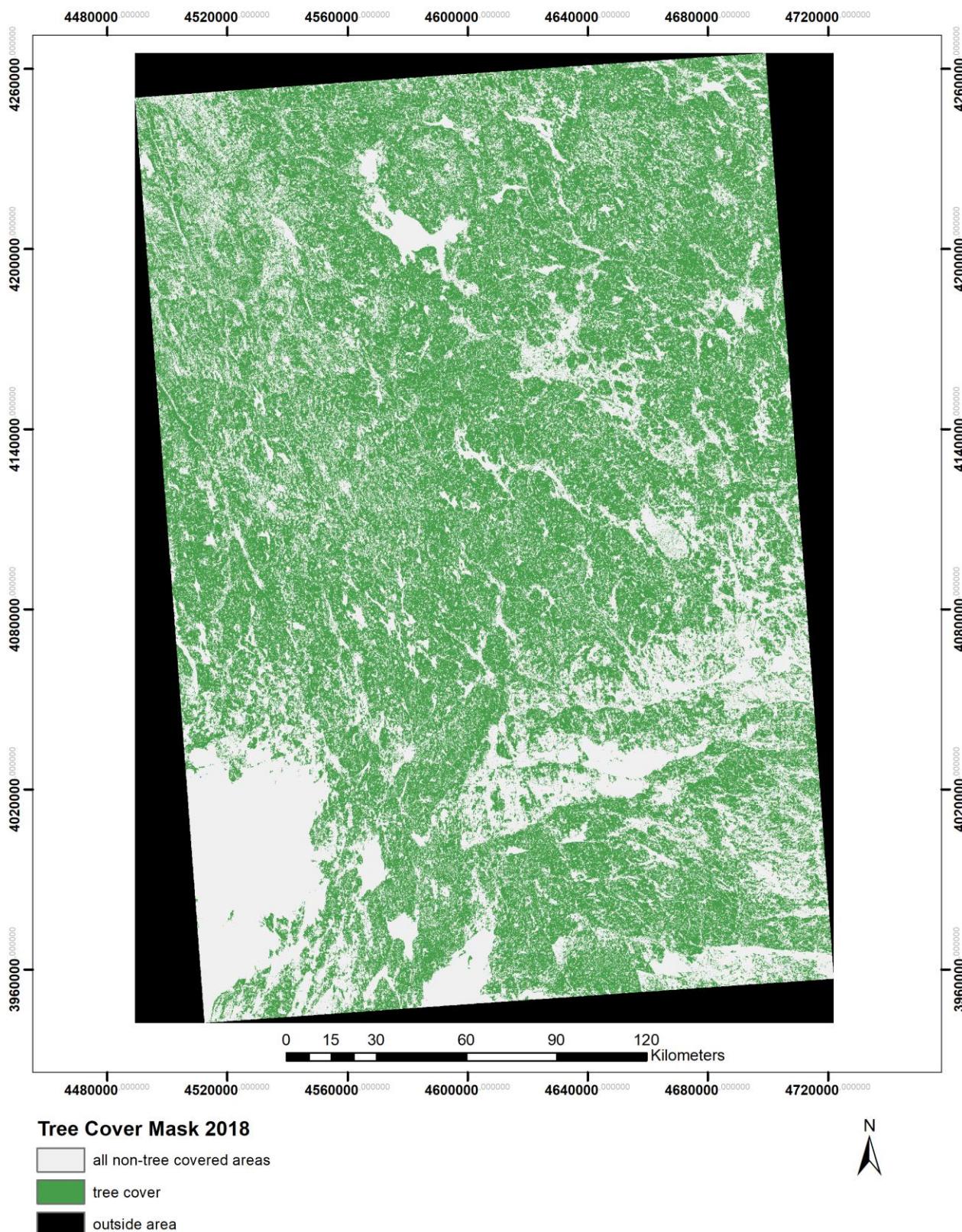
The following figures show the results of the improved Tree Cover Masks (basic products), as well as the improved DLT and TCD status layers at 10 m spatial resolution.

## Tree Cover Mask 2017 010m - Demo Site North



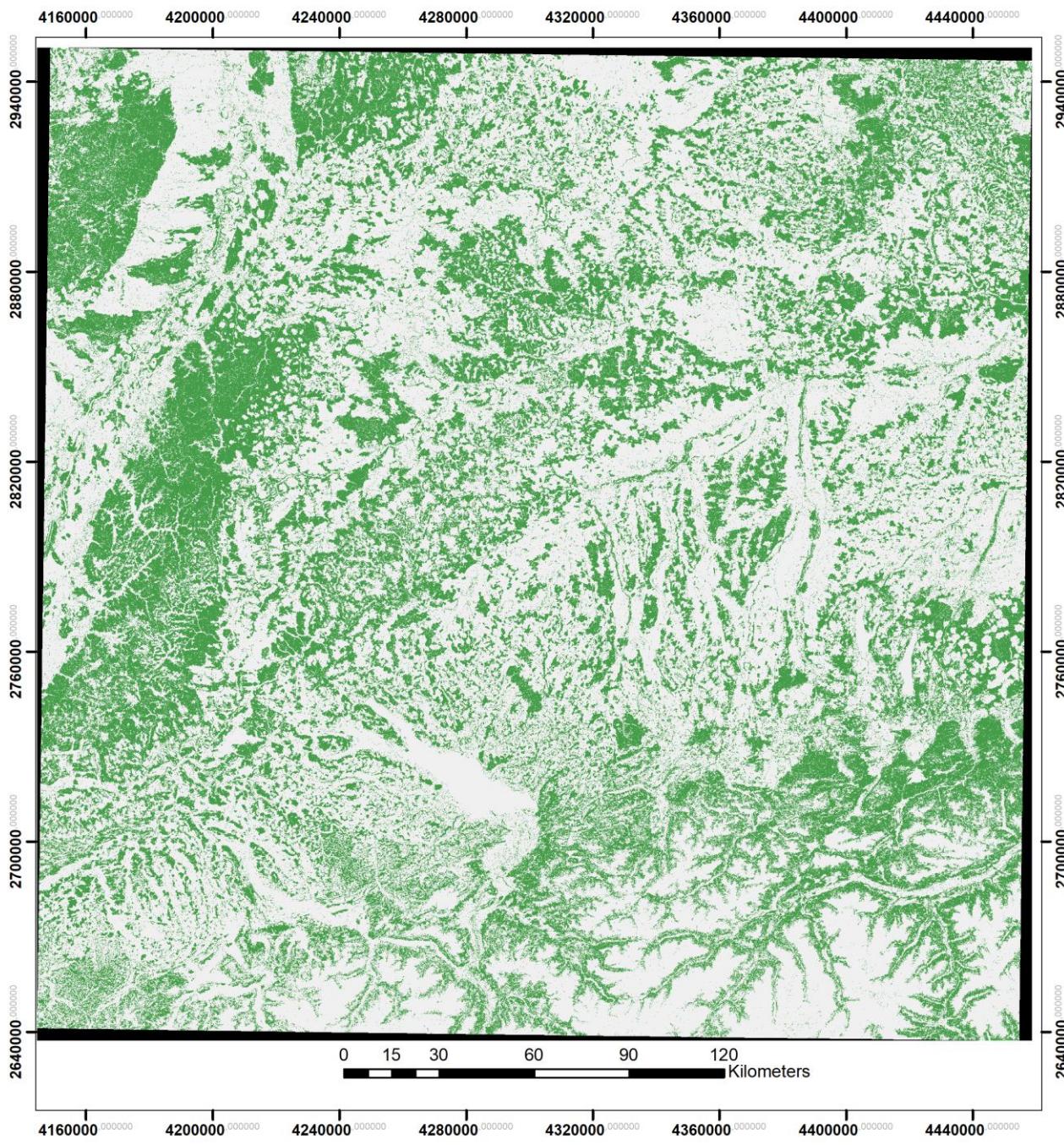
**Figure 5-32: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demonstration site North**  
Produced using Copernicus Sentinel data [2017]

## Tree Cover Mask 2018 010m - Demo Site North



**Figure 5-33: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demonstration site North**  
Produced using Copernicus Sentinel data [2018]

## Tree Cover Mask 2017 010m - Demo Site Central



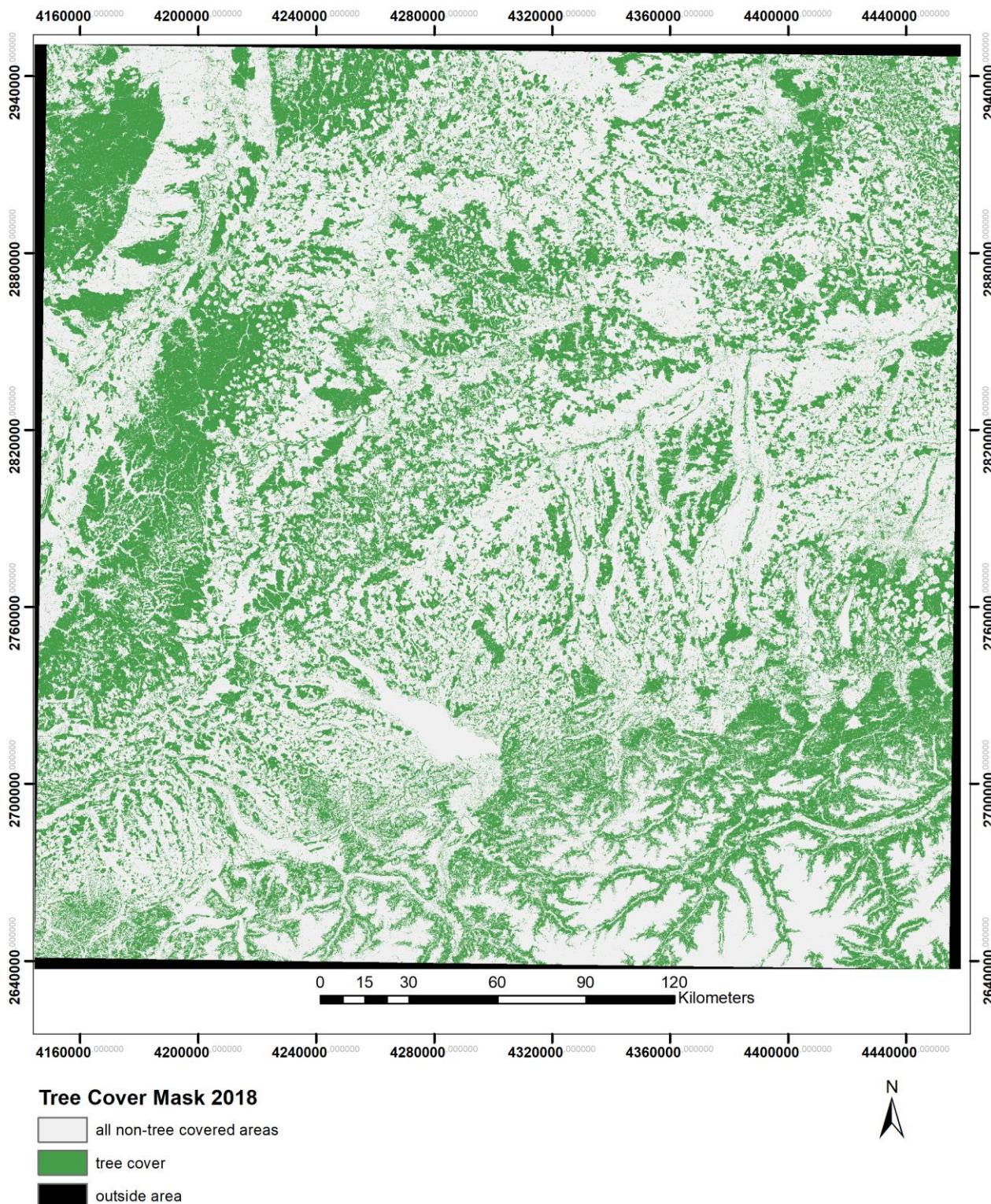
### Tree Cover Mask 2017

- all non-tree covered areas
- tree cover
- outside area



**Figure 5-34: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demonstration site Central**  
Produced using Copernicus Sentinel data [2017]

## Tree Cover Mask 2018 010m - Demo Site Central



**Figure 5-35: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demonstration site Central**  
Produced using Copernicus Sentinel data [2018]

## Tree Cover Mask 2017 010m - Demo Site South-East

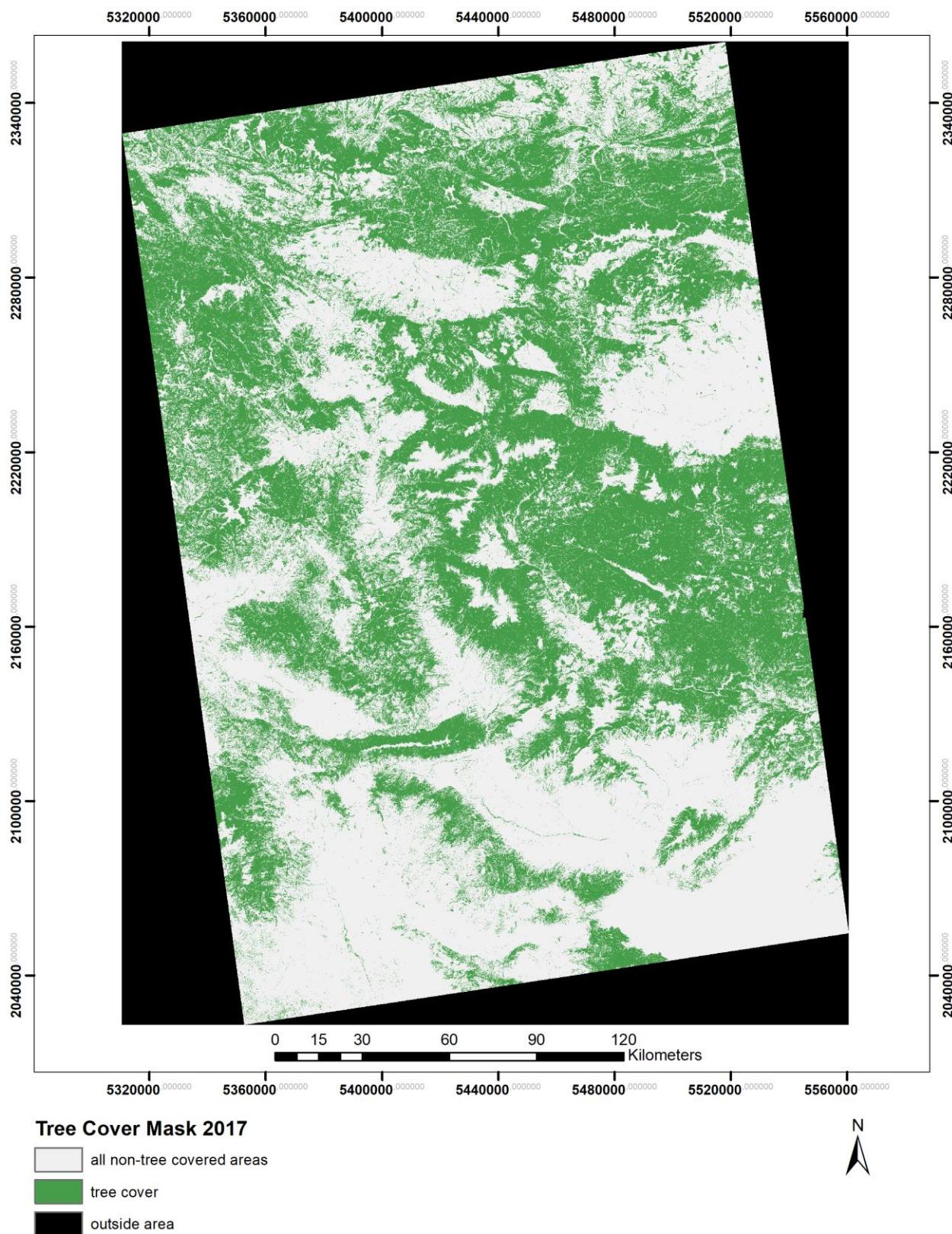


Figure 5-36: Improved Tree Cover Mask 2017 in 10m spatial resolution for the demonstration site South-East  
Produced using Copernicus Sentinel data [2017]

## Tree Cover Mask 2018 010m - Demo Site South-East

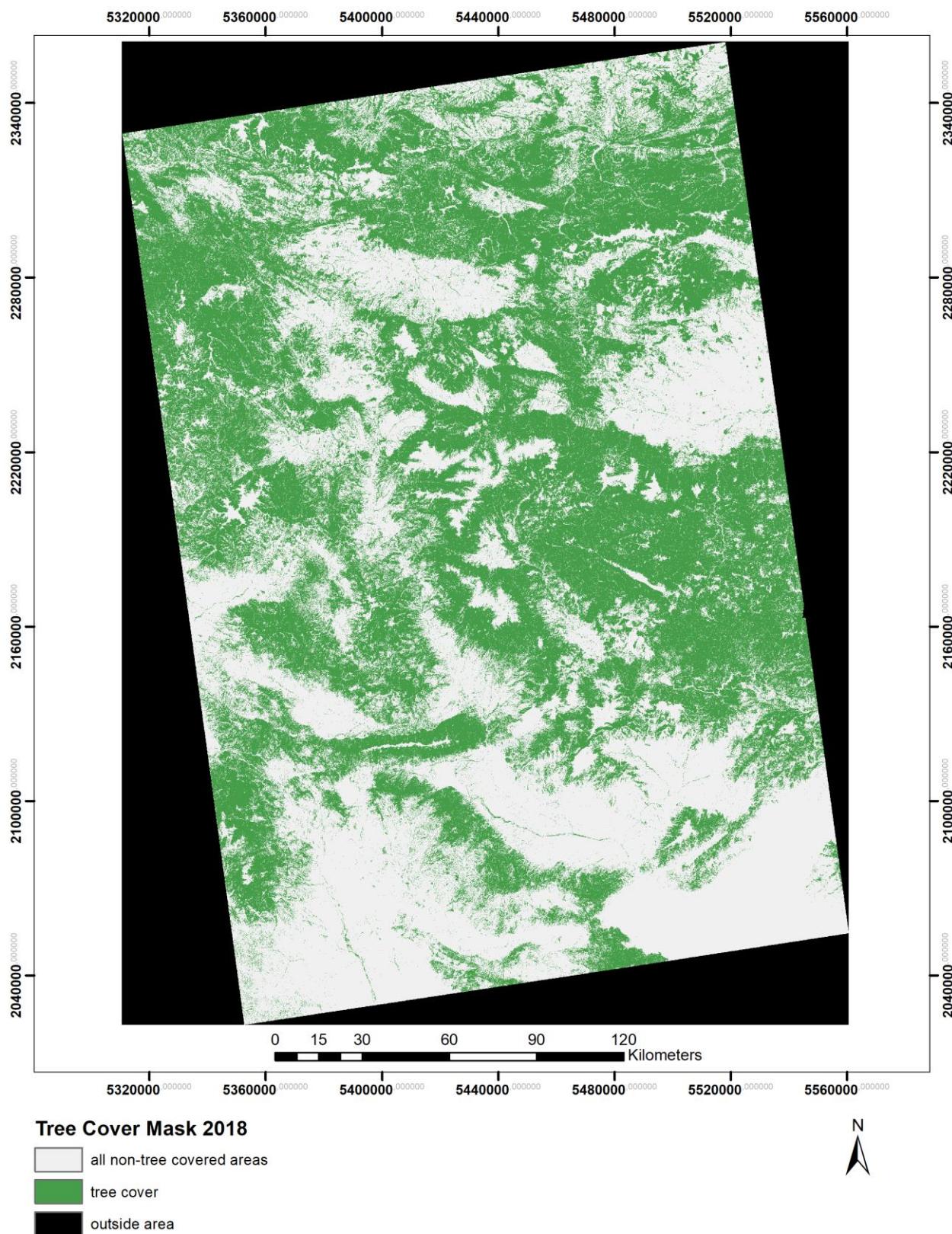
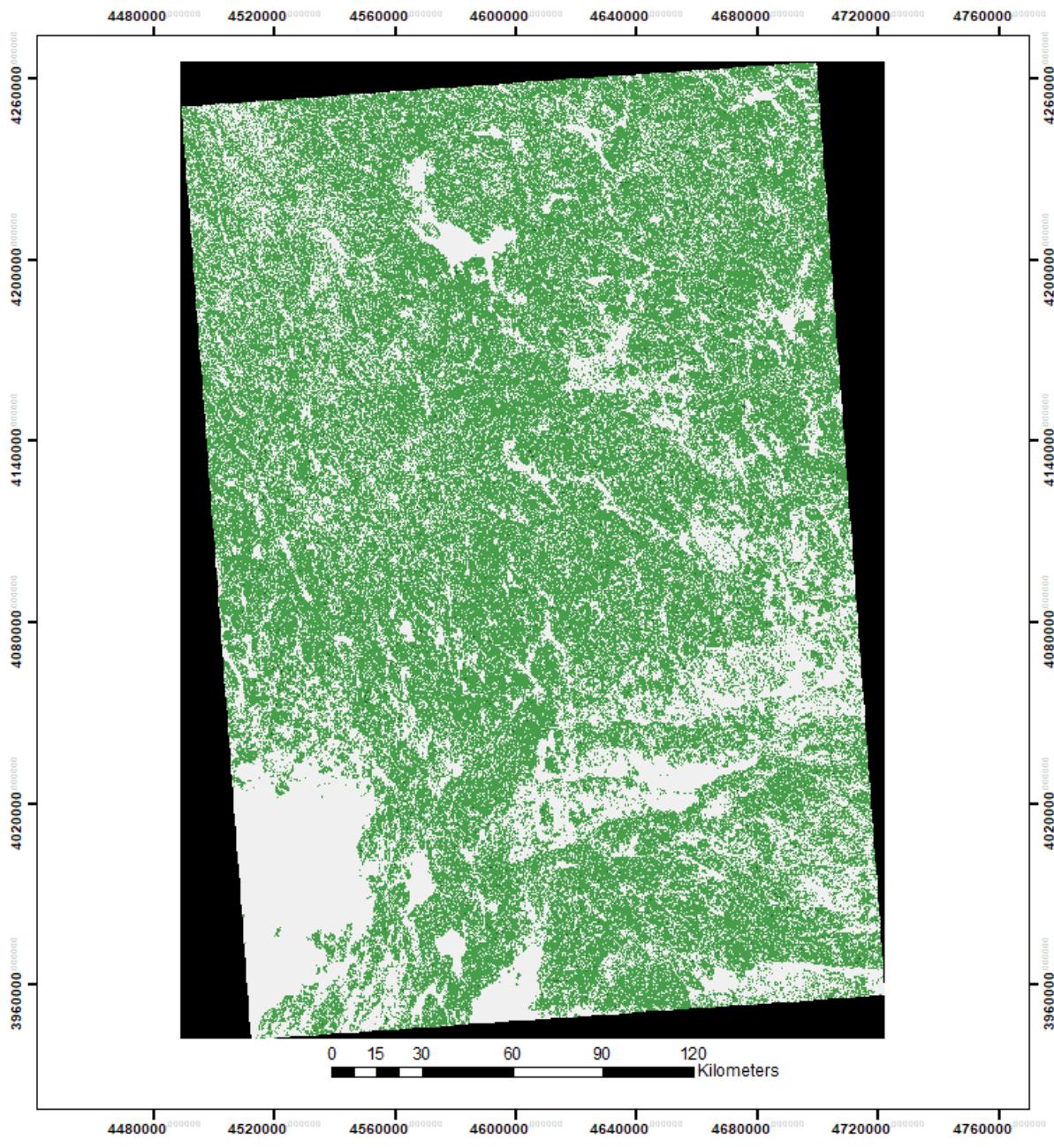


Figure 5-37: Improved Tree Cover Mask 2018 in 10m spatial resolution for the demonstration site South-East  
Produced using Copernicus Sentinel data [2018]

## Dominant Leaf Type 2018 010m - Demo Site North

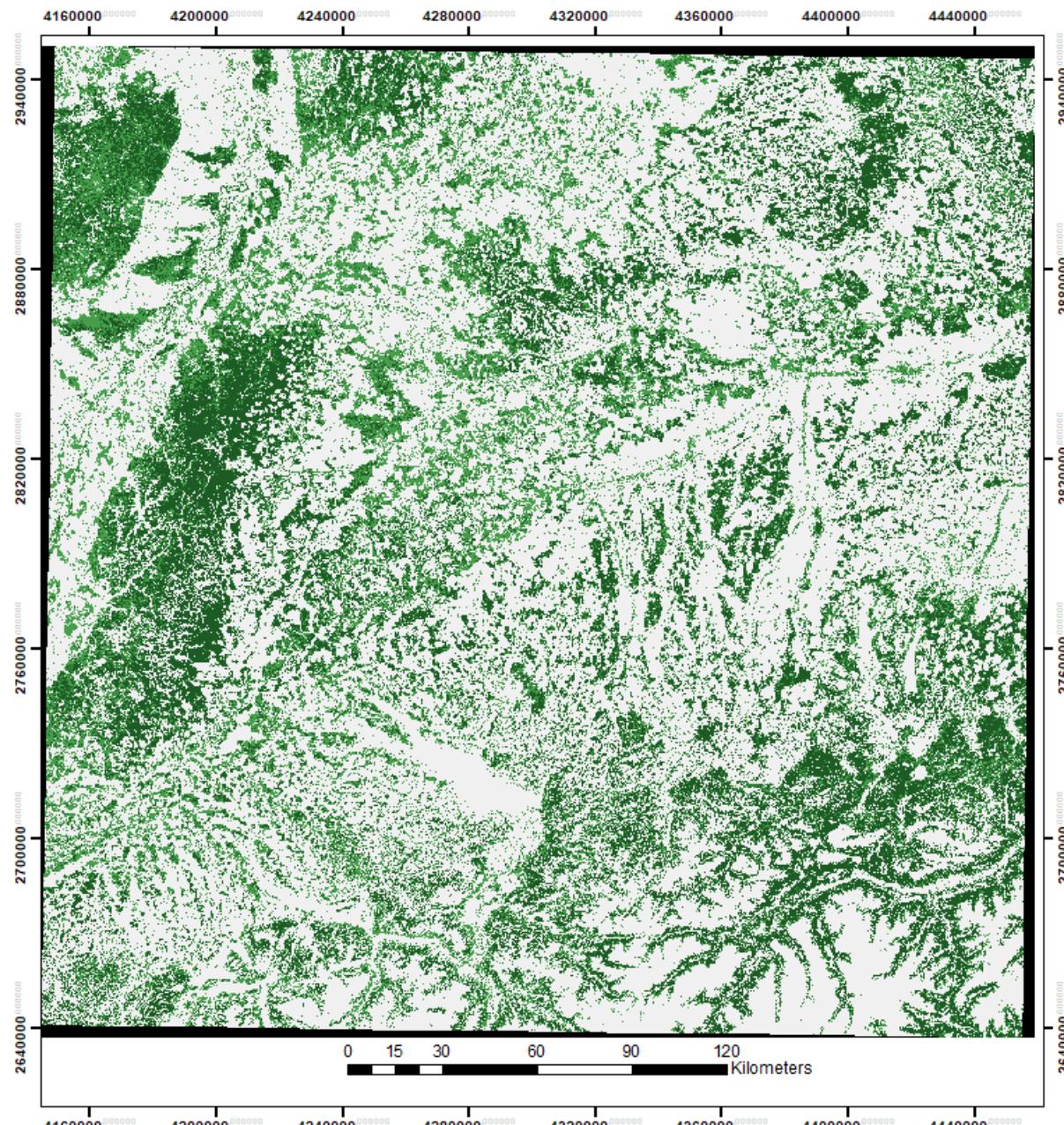


### Dominant Leaf Type 2018

- all non-tree covered areas
- broadleaved trees
- coniferous trees
- outside area

**Figure 5-38: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demonstration site North**  
*Produced using Copernicus Sentinel data [2018]*

## Dominant Leaf Type 2018 010m - Demo Site Central



### Dominant Leaf Type 2018

- [Light Gray Box] all non-tree covered areas
- [Dark Green Box] broadleaved trees
- [Medium Green Box] coniferous trees
- [Black Box] outside area



Figure 5-39: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demonstration site Central  
Produced using Copernicus Sentinel data [2018]

## Dominant Leaf Type 2018 010m - Demo Site South-East

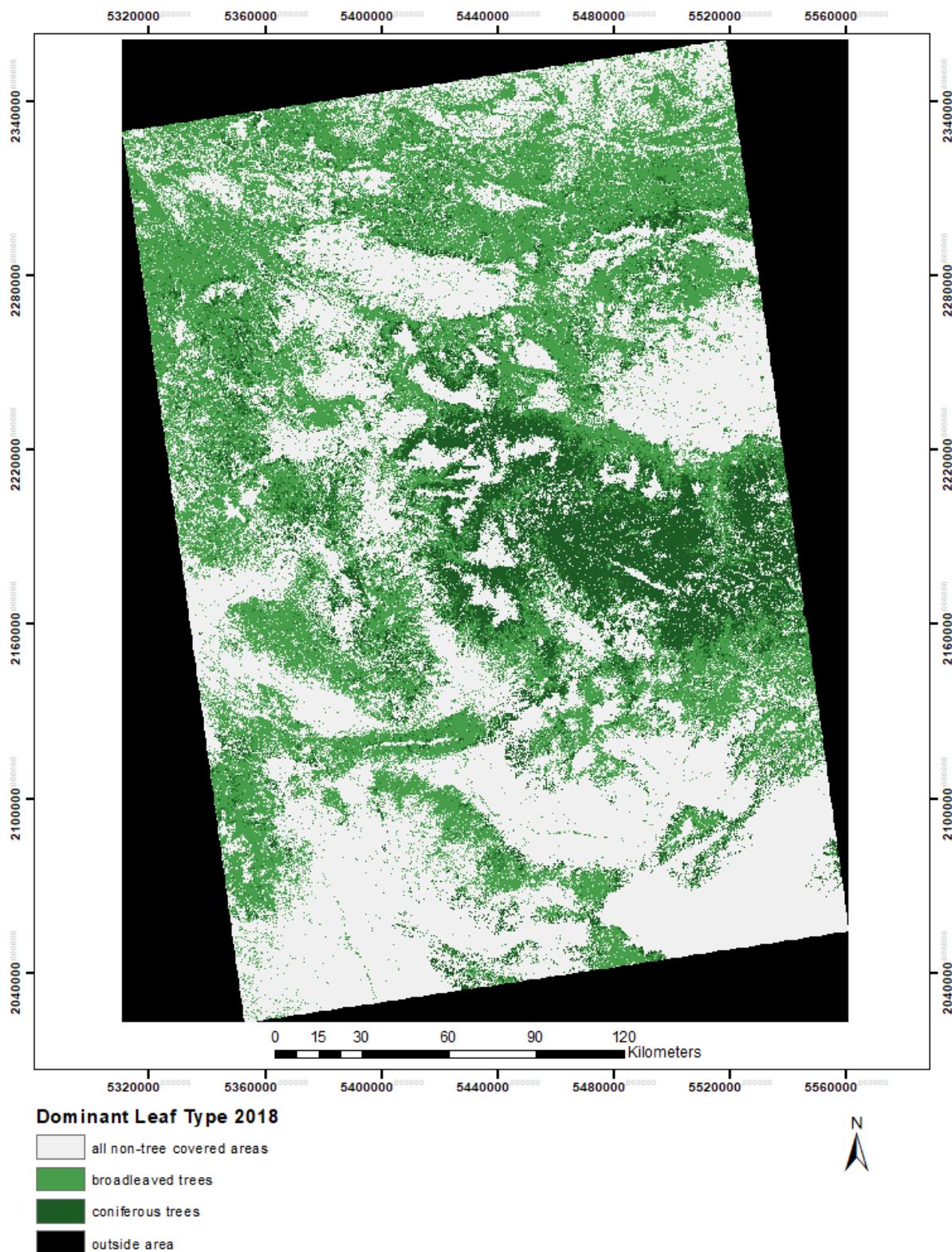


Figure 5-40: Improved Dominant Leaf Type 2018 in 10m spatial resolution for the demonstration site South-East  
Produced using Copernicus Sentinel data [2018]

## Tree Cover Density 2018 010m - Demo Site North

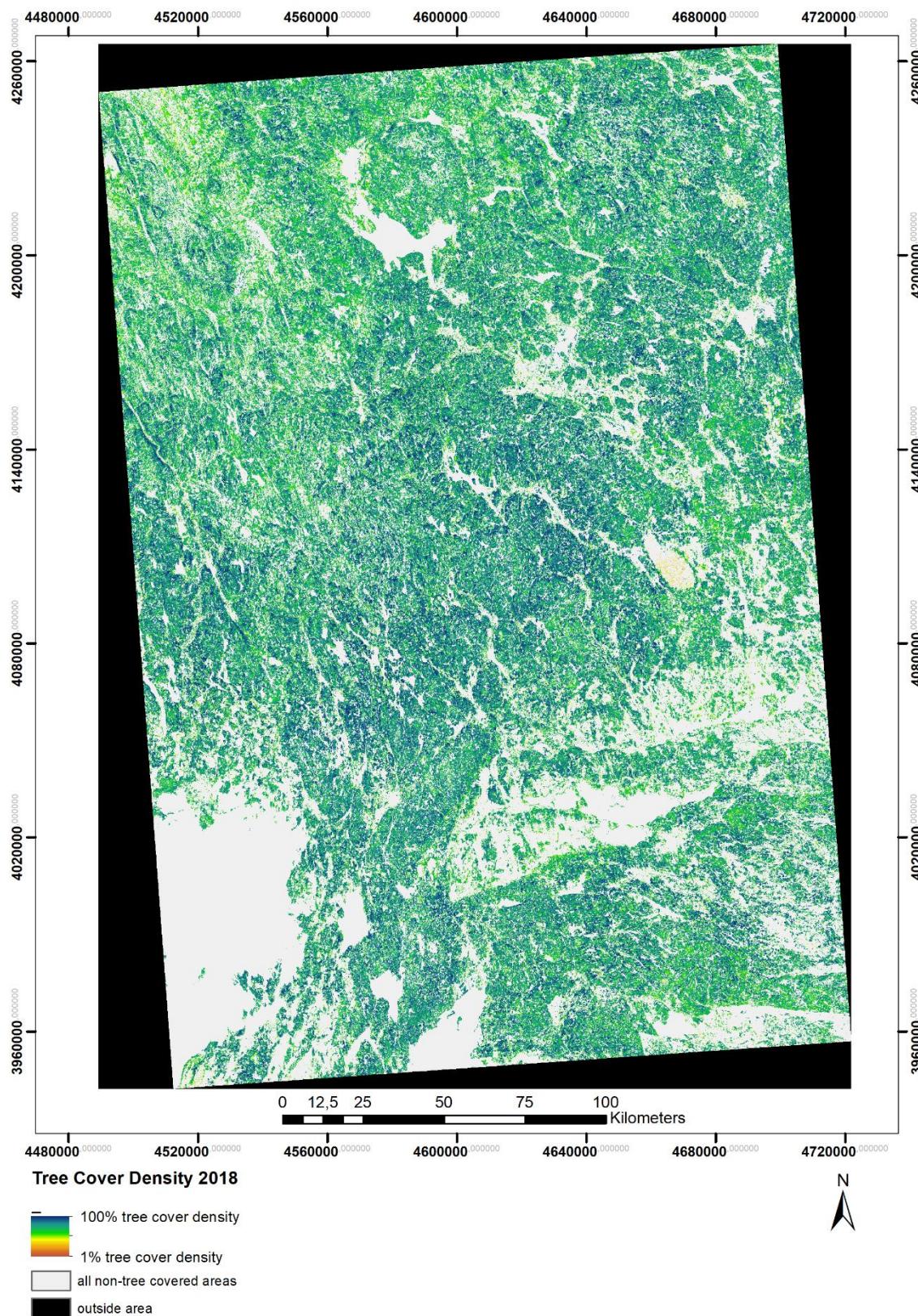
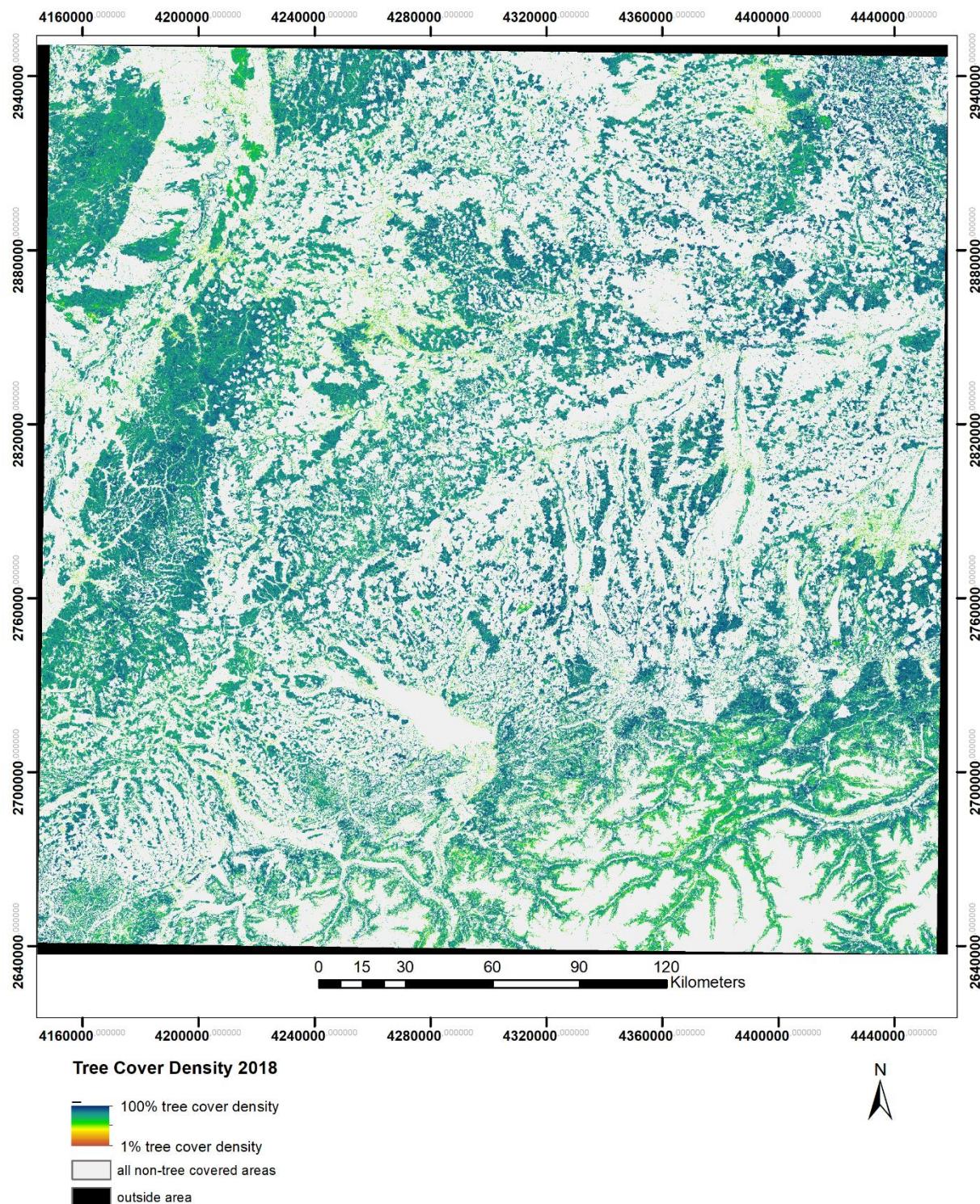


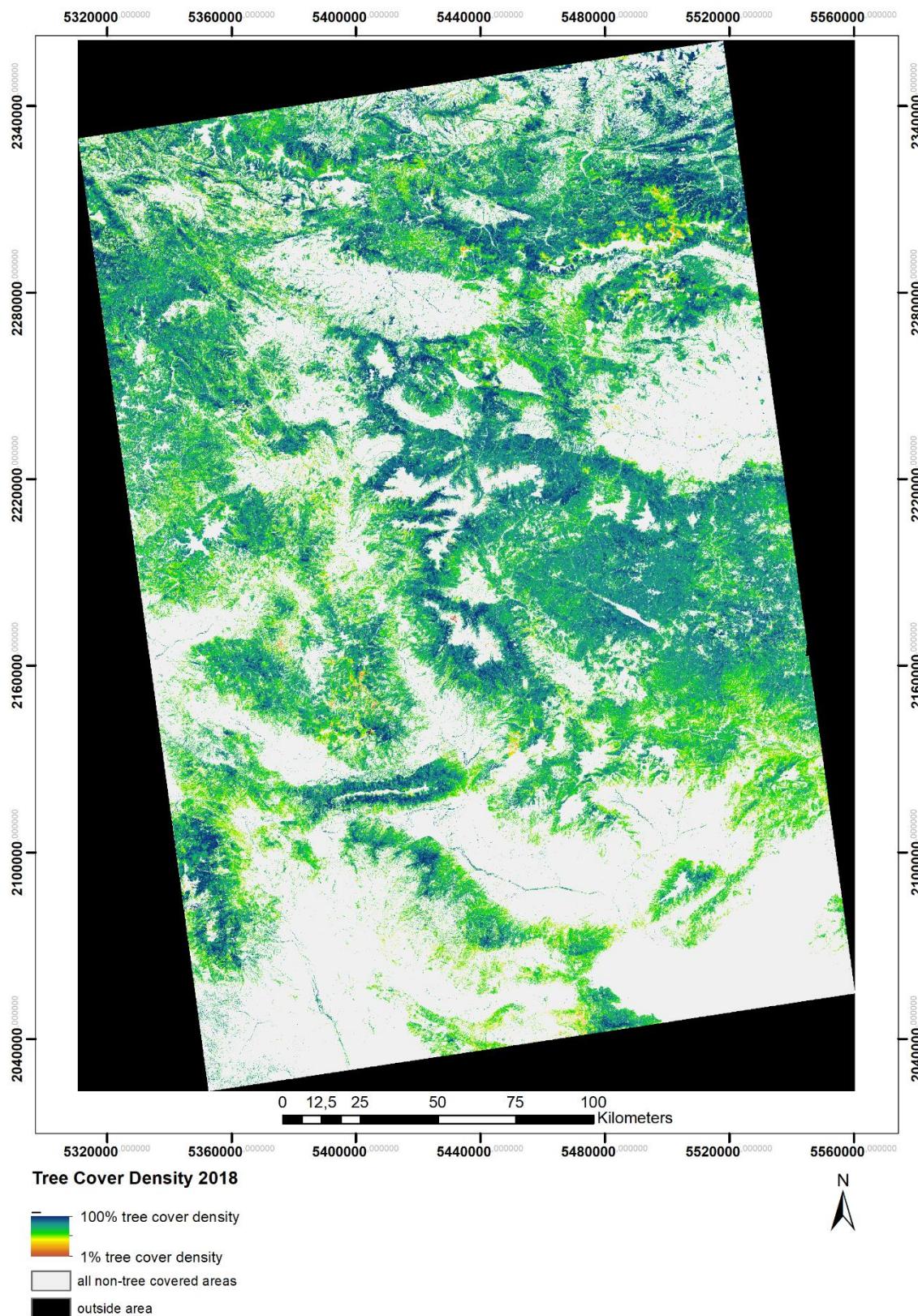
Figure 5-41: Tree Cover Density 2018 in 10m spatial resolution for the demonstration site North  
Produced using Copernicus Sentinel data [2018]

## Tree Cover Density 2018 010m - Demo Site Central



**Figure 5-42: Tree Cover Density 2018 in 10m spatial resolution for the demonstration site Central**  
Produced using Copernicus Sentinel data [2018]

## Tree Cover Density 2018 010m - Demo Site South-East



**Figure 5-43: Tree Cover Density 2018 in 10m spatial resolution for the demonstration site South-East**  
Produced using Copernicus Sentinel data [2018]

### 5.2.2.1 Thematic accuracy

A validation of the TCM 2017, TCM 2018 and the DLT 2018 was performed for all demonstration sites using the best available VHR reference data and HRguiding data . Guiding data are those that have been used in the production process and therefore act as a temporal reference in case of land cover changes, as compared to the VHR reference data. In the case of the DLT and TCM, guiding data consist of:

- Sentinel-2A+B scenes at 10m spatial resolution acquired between March and September 2018.

In turn, reference data consist of available suitable VHR data sources, namely:

- DWH VHR\_IMAGE\_2015
- DWH VHR\_IMAGE\_2018
- DWH D2\_MG2b\_ECOL\_011a (Archive\_standard\_Optical\_VHR1)
- DWH D2\_MG2b\_ECOL\_012a (New\_Acquisitions\_Optical\_VHR1)
- Bing Maps
- Google Earth Pro

The accuracy assessment has been performed using the LUCAS 2018 points, as the ground-truth dataset, after reclassifying their classes to conform to the ones found in the Dominant Leaf Type raster product. Points have been visually interpreted using the above-mentioned guiding reference data.

Thematic accuracy is presented in the form of an error matrix. Unequal sampling intensity resulting from the stratified systematic sampling approach was accounted for by applying a weight factor ( $p$ ) to each sample unit based on the ratio between the number of samples and the size of the stratum considered, according to the accuracy guidelines presented in WP 33 [AD 07]:

$$\hat{p}_{ij} = \left( \frac{1}{N} \right) \sum_{x \in (i,j)} \frac{1}{\pi_{uh}^*}$$

Where  $i$  and  $j$  are the columns and rows in the matrix,  $N$  is the total number of possible units (population) and  $\pi$  is the sampling intensity for a given stratum. Overall accuracy and user's and producer's accuracies were computed for all thematic classes and 95% confidence intervals were calculated for each accuracy.

The standard error of the error rate was calculated as follows:  $\sigma_h = \sqrt{\frac{p_h(1-p_h)}{n_h}}$  where  $n_h$  is the sample size for stratum  $h$  and  $p_h$  is the expected error rate. The standard error was calculated for each stratum and an overall standard error was calculated based on the following formula:

$$\sigma = \sqrt{\sum w_h^2 \cdot \sigma_h^2}$$

In which  $w_h$  is the proportion of the total area covered by each stratum. The 95% confidence interval is  $\pm 1.96 \sigma$ .

Results of the accuracy assessment are presented in the following tables for the North, Central and South-East demonstration sites respectively. **Figure 5-44** is exemplarily showing the distribution of LUCAS 2018 validation samples, following a 2 km grid. LUCAS points serve for validation of both, TCM and DLT products. In case of the demonstration site North, LUCAS points have been complemented by 200 additional samples, randomly distributed for the leaf types, since the number of available LUCAS points was not sufficient for a meaningful assessment of the leaf types.

## Dominant Leaf Type 2018 010m - Demo Site Southeast

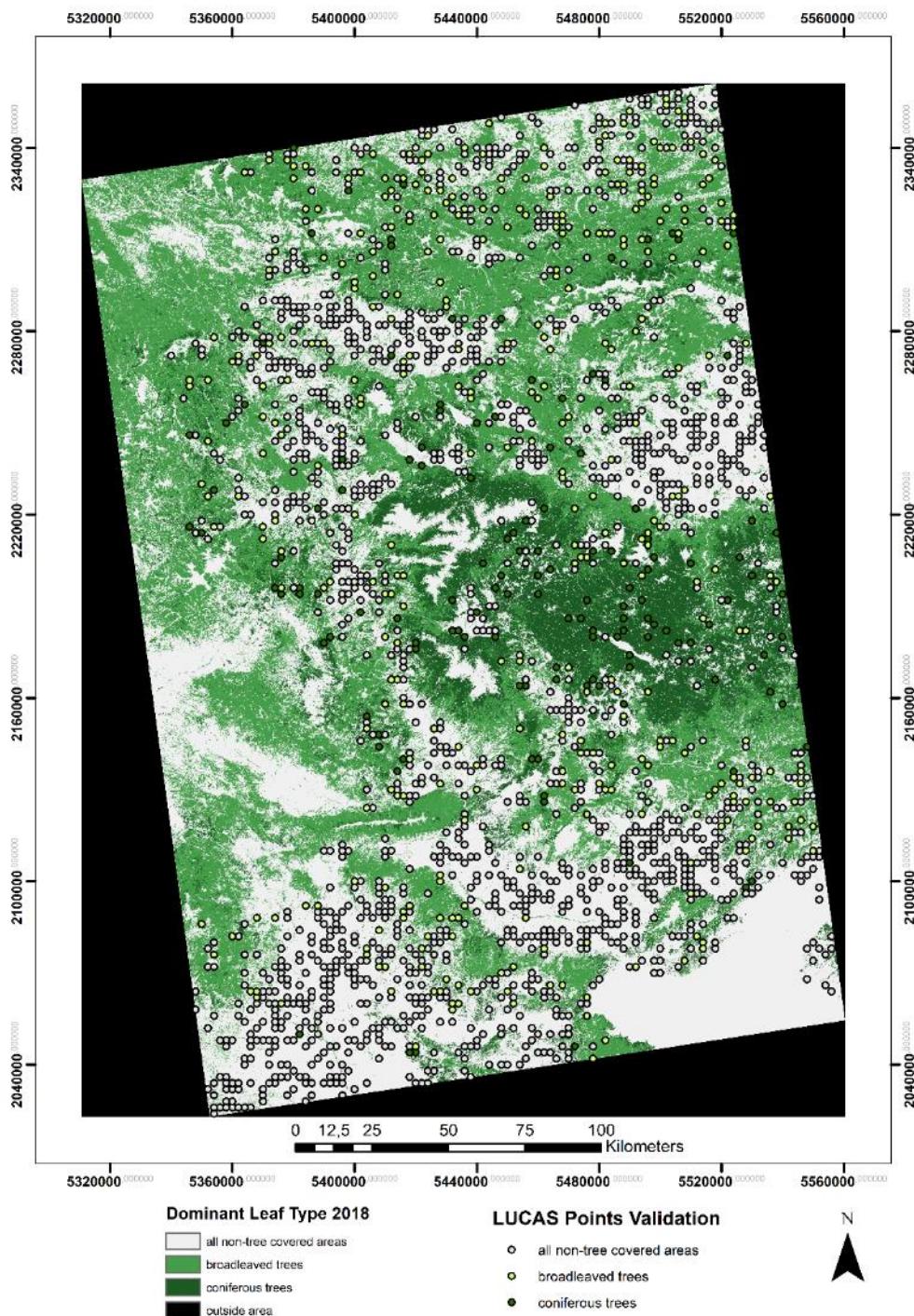


Figure 5-44: Distribution of LUCAS 2018 validation samples within the demonstration site South-East

### TREE COVER MASKS

Different accuracy levels can be observed when comparing the trained model with the validation results (see Table 5-27 to Table 5-32). In general, accuracies of the 2017 products are lower than for the 2018 products and the lowest overall accuracy (89.94%) is reported for the demonstration site North. This is explained by the low density of optical data in 2017. Despite the integration of Sentinel-1 time features, optical features from Sentinel-2 had a greater importance in the classification process. The high rate of

commission and omission errors can be attributed to the poor quality of Sentinel-2 time features and/or a low density of cloud-free observations per pixel.

**Table 5-27: Error matrix for the improved TCM 2017 of the demonstration site NORTH**

| TCM_2017_010m_NO_03035 |               | REFERENCE     |                     |                | User Accuracy  | Confidence Interval                |
|------------------------|---------------|---------------|---------------------|----------------|----------------|------------------------------------|
|                        |               | No Tree Cover | Tree Cover          | Total          |                |                                    |
| PRODUCT                | No Tree Cover | 865           | 93                  | 958            | 90.29 %        | 88.37 – 92.22%                     |
|                        | Tree Cover    | 78            | 664                 | 742            | 89.49%         | 87.21 – 91.76%                     |
|                        | Total         | 943           | 757                 | 1700           |                |                                    |
|                        |               |               | Producer Accuracy   | 91.73%         | 87.71%         | 89.94% Overall Accuracy            |
|                        |               |               | Confidence Interval | 89.92 – 93.54% | 85.31 – 90.12% | 88.48 – 91.40% Confidence Interval |
|                        |               |               |                     |                |                | 0.910 F-Score No Tree Cover        |
|                        |               |               |                     |                |                | 0.886 F-Score Tree Cover           |
|                        |               |               |                     |                |                | 0.796 Kappa                        |

**Table 5-28: Error matrix for the improved TCM 2018 of the demonstration site NORTH**

| TCM_2018_010m_NO_03035 |               | REFERENCE     |                     |                | User Accuracy  | Confidence Interval                |
|------------------------|---------------|---------------|---------------------|----------------|----------------|------------------------------------|
|                        |               | No Tree Cover | Tree Cover          | Total          |                |                                    |
| PRODUCT                | No Tree Cover | 934           | 77                  | 1011           | 92.38%         | 90.70 – 94.07%                     |
|                        | Tree Cover    | 23            | 752                 | 775            | 97.03%         | 95.77 – 98.29%                     |
|                        | Total         | 957           | 829                 | 1786           |                |                                    |
|                        |               |               | Producer Accuracy   | 97.60%         | 90.71%         | 94.40% Overall Accuracy            |
|                        |               |               | Confidence Interval | 96.57 – 98.62% | 88.68 – 92.75% | 93.31 – 95.50% Confidence Interval |
|                        |               |               |                     |                |                | 0.949 F-Score No Tree Cover        |
|                        |               |               |                     |                |                | 0.937 F-Score Tree Cover           |
|                        |               |               |                     |                |                | 0.886 Kappa                        |

**Table 5-29: Error matrix for the improved TCM 2017 of the demonstration site CENTRAL**

| TCM_2017_010m_CE_03035 |               | REFERENCE     |                     |                | User Accuracy   | Confidence Interval               |
|------------------------|---------------|---------------|---------------------|----------------|-----------------|-----------------------------------|
|                        |               | No Tree Cover | Tree Cover          | Total          |                 |                                   |
| PRODUCT                | No Tree Cover | 2845          | 96                  | 2941           | 96.74%          | 96.08 - 97.40%                    |
|                        | Tree Cover    | 140           | 724                 | 864            | 83.80%          | 81.28 - 86.31%                    |
|                        | Total         | 2985          | 820                 | 3805           |                 |                                   |
|                        |               |               | Producer Accuracy   | 95.31%         | 88.29%          | 93.80% Overall Accuracy           |
|                        |               |               | Confidence Interval | 94.53 - 96.09% | 86.03 - 90.55 % | 93.02 – 94.8% Confidence Interval |
|                        |               |               |                     |                |                 | 0.960 F-Score No Tree Cover       |
|                        |               |               |                     |                |                 | 0.859 F-Score Tree Cover          |
|                        |               |               |                     |                |                 | 0.82 Kappa                        |

**Table 5-30: Error matrix for the improved TCM 2018 of the demonstration site CENTRAL**

| TCM_2018_010m_CE_03035 |               | REFERENCE     |                     |                | User Accuracy  | Confidence Interval                                |
|------------------------|---------------|---------------|---------------------|----------------|----------------|--|
|                        |               | No Tree Cover | Tree Cover          | Total          |                |  |
| PRODUCT                | No Tree Cover | 2819          | 45                  | 2864           | 98.43%         | 97.96 - 98.90%                                     |
|                        | Tree Cover    | 22            | 657                 | 679            | 96.76%         | 95.35 - 98.17%                                     |
|                        | Total         | 2841          | 702                 | 3543           |                |  |
|                        |               |               | Producer Accuracy   | 99.23%         | 93.59%         | 98.11%<br>Overall Accuracy                         |
|                        |               |               | Confidence Interval | 98.89 - 99.57% | 91.71 - 95.47% | 97.65 – 98.57%<br>F-Score No Tree Cover            |
|                        |               |               |                     |                |                | 0.988<br>0.951<br>0.939<br>Kappa                   |
|                        |               |               |                     |                |                | Confidence Interval<br>F-Score Tree Cover<br>Kappa |

**Table 5-31: Error matrix for the improved TCM 2017 of the demonstration site SOUTH-EAST**

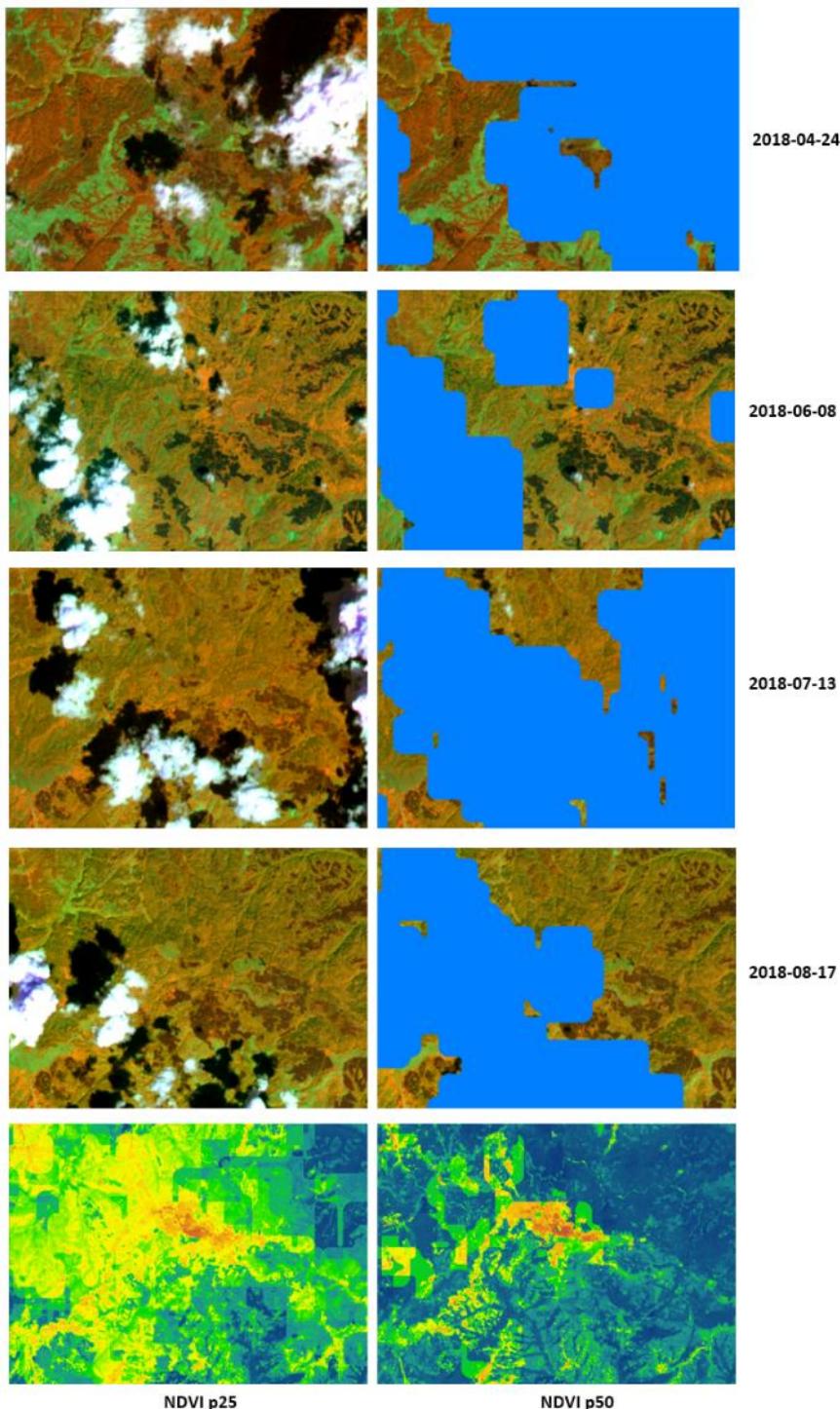
| TCM_2017_010m_SE_03035 |               | REFERENCE     |                     |                | User Accuracy  | Confidence Interval                                |
|------------------------|---------------|---------------|---------------------|----------------|----------------|--|
|                        |               | No Tree Cover | Tree Cover          | Total          |                |  |
| PRODUCT                | No Tree Cover | 1461          | 143                 | 1604           | 91.08%         | 89.66 – 92.51%                                     |
|                        | Tree Cover    | 46            | 345                 | 391            | 88.24%         | 84.91 – 91.56%                                     |
|                        | Total         | 1507          | 488                 | 1995           |                |  |
|                        |               |               | Producer Accuracy   | 96.95%         | 70.70%         | 90.53%<br>Overall Accuracy                         |
|                        |               |               | Confidence Interval | 96.05 – 97.85% | 66.56 – 74.84% | 89.22 – 91.84%<br>F-Score No Tree Cover            |
|                        |               |               |                     |                |                | 0.939<br>0.785<br>0.725<br>Kappa                   |
|                        |               |               |                     |                |                | Confidence Interval<br>F-Score Tree Cover<br>Kappa |

**Table 5-32: Error matrix for the improved TCM 2018 of the demonstration site SOUTH-EAST**

| TCM_2018_010m_SE_03035 |               | REFERENCE     |                     |                | User Accuracy  | Confidence Interval                                |
|------------------------|---------------|---------------|---------------------|----------------|----------------|--|
|                        |               | No Tree Cover | Tree Cover          | Total          |                |  |
| PRODUCT                | No Tree Cover | 1430          | 43                  | 1473           | 97.08%         | 96.19 – 97.97%                                     |
|                        | Tree Cover    | 17            | 358                 | 375            | 95.47%         | 93.23 - 97.71%                                     |
|                        | Total         | 1447          | 401                 | 1848           |                |  |
|                        |               |               | Producer Accuracy   | 98.83%         | 89.28%         | 96.75%<br>Overall Accuracy                         |
|                        |               |               | Confidence Interval | 98.24 – 99.41% | 86.12 – 92.43% | 95.92 – 97.59%<br>F-Score No Tree Cover            |
|                        |               |               |                     |                |                | 0.979<br>0.923<br>0.902<br>Kappa                   |
|                        |               |               |                     |                |                | Confidence Interval<br>F-Score Tree Cover<br>Kappa |

Even though the overall accuracies for the Tree Cover Mask in the demonstration site South-East were above 90% for the years 2017 and 2018, the omission errors range between 11% and 29%. This is due to the low availability of Sentinel-2 imagery for this demonstration site and serious issues of the cloud

masking algorithm implemented in the MACCS processor, which is producing characteristic block structures (**Figure 5-45**).



**Figure 5-45: Cloud masking issues in the South-East demonstration site. Artefacts are inherited in the derived time features, representing the input for the thematic classification.**

Although the comparison of different methods for cloud masking is out of scope in this report, it should be noted that the extent and pattern of artefacts in the other demonstration sites (North and Central) was less severe than in the South-East demonstration site. Artefacts are directly influencing the thematic quality and look and feel of the products. This is especially valid for the continuous-scale Tree Cover

Density product and the Tree Cover Change prototype, where the applied NDVI plausibility approach failed to a certain extent.

### DOMINANT LEAF TYPE

The achieved thematic accuracy of the DLT products exceed the expected minimum Overall Accuracy of 90% in each demonstration site. Producer accuracies range between 83% and 99% and the lowest values are reported for the broadleaved class in the North (84.23%) and South-East (83.04) demonstration sites. This is mainly explained by the relatively high number of commission errors within the corresponding Tree Cover Masks and is less an issue of the leaf type discrimination. User's accuracies for the two leaf type classes are in a range of 88.6–96.8% and exceed the minimum requirements of 90% when taking the confidence interval into account. Best results with an Overall Accuracy of 97.80% have been achieved in the Central site, in which the overall data situation has been rated as the best. Surprisingly, the strong occurrence of artefacts in the South-East site had a relatively small impact on the leaf type discrimination. Here, an overall accuracy of 95.83% can be reported.

**Table 5-33: Error matrix for the improved DLT 2018 status layer of the demonstration site NORTH**

| DLT_2018_010m_NO_03035 |                     | REFERENCE      |                |                |       | User Accuracy  | Confidence Interval   |
|------------------------|---------------------|----------------|----------------|----------------|-------|----------------|-----------------------|
|                        |                     | No Tree Cover  | Broadleaved    | Coniferous     | Total |                |                       |
| PRODUCT                | No Tree Cover       | 934            | 38             | 39             | 1011  | 92.38%         | 90.70 – 94.07%        |
|                        | Broadleaved         | 9              | 251            | 20             | 280   | 89.64%         | 86.86 – 92.43%        |
|                        | Coniferous          | 14             | 9              | 472            | 495   | 95.35%         | 93.40 – 97.31%        |
|                        | Total               | 957            | 298            | 531            | 1786  |                |                       |
|                        | Producer Accuracy   | 97.60%         | 84.23%         | 88.89%         |       | 92.78%         | Overall Accuracy      |
|                        | Confidence Interval | 96.57 – 98.62% | 79.92 – 87.42% | 86.12 – 91.66% |       | 91.55 – 94.01% | Confidence Interval   |
|                        |                     |                |                |                |       | 0.949          | F-Score No Tree Cover |
|                        |                     |                |                |                |       | 0.869          | F-Score Broadleaved   |
|                        |                     |                |                |                |       | 0.920          | F-Score Coniferous    |
|                        |                     |                |                |                |       | 0.877          | Kappa                 |

**Table 5-34: Error matrix for the improved DLT 2018 status layer of the demonstration site CENTRAL**

| DLT_2018_10m_CE_03035 |                     | REFERENCE      |                |                |       | User Accuracy  | Confidence Interval   |
|-----------------------|---------------------|----------------|----------------|----------------|-------|----------------|-----------------------|
|                       |                     | No Tree Cover  | Broadleaved    | Coniferous     | Total |                |                       |
| PRODUCT               | No Tree Cover       | 2819           | 40             | 5              | 2864  | 98.43%         | 97.96 – 98.90%        |
|                       | Broadleaved         | 19             | 337            | 4              | 360   | 93.61%         | 90.77 – 96.45%        |
|                       | Coniferous          | 3              | 7              | 309            | 319   | 96.87%         | 94.80 – 98.93%        |
|                       | Total               | 2841           | 384            | 318            | 3543  |                |                       |
|                       | Producer Accuracy   | 99.23%         | 87.76%         | 97.17%         |       | 97.80%         | Overall Accuracy      |
|                       | Confidence Interval | 98.89 – 99.57% | 84.35 – 91.52% | 95.19 – 99.15% |       | 97.30 – 98.30% | Confidence Interval   |
|                       |                     |                |                |                |       | 0.988          | F-Score No Tree Cover |
|                       |                     |                |                |                |       | 0.906          | F-Score Broadleaved   |
|                       |                     |                |                |                |       | 0.971          | F-Score Coniferous    |
|                       |                     |                |                |                |       | 0.934          | Kappa                 |

**Table 5-35: Error matrix for the improved DLT 2018 status layer of the demonstration site SOUTH-EAST**

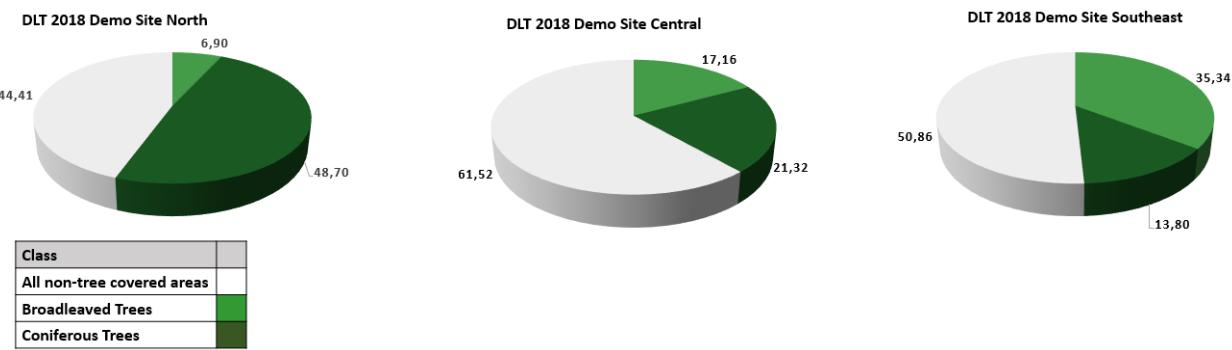
| DLT_2018_10m_SE_03035 |                     | REFERENCE      |                |                |       | User Accuracy  | Confidence Interval   |
|-----------------------|---------------------|----------------|----------------|----------------|-------|----------------|-----------------------|
|                       |                     | No Tree Cover  | Broadleaved    | Coniferous     | Total |                |                       |
| PRODUCT               | No Tree Cover       | 1430           | 39             | 4              | 1473  | 97.08%         | 96.19 – 97.97%        |
|                       | Broadleaved         | 14             | 240            | 7              | 261   | 91.95%         | 86.52 – 97.39%        |
|                       | Coniferous          | 3              | 10             | 101            | 114   | 88.60%         | 82.32 – 94.87%        |
|                       | Total               | 1447           | 289            | 112            | 1848  |                |                       |
|                       | Producer Accuracy   | 98.83 %        | 83.04 %        | 90.18          |       | 95.83%         | Overall Accuracy      |
|                       | Confidence Interval | 98.24 – 99.41% | 78.55 – 90.44% | 84.22 – 96.14% |       | 94.90 – 96.77% | Confidence Interval   |
|                       |                     |                |                |                |       | 0.979          | F-Score No Tree Cover |
|                       |                     |                |                |                |       | 0.873          | F-Score Broadleaved   |
|                       |                     |                |                |                |       | 0.894          | F-Score Coniferous    |
|                       |                     |                |                |                |       | 0.881          | Kappa                 |

Regarding the distribution of the different leaf types, it is observed that broadleaved forests decrease at higher latitudes. Table 5-36 provides the class distribution for the dominant leaf types per demonstration site.

**Table 5-36: Distribution of leaf type classes per demonstration site**

| Product                  | No Tree Cover (km <sup>2</sup> ) | Broadleaved (km <sup>2</sup> ) | Coniferous (km <sup>2</sup> ) |
|--------------------------|----------------------------------|--------------------------------|-------------------------------|
| DLT 2018 010m NORTH      | 28,890.42                        | 6.89                           | 48.70                         |
| DLT 2018 010m CENTRAL    | 59,066.53                        | 16,476.66                      | 20,470.09                     |
| DLT 2018 010m SOUTH-EAST | 33,034.06                        | 22,952.34                      | 8,965.03                      |

Coniferous trees are the predominant leaf type in the North demonstration site. For the Central and South-East demonstration sites there is a more balanced occurrence of broadleaved and coniferous trees, covering around 40 % of the total area (Figure 5-46).

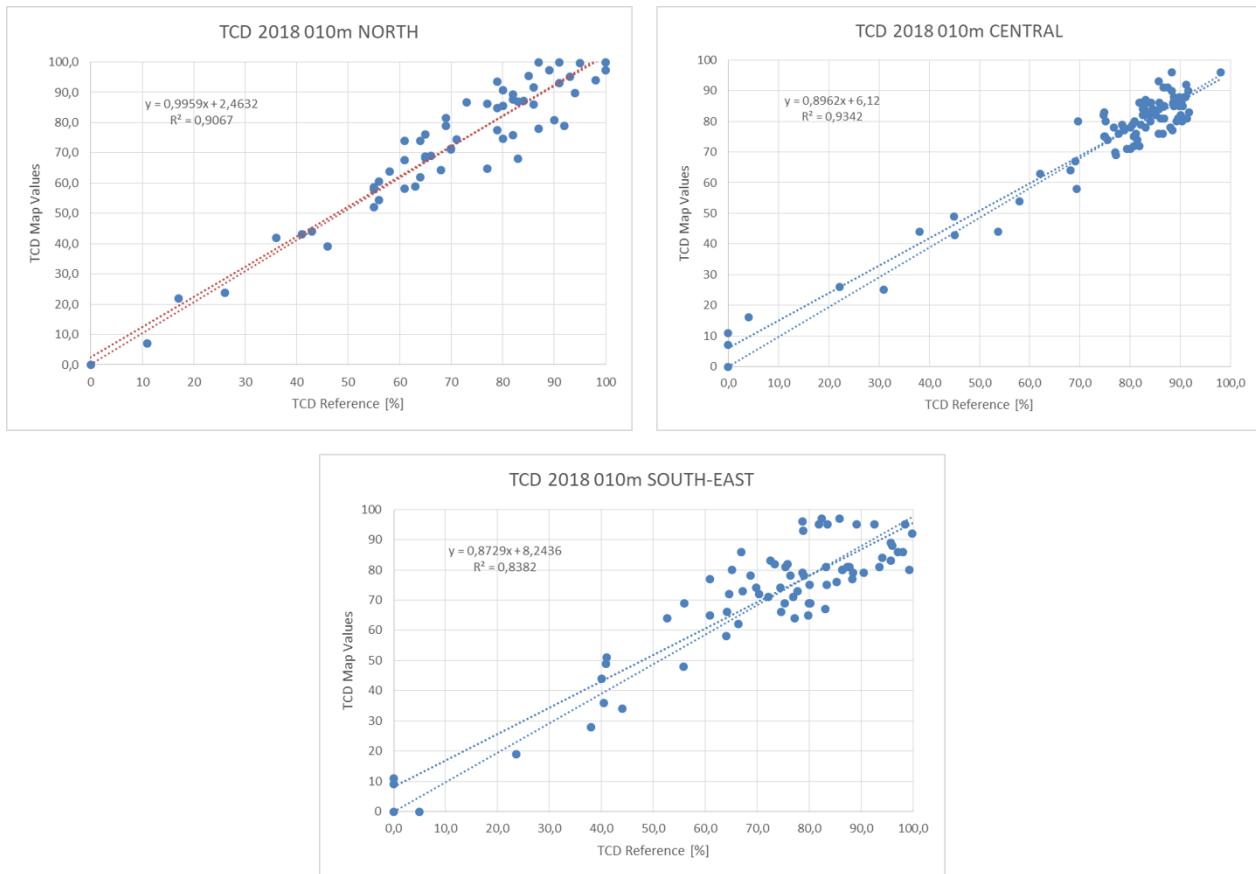


**Figure 5-46: Proportion of DLT 2018 land cover classes per demonstration site**

### TREE COVER DENSITY

Similar to the Imperviousness Degree prototypes scatter plots are used to produce a coefficient of determination  $R^2$  which provides information about the goodness of fit of the estimated regression model for each TCD 2018 prototype (see also section 5.1.2.1). Based on the independent TCD reference dataset

for validation, the relationship between the model prediction (map densities) and the ground-truth reference has been assessed. The resulting scatter plots are presented in **Figure 5-47** below.



**Figure 5-47: Scatter plots for validation of the TCD 2018 prototypes**

In all demonstration sites regression coefficients greater than 0.83 could be achieved, which indicates a very good relationship between the reference dataset and the calculated tree cover density values. The lowest value is observed in the demonstration site South-East with a  $R^2$  of 0.83. Here, artefacts introduced by the insufficient cloud masking through the MACCS processor result in relative strong deviations from the model prediction in both directions. Furthermore, dry-related effects (drought stress) led to a slight underestimation of density values. However, this is to a certain extent the result of the independent distribution of the validation samples. Regression coefficients are greater than 0.90 for the demonstration sites North and Central. The latter shows an  $R^2$  of 0.93 and is almost fully in line with the model prediction ( $R^2 = 0.94$ ), whereas the  $R^2$  of the North demonstration site is significantly lower than the prediction model (0.90 vs 0.95). Density values of broadleaved samples are generally underestimated in the demonstration site North, presumably as a result of drought stress in summer 2018.

### 5.2.2.2 Discussion of the validation results

The ECoLaSS consortium is very pleased with the results of the forest prototypes DLT and TCD as well as the tree cover classification results in general. It could be demonstrated that high quality Forest products can be derived for different bio-geographic regions across Europe using spatio-temporal features of Sentinel-2 and Sentinel-1.

In all three demonstration sites, Tree Cover Masks in high quality could be produced in a combined optical Sentinel-2 and Sentinel-1 SAR approach for the reference years 2017 and 2018. Even though data density in 2017 was significantly lower as in 2018, the achieved Overall Accuracy exceeded the 90% threshold in each single case, when taking the confidence interval into account. The derived masks represent basic

input datasets for all FOR prototypes and thus contribute significantly to the generally high quality of the prototypes. Issues are mainly related to the quality of the cloud masks (Sen2Cor in the North and Central demonstration sites and MACCS in the South-East demonstration sites). In case of frequent cloud cover and cloud-shadow overestimations, artefacts and even nodata gaps can be introduced when calculating relevant time features. In the worst case, these are inherited in the produced Tree Cover Masks and thus present in every Forest prototype.

The improved Dominant Leaf Type products show also very high Overall Accuracies above 90% within each FOR demonstration site. Optical data has a high importance for the discrimination of the leaf types, whereas Sentinel-1 SAR features show almost no significance in the classification process. The observed omission errors for the broadleaved trees are largely explained by overestimations of tree cover (agricultural land, shrubland) and are less an issue of the leaf type discrimination.

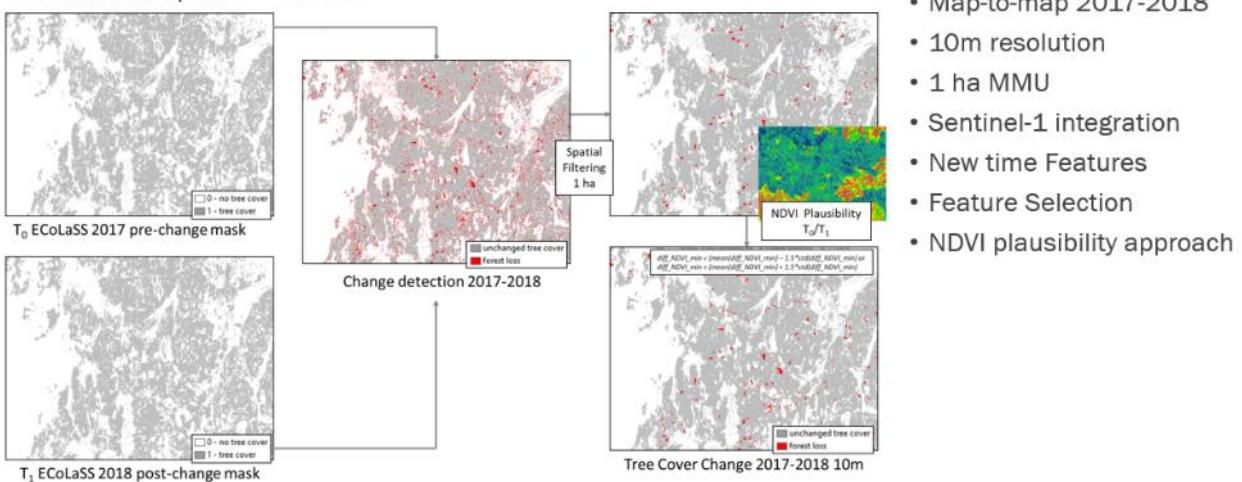
Tree Cover Density 2018 products at 10 m spatial resolution provide much more spatial detail than its 2015 precursor at 20 m resolution, which has been produced on a mono-temporal basis and certain sensors. With Sentinel-2 band-specific time features the TCD could be consistently produced without any density leaps between tile borders. The features calculated for the summer period (01.06.-31.08.2018) proved to be suitable to generate cloudless mosaics from which seamless density rasters have been produced. All regression coefficients are greater than 0.80, indicating a very good relationship between the reference dataset and the produced map density values.

### **5.2.3 Change Detection and Incremental Update Results and Validation**

In project phase 1, an Incremental Update layer has been simulated at 20 m spatial resolution for the period 2015-2017, based on a map-to-map change detection approach using the HRL2015 Tree Cover Mask (20m) and a newly classified Tree Cover Mask 2017 (10m). The selected map-to-map approach worked well, but results were below expectations as the input masks from 2015 and 2017 differs significantly in terms of spatial resolution and accuracy (error propagation through omission and commission errors in 2015/2017). Geometric differences between the TCM 2015 (derived from 20m satellite imagery), and the TCM 2017 (derived from 10m Sentinel-2 Level-2A time series) as well as cloud cover issues could be observed, resulting in a significant overestimation of forest loss. Although an MMU of 3 ha was proven to be the best-suited MMU size to capture real forest loss, 8% of forest loss were reported for the demonstration site North in the end. Those changes were unrealistic and not justifiable by intensive forest management activities. Furthermore, the selected MMU was not satisfying from a user perspective.

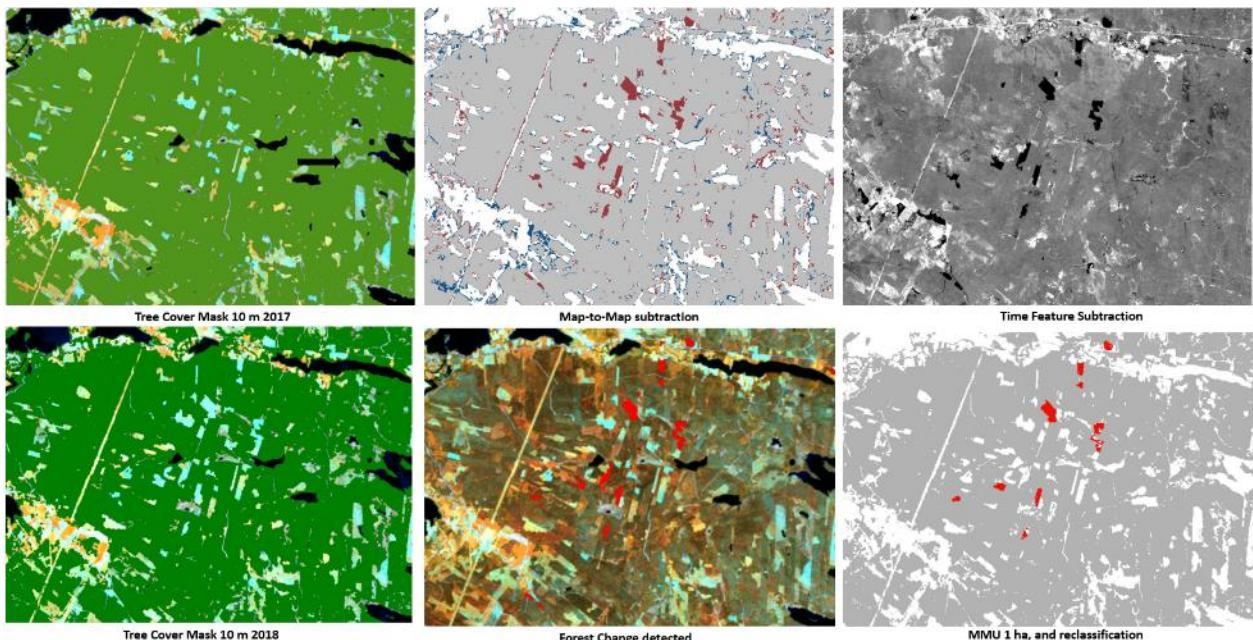
Based on the experiences and lessons learned from project phase 1, several measures have been taken to improve the quality of the Incremental Update. A modified version of the map-to-map change detection approach from phase 1 [see AD08, AD09] has been developed and finally applied using the improved TCM 2017 and TCM 2018 at 10 m spatial resolution in project phase 2. This modified version is shown in Figure 5-48 and includes the improved and consistently produced tree cover masks with integration of SAR Sentinel-1 time features and a subsequently applied NDVI plausibility approach of detected changes. A MMU of 1 ha has been defined in line with the HRL2018 specifications and finally applied in order to capture small changes.

### Incremental Update – Phase 2



**Figure 5-48: Modified Change detection workflow applied to the forest prototypes**

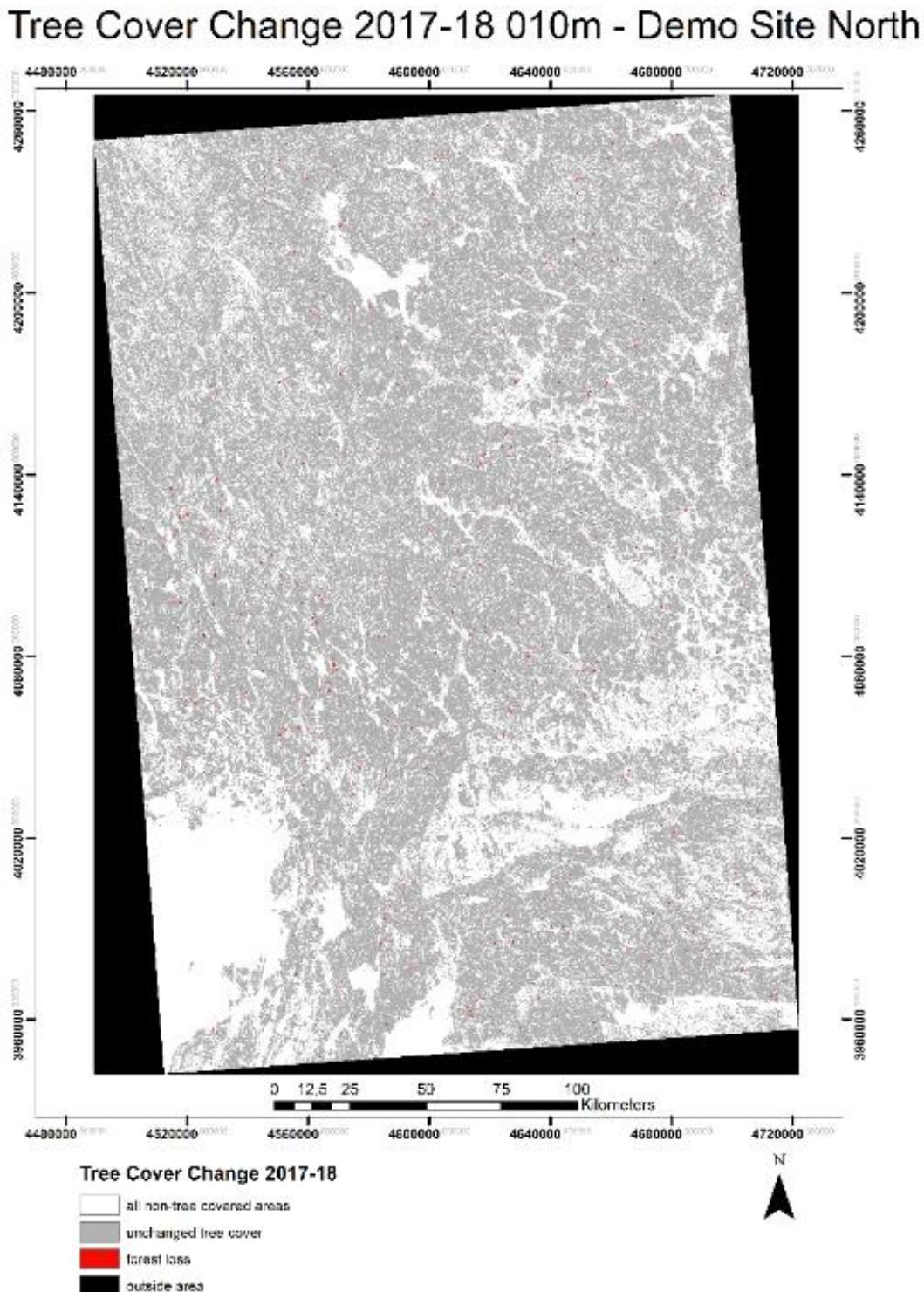
Similarly to the method applied to the Grassland product in the Central demonstration site, only areas of forest loss, and with a value above a 1.5 standard deviations from the NDVI mean value of the subtracted NDVI time feature layer remain as negative forest change. Areas of forest regrowth are reclassified to unchanged areas of tree cover. Figure 5-49 shows the relevant intermediate results of the applied change detection workflow for generation of the Incremental Update layers.



**Figure 5-49: Intermediate Products of the modified change detection workflow**

From a technical point of view, any so “measured” short-term tree cover increase would represent predominantly “technical changes” (i.e. either tree cover omissions in 2017; or commissions in 2018) rather than real forest regrowth. For the abovementioned reasons, forest increase is not taken into account for the present prototype, but could be potentially used to improve historic products in a focused enhancement step (e.g. by visual/manual enhancement steps of omission and commission errors in the reference years of the two-time steps steps or as basis for a re-classification based on probability approach).

As already stated in [AD07], for capturing meaningful increases in tree cover (density), longer observation periods of actually  $\geq 5$  years should be considered. However, this is actually not feasible within the runtime of the ECoLaSS project.

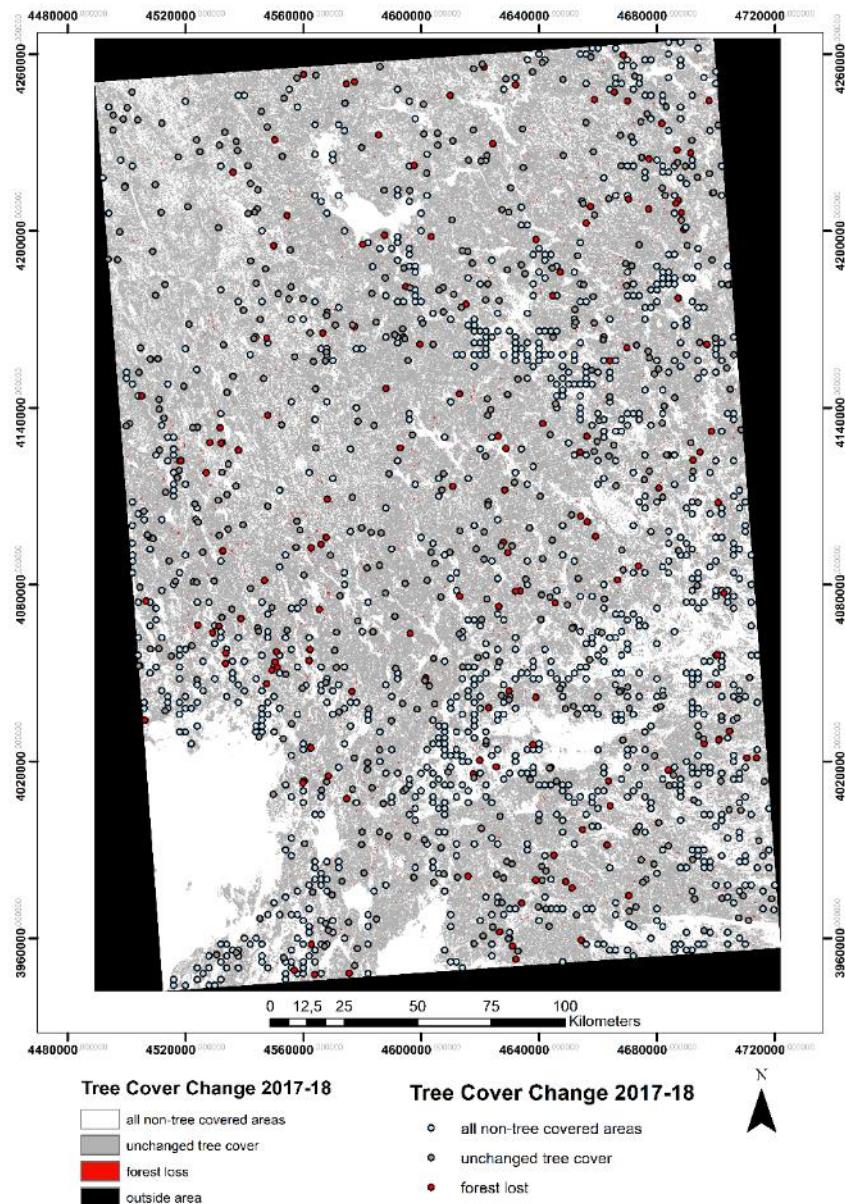


**Figure 5-50: Incremental Update Layer Tree Cover Change 2017-2018 at 10m spatial resolution for the demonstration site North**

The final TCC 1718 layers have been validated using the best available reference and guiding data (see section 5.2.2.1). The accuracy assessment has been performed using LUCAS 2018 points to assess the

classes 0 and 10 (recoded to conform to the classes in the TCC 1718), and a set of random points located in the forest loss class (12). Figure 5-51 shows exemplarily the TCC 1718 product for the demonstration site North overlaid with the validation points. Observable grid sampling effects are introduced by the original LUCAS samples.

### Tree Cover Change 2017-18 010m - Demo Site North

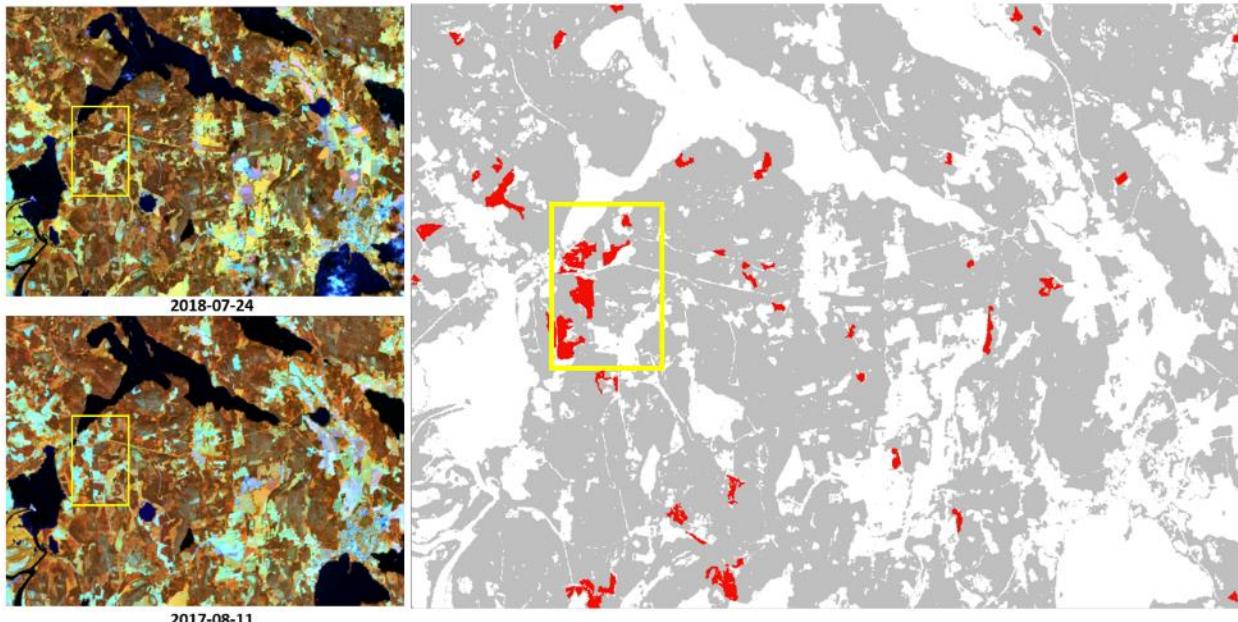


**Figure 5-51: Distribution of LUCAS 2018 points and additional random points for validation of the TCC 1718 layer**

Table 5-37 provides the error matrix for the Incremental Update TCC 1718 of the demonstration site North. A very high Overall Accuracy of 94.34 % could be achieved, which is due to the overall high proportion of (correctly classified) non-tree covered areas (class 0) and the high recognition rate of forest loss (class 12). Unchanged tree covered areas (class 10) are overestimated in the product, but shows a good level of reliability (above 80.00%). Commissions in this class are mainly caused by the error propagation from the TCM 2017 input mask. However, the producer accuracies of all classes are greater than 90% and in view of the forest loss, the producer's and user's accuracy is 95.54 % and 93.67 respectively. Consequently, the majority of forest losses has been correctly captured. Figure 5-55 provides an example of the high quality of the TCC 1718 prototype.

**Table 5-37: Error matrix for the Incremental Update Layer TCC 1718 of the demonstration site NORTH.**

| TCC_1718_010m_NO_03035 |                            | REFERENCE                  |                      |                |       | User Accuracy  | Confidence Interval          |
|------------------------|----------------------------|----------------------------|----------------------|----------------|-------|----------------|------------------------------|
|                        |                            | All non-tree covered areas | Unchanged tree cover | Forest loss    | Total |                |                              |
| PRODUCT                | All non-tree covered areas | 908                        | 1                    | 0              | 909   | 99.89%         | 99.62 – 100.16%              |
|                        | Unchanged tree cover       | 155                        | 439                  | 0              | 594   | 82.33%         | 82.33 – 88.19%               |
|                        | Forest loss                | 5                          | 0                    | 150            | 155   | 93.67%         | 93.67 – 99.88%               |
|                        | Total                      | 1068                       | 440                  | 150            | 1658  |                |                              |
|                        | Producer Accuracy          | 91.35%                     | 99.80%               | 95.54 %        |       | 94.34%         | Overall Accuracy             |
|                        | Confidence Interval        | 89.55 – 93.15%             | 99.32 – 100.29%      | 91.99 – 99.09% |       | 93.20 – 95.48% | Confidence Interval          |
|                        |                            |                            |                      |                |       | 0.954          | F-Score No Change non-forest |
|                        |                            |                            |                      |                |       | 0.919          | F-Score No Change Forest     |
|                        |                            |                            |                      |                |       | 0.962          | F-Score Loss                 |
|                        |                            |                            |                      |                |       | 0.898          | Kappa                        |

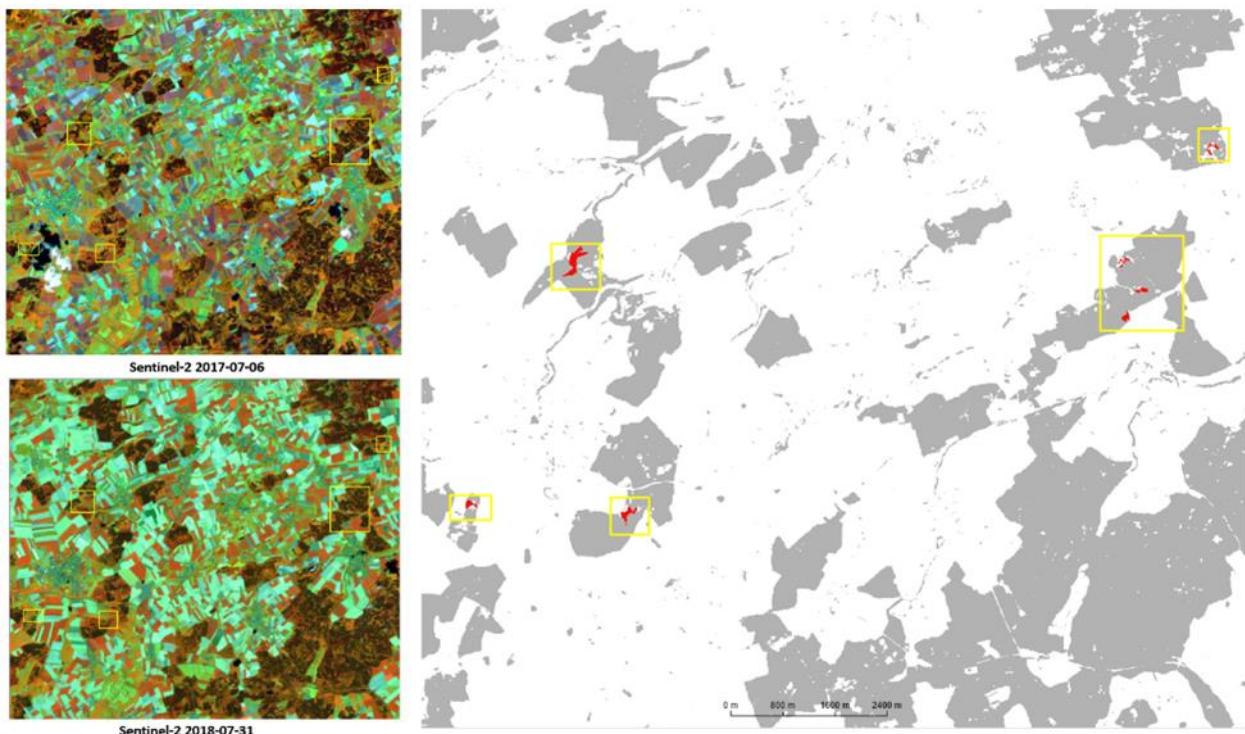


**Figure 5-52: Examples of forest loss provided by the TCC 1718 layer for the demonstration site North**  
 Left: Sentinel-2A scenes from July 2018 and August 2017; right: Tree Cover Change 2017-2018 10m

Likewise, the Incremental Update layer of the Central demonstration site shows a very high Overall Accuracy of 96.81% (see Table 5-38). A few commission and omission errors were found for the forest loss class, and its producer's and user's accuracy are 98.21% and respectively 96.48%. Both, producer's and user's accuracies are higher than in the North demonstration site and the user accuracy of the unchanged tree cover class exceeds the 85% threshold. Again, commission errors of tree cover are introduced by issues in the input Tree Cover Masks of 2017 and 2018. Figure 5-53 provides some details of the TCC 1718 prototype.

**Table 5-38: Error matrix for the Incremental Update Layer TCC 1718 of the demonstration site CENTRAL**

| TCC_1718_010m_CE_03035 |                            | REFERENCE                  |                      |                 |       | User Accuracy  | Confidence Interval          |
|------------------------|----------------------------|----------------------------|----------------------|-----------------|-------|----------------|------------------------------|
|                        |                            | All non-tree covered areas | Unchanged tree cover | Forest loss     | Total |                |                              |
| PRODUCT                | All non-tree covered areas | 2817                       | 0                    | 0               | 2817  | 100.00%        | 99.98 – 100.02%              |
|                        | Unchanged tree cover       | 113                        | 754                  | 4               | 871   | 86.57%         | 84.25 – 88.89%               |
|                        | Forest loss                | 4                          | 4                    | 219             | 227   | 96.48%         | 93.86 – 99.09%               |
|                        | Total                      | 2934                       | 758                  | 223             | 3915  |                |                              |
|                        | Producer Accuracy          | 96.01%                     | 99.47%               | 98.21%          |       | 96.81%         | Overall Accuracy             |
|                        | Confidence Interval        | 95.29 – 96.74%             | 98.89 – 100.5%       | 96.24 – 100.17% |       | 96.24 – 97.37% | Confidence Interval          |
|                        |                            |                            |                      |                 |       | 0.979          | F-Score No Change non-forest |
|                        |                            |                            |                      |                 |       | 0.926          | F-Score No Change Forest     |
|                        |                            |                            |                      |                 |       | 0.973          | F-Score Loss                 |
|                        |                            |                            |                      |                 |       | 0.922          | Kappa                        |



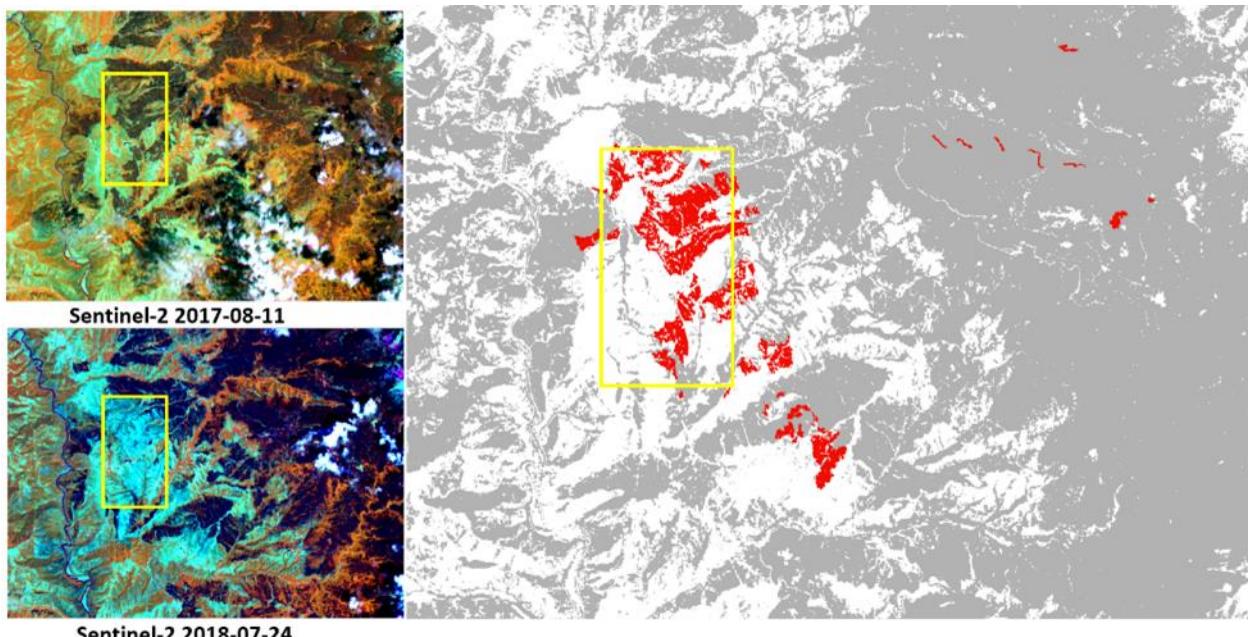
**Figure 5-53: Examples of forest loss provided by the TCC 1718 layer for the demonstration site Central**  
 Left: Sentinel-2A scenes from July 2018 and July 2017; right: Tree Cover Change 2017-2018 10m

The Incremental Update for the South-East demonstration site shows also a remarkable high Overall Accuracy which is at 96.78% (see Table 5-39). Producer accuracies are basically above 95%, but the user's accuracy of the forest loss class was found to be at 80.26%, which is the lowest accuracy value in all three sites. The latter comes from the relatively high commission errors, where non-tree covered areas (class 0) and unchanged tree covered areas (class 10) were classified as forest loss. The reason for this is in the deficient cloud masks produced by the MACCS processor, influencing each kind of calculated time features through disturbing block structures and, in case of frequent cloud cover, even through no data gaps. These

effects are negatively influencing the tree cover detection in 2017 and 2018 as well as the NDVI plausibility approach (see section 5.2.2).

**Table 5-39: Error matrix for the Incremental Update Layer TCC 1718 of the demonstration site SOUTH-EAST**

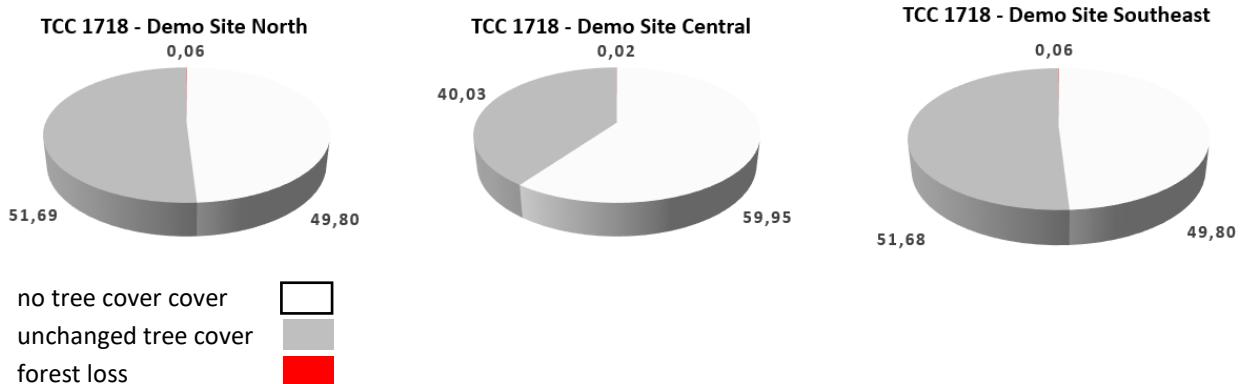
| TCC_1718_010m_SE_03035 |                            | REFERENCE                  |                      |                 |       | User Accuracy  | Confidence Interval          |
|------------------------|----------------------------|----------------------------|----------------------|-----------------|-------|----------------|------------------------------|
|                        |                            | All non-tree covered areas | Unchanged tree cover | Forest loss     | Total |                |                              |
| PRODUCT                | All non-tree covered areas | 2107                       | 0                    | 0               | 2107  | 100.00%        | 99.98 – 100.02%              |
|                        | Unchanged tree cover       | 60                         | 446                  | 0               | 539   | 88.87%         | 86.12 – 91.62%               |
|                        | Forest loss                | 16                         | 14                   | 122             | 152   | 80.26%         | 73.61 – 86.92%               |
|                        | Total                      | 2183                       | 493                  | 122             | 2798  |                |                              |
|                        | Producer Accuracy          | 96.52%                     | 97.16%               | 100.00%         |       | 96.78%         | Overall Accuracy             |
|                        | Confidence Interval        | 97.73 – 97.31%             | 95.59 – 98.73%       | 99.59 – 100.41% |       | 96.11 – 97.46% | Confidence Interval          |
|                        |                            |                            |                      |                 |       | 0.982          | F-Score No Change non-forest |
|                        |                            |                            |                      |                 |       | 0.928          | F-Score No Change Forest     |
|                        |                            |                            |                      |                 |       | 0.891          | F-Score Loss                 |
|                        |                            |                            |                      |                 |       | 0.9145         | Kappa                        |



**Figure 5-54: Examples of forest loss provided by the TCC 1718 layer for the demonstration site South-East**  
**Left: Sentinel-2A scenes from July 2018 and August 2017; right: Tree Cover Change 2017-2018 10m**

Independent from the shortcomings of the TCC in certain parts (deficient cloud masks, artefacts, data gaps) most of the detected losses confirm intensive forest management in the demonstration sites, especially within the northern site in Sweden. The high accuracy levels achieved within a one-year cycle provide reliable information on forest loss and show the potential of the suggested methodology to be applied at larger scale.

In summary, a forest loss of 280.16 km<sup>2</sup> has been calculated for the demonstration site North, 16.67 km<sup>2</sup> for the Central demonstration site and 36.33 km<sup>2</sup> for the South-East demonstration site (see **Figure 5-55**). These numbers, however, compared with the total study area, implied less than a 0.1% of the total area, which is plausible in view of the estimated annual change rates.



**Figure 5-55: Final statistics of the Incremental Update layers for all FOR demonstration sites**

In general, the selected map-to-map change approach is working well but is partially and significantly dependent on the quality of the available input maps/masks. It is also important to note that these results are based on the TCM basic products, which are fully obtained automatically without manual enhancement, post-processing or any kind of reclassification based on the change detection approach. So, user's and producer's accuracies could be largely improved in principle.

## 5.3 Prototype Specifications

This section provides a description of the dataset properties and metadata for the implemented prototypes, also referring to “P42.2 - Data Sets of HR Layer Incremental Updates”.

The raster products of the ECoLaSS prototypes are delivered as GeoTIFF (\*.tif) with world file (\*.tfw), pyramids (\*.ovr), attribute table (\*.dbf) and statistics (\*.aux.xml), enabling an instant illustration and analysis of the products within Geographic Information System (GIS) software. Each product is accompanied with a product-specific colour table (\*.clr) and INSPIRE-compliant metadata in XML format, and includes the probability layer layer as an additional quality assessment product.

Metadata are provided together with the products as INSPIRE-compliant XML files according to the EEA Metadata Standard for Geographic Information (EEA-MSGI). EEA-MSGI has been developed by EEA to meet needs and demands for inter-operability of metadata. EEA’s standard for metadata is a profile of the ISO 19115 standard for geographic metadata and contains more elements than the minimum required to comply the INSPIRE metadata regulation. Detailed conceptual specifications on EEA-MSGI and other relevant information on metadata can be found at <http://www.eionet.europa.eu/gis>.

Within ECoLaSS, a standardised and harmonised product file naming convention for all prototypes has been developed which is oriented along the already existing naming convention of the CLMS High Resolution Layers. This file naming convention has been applied to all raster prototypes and associated reference files and is documented in the Deliverables of Task 4. The naming convention consists of the following 7 descriptors:

THEME      YEAR      RESOLUTION      EXTENT      EPSG      TYPE      VERSION  
as follows:

### THEME

3-4 letter abbreviation for main products (TCM (Tree Cover Mask), DLT (Dominant Leaf Type), TCD (Tree Cover Density), TCC (Tree Cover Change), GRA (grassland), IBU (Built-up status layer), IMD (imperviousness degree), IMC (imperviousness change), IMCC (Imperviousness Change Classified), CRT (crop type), CRM (crop mask) and NLC (new land cover products).

### REFERENCE YEAR

2018 in four digits; change products in four digits (e.g. 1718)

### RESOLUTION

Four-digit (020m and 010m)

### EXTENT

2-digit code for demonstration-sites (CE (Central), NO (North), WE (West), SW (South West), SE (South East), SA (South Africa), ML (Mali))

### EPSG

5-digit EPSG code (geodetic parameter dataset code by the European Petroleum Survey Group) “03035” for the European LAEA projection

### TYPE

prototype

### VERSION

3-digit code “v01”

**EXAMPLE:**

"DLT\_2017\_010m\_NO\_03035\_prototype\_v01.tif" stands for: Dominant Leaf Type, 2017 reference year, 10m, Demonstration-site North, European projection (EPSG: 3035), prototype, version 01

"TCC\_1517\_020m\_NO\_03035\_prototype\_v01.tif" stands for: Tree Cover Change, 2015-2017 change period, 20m, Demonstration-site North, European projection (EPSG: 3035), prototype, version 01

The following 24 prototypes files as part of **D12.4 – P42.2b – Data Sets HR Layer Incremental Updates (Issue 2)** for Imperviousness in the demosites Central, South-East and South-West and for Forest in the demosites Central, South-East and North were submitted:

- IMC\_1518\_020m\_CE\_03035\_prototype\_v01.tif
- IMC\_1518\_020m\_SE\_03035\_prototype\_v01.tif
- IMC\_1718\_010m\_SW\_03035\_prototype\_v01.tif
- IMCC\_1518\_020m\_CE\_03035\_prototype\_v01.tif
- IMCC\_1518\_020m\_SE\_03035\_prototype\_v01.tif
- IMCC\_1718\_010m\_SW\_03035\_prototype\_v01.tif
- IMD\_2018\_010m\_CE\_03035\_prototype\_v01.tif
- IMD\_2018\_010m\_SE\_03035\_prototype\_v01.tif
- IMD\_2018\_010m\_SW\_03035\_prototype\_v01.tif
- IBU\_2018\_010m\_CE\_03035\_prototype\_v01.tif
- IBU\_2018\_010m\_SE\_03035\_prototype\_v01.tif
- IBU\_2018\_010m\_SW\_03035\_prototype\_v01.tif
- TCM\_2017\_010m\_CE\_03035\_prototype\_v01.tif
- TCM\_2018\_010m\_CE\_03035\_prototype\_v01.tif
- TCM\_2017\_010m\_SE\_03035\_prototype\_v01.tif
- TCM\_2018\_010m\_SE\_03035\_prototype\_v01.tif
- TCM\_2017\_010m\_NO\_03035\_prototype\_v02.tif
- TCM\_2018\_010m\_NO\_03035\_prototype\_v01.tif
- DLT\_2018\_010m\_CE\_03035\_prototype\_v01.tif
- DLT\_2018\_010m\_SE\_03035\_prototype\_v01.tif
- DLT\_2018\_010m\_NO\_03035\_prototype\_v01.tif
- TCD\_2018\_010m\_CE\_03035\_prototype\_v01.tif
- TCD\_2018\_010m\_SE\_03035\_prototype\_v01.tif
- TCD\_2018\_010m\_NO\_03035\_prototype\_v01.tif

The prototype specifications for the IMD, IBU, IMCC and IMC layers are respectively listed hereafter in Table 5-40, Table 5-41, Table 5-42 and Table 5-43, followed by the specifications for the TCM (Table 5-44), DLT (Table 5-45), TCD (Table 5-46) and TCC (Table 5-47). Analogous specifications are adopted for the 2018 prototypes. Thematic status layers are provided with the corresponding reliability layer (classification probability layer ranging from 0-100 %).

**Table 5-40: Detailed specifications for the improved primary status layer Imperviousness 2017/2018**

| <b>Imperviousness Degree 10m</b>  | <b>Acronym</b> | <b>Product category</b>       |
|---|----------------|-------------------------------|
|   | IMD            | Improved Primary Status Layer |
| <b>Reference year</b>   |                |                               |
| 2017/2018   |                |                               |
| <b>Extent</b>   |                |                               |
| Demonstration site South-West, Central and South-East   |                |                               |
| <b>Geometric resolution</b>   |                |                               |
| Pixel resolution 10m x 10m, fully conform with the EEA reference grid   |                |                               |
| <b>Coordinate Reference System</b>  |                |                               |
| European ETRS89 LAEA projection   |                |                               |
| <b>Geometric accuracy (positioning scale)</b>   |                |                               |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by Theia.   |                |                               |
| <b>Thematic accuracy</b>  |                |                               |
| Minimum 90% user's / producer's accuracy in general for status layers for a (derived) built-up/non built up map. Threshold to be applied in transforming imperviousness to built-up mask at 1%. |                |                               |
| <b>Data type</b>  |                |                               |
| 8bit unsigned integer raster with LZW compression   |                |                               |
| <b>Minimum Mapping Unit (MMU)</b>   |                |                               |
| Pixel-based (no MMU)  |                |                               |
| <b>Necessary attributes</b>   |                |                               |
| Raster value, count, class name   |                |                               |
| <b>Raster coding (thematic pixel values)</b>  |                |                               |
| 0: all non-impervious areas   |                |                               |
| 1-100: imperviousness values  |                |                               |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)   |                |                               |
| 255: outside area   |                |                               |
| <b>Metadata</b>   |                |                               |
| XML metadata files according to INSPIRE metadata standards  |                |                               |
| <b>Delivery format</b>  |                |                               |
| GeoTIFF   |                |                               |
| <b>Colour table</b>   |                |                               |
| ArcGIS *.clr format   |                |                               |

| Class Code | Class Name   | Red | Green | Blue |   |
|------------|--|-----|-------|------|---|
| 0          | all non-impervious areas   | 240 | 240   | 240  |  |
| 1          | 1% imperiousness value   | 255 | 237   | 195  |  |
| 50         | 50% imperiousness value  | 175 | 74    | 51   |  |
| 100        | 100% imperiousness value   | 113 | 12    | 2    |  |
| 254        | unclassifiable (no satellite image available, or clouds, or shadows) | 153 | 153   | 153  |  |
| 255        | outside area   | 0   | 0     | 0    |  |

**Table 5-41: Detailed specifications for the built-up status layer 2018**

| <b>Imperviousness Degree 10m</b>  | <b>Acronym</b>   | <b>Product category</b> |              |             |
|---|--|-------------------------|--------------|-------------|
|   | IBU  | Primary Status Layer    |              |             |
| <b>Reference year</b>   |  |                         |              |             |
| 2018  |  |                         |              |             |
| <b>Extent</b>   |  |                         |              |             |
| Demonstration sites South-West, Central and South-East  |  |                         |              |             |
| <b>Geometric resolution</b>   |  |                         |              |             |
| Pixel resolution 10m x 10m, fully conform with the EEA reference grid   |  |                         |              |             |
| <b>Coordinate Reference System</b>  |  |                         |              |             |
| European ETRS89 LAEA projection   |  |                         |              |             |
| <b>Geometric accuracy (positioning scale)</b>   |  |                         |              |             |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by Theia.   |  |                         |              |             |
| <b>Thematic accuracy</b>  |  |                         |              |             |
| Minimum 90% user's / producer's accuracy in general for status layers for a (derived) built-up/non built up map. Threshold to be applied in transforming imperviousness to built-up mask at 1%. |  |                         |              |             |
| <b>Data type</b>  |  |                         |              |             |
| 8bit unsigned integer raster with LZW compression   |  |                         |              |             |
| <b>Minimum Mapping Unit (MMU)</b>   |  |                         |              |             |
| Pixel-based (no MMU)  |  |                         |              |             |
| <b>Necessary attributes</b>   |  |                         |              |             |
| Raster value, count, class name   |  |                         |              |             |
| <b>Raster coding (thematic pixel values)</b>  |  |                         |              |             |
| 0: all non-built-up areas   |  |                         |              |             |
| 1: built-up areas   |  |                         |              |             |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)   |  |                         |              |             |
| 255: outside area   |  |                         |              |             |
| <b>Metadata</b>   |  |                         |              |             |
| XML metadata files according to INSPIRE metadata standards  |  |                         |              |             |
| <b>Delivery format</b>  |  |                         |              |             |
| GeoTIFF   |  |                         |              |             |
| <b>Colour table</b>   |  |                         |              |             |
| ArcGIS *.clr format   |  |                         |              |             |
| <b>Class Code</b>   | <b>Class Name</b>  | <b>Red</b>              | <b>Green</b> | <b>Blue</b> |
| 0   | all non built up areas   | 240                     | 240          | 240         |
| 1   | Built up areas   | 255                     | 178          | 0           |
| 254   | unclassifiable (no satellite image available, or clouds, shadows, or snow) | 153                     | 153          | 153         |
| 255   | outside area   | 0                       | 0            | 0           |

**Table 5-42: Detailed specifications for the Incremental Update Layer Imperviousness Change Classified**

| <b>Imperviousness Change Classified 20m</b>   | <b>Acronym</b> | <b>Product category</b>  |
|---|----------------|--------------------------|
|   | IMCC           | Incremental Update Layer |
| <b>Reference year</b>   |                |                          |
| 2015/16-2017/18   |                |                          |
| <b>Extent</b>   |                |                          |
| Demonstration site South-West   |                |                          |
| <b>Geometric resolution</b>   |                |                          |
| Pixel resolution 20m x 20m, fully conform with the EEA reference grid                         |                |                          |
| <b>Coordinate Reference System</b>  |                |                          |
| European ETRS89 LAEA projection   |                |                          |
| <b>Geometric accuracy (positioning scale)</b>   |                |                          |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by Theia. |                |                          |
| <b>Thematic accuracy</b>  |                |                          |
| 90% user's / producer's accuracy for derived IMD changes                                      |                |                          |
| <b>Data type</b>  |                |                          |
| 8bit unsigned integer raster with LZW compression   |                |                          |
| <b>Minimum Mapping Unit (MMU)</b>   |                |                          |
| Pixel-based (no MMU)  |                |                          |
| <b>Necessary attributes</b>   |                |                          |
| Raster value, count, class name   |                |                          |
| <b>Raster coding (thematic pixel values)</b>  |                |                          |
| 0: unchanged areas with imperviousness degree of 0  |                |                          |
| 1: new cover - increased imperviousness density, zero IMD at first reference date             |                |                          |
| 2: loss of cover - decreasing imperviousness density, zero IMD at second reference date       |                |                          |
| 10: unchanged areas, IMD>0 at both reference dates  |                |                          |
| 11: increased imperviousness density, IMD>0 at both reference dates                           |                |                          |
| 12: decreased imperviousness density, IMD>0 at both reference dates                           |                |                          |
| 254: unclassifiable in any of parent status layers  |                |                          |
| 255: outside area   |                |                          |
| <b>Metadata</b>   |                |                          |
| XML metadata files according to INSPIRE metadata standards                                    |                |                          |
| <b>Delivery format</b>  |                |                          |
| GeoTIFF   |                |                          |
| <b>Colour table</b>   |                |                          |
| ArcGIS *.clr format   |                |                          |

| Class Code | Class Name   | Red | Green | Blue |   |
|------------|--|-----|-------|------|---|
| 0          | unchanged areas with imperviousness degree of 0                                      | 240 | 240   | 240  |  |
| 1          | new cover - increased imperviousness density, zero IMD at first reference date       | 255 | 0     | 0    |  |
| 2          | loss of cover - decreasing imperviousness density, zero IMD at second reference date | 0   | 100   | 0    |  |
| 10         | unchanged areas, IMD>0 at both reference dates                                       | 156 | 156   | 156  |  |
| 11         | increased imperviousness density, IMD>0 at both reference dates                      | 255 | 191   | 0    |  |
| 12         | decreased imperviousness density, IMD>0 at both reference dates                      | 64  | 178   | 0    |  |
| 254        | unclassifiable in any of parent status layers  | 153 | 153   | 153  |  |
| 255        | outside area   | 0   | 0     | 0    |  |

Table 5-43: Detailed specifications for the Incremental Update Layer Imperviousness Change

| Imperviousness Change 20m   | Acronym | Product category         |
|---|---------|--------------------------|
|   | IMC     | Incremental Update Layer |
| <b>Reference year</b>   |         |                          |
| 2015/16-2017/18   |         |                          |
| <b>Extent</b>   |         |                          |
| Demonstration site South-West   |         |                          |
| <b>Geometric resolution</b>   |         |                          |
| Pixel resolution 20m x 20m, fully conform with the EEA reference grid                         |         |                          |
| <b>Coordinate Reference System</b>  |         |                          |
| European ETRS89 LAEA projection   |         |                          |
| <b>Geometric accuracy (positioning scale)</b>   |         |                          |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by Theia. |         |                          |
| <b>Thematic accuracy</b>  |         |                          |
| 90% user's / producer's accuracy for derived IMD changes                                      |         |                          |
| <b>Data type</b>  |         |                          |
| 8bit unsigned integer raster with LZW compression   |         |                          |
| <b>Minimum Mapping Unit (MMU)</b>   |         |                          |
| Pixel-based (no MMU)  |         |                          |
| <b>Necessary attributes</b>   |         |                          |
| Raster value, count, class name   |         |                          |
| <b>Raster coding (thematic pixel values)</b>  |         |                          |
| 0-100: imperviousness decrease  |         |                          |
| 100: Sealed in both years (stable built-up)   |         |                          |
| 101-200: imperviousness INcrease  |         |                          |
| 201: Non-sealed in both years (stable non-built up)   |         |                          |

| 254: unclassifiable in any of parent status layers         |  |     |       |      |   |
|--|--|-----|-------|------|---|
| 255: outside area  |  |     |       |      |   |
| <b>Metadata</b>  |  |     |       |      |   |
| XML metadata files according to INSPIRE metadata standards |  |     |       |      |   |
| <b>Delivery format</b>                                     |  |     |       |      |   |
| Geotiff  |  |     |       |      |   |
| <b>Colour table</b>  |  |     |       |      |   |
| ArcGIS *.clr format  |  |     |       |      |   |
| Class Code   | Class Name   | Red | Green | Blue |   |
| 0  | 100% imperviousness decrease   | 3   | 102   | 0    |    |
| 10   | 90% imperviousness decrease  | 12  | 114   | 0    |    |
| 50   | 50% imperviousness decrease  | 63  | 178   | 0    |    |
| 100  | Sealed in both years (stable built-up)                                     | 178 | 178   | 178  |    |
| 150  | 50% imperviousness increase  | 255 | 191   | 0    |    |
| 200  | 100% imperviousness increase   | 255 | 0     | 0    |  |
| 201  | Non-sealed in both years (stable non-built up)                             | 240 | 240   | 240  |  |
| 254  | unclassifiable (no satellite image available, or clouds, shadows, or snow) | 168 | 0     | 229  |  |
| 255  | outside area   | 0   | 0     | 0    |  |

The prototype specifications for the TCM, DLT, TCD and TCC layers are listed hereafter in Table 5-44, Table 5-45, Table 5-46 and Table 5-47 respectively follow. Analogous specifications are adopted for the 2018 prototypes.

**Table 5-44: Detailed specifications for the improved Tree Cover Mask**

| <b>Tree Cover Mask 10m</b>  | <b>Acronym</b>   | <b>Product category</b> |       |      |   |
|---|--|-------------------------|-------|------|---|
|   | TCM  | Basic Product           |       |      |   |
| <b>Reference year</b>   |  |                         |       |      |   |
| 2017 & 2018   |  |                         |       |      |   |
| <b>Extent</b>   |  |                         |       |      |   |
| Demonstration sites North, Central & South-East   |  |                         |       |      |   |
| <b>Geometric resolution</b>   |  |                         |       |      |   |
| Pixel resolution 10m x 10m, fully conform with the EEA reference grid   |  |                         |       |      |   |
| <b>Coordinate Reference System</b>  |  |                         |       |      |   |
| European ETRS89 LAEA projection   |  |                         |       |      |   |
| <b>Geometric accuracy (positioning scale)</b>   |  |                         |       |      |   |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.                         |  |                         |       |      |   |
| <b>Thematic accuracy</b>  |  |                         |       |      |   |
| Minimum 90% user's / producer's accuracy in general for status layers   |  |                         |       |      |   |
| <b>Data type</b>  |  |                         |       |      |   |
| 8bit unsigned raster with LZW compression   |  |                         |       |      |   |
| <b>Minimum Mapping Unit (MMU)</b>   |  |                         |       |      |   |
| Pixel-based (no MMU)  |  |                         |       |      |   |
| <b>Necessary attributes</b>   |  |                         |       |      |   |
| Raster value, count, class name, area (in km <sup>2</sup> ), area percentage (taking outside area not into account) |  |                         |       |      |   |
| <b>Raster coding (thematic pixel values)</b>  |  |                         |       |      |   |
| 0: all non-tree covered areas   |  |                         |       |      |   |
| 1: broadleaved trees  |  |                         |       |      |   |
| 2: coniferous trees   |  |                         |       |      |   |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)   |  |                         |       |      |   |
| 255: outside area   |  |                         |       |      |   |
| <b>Metadata</b>   |  |                         |       |      |   |
| XML metadata files according to INSPIRE metadata standards  |  |                         |       |      |   |
| <b>Delivery format</b>  |  |                         |       |      |   |
| GeoTIFF   |  |                         |       |      |   |
| <b>Colour table</b>   |  |                         |       |      |   |
| ArcGIS *.clr format   |  |                         |       |      |   |
| Class Code  | Class Name   | Red                     | Green | Blue |   |
| 0   | all non-tree covered areas   | 240                     | 240   | 240  |  |
| 1   | tree cover   | 70                      | 158   | 74   |  |
| 254   | unclassifiable (no satellite image available, or clouds, or shadows) | 153                     | 153   | 153  |  |
| 255   | outside area   | 0                       | 0     | 0    |  |

**Table 5-45: Detailed specifications for the improved primary status layer Dominant Leaf Type**

| <b>Dominant Leaf Type 10m</b>  | <b>Acronym</b> | <b>Product category</b>       |
|--|----------------|-------------------------------|
|  | DLT            | Improved Primary Status Layer |
| <b>Reference year</b>  |                |                               |
| 2018   |                |                               |
| <b>Extent</b>  |                |                               |
| Demonstration sites North, Central & South-East  |                |                               |
| <b>Geometric resolution</b>  |                |                               |
| Pixel resolution 10m x 10m, fully conform with the EEA reference grid                                  |                |                               |
| <b>Coordinate Reference System</b>   |                |                               |
| European ETRS89 LAEA projection  |                |                               |
| <b>Geometric accuracy (positioning scale)</b>  |                |                               |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.            |                |                               |
| <b>Thematic accuracy</b>   |                |                               |
| Minimum 90% user's / producer's accuracy in general for status layers                                  |                |                               |
| <b>Data type</b>   |                |                               |
| 8bit unsigned raster with LZW compression  |                |                               |
| <b>Minimum Mapping Unit (MMU)</b>  |                |                               |
| Pixel-based (no MMU)   |                |                               |
| <b>Necessary attributes</b>  |                |                               |
| Raster value, count, class name, area (in km2), area percentage (taking outside area not into account) |                |                               |
| <b>Raster coding (thematic pixel values)</b>   |                |                               |
| 0: all non-tree covered areas  |                |                               |
| 1: broadleaved trees   |                |                               |
| 2: coniferous trees  |                |                               |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)                              |                |                               |
| 255: outside area  |                |                               |
| <b>Metadata</b>  |                |                               |
| XML metadata files according to INSPIRE metadata standards   |                |                               |
| <b>Delivery format</b>   |                |                               |
| GeoTIFF  |                |                               |
| <b>Colour table</b>  |                |                               |
| ArcGIS *.clr format  |                |                               |

| Class Code | Class Name   | Red | Green | Blue |   |
|------------|--|-----|-------|------|---|
| 0          | all non-tree covered areas   | 240 | 240   | 240  |  |
| 1          | broadleaved trees  | 70  | 158   | 74   |  |
| 2          | coniferous trees   | 28  | 92    | 36   |  |
| 254        | unclassifiable (no satellite image available, or clouds, or shadows) | 153 | 153   | 153  |  |
| 255        | outside area   | 0   | 0     | 0    |  |

Table 5-46 : Detailed specifications for the improved primary status layer Tree Cover Density

| Tree Cover Density 10m   | Acronym | Product category              |
|--|---------|-------------------------------|
|  | TCD     | Improved Primary Status Layer |
| <b>Reference year</b>  |         |                               |
| 2018   |         |                               |
| <b>Extent</b>  |         |                               |
| Demonstration sites North, Central & South-East  |         |                               |
| <b>Geometric resolution</b>  |         |                               |
| Pixel resolution 10m x 10m, fully conform with the EEA reference grid                                  |         |                               |
| <b>Coordinate Reference System</b>   |         |                               |
| European ETRS89 LAEA projection  |         |                               |
| <b>Geometric accuracy (positioning scale)</b>  |         |                               |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.            |         |                               |
| <b>Thematic accuracy</b>   |         |                               |
| Minimum 90% user's / producer's accuracy in general for status layers                                  |         |                               |
| <b>Data type</b>   |         |                               |
| 8bit unsigned raster with LZW compression  |         |                               |
| <b>Minimum Mapping Unit (MMU)</b>  |         |                               |
| Pixel-based (no MMU)   |         |                               |
| <b>Necessary attributes</b>  |         |                               |
| Raster value, count, class name, area (in km2), area percentage (taking outside area not into account) |         |                               |
| <b>Raster coding (thematic pixel values)</b>   |         |                               |
| 0: all non-tree covered areas  |         |                               |
| 1 - 100: tree cover density in percent   |         |                               |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)                              |         |                               |
| 255: outside area  |         |                               |
| <b>Metadata</b>  |         |                               |
| XML metadata files according to INSPIRE metadata standards   |         |                               |
| <b>Delivery format</b>   |         |                               |
| GeoTIFF  |         |                               |

| Colour table        |  |     |       |      |   |
|---------------------|--|-----|-------|------|---|
| ArcGIS *.clr format |  |     |       |      |   |
| Class Code          | Class Name   | Red | Green | Blue |   |
| 0                   | all non-tree covered areas   | 240 | 240   | 240  |  |
| 1                   | 1% tree cover density  | 194 | 85    | 60   |  |
| 50                  | 50% tree cover density   | 126 | 237   | 0    |  |
| 100                 | 100% tree cover density  | 12  | 51    | 122  |  |
| 254                 | unclassifiable (no satellite image available, or clouds, or shadows) | 153 | 153   | 153  |  |
| 255                 | outside area   | 0   | 0     | 0    |  |

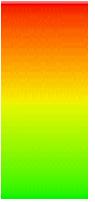
Table 5-47: Detailed specifications for the Incremental Update Layer Tree Cover Change

| Tree Cover Change 10m   | Acronym | Product category         |
|---|---------|--------------------------|
| Reference year  | TCC     | Incremental Update Layer |
| 2017/2018   |         |                          |
| Extent  |         |                          |
| Demonstration sites North, Central & South-East   |         |                          |
| Geometric resolution  |         |                          |
| Pixel resolution 20m x 20m, fully conform with the EEA reference grid                             |         |                          |
| Coordinate Reference System   |         |                          |
| European ETRS89 LAEA projection   |         |                          |
| Geometric accuracy (positioning scale)  |         |                          |
| Less than half a pixel. According to ortho-rectified satellite image base delivered by ESA.       |         |                          |
| Thematic accuracy   |         |                          |
| 80-85% overall accuracy   |         |                          |
| Data type   |         |                          |
| 8bit unsigned integer raster with LZW compression   |         |                          |
| Minimum Mapping Unit (MMU)  |         |                          |
| 1 ha  |         |                          |
| Necessary attributes  |         |                          |
| Raster value, count, class name, area (in km2), percentage (taking outside area not into account) |         |                          |

| <b>Raster coding (thematic pixel values)</b>  |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
|---|--|------------|-------|-------|---|--|---|----------------------------|-----|-----|-----|---|----|----------------------|-----|-----|-----|---|----|-----------------|----|----|-----|---|----|-------------|-----|---|---|---|-----|--|-----|-----|-----|---|-----|--------------|---|---|---|---|
| 0: all non-tree covered areas   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 10: unchanged tree cover  |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 11: forest regrowth (not relevant for this implementation of the TCC)   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 12: forest loss   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 254: unclassifiable (no satellite image available, or clouds, or shadows)   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 255: outside area   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| <b>Metadata</b>   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| XML metadata files according to INSPIRE metadata standards  |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| <b>Delivery format</b>  |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| GeOTIFF   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| <b>Colour table</b>   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| ArcGIS *.clr format   |  |            |       |       |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| <table border="1"> <thead> <tr> <th>Class Code</th> <th>Class Name</th> <th>Red</th> <th>Green</th> <th>Blue</th> <th></th> </tr> </thead> <tbody> <tr> <td>0</td> <td>all non-tree covered areas</td> <td>240</td> <td>240</td> <td>240</td> <td></td> </tr> <tr> <td>10</td> <td>unchanged tree cover</td> <td>240</td> <td>240</td> <td>240</td> <td></td> </tr> <tr> <td>11</td> <td>forest regrowth</td> <td>28</td> <td>72</td> <td>201</td> <td></td> </tr> <tr> <td>12</td> <td>forest loss</td> <td>255</td> <td>0</td> <td>0</td> <td></td> </tr> <tr> <td>254</td> <td>unclassifiable (no satellite image available, or clouds, or shadows)</td> <td>153</td> <td>153</td> <td>153</td> <td></td> </tr> <tr> <td>255</td> <td>outside area</td> <td>0</td> <td>0</td> <td>0</td> <td></td> </tr> </tbody> </table> | Class Code   | Class Name | Red   | Green | Blue  |  | 0 | all non-tree covered areas | 240 | 240 | 240 |  | 10 | unchanged tree cover | 240 | 240 | 240 |  | 11 | forest regrowth | 28 | 72 | 201 |  | 12 | forest loss | 255 | 0 | 0 |  | 254 | unclassifiable (no satellite image available, or clouds, or shadows) | 153 | 153 | 153 |  | 255 | outside area | 0 | 0 | 0 |  |
| Class Code  | Class Name   | Red        | Green | Blue  |   |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 0   | all non-tree covered areas   | 240        | 240   | 240   |    |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 10  | unchanged tree cover   | 240        | 240   | 240   |    |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 11  | forest regrowth  | 28         | 72    | 201   |  |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 12  | forest loss  | 255        | 0     | 0     |  |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 254   | unclassifiable (no satellite image available, or clouds, or shadows) | 153        | 153   | 153   |  |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |
| 255   | outside area   | 0          | 0     | 0     |  |  |   |                            |     |     |     |   |    |                      |     |     |     |   |    |                 |    |    |     |   |    |             |     |   |   |   |     |  |     |     |     |   |     |              |   |   |   |   |

In addition to the prototype layers, probability layers are provided as a by-product. This additional file serves as one of the accuracy parameters that are described in detail in the WP 33 final report [AD07] and range from 0 to 100%. The higher the percentage the higher the probability that the respective pixel belongs to the depicted class. In this manner, the probability band depicts the error map at pixel level. Furthermore, areas excluded by the referring mask are coded with the value “101”. An overview of the probabilities’ colour palette is given in Table 5-48.

**Table 5-48: Colour palette for the probability layers**

| Class Code | Class Name                    | Red | Green | Blue |   |
|------------|-------------------------------|-----|-------|------|---|
| 0          | 0% probability                | 245 | 0     | 0    |  |
| 50         | 50% probability               | 245 | 241   | 0    |   |
| 100        | 100% probability              | 20  | 245   | 0    |   |
| 101        | areas excluded by binary mask | 128 | 128   | 128  |  |
| 255        | outside area                  | 0   | 0     | 0    |  |

## 6 Conclusions and Outlook

The implementation of the prototypes on the **HRL Imperviousness** led to methods now fully operational (fully obtained automatically without manual enhancement and post-processing) with a change layer (IMC), a classified change layer (IMCC) produced at 20m and a imperviousness status layer (IMD) and a built-up status layer (BU) at 10m.

In phase 2, it has been demonstrated that performant supervised classification algorithm based on the exploitation of the Sentinel-1+2 constellation (and derived time features) associated to active learning and textural indices leads to reliable products. Thus, the active learning algorithm also shows great classification performances and was very computation efficient. The optical series were not dense enough to highlight the inter-yearly and intra-yearly phenology dynamics in order to highlight non-urban areas and so was associated to Sentinel-1 time series.

It also has been demonstrated the high adding value of the Pantex index for the separation of building (built-up features) from flat impervious surfaces.

With the applied change detection approach, it has been demonstrated that the re-classification procedure based on a reduction- of- bias algorithm to ensure temporal and spatial consistency of changes or the mask permits isolating actual change areas from technical errors and so an updated and improved sealed surface masks can be generated. It should be stressed out that the specifications of the HRL 2015 layer are different from the specifications and the input data situation for the HRL 2018 (10m vs 20m) which explains most of the omission errors.

In turn, in view of a future **HRL Forest** prototype, it can be stated that the full exploitation of the Sentinel-2A+B time series of the spring period using time features has generally proven a stable and reliable approach in phase 1, capable to successfully differentiate between tree covered and non-tree covered areas as well as broadleaved and coniferous tree/forest stands in a seamless and consistent manner. With the selected innovative approach covering a time period of three months, it could be demonstrated that an almost fully automatic DLT status layer generation is generally possible with satellite data acquired within one year and in 10m spatial resolution. However, dependent on the overall data situation, cloud coverage might be an issue. Therefore, the integration of SAR data and an extension of the observation period have been recommended as main topics regarding a further improvement of the status layers and the thereof derived Incremental Update.

In phase 2, these recommendations have been considered in all three FOR demonstration sites. SAR data has been successfully integrated and cloud coverage issues could be partially addressed by an extension of the observation period to six month. However, this procedure has some limitations for areas with generally very high cloud cover rates (e.g. British Isles, Scandinavia and French-Guiana), as it could be observed for two out of the six tiles within the demonstration site North. Consequently, the integration of Sentinel-1 SAR data is not a trivial matter towards a consistent pan-European target. Further improvements implemented in phase 2 include plausibility checks, and a refinement of the classification workflow. In general, the achieved higher degree of automation in the classification process might contribute to shorter production times in the future. It should also be noted, that the improved primary status layers DLT and TCD could be produced without manual enhancements, providing a high level of detail with promising accuracy figures.

With the applied image-to-image change detection method, using the TCM 2015 mask as benchmark, an incremental update layer could be simulated within the confines of the previous (historic) 2015 inventory at 20m spatial resolution, providing information on forest cover loss within a relatively short time frame of  $\leq 12$  months. Due to the reasons discussed in section 5.2.2, the achieved accuracy of the incremental update layer in phase 1 was still not sufficient, but was significantly improved in phase 2 by producing a

more consistent product derived from the consistently produced Tree Cover Masks 2017 and 2018 at 10m spatial resolution with a subsequently applied NDVI plausibility analysis of detected changes. The selected approach is to a certain degree dependent on the initial thematic quality of the input data (e.g. in form of Tree Cover Masks and NDVI features for certain reference years), but the spatial and temporal consistency of the input layers make this cost-efficient approach attractive for application at larger scales.

At this stage, it should be noted, that the quite ambitious and experimental 20m Dominant Leaf Type Change product of the Copernicus HRL Forest 2015 product portfolio (covering the time period 2012 to 2015; providing 14 thematic classes, thereof 10 change classes) taking different combinations of leaf type and tree cover density changes into account, has been assessed as too complex for users. The initially defined target accuracy of 90% for producer's and user's accuracy for detected changes could by far not be achieved within the HRL 2015 production due to the complexity of the product and the specific lineage of the input data (HRL Forest products 2012 and 2015). Therefore, it has been recommended to revise (simplify) the current product specifications to achieve more reliable product characteristics and accuracies for a forest change product in the HR domain.

This has been finally considered by the responsible Entrusted Entity in the HRL2018 ITT and the proposed Incremental Update in form of the Tree Cover Change layer within ECoLaSS is going towards such a simplification of the change aspect, strongly highlighting the temporal aspect. From a data perspective, an operational roll-out of the Incremental Update at pan-European level seems to be feasible in a 1-year update cycle with the Sentinel-2 constellation, regionally supported by Sentinel-1 SAR data. The selected approach has been successfully transferred to areas of different geographic conditions and seasonal patterns, providing convincing accuracy figures.

Collectively, we consider the forest topic has covered the most relevant and representative conditions for a potential pan-European roll-out.

## References

- Alajlan N., Bazi, Y., Al Hichri H., Melgani F., Yager R., (2013)."Using OWA Fusion Operators for the Classification of Hyperspectral Images" in Journal of Selected Topics in Applied Earth Observations and Remote Sensing v. 6, n. 2 (2013), p. 602-614.
- Breiman, L. (2001). Random Forests. Machine Learning, 4(1), 5-32.
- Breiman, L. (1984). Classification and Regression Trees. Chapman & Hall/CRC.
- Camp-Valls, G., & Bruzzone, L. (2009). Kernek methods for remote sensing data analysis. John Wiley & Sons.
- Carlinet, E., & Géraud, T. (2014). A Comparative Review of Component TreeComputation Algorithms. IEEE Transactions on Image Processing, 23(9), 3885-3895.
- Crawford, M., Tuia, D., & Ynag, H. (2013). Active Learning: Any Value for Classification of Remotely Sensed Data? Proceedings of the IEEE, 593-608.
- Dalla Mura, M., Benediktsson, J., Waske, B., & Bruzzone, L. (2010, October). Morphological Attribute Profiles for the Analysisof Very High Resolution Images. IEEE Transactions on Geoscience and Remote Sensing, 48(10), 3747-3761.
- Florczyk, A., Ferri, S., Vasileios, S., Kemper, T., Halkia, M., Soille, P., & Pesaresi, M. (2015). A New European Settlement Map From Optical Remotely Sensed Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1-15.
- Gallaun, H., Schardt, M., Linser, S., 2007. Remote sensing based forest map of Austria and derived environmental indicators. Proceedings of ForestSat.
- Henrich, V., Krauss, G., Götze, C., Sandow, C., 2012. IDB. URL: <http://www.indexdatabase.de> (Entwicklung einer Datenbank für Fernerkundungsindizes. AK Fernerkundung, Bochum)
- Lefebvre, A., Sannier, C., & Corpetti, T. (2016). Monitoring Urban Areas with Sentinel-2A Data: Application to the Update of the Copernicus High Resolution Layer Imperviousness Degree. Remote Sensing, 8(606), 1-21.
- Mas, J., & Flores, J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. Internation Journal of Remote Sensing, 617-663.
- Mountrakis, G., Im, J., & Ogola, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and remote Sensing, 66(3), 247-259.
- Pesaresi, M., & Benediktsson, J. A. (2001, March). A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE Transactions on Geoscience and Remote Sensing, 39(2), 309-319.
- Pesaresi, M., G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu, M. A. Marin-Herrera, G. K. Ouzounis, M. Scavazzon, P. Soille, V. Syrris & L. Zanchetta (2013) A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6, 2102-2131.
- Stehman, S. V., and Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. Remote sensing of environment, 64(3), 331-344.
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). Introduction to Data Mining. Pearson.
- Tuia D., Ratle F., Pacifici F., Kanevski M.F., Emery W.J., (2009). "Active Learning Methods for Remote Sensing Image Classification,"., IEEE Transactions on Geoscience and Remote Sensing, vol.47, no.7, pp.2218,2232, July 2009 doi: 10.1109/TGRS.2008.2010404.

Tuia D., Volpi M., Copa L., Kanevski M., Munoz-Mar Jordi, (2011). A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal of Selected Topics in Signal Processing* 5(3):606 - 617 · July 2011.

Wulder, M., Franklin, S. E., 2012. Remote sensing of forest environments: concepts and case studies. Springer Science & Business Media.

Zhang, L., Zhu, X., Zhang, L., & Du, B. (2016). Multidomain Subspace Classification for Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 1-13.