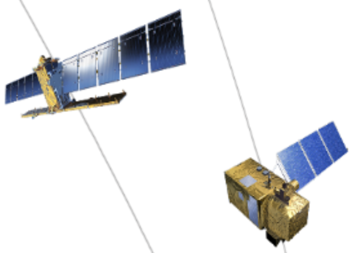

Earth Observation for
Sustainable Development

urban development
eo4sd

Urban Development Project

EO4SD-Urban Project: Campeche City Report

ESA Ref:	AO/1-8346/15/I-NB
Doc. No.:	City Operations Report
Issue/Rev.:	1.1
Date:	19.11.2019

GAF AG · SIRS · gisat · egis · DLR · NEO · JOANNEUM RESEARCH DIGITAL · GISBOX · esa

Consortium Partners

No.	Name	Short Name	Country
1	GAF AG	GAF	Germany
2	Système d'Information à Référence Spatiale SAS	SIRS	France
3	GISAT S.R.O.	GISAT	Czech Republic
4	Egis SA	EGIS	France
5	Deutsche Luft- und Raumfahrt e. V	DLR	Germany
6	Netherlands Geomatics & Earth Observation B.V.	NEO	The Netherlands
7	JOANNEUM Research Forschungsgesellschaft mbH	JR	Austria
8	GISBOX SRL	GISBOX	Romania

Summary

This document contains information related to the provision of geo-spatial products from the European Space Agency (ESA) supported project "Earth Observation for Sustainable Development" Urban Applications (EO4SD-Urban) to the Inter-American Development Bank (IADB) for supporting its funding project called "Emerging and Sustainable Cities Initiative" (ESCI) for the city of Campeche (Mexico), linked to the GEF funded project called "Global Platform for Sustainable Cities" (GPSC).

	Affiliation/Function	Name	Date
Prepared	SIRS NEO GISBOX	S. Delbour, D. Fretin F. Fang R. Ocoliseanu	29/07/2019
Reviewed	SIRS	C. Sannier	02/08/2019
Approved	GAF AG, Project Coordinator	T. Haeusler	02/08/2019

The document is accepted under the assumption that all verification activities were carried out correctly and any discrepancies are documented properly.

Distribution

Affiliation	Name	Copies
ESA	Z. Bartalis	electronic copy
IADB	A. Ruiz D. Arcia	electronic copy

Document Status Sheet

Issue	Date	Details
1.0	02/08/2019	First Document Issue
1.1	19/11/2019	Addition of section 4.4 Sustainable Development Goal 11 Indicators

Document Change Record

#	Date	Request	Location	Details
1	02/08/2019			Initial version

This page is intentionally left blank!

Executive Summary

The European Space Agency (ESA) has been working closely together with the International Finance Institutes (IFIs) and their client countries to demonstrate the benefits of Earth Observation (EO) in the IFI development programmes. Earth Observation for Sustainable Development (EO4SD) is a new ESA initiative, which aims to achieve an increase in the uptake of satellite-based information in the regional and global IFI programmes. The overall aim of the EO4SD Urban project is to integrate the application of satellite data for urban development programmes being implemented by the IFIs or Multi-Lateral Development Banks (MDBs) with the developing countries. The overall goal will be achieved via implementation of the following main objectives:

- To provide a service portfolio of Baseline and Derived urban-related geo-spatial products
- To provide the geo-spatial products and services on a geographical regional basis
- To ensure that the products and services are user-driven

This Report describes the generation and the provision of EO-based information products to the Inter-American Development Bank (IADB) for supporting its funding project called "Emerging and Sustainable Cities Initiative" (ESCI) for Mexico and the counterpart City authorities in Campeche. This is linked to the GEF funded project called "Global Platform for Sustainable Cities" (GPSC).

The Report provides a Service Description by referring to the user driven service requirements and the associated product list with the detailed product specifications. The following products were requested:

- Urban Land Use / Land Cover (LU/LC) & Changes
- Settlement Extent and Imperviousness and Change
- Urban Green Areas & Changes

The current Version of this Report contains the description of the generation and delivery of each requested product. The Land Use / Land Cover (LU/LC) and Urban Green Areas products have been generated for two reference dates (current status and historical one, respectively 2006 and 2018) and thus include also the changes occurred during this period.

This City Operations Report for Campeche systematically reviews the main production steps involved and importantly highlights the Quality Control (QC) mechanisms involved; the steps of QC and the assessment of quality is provided in related QC forms in the Annexe of this Report. There is also the provision of standard analytical work undertaken with the products which can be further included as inputs into further urban development assessments, modelling and reports.

This page is intentionally left blank!

Table of Contents

1 GENERAL BACKGROUND OF EO4SD-URBAN 1
2 SERVICE DESCRIPTION 1
2.1 STAKEHOLDERS AND REQUIREMENTS 1
2.2 SERVICE AREA SPECIFICATION 2
2.3 PRODUCT LIST AND PRODUCT SPECIFICATIONS 4
2.4 LAND USE/LAND COVER NOMENCLATURE 5
2.5 WORLD SETTLEMENT EXTENT 8
2.6 PERCENTAGE IMPERVIOUS SURFACE 8
2.7 URBAN GREEN AREAS NOMENCLATURE 9
2.8 TERMS OF ACCESS 9
3 SERVICE OPERATIONS 10
3.1 SOURCE DATA 10
3.2 PROCESSING METHODS 11
3.3 ACCURACY ASSESSMENT OF MAP PRODUCTS 11
3.3.1 Accuracy Assessment of LU/LC Product 11
3.3.2 Accuracy Assessment of World Settlement Extent Product 17
3.3.3 Accuracy Assessment of the Percentage Impervious Surface Product 19
3.3.4 Accuracy Assessment of Urban Green Areas Product 21
3.4 QUALITY CONTROL/ASSURANCE 24
3.5 METADATA 25
4 ANALYSIS OF MAPPING RESULTS 26
4.1 SETTLEMENT EXTENT – DEVELOPMENTS 1985, 1990, 1995, 2000, 2005, 2010 AND 2015 26
4.2 LAND USE / LAND COVER FOR 2006 AND 2018 27
4.2.1 LU/LC Mapping for Core City Area 27
4.2.2 Spatial Distribution of Main LU/LC Change Categories for Core City Area 31
4.2.3 LU/LC Mapping for Larger Urban Area 34
4.2.4 Spatial Distribution of Main LU/LC Change Categories for Larger Urban Area 37
4.3 URBAN GREEN AREAS 39
4.4 SUSTAINABLE DEVELOPMENT GOAL 11 INDICATORS 41
4.4.1 SDG 11 Indicator 11.2.1 42
4.4.2 SDG 11 Indicator 11.3.1 43
4.4.3 SDG 11 Indicator 11.7.1 45
4.5 CONCLUDING POINTS 46
5 REFERENCES 47

Annex 1: AOI Calculation based on the DG Regio approach
Annex 2: Processing Methods for EO4SD-Urban Products
Annex 3: Filled Quality Control Sheets

List of Figures

Figure 1: Illustration of Core City and Larger Urban Areas of Campeche. 4
Figure 2 : Mapping result of the Core City area of Campeche of the year 2018 overlaid with randomly distributed sample points used for accuracy assessment. 14
Figure 3: Mapping result of the Larger Urban area of Campeche of the year 2018 overlaid with randomly distributed sample points used for accuracy assessment. 15
Figure 4: Result of the Urban Green Area mapping in Campeche (change product) with sampling points used for product validation. 22
Figure 5: Quality Control process for EO4SD-Urban product generation. At each intermediate processing step output properties are compared against pre-defined requirements. 24
Figure 6: Settlement Extent developments in the epochs 1985 to 1990, 1990 to 1995, 1995 to 2000, 2000 to 2005, 2005 to 2010 and 2010 to 2015 in Campeche and surrounding region. 26
Figure 7: Settlement Extent developments in the epochs 1985 to 1990, 1990 to 1995, 1995 to 2000, 2000 to 2005, 2005 to 2010 and 2010 to 2015 in Campeche within the High Density Area. 27
Figure 8: Core City Area - Detailed Land Use Land Cover 2018 over Campeche. 28
Figure 9: Core City Area - Insight on the detailed Land Use Land Cover 2018 inside the city. 28
Figure 10: Core City Area - Detailed Land Use Land Cover 2018 structure presented in % (left) and km^2 (right). 29
Figure 11: Core City Area - Detailed Land Use Land Cover 2006 structure presented in % (left) and km² (right). 29
Figure 12: Core City Area - Land Use Land Cover change types and spatial distribution. 32
Figure 13: Core City Area - Land Use Land Cover Change Types 2006-2018 presented in % (left) and km² (right). 33
Figure 14: Larger Urban Area - Land Use Land Cover 2018 over Campeche. 34
Figure 15: Larger Urban Area - Insight on the Land Use Land Cover 2018 outside the city. 34
Figure 16: Larger Urban Area - Land Use Land Cover 2006 structure presented in % (left) and km² (right). 35
Figure 17: Larger Urban Area - Land Use Land Cover 2018 structure presented in % (left) and km^2 (right). 35
Figure 18: Larger Urban Area - Land Use Land Cover change types and spatial distribution. 37
Figure 19: Larger Urban Area - Land Use Land Cover Change Types 2006-2018 presented in % (left) and km² (right). 38
Figure 20: Urban Green Areas changes and spatial distribution. 39
Figure 21: Status and change of urban green areas in-between 2006 and 2018 expressed in %. 40
Figure 22: Status and change of urban green areas in-between 2006 and 2018 expressed in area. 40
Figure 23: Proportion of population with convenient access to public transport. 42
Figure 24: Ratio of land consumption rate to population growth rate between 2005 and 2015. 43
Figure 25: Percentage change of population and land consumption between 2005 and 2015. 44
Figure 26: Average share of the built-up area that is open space for public use. 46
Figure 27: Satellite image showing Melaka and the surrounding area. 52
Figure 28: Global Human Settlement Population Layer (spatial resolution of 1 km). 52
Figure 29: DLR population layer (spatial resolution of 10 m). 52
Figure 30: Aggregated DLR population layer (spatial resolution of 1 km). 53
Figure 31: High Density Core area of Melaka calculated based on the aggregated DLR population layer. 53
Figure 32: Urban Cluster area of Melaka calculated based on the aggregated DLR population layer.. 54

List of Tables

Table 1: LU/LC Nomenclature for GPSC Cities (High Density AOI)... 6
Table 2: LU/LC Nomenclature for GPSC Cities (Urban Cluster Area)... 6
Table 3: Nomenclature used for the mapping and identification of Urban Green Areas. 9
Table 4: Number of sampling points for the EO4SD-Urban mapping classes after applied sampling design with information on overall land cover by class. 13
Table 5: Number of sampling points for the Urban Cluster classes after applied sampling design with information on overall land cover by class... 13
Table 6: Accuracies exhibited by the WSF2015 according to the three considered agreement criteria for different definitions of settlement... 18
Table 7: Acquisition dates and size of the WV2 images available for the 5 test sites analysed in the validation exercise along with the number of corresponding 30x30m validation samples. 20
Table 8: Calculation of the minimum number of samples according Goodchild et al. (1994). 21
Table 9: Results of the Accuracy Assessment of Urban Green Areas in Campeche, 2006.................. 22
Table 10: Results of the Accuracy Assessment of Urban Green Areas in Campeche, 2018. 23
Table 11: Core City Area - Detailed information on area and percentage of total area for each class for 2006 and 2018 as well as the changes. 30
Table 12: Core City Area - Overall Main LU/LC Changes Statistics. 33
Table 13: Larger Urban Area - Detailed information on area and percentage of total area for each class for 2006 and 2018 as well as the changes. 36
Table 14: Larger Urban Area - Overall Main LU/LC Changes Statistics. 38

List of Abbreviations

AOI	Area of Interest
CDS	City Development Strategy
CS	Client States
DEM	Digital Elevation Model
DLR	German Space Agency
EEA	European Environmental Agency
EGIS	Consulting Company for Environmental Impact Assessment and Urban Planning, France
EO	Earth Observation
ESA	European Space Agency
ESCI	Emerging and Sustainable Cities Initiative
EU	European Union
GAF	GAF AG, Geospatial Service Provider, Germany
GIS	Geographic Information System
GISAT	Geospatial Service Provider, Czech Republic
GISBOX	Romanian company with activities of Photogrammetry and GIS
GPSC	Global Platform for Sustainable Cities
HR	High Resolution
HRL	High Resolution Layer
IADB	Inter-American Development Bank
IFI	International Financing Institute
INSPIRE	Infrastructure for Spatial Information in the European Community
INFOCAM	Instituto de Información Estadística Geográfica y Catastral del Estado de Campeche
ISO/TC 211	Standardization of Digital Geographic Information
JR	JOANNEUM Research, Austria
LULC	Land Use / Land Cover
LULCC	Land Use and Land Cover Change
MMU	Minimum Mapping Unit
NDVI	Normalized Difference Vegetation Index
NEO	Geospatial Service Provider, The Netherlands
PIS	Percentage Impervious Surface
QA	Quality Assurance
QC	Quality Control
QM	Quality Management
SIRS	Geospatial Service Provider, France
SP	Service Provider
VHR	Very High Resolution
WB	World Bank
WSF	World Settlement Footprint

This page is intentionally left blank!

1 General Background of EO4SD-Urban

Since 2008 the European Space Agency (ESA) has worked closely together with the International Finance Institutes (IFIs) and their client countries to harness the benefits of Earth Observation (EO) in their operations and resources management. Earth Observation for Sustainable Development (EO4SD) is a new ESA initiative, which aims to achieve an increase in the uptake of satellite-based information in the regional and global IFI programmes. The EO4SD-Urban project initiated in May 2016 (with a duration of 3 years) has the overall aim to integrate the application of satellite data for urban development programmes being implemented by the IFIs with the developing countries. The overall goal will be achieved via implementation of the following main objectives:

- To provide the services on a regional basis (i.e. large geographical areas); in the context of the current proposal with a focus on S. Asia, SE Asia and Africa, for at least 35-40 cities.
- To ensure that the products and services are user-driven; i.e. priority products and services to be agreed on with the MDBs in relation to their regional programs and furthermore to implement the project with a strong stakeholder engagement especially in context with the validation of the products/services on their utility.
- To provide a service portfolio of Baseline and Derived urban-related geo-spatial products that have clear technical specifications and are produced on an operational manner that are stringently quality controlled and validated by the user community.
- To provide a technology transfer component in the project via capacity building exercises in the different regions in close co-operation with the MDB programmes.

This Report supports the fulfilment of the third objective which requires the provision of geo-spatial Baseline and Derived geo-spatial products to various stakeholders in the IFIs and counterpart City authorities. The Report provides a Service Description, and then in Chapter 3 systematically reviews the main production steps involved and importantly highlights whenever there are Quality control (QC) mechanisms involved with the related QC forms in the Annexe of this Report. The description of the processes is kept intentionally at a top leave and avoiding technical details as the Report is considered mainly for non-technical IFI staff and experts and City authorities. Finally, Chapter 4 presents the standard analytical work undertaken with the products which can be inputs into further urban development assessments, modelling and reports.

2 Service Description

The following Section summarises the service as it has been performed for the city of Campeche, Mexico, within the EO4SD-Urban Project and as it has been delivered to Avelina Ruiz, Climate Change Consultant at IADB, and Diego Arcia, Urban Development Expert at IADB.

2.1 Stakeholders and Requirements

The EO4SD-Urban products described in this Report has been provided for Campeche, Mexico, upon request of the Inter-American Development Bank (IADB) for supporting its funding project called “Emerging and Sustainable Cities Initiative” (ESCI). This is linked to the GEF funded project called “Global Platform for Sustainable Cities” (GPSC).

The main local stakeholder is the Instituto de Información Estadística Geográfica y Catastral del Estado de Campeche (INFOCAM) who is in charge of urban planning policies, along with the production of supporting studies and analysis. INFOCAM also undertakes the development and implementation of an online platform guaranteeing the availability of statistical, territorial and cadastral information.

The ESCI project aims at providing useful tools to the city of Campeche as it develops its urban code. EO4SD products are meant to enhance the urban growth and vulnerability assessment currently taking place in the city as part of IADB's Emerging and Sustainable Cities program.

The initiative's methodology relies on the use of several studies describing the city's reality as it enters the ESCI programme.

The first stage consists in a general diagnosis of the city and its sustainability, conducted in a concerted manner with local, state and national authorities to establish the baseline understanding of how the city has developed and what efforts are currently undertaken. This is a critical step in the establishement of a coherent Action Plan for the city, in order to make a proposal that takes into account pre-existing plans and the strategies to carry them out.

Once the indicators and critical sectors are highlighted, the following filters are applied, allowing the ICES team to have a clearer vision of which sector should be addressed immediately:

- Public Opinion
- Economic Impact
- Climate Change
- Expert opinion of the Bank.

Prioritized issues sectors are synthetized in a high-level Action Plan, here the *Campeche Sostenible Plan de Accion*, written in 2015 by the municipality and a number of bank partners[1].

The second stage of the ESCI methodology consists in the execution plan, starting with a Pre-investment phase of studies aiming at evaluating the feasibility level and the financial resources regarding the implementation of the measures described in the Action Plan. Support of an existant or creation of a new citizen monitoring system is also implemented early on, in order to monitor the city's progress in terms of sustainability.

The analysis of the 117 basic indicators of the ICES pinpointed the critical themes to be addressed in the case of Campeche:

- Sanitation and drainage
- Urban inequality
- Participatory public management
- Modern public management

In addition, the geospatial products from EO4SD-Urban can be useful for the IADB regarding internal analytics improving the level of urban information to counterparts and the Bank itself, knowledge of utilities, better and cheaper techniques, homogeneous data and rather difficult to obtain, quick assessment of the land use evolution over time.

2.2 Service Area Specification

So far, no internationally accepted definition for the term "Urban Area" and the related Core and Peri-Urban areas exists. Different initiatives are currently trying to address a standardised approach for defining the terms "Urban Area". During discussions with the GPSC coordinator it was considered important to use a uniform definition for the GPSC cities in order for the cities to exchange information and share products/experiences and conduct potential comparative studies.

In this context, it was decided to use an international approach for the demarcation of the Area of Interest (AOI) for mapping the GPSC cities in terms of Core Urban area and Peri-Urban area. Thus,

1 On line at http://www.ccpy.gob.mx/agenda-campeche/campeche-sostenible.php ; in Spanish.

the approach is based on the European Union's Directorate-General for Regional and Urban Policy (DG REGIO) method and the definitions are described in the Regional Working Paper 2014 from the European Commission on "A harmonised definition of cities and rural areas: the new degree of urbanisation" (European Commission, 2014). Following the naming of the DG Regio approach, the Urban Core is named as "High Density Core" and the Peri-Urban area is termed as "Urban Cluster". Within the DG REGIO approach, the High Density Core area is defined as contiguous grid cells of 1 km^2 with a density of at least 1 500 inhabitants per km^2 and a minimum population of 50 000. The Urban Cluster is defined as clusters of contiguous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5 000.

The DG REGIO methodology used in the EO4SD-Urban project was slightly adjusted to Non-European countries. For the first three GPSC cities (namely Bhopal, Vijayawada and Saint-Louis) produced within the project the Global Human Settlement Population (GHSP) grid with a spatial resolution of 1 km were used for the classification into "High Density Core" and "Urban Cluster". The raster dataset is available for the years 1975, 1990, 2000, 2015. This dataset depicts the distribution and density of population, expressed as the number of people per cell. The data can be downloaded under following link http://data.jrc.ec.europa.eu/dataset/jrc-ghsl-ghs_pop_gpw4_globe_r2015a.

In 2019, a higher resolution population layer (spatial resolution of 10 m) produced by the German Aerospace Centre (DLR) became available. The AOIs for the remaining GPSC cities (namely Melaka, Abidjan, Dakar and Campeche) were produced based on the DLR population layer.

The High Density Core AOI for a city is created by merging the contiguous grid cells of 1 km^2 with a density of at least 1500 inhabitants per km^2 and a minimum population of 50 000. In the definition of the High Density Core the contiguity is only allowed via a vertical or horizontal connection. In a next step, gaps are filled. Due to the coarse resolution of the population grid cells additional grid cells were in a last step added for under estimated settlement areas. The same was done for over estimations, here grid cells were removed. The GHSP layer can be directly used for the calculation, while the DLR population has to be aggregated to a resolution of 1 km^2 before being used for the AOI definition. In this aggregation step, each output cell contains the sum of the input cells that are encompassed by the extent of that new cell.

The Urban Cluster is created very similar to the High Density Core. Continuous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5 000 are merged together to form the Urban Cluster. The contiguity within the Urban Cluster can also be diagonal. After gaps are filled, areas, which were over or under estimated by the population grid were removed or added to the AOI. The GHSP layer was directly used, the DLR population layer had to undergo an aggregation step in order to reduce the spatial resolution to 1 km^2.

For Bhopal and Vijayawada a buffer of 1 km was calculated around the High Density Core AOI and the Urban Cluster AOI to smoothen the border of the AOIs.

In all remaining GPSC cities, the border was not smoothed, but when the population grid was under or over estimating the real settlement extent, grid cells were added or removed.

In some cases, the city counterparts requested that the AOIs for the High Density Core and the Urban Cluster follow the municipal or administrative boundary of the city. In this case, the municipal/administrative boundary was used, but enlarged in areas where the AOI created according to the adjusted DG Regio approach was bigger. This adjustment of the DG Regio AOI was done for Melaka, Abidjan, Dakar and Campeche. These further adjusted DG Regio AOIs are in the following report named as Core City Area and Larger Urban Area (see Figure 1).

A more detailed description on how the AOIs are calculated is provided in Annex 1.

The AOI were presented in a power point, and sent to the Users for verification. Figure 1 shows the created AOIs after combining the DG Regio AOIs with the municipal/administrative boundaries of the cities.

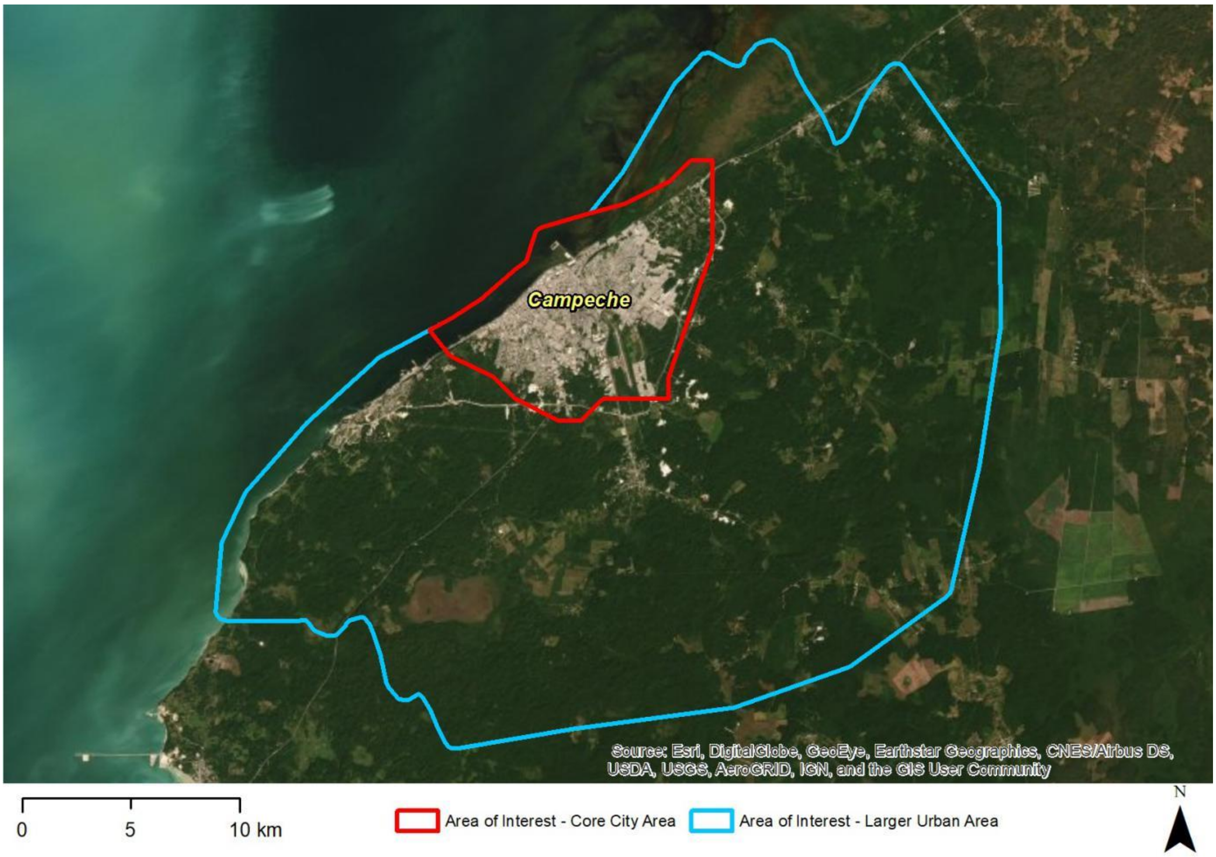

Figure 1: Illustration of Core City and Larger Urban Areas of Campeche.

The Core City has an area of 88 km^2 and the Larger Urban has an area of 741 km^2.

2.3 Product List and Product Specifications

During the discussions related to the AOIs the potential geo-spatial products that could be provided for the Cities were also reviewed with the WB Team and Users. It was noted that the Baseline Land Use/Land Cover (LU/LC) products (for the Core and Peri-Urban areas) were a standard product that would be provided for all Cities as it is required for the derived products. In the case of Campeche, the full list of products for both the Core and Peri-Urban areas are as follows:

- Settlement Extent & Change (producer: DLR)
- Percentage Impervious Surface & Change (producer: DLR)
- Urban Land Use / Land Cover (LU/LC) & Change (producer: SIRS)
- Urban Green Areas & Change (producer: NEO)

The first two products have been generated by the German Aerospace Agency (DLR) over the full metropolitan area for seven reference years: 1985 - 1990 - 1995 - 2000 - 2005 - 2010 - 2015.

For each of the other products two time slots were used to provide historic and recent information. In detail, LU/LC and Urban Green Areas products have been generated for 2006 and 2018 depending on the EO data availability. The current Report will focus on the provision of the Baseline LU/LC product, the Settlement Extent and Impervious Surface product, as well as the Urban Green Area product for Campeche.

2.4 Land Use/Land Cover Nomenclature

A pre-cursor to starting production was the establishment with the stakeholders of the relevant Land Use/Land Cover (LU/LC) nomenclature as well as class definitions. The approach taken was to use a standard remote sensing based LU/LC nomenclature i.e. the European Urban Atlas Nomenclature (European Union, 2011) and adapt it to the User's LU requirements. Thus, the remote-sensing based LU/LC classes in the urban context can be grouped into five Level 1 classes, which are Artificial Surfaces, Natural/ Semi Natural Areas, Agricultural Areas, Wetlands, and Water. These classes can then be sub-divided into several different more detailed classes such that the dis-aggregation can be down to Level 2-4. This hierarchical classification system is often used in operational Urban mapping programmes and is the basis for example of the European Commission's Urban Atlas programme which provides pan-European comparable LU/LC data with regular updates. A depiction of the way the levels and classes in the Urban Atlas programme are structured is presented as follows:

Level I Artificial surfaces

- Level II: Urban Fabric

Level III

- Continuous Urban Fabric (Sealing Layer-S.L. > 80%)
- Discontinuous Urban Fabric (S.L. 10% - 80%)
  Level IV
  - Discontinuous Dense Urban Fabric (S.L. 50% - 80%)
  - Discontinuous Medium Density Urban Fabric (S.L. 30% - 50%)
  - Discontinuous Low Density Urban Fabric (S.L. 10% - 30%)
  - Discontinuous Very Low Density Urban Fabric (S.L. < 10%)

- Level II: Industrial, Commercial, Public, Military, Private Units and Transport

Level III

- Industrial, Commercial, Public, Military and Private Units
- Transport Infrastructure
  Level IV
  - Fast Transit Roads
  - Other Roads
  - Railway
- Port and associated land
- Airport and associated land

- Level II: Mine, Dump and Construction Sites

Level III

- Mineral Extraction and Dump Sites
- Construction Sites
- Land Without Current Use

- Level II: Artificial Non-Agricultural Vegetated Areas

Level III

- Green Urban Areas
- Sports and Leisure Facilities

(Reference: European Union, 2011)

It should be noted that in the current project, the Level 1 classes were used as the basis for classification of the Urban Cluster areas using the High Resolution (HR) data such as Landsat or

Sentinel. However, for the High Density Core areas using the Very High Resolution (VHR) data it was possible to go down to Level III and IV.

The different levels, classes and sub-classes from the remote sensing based urban classification, were harmonised within the GPSC cities. The following tables give the nomenclature for the High Density Core and the Urban Cluster region (see Table 1 and Table 2).

Table 1: LU/LC Nomenclature for GPSC Cities (High Density AOI).

Actual and Historic Nomenclature High Density Core			
Level I	**Level II**	**Level III**	**Level IV**
1000 Artificial Surfaces	1100 Residential	1110 Continuous Urban Fabric (80 - 100 % Sealed)	
		1120 Discontinuous Urban Fabric	1121 Discontinuous dense urban fabric (50 - 80 % Sealed)
			1122 Discontinuous medium density urban fabric (30 - 50 % Sealed)
			1123 Discontinuous low density urban fabric (10 - 30 % Sealed)
			1124 Discontinuous very low density urban fabric (0 - 10 % Sealed)
	1200 Industrial, Commercial, Public, Military, Private Units and Transport	1210 Industrial, Commercial, Public, Military and Private Units	
		1220 Transport Infrastructure	1221 Arterial Roads
			1222 Collector Roads
			1223 Railway
		1230 Port Area	
		1240 Airport	
	1300 Mine, Dump and Construction Sites	1310 Mineral Extraction and Dump Sites	
		1330 Construction Sites	
		1340 Land Without Current Use	
	1400 Artificial Non-Agricultural Vegetated Areas	1410 Green Urban Areas	
		1420 Sports and Leisure Facilities	
2000 Agricultural Area			
3000 Natural and Semi-natural Areas	3100 Forest and Shrublands		
	3200 Natural Areas (Grassland)		
	3300 Bare Soil		
4000 Wetlands			
5000 Water	5100 Inland Water		
	5200 Marine Water		

Table 2: LU/LC Nomenclature for GPSC Cities (Urban Cluster Area).

Actual and Historic Nomenclature Urban Cluster Area			
Level I	**Level II**	**Level III**	**Level IV**
1000 Artificial Surfaces			
2000 Agricultural Area			
3000 Natural and Semi-natural Areas	3100 Forest and Shrublands		
	3200 Natural Areas (Grassland)		
	3300 Bare Soil		
4000 Wetlands			
5000 Water	5100 Inland Water		
	5200 Marine Water		

It is important to note that the possibility to classify at Level IV is highly dependent on the availability of reliable reference datasets from the City or sources such as Google Earth. This aspect is further discussed in Chapter 3.

Especially regarding the road hierarchy used in the classification at Level IV, the international road classification standards have been followed; this is for example defined by the European Commission (https://ec.europa.eu/transport/road_safety/specialist/-knowledge/road/designing_for_road_function/road_classification_en).

Roads are divided into three groups: arterial or through traffic flow routes (in our case **Arterial Roads**), distributor roads (in our case **Collector Roads**), and access roads (or **Local Roads**). The three road types are defined as follows:

Arterial Roads:

Roads with a flow function allow efficient throughput of (long distance) motorized traffic. All motorways and express roads as well as some urban ring roads have a flow function. The number of access and exit points is limited. (https://ec.europa.eu/transport/road_safety/specialist/knowledge/-road/designing_for_road_function/road_classification_en)

Collector Roads:

Roads with an area distributor function allow entering and leaving residential areas, recreational areas, industrial zones, and rural settlements with scattered destinations. Junctions are for traffic exchange (allowing changes in direction etc.); road sections between junctions should facilitate traffic in flowing.
(https://ec.europa.eu/transport/road_safety/specialist/knowledge/road/designing_for_road_function/road_classification_en)

Local Roads:

Roads with an access function allow actual access to properties alongside a road or street. Both junctions and the road sections between them are for traffic exchange. (https://ec.europa.eu/transport/-road_safety/specialist/knowledge/road/designing_for_road_function/road_classification_en).

2.5 World Settlement Extent

Reliably outlining settlements is of high importance since an accurate characterization of their extent is fundamental for accurately estimating, among others, the population distribution, the use of resources (e.g. soil, energy, water, and materials), infrastructure and transport needs, socioeconomic development, human health and food security. Moreover, monitoring the change in the extent of settlements over time is of great support for properly modelling the temporal evolution of urbanization and thus, better estimating future trends and implementing suitable planning strategies.

At present, no standard exists for defining settlements and worldwide almost each country applies its own definition either based on population, administrative or geometrical criteria. The German Space Agency (DLR) were responsible for the provision of the "Settlement Extent" product; when generating the settlement extent maps from HR imagery, pixels are labelled as **settlement** if they *intersect any building, lot or – just within urbanized areas – roads and paved surface* where we define:

- **building** as any structure having a roof supported by columns or walls and intended for the shelter, housing, or enclosure of any individual, animal, process, equipment, goods, or materials of any kind;
- **lot** as the area contained within an enclosure (wall, fence, hedge) surrounding a building or a group of buildings. In cases where there are many concentric enclosures around a building, the lot is considered to stop at the inner most enclosure;
- **road** as any long, narrow stretch with a smoothed or paved surface, made for traveling by motor vehicle, carriage, etc., between two or more points;
- **paved surface** as any level horizontal surface covered with paving material (i.e., asphalt, concrete, concrete pavers, or bricks but excluding gravel, crushed rock, and similar materials).

Instead, pixels not satisfying this condition are marked as **non-settlement**.

The settlement extent product is a binary mask outlining - in the given Area of Interest (AOI) – settlements in contrast to all other land-cover classes merged together into a single information class. The settlement class and the non-settlement class are associated with values "255" and "0", respectively.

2.6 Percentage Impervious Surface

Settlement growth is associated not only to the construction of new buildings, but – more in general – to a consistent increase of all the impervious surfaces (hence also including roads, parking lots, squares, pavement, etc.), which do not allow water to penetrate, forcing it to run off. To effectively map the percentage impervious surface (PIS) is then of high importance being it related to the risk of urban floods, the urban heat island phenomenon as well as the reduction of ecological productivity. Moreover, monitoring the change in the PIS over time is of great support for understanding, together with information about the spatiotemporal settlement extent evolution, also more details about the type of urbanization occurred (e.g. if areas with sparse buildings have been replaced by highly impervious densely built-up areas or vice-versa).

In the framework of the EO4SD-Urban project, one pixel in the generated PIS maps is associated with the estimated percentage of the corresponding surface at the ground covered by buildings or paved surfaces, are defined as:

- **building** as any structure having a roof supported by columns or walls and intended for the shelter, housing, or enclosure of any individual, animal, process, equipment, goods, or materials of any kind;
- **paved surface** as any level horizontal surface covered with paving material (i.e. asphalt, concrete, concrete pavers, or bricks but excluding gravel, crushed rock, and similar materials).

The product provides for each pixel in the considered AOI the estimated PIS. Specifically, values are integer and range from 0 (no impervious surface in the given pixel) to 100 (completely impervious surface in the given pixel) with step 5.

2.7 Urban Green Areas Nomenclature

Developing cities in a sustainable way implies to preserve and promote green areas also and especially within the urban extent. Green areas refer to any surfaces covered by vegetation (grass, bushes, trees).

Table 3: Nomenclature used for the mapping and identification of Urban Green Areas.

Single date		
Code 0	Non-urban green area	
Code 1	Urban green area	
Code 255	Non-urban areas. All areas that do not fall in "Artificial Surfaces" Level 1 class of the Land Use Land Cover product (See Table 1).	
Change product		
Code 0	Non-urban green area. No vegetated surfaces occurring on "Artificial Surfaces", Level I, at both points in time.	
Code 1	Permanent urban green area. Vegetated surfaces in historic and recent year.	
Code 2	Loss of urban green area. Vegetated areas in historic year, which changed to non-vegetated areas in recent year.	
Code 3	New urban green area. Non-vegetated surfaces in historic year with vegetation cover in recent year.	
Code 255	Non-Urban Areas. All areas that do not fall in "Artificial Surfaces" Level 1 class of the Land Use Land Cover product.	

2.8 Terms of Access

The Dissemination of the digital data and the Report was undertaken via FTP.

3 Service Operations

The following Sections present all steps of the service operations including the necessary input data, the processing methods, the accuracy assessment and the Quality Control procedures. Methods are presented in a top-level and standardised manner for all the EO4SD-Urban City Reports.

3.1 Source Data

This Section presents a summary of the remote sensing and ancillary datasets that were used. Different types of data from several data providers have been acquired. A complete list of source data as well as a quality assessment is provided in Annex 2.

High Resolution Optical EO Data

The major data sources for the peri-urban current and historic mapping of urban LULC, urban extent and imperviousness were Landsat and Sentinel-2 data which were accessible and downloadable free of charge.

- **Sentinel-2:** The most recent data coverage comprises one Sentinel-2 data set from the 6th of August 2018. The data was downloaded and processed at Level 1C.
- **Landsat 7:** As a source of historical data two scenes of Landsat TM 7 from the 3rd of March 2006 and from the 20th of April 2006 have been acquired which cover the whole area of interest.

Very High Resolution Optical EO Data

The VHR data for the core urban area mapping had to be acquired and purchased through commercial EO Data Providers such as Airbus Defence and European Space Imaging.

It has to be noted that under the current collaboration project the VHR EO data had to be purchased under **mono-license agreemen**ts between GAF AG and the EO Data Providers. If EO data would have to be distributed to other stakeholders then further licences for multiple users would have to be purchased.

The following VHR sensor data have been acquired to cover the defined Areas of Interest:

- **Pléiades-1A:**
  - 1 scene from August 2018
- **Quickbird-2:**
  - 2 scenes, from August 2004 and January 2006

Ancillary Data

- Open Street Map (OSM) data: OSM data is freely available and generated by volunteers across the globe. The so called crowd sourced data is not always complete, but has for the most parts of the world valuable spatial information. Data was downloaded to complement the Transport Network layer and further enhanced. The spatial location of the OSM based streets was used a geospatial reference.

Detailed lists of the used EO and ancillary data as well as their quality is documented in the attached Quality Control Sheets in Annex 2.

3.2 Processing Methods

Data processing starts at an initial stage with quality checks and verification of all incoming data. This assessment is performed in order to guarantee the correctness of data before geometric or radiometric pre-processing is continued. These checks follow defined procedures in order to detect anomalies, artefacts and inconsistencies. Furthermore, all image and statistical data were visualised and interpreted by operators.

The main techniques and standards used for data analysis, processing and modelling for each product are described in Annex 1.

3.3 Accuracy Assessment of Map Products

Data and maps derived from remote sensing contain - like any other map - uncertainties which can be caused by many factors. The components, which might have an influence on the quality of the maps derived from EO include quality and suitability of satellite data, interoperability of different sensors, radiometric and geometric processing, cartographic and thematic standards, and image interpretation procedures, post-processing of the map products and finally the availability and quality of reference data. However, the accuracy of map products has a major impact on secondary products and its utility and therefore an accuracy assessment was considered as a critical component of the entire production and products delivery process. The main goal of the thematic accuracy assessment was to guarantee the quality of the mapping products with reference to the accuracy thresholds set by the user requirements.

The applied accuracy assessments were based on the use of reference data and applying statistical sampling to deduce estimates of error in the classifications. In order to provide an efficient, reliable and robust method to implement an accuracy assessment, there are three major components that had to be defined: the **sampling design**, which determines the spatial location of the reference data, the **response design** that describes how the reference data is obtained and an **analysis design** that defines the accuracy estimates. These steps were undertaken in a harmonised manner for the validation of all the geo-spatial products.

3.3.1 Accuracy Assessment of LU/LC Product

Sampling Design

The sampling design specifies the sample size, sample allocation and the reference assessment units (i.e. pixels or image blocks). Generally, different sampling schemes can be used in collecting accuracy assessment data including: simple random sampling, systematic sampling, stratified random sampling, cluster sampling, and stratified systematic unaligned sampling. In the current project a **single stage stratified random sampling** based on the method described by Olofson et al (2013[2]) was applied which used the map product as the basis for stratification. This ensured that all classes, even very minor ones were included in the sample.

The sampling design is applied separately for the High Density and for the Urban Cluster classification.

2 Olofsson, P., Foody, G. M., Stehman, S. V., & Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment, 129, 122–131. doi:10.1016/j.rse.2012.10.031

In the complex LU/LC product with many classes, this usually results in a large number of strata (one stratum per LU/LC classes), of which some classes cover only very small areas (e.g. sport fields, cemeteries) and not being adequately represented in the sampling. In order to achieve a representative sampling for the statistical analyses of the mapping accuracy it was decided to extend the single stage stratified random sampling. Slightly different approaches were used for the High Density and the Urban Cluster classification.
The first step is the same for both classifications: the number of required samples is allocated within each of the Level I strata (1000 Artificial Surfaces, 2000 Agricultural Area, 3000 Natural and Semi-natural Areas, 4000 Wetlands, 5000 Water).
In the second step, all Level III classes that were not covered by the first sampling were grouped into one new stratum for the High Density classification. For the Urban Cluster classification all Level II classes that were not covered by the first sampling were grouped into one new stratum.

Within that stratum the same number of samples was randomly allocated as the Level I strata was received. To avoid a clustering of point samples within classes and to minimise the effect of spatial autocorrelation a minimum distance in between the sample points was set to be 150 m. The final sample size for each class can be considered to be as close as possible to the proportion of the area covered by each stratum considering that the target was to determine the overall accuracy of the entire map.

The total sample size per stratum was determined by the expected standard error and the estimated error rate based on the following formula, which assumes a simple random sampling (i.e. the stratification is not considered):

$$n \frac{P*q}{\left(\frac{E}{z}\right)^2}$$

n = number of samples per strata / map class
p = expected accuracy
q = 1 – *p*
E = Level of acceptable (allowable) sample error
Z = z-value (the given level of significance)

Hence, with an expected accuracy of *p* = 0.85, a 95% confidence level and an acceptable sampling error of 5%, the minimum sample size is 196. A 10% oversampling was applied to compensate for stratification inefficiencies and potentially inadequate samples (e.g. in case of cloudy or shady reference data). For each Level I strata 215 samples have been randomly allocated. Afterwards, for all classes of Level III of the High Density classification that did not received samples in the first run, additionally 215 samples were randomly drawn across all these classes. A summary of the number of sample point for each High Density class is given in Table 4.

The same applies for the Urban Cluster classification: All Level II classes that did not receive samples in the first run, additionally 215 samples were randomly drawn across all these classes. A summary of the number of sample point for each High Density class is given in Table 5.

The main difference of the sampling design for the two areas is that the resampling is done at Level III for the High Density areas and at Level II for the Urban Cluster areas.

Table 4: Number of sampling points for the EO4SD-Urban mapping classes after applied sampling design with information on overall land cover by class.

Class Name	Class ID	No. of Sampling Points	Area coverage (sqkm)
Continuous Urban Fabric (80% -100% Sealed)	1110	126	15.1
Discontinuous Urban fabric	1120	77	10.3
Industrial, Commercial, Public, Military and Private Units	1210	45	5.5
Transport Infrastructure	1220	19	2.6
Port area	1230	2	0.02
Airport	1240	31	3.7
Mineral Extraction and Dump Sites	1310	8	1.0
Construction Sites	1330	7	0.8
Land without current use	1340	10	1.3
Green Urban Areas	1410	31	3.4
Sport and Leisure Facilities	1420	8	1.3
Agricultural Area	2000	5	0.6
Forest and Shrublands	3100	185	23.1
Natural Areas (Grassland)	3200	28	3.4
Bare Soil	3300	2	0.05
Wetlands	4000	15	1.8
Inland Water	5100	3	0.4
Marine Water	5200	107	13.3
Total	--	**709**	**87.6 sqkm**

Table 5: Number of sampling points for the Urban Cluster classes after applied sampling design with information on overall land cover by class.

Class Name	Class ID	No. Of Sampling Points	Area coverage (sqkm)
Artificial Surfaces	1000	215	71.8
Agriculture	2000	215	93.7
Forest and Shrublands	3100	193	457.6
Natural Areas	3200	18	45.0
Bare Soil	3300	4	12.2
Wetlands	4000	215	18.8
Inland Water	5100	4	0.6
Marine Water	5200	211	40.8
Total	-	**1075**	**740.5 sqkm**

Response Design

The response design determines the reference information for comparing the map labels to the reference labels. Collecting reference data on the ground by means of intensive fieldwork is both costly and time consuming and in most projects not feasible. The most cost effective reference data sources are VHR satellite data with 0.5 m to 1 m spatial resolution. Czaplewski (2003)[3] indicated that visual interpretation of EO data is acceptable if the spatial resolution of EO data is sufficiently better

3 Czaplewski, R. L. (2003). Chapter 5: accuracy assessment of maps of forest condition: statistical design and methodological considerations, pp. 115–140. In Michael A.Wulder, & Steven E. Franklin (Eds.), Remote sensing of forest environments: concepts and case studies. Boston: Kluwer Academic Publishers (515 pp.).

compared to the thematic classification system. However, if there are no EO data with better spatial resolution available, the assessment results need to be checked against the imagery used in the production process.

The calculated number of necessary sampling points for each mapping category was randomly distributed among the strata and overlaid onto the two LULC mapping products. The following two Figures (see Figure 2 and Figure 3) show the mapping results with the overlaid sample points.

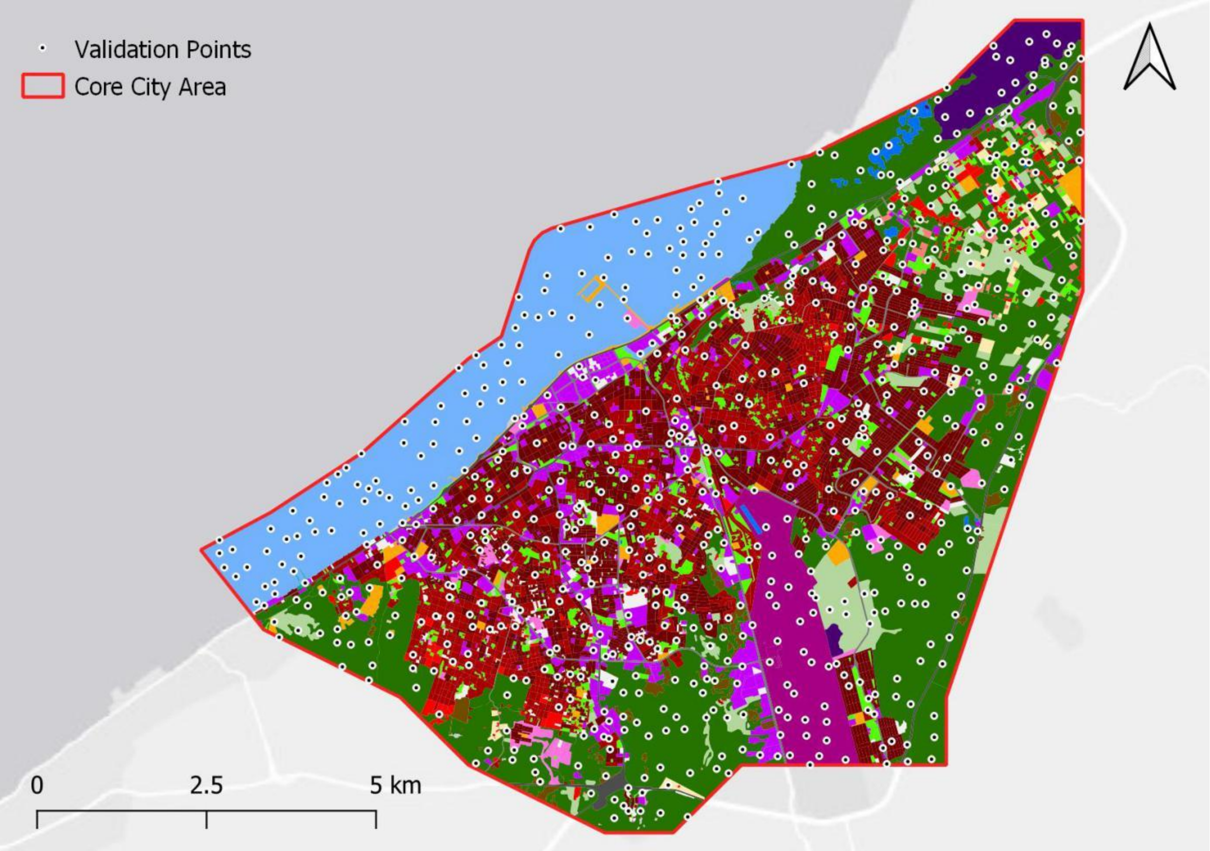

Figure 2 : Mapping result of the Core City area of Campeche of the year 2018 overlaid with randomly distributed sample points used for accuracy assessment.

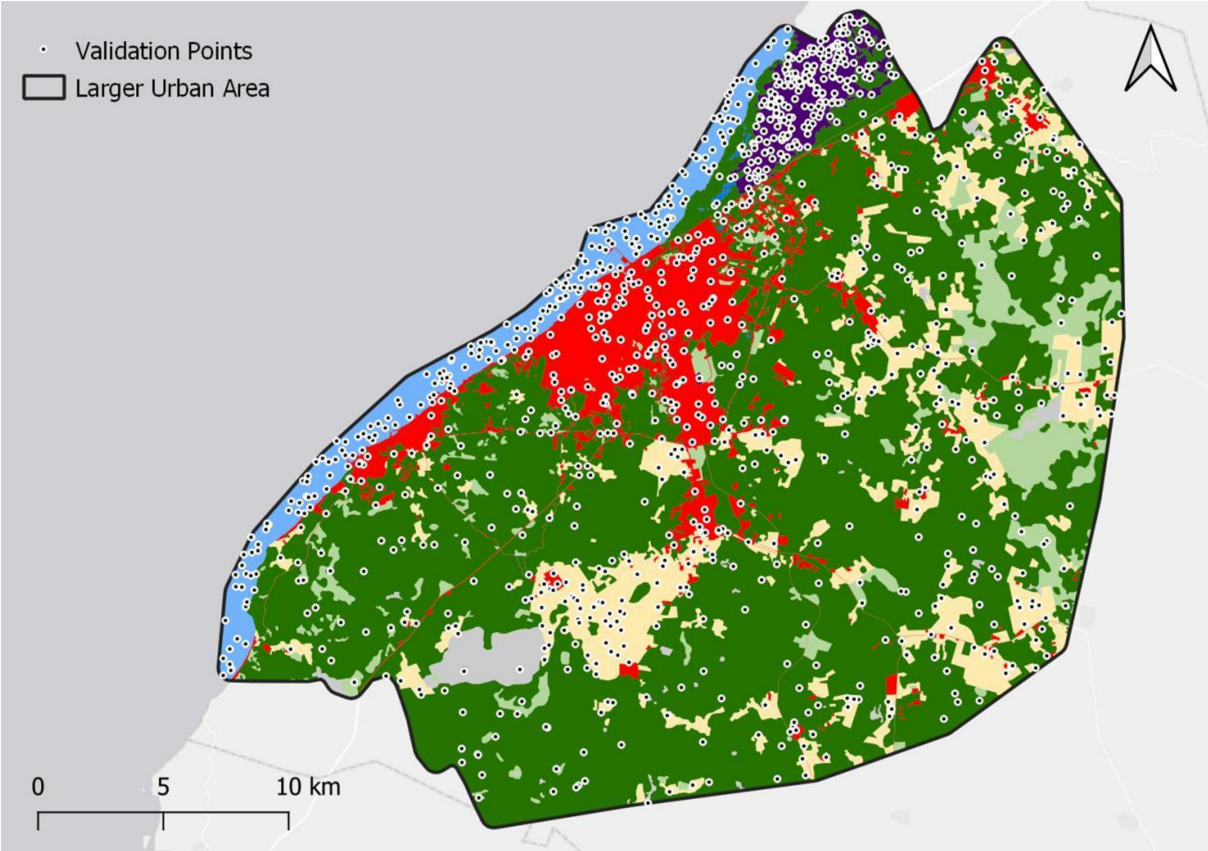

Figure 3: Mapping result of the Larger Urban area of Campeche of the year 2018 overlaid with randomly distributed sample points used for accuracy assessment.

In this way a reference information could be extracted for each sample point by visual interpretation of the VHR data for all mapped classes. The size of the area to be observed had to be related to the Minimum Mapping Unit (MMU) of the map product to be assessed. The reference information of each sampling point was compared with the mapping results and the numbers of correctly and not-correctly classified observations were recorded for each class. From this information the specific error matrices and statistics were computed (see next Section).

Analysis

Each class usually has errors of both omission and commission, and in most situations, these errors for a class are not equal. In order to calculate these errors as well as the uncertainties (confidence intervals) for the area of each class a statistically sound accuracy assessment was implemented.

The confusion matrix is a common and effective way to represent quantitative errors in a categorical map, especially for maps derived from remote sensing data. The matrices for each assessment epoch were generated by comparing the "reference" information of the samples with their corresponding classes on the map. The *Reference* represented the "truth", while the *Map* provided the data obtained from the map result. Thematic accuracy for each class and overall accuracy is then presented in error matrices (see Tables below). Unequal sampling intensity resulting from the random 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:

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

Where *i* and *j* are the columns and rows in the matrix, M is the total number of possible units (population) and π is the sampling intensity for a given sample unit *u* in stratum *h*.

Overall accuracy and User and producer accuracy were computed for all thematic classes and 95% confidence intervals were calculated for each accuracy metric.

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:

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

In which w_h is the proportion of the total area covered by each stratum. The 95% Confidence Interval (CI) is +/- 1.96*σ.

Results

The confusion matrices are provided within the Annex 2 and showing the mapping error for each relevant class. For each class the number of samples which are correctly and not correctly classified are listed, which allows the calculation of the user and producer accuracies for each class as well as the confidence interval at 95% confidence levels based on the formulae above.

The Land Use/Land Cover product for Campeche in 2018 in Core City has an overall mapping accuracy of 95.85% with a CI ranging from 94.38% to 97.32% at a 95% CI.
The Larger Urban Area has and overall accuracy of 97.93% with a CI ranging from 97.08% to 98.78% at a 95% CI. The specific class accuracies for Core City and Larger Urban products are given in Annex 2.

3.3.2 Accuracy Assessment of World Settlement Extent Product

In the following, the strategy designed for validating the World Settlement Extent (WSE) or World Settlement Footprint (WSF) 2015, i.e. a global settlement extent layer obtained as a mosaic of ~18.000 tiles of 1x1 degree size where the same technique employed in the EO4SD-Urban project is presented. In particular, specific details are given for all protocols adopted for each of the accuracy assessment components, namely response design, sampling design, and analysis; final results are discussed afterwards. In the light of the quality and amount of validation points considered, it can be reasonably assumed that the corresponding quality assessment figures are also representative for any settlement extent map generated in the framework of EO4SD-Urban.

Response Design

The response design encompasses all steps of the protocol that lead to a decision regarding agreement of the reference and map classifications. The four major features of the response design are the source of information used to determine the source of reference data, the spatial unit, the labelling protocol for the reference classification, and a definition of agreement.

- **Source of Reference Data:** Google Earth (GE) satellite/aerial VHR imagery has been used given its free access and the availability for all the project test sites in the period 2014-2015. In particular, GE automatically displays the latest available data, but it allows to browse in time over all past historical images. The spatial resolution varies depending on the specific data source; in the case of SPOT imagery it is ~1.5m, for Digital Globe's WorldView-1/2 series, GeoEye-1, and Airbus' Pleiades it is in the order of ~0.5m resolution, whereas for airborne data (mostly available for North America, Europe and Japan) it is about 0.15m.
- **Spatial Assessment Unit:** A 3x3 block spatial assessment unit composed of 9 cells of 10x10m size has been used. Specifically, this choice is justified one the one hand by the fact that input data with different spatial resolutions have been used to generate the WSF2015 (i.e. 30m Landsat-8 and 10m S1). On the other hand, GE imagery exhibited in some cases a mis-registration error of the order of 10-15m, hence using a 3x3 block allows defining an agreement e.g. based on statistics computed over 9 pixels, thus reducing the impact of such shift.
- **Reference Labelling Protocol:** For each spatial assessment block any cell is finally labelled as **settlement** if *it intersects any building, lot or – just within settlements – roads and paved surface*. Instead, pixels not satisfying this condition are marked as **non-settlement**.
- **Definition of Agreement:** Given the classification and the reference labels derived as described above, three different agreement criteria have been defined:
  1) for each pixel, positive agreement occurs only for matching labels between the classification and the reference;
  2) for each block, a majority rule is applied over the corresponding 9 pixels of both the classification and the reference; if the final labels match, then the agreement is positive;
  3) for the classification a majority rule is applied over each assessment block, while for the reference each block is labelled as “settlement” only in the case it contains at least one pixel marked as “settlement”; if the final labels match, then the agreement is positive.

Crowd-sourcing was performed internally at Google. In particular, by means of an ad-hoc tool, operators have been iteratively prompted a given cell on top of the available Google Earth reference VHR scene closest in time to the year 2015 and given the possibility of assigning to each cell a label among: “building”, “lot”, “road/paved surface” and “other”. For training the operators, a representative set of 100 reference grids was prepared in collaboration between Google and DLR.

Sampling Design

The stratified random sampling design has been applied since it satisfies the basic accuracy assessment objectives and most of the desirable design criteria. In particular, stratified random sampling is a probability sampling design and it is one of the easier to implement; indeed, it involves first the division of the population into strata within which random sampling is performed afterwards. To include a representative population of settlement patterns, 50 out of the ~18.000 tiles of 1x1 degree size considered in the generation of the WSF2015 have been selected based on the ratio between the number of estimated settlements (i.e. disjoint clusters of pixels categorized as settlement in the WSF2015) and their area. In particular, the *i*-th selected tile has been chosen randomly among those whose ratio belongs to the interval $P_{2(i-1)}; P_{2i}, i \in [1; 50] \subset N$ (where P_x denotes the *x*-th percentile of the ratio).

Table 6: Accuracies exhibited by the WSF2015 according to the three considered agreement criteria for different definitions of settlement.

Settlement =	Accuracy Measure	Agreement Criterion 1		Agreement Criterion 2		Agreement Criterion 3	
buildings	*OA%*	86.96		87.86		91.15	
	AA%	88.57		90.35		88.91	
	Kappa	0.6071		0.6369		0.7658	
	$UA_{NS}\%$ - $UA_S\%$	98.11	54.69	98.73	56.76	94.84	80.58
	$PA_{NS}\%$ - $PA_S\%$	86.24	90.90	86.72	93.98	93.32	84.51
buildings + lots	*OA*	88.08		88.94		91.26	
	AA%	88.64		90.19		88.71	
	Kappa	0.6510		0.6784		**0.7716**	
	$UA_{NS}\%$ - $UA_S\%$	97.54	60.71	98.13	62.66	94.29	82.62
	$PA_{NS}\%$ - $PA_S\%$	87.79	89.49	88.26	92.12	93.95	83.48
buildings + lots + roads / paved surface	*OA*	88.77		90.09		88.51	
	AA%	86.34		88.28		84.27	
	Kappa	**0.6938**		**0.7317**		0.7219	
	$UA_{NS}\%$ - $UA_S\%$	94.49	72.20	95.35	75.06	88.13	89.60
	$PA_{NS}\%$ - $PA_S\%$	90.78	81.91	91.62	84.94	96.04	72.51

As the settlement class covers a sensibly small proportion of area compared to the merger of all other non-settlement classes (~1% of Earth's emerged surface), an equal allocation reduces the standard error of its class-specific accuracy. Moreover, such an approach allows to best address user's accuracy estimation, which corresponds to the map "reliability" and is indicative of the probability that a pixel classified on the map actually represents the corresponding category on the ground. Accordingly, in this framework for each of the 50 selected tiles we randomly extracted 1000 settlement and 1000 non-settlement samples from the WSF2015 and used these as center cells of the 3x3 reference block assessment units to label by photointerpretation. Such a strategy resulted in an overall amount of $(1000 + 1000) \times 9 \times 50 = 900.000$ cells labelled by the crowd.

Analysis

As measures for assessing the accuracy of the settlement extent maps, we considered:

- the percentage overall accuracy *OA%*;
- the Kappa coefficient;
- the percentage producer's ($PA_S\%$, $PA_{NS}\%$) and user's ($UA_S\%$, $UA_{NS}\%$) accuracies for both the settlement and non-settlement class;
- the percentage average accuracy *AA%* (i.e., the average between $PA_S\%$ and $PA_{NS}\%$).

Results

Table 6 reports the accuracies exhibited by the WSF2015 according to the three considered agreement criteria for different definitions of settlement; specifically, we considered as "settlement" all areas covered by: i) buildings; ii) buildings or building lots; or iii) buildings, building lots or roads / paved surfaces. As one can notice, accuracies are always particularly high, thus confirming the effectiveness of the employed approach and the reliability of the final settlement extent maps. The best performances in terms of kappa are obtained when considering settlements as composed by buildings, building lots and roads / paved surfaces for criteria 1 and 2 (i.e., 0.6938 and 0.7317, respectively) and by buildings and building lots for criteria 3 (0.7716); the OA% follows a similar trend. This is in line with the adopted settlement definition. Moreover, agreement criteria 3 results in accuracies particularly high with respect to criteria 1 and 2 when considering as settlement just buildings or the combination of buildings and lots. This can be explained by the fact that when the detection is mainly driven by Landsat data then the whole 3x3 assessment unit tends to be labelled as settlement if a building or a lot intersect the corresponding 30m resolution pixel.

3.3.3 Accuracy Assessment of the Percentage Impervious Surface Product

In the following, the strategy designed for validating the PIS product is presented; specifically, details are given for all protocols adopted for each of the accuracy assessment components, namely response design, sampling design, and analysis. Results are discussed afterwards.

Response Design

The response design encompasses all steps of the protocol that lead to a decision regarding agreement of the reference and map classifications. The four major features of the response design are the source of information used to determine the source of reference data, the spatial unit, the labelling protocol for the reference classification, and a definition of agreement.

- **Source of Reference Data:** Cloud-free VHR multi-spectral imagery (Visible + Near Infrared) acquired at 2m spatial resolution (or higher) covering a portion of the AOI for which the Landsat-based PIS product has been generated;
- **Spatial Assessment Unit:** A 30x30m size unit has been chosen according to the spatial resolution of the Landsat imagery employed to generate the PIS product;
- **Reference Labelling Protocol:** As a first step, the NDVI is computed for each VHR scene followed by a manual identification of the most suitable threshold that allows to exclude all the vegetated areas (i.e. non-impervious). Then, the resulting mask is refined by extensive photointerpretation.
- **Definition of Agreement:** The above-mentioned masks are aggregated at 30m spatial resolution and compared per-pixel with the resulting VHR-based reference PIS to the corresponding portion of the Landsat-based PIS product.

Sampling Design

The entirety of pixels covered by the available VHR imagery over the given AOI is employed for assessing the quality of the Landsat-based PIS product.

Analysis

As measures for assessing the accuracy of the PIS maps, following indices are computed:

- the *Pearson's Correlation coefficient*: it measures the strength of the linear relationship between two variables and it is defined as the covariance of the two variables divided by the product of their standard deviations; in particular, it is largely employed in the literature for validating the output of regression models;
- The *Mean Error (ME)*: it is calculated as the difference between the estimated value (i.e., the Landsat-based PIS) and the reference value (i.e., the VHR-based reference PIS) averaged over all the pixels of the image;
- The *Mean Absolute Error (MAE)*: it is calculated as the absolute difference between the estimated value (i.e., the Landsat-based PIS) and the reference value (i.e., the VHR-based reference) averaged over all the pixels of the image.

Results

To assess the effectiveness of the method developed to generate the PIS maps, its performances over 5 test sites is analysed (i.e. Antwerp, Helsinki, London, Madrid and Milan) by means of WorldView-2 (WV2) scenes acquired in 2013-2014 at 2m spatial resolution. In particular, given the spatial detail offered by WV2 imagery, it was possible to delineate with a very high degree of confidence all the buildings and other impervious surfaces included in the different investigated areas. Details about acquisition date and size are reported in Table 7, along with the overall number of final 30x30m validation samples derived for the validation exercise. Such a task demanded a lot of manual interactions and transferring it to other AOIs would require extensive efforts; however, it can be reasonably assumed that the final quality assessment figures (computed on the basis of more than 1.9 million validation samples) shall be considered representative also for PIS maps generated in the framework of EO4SD-Urban. Table 7 reports the quantitative results of the comparison between the PIS maps generated using Landsat-7/8 data acquired in 2013-2014 and the WV2-based reference PIS maps. In particular, the considered approach allowed to obtain a mean correlation of 0.8271 and average ME and MAE equal to -0.09 and 13.33, respectively, hence assessing the great effectiveness of the Landsat-based PIS products. However, it is worth also pointing out that due to the different acquisition geometries, WV2 and LS8 images generally exhibit a very small shift. Nevertheless, despite limited, such displacement often results in a one-pixel shift between the Landsat-based PIS and the WV2-based reference PIS aggregated at 30m resolution. This somehow affects the computation of the MAE and of the correlation coefficient (which however yet resulted in highly satisfactory values). Instead, the bias does not alter the ME, which always exhibited values close to 0, thus confirming the capabilities of the technique and the reliability of the final products.

Table 7: Acquisition dates and size of the WV2 images available for the 5 test sites analysed in the validation exercise along with the number of corresponding 30x30m validation samples.

	Acquisition Date [DD.MM.YYYY]	Original Size [2x2m pixel]	Validation Samples [30x30m unit]
Antwerp	31.07.2014	5404 x 7844	188.280
Helsinki	21.04.2014	12468 x 9323	516.882
London	28.08.2013	7992 x 8832	313.937
Madrid	20.12.2013	10094 x 13105	588.202
Milan	14.05.2014	8418 x 7957	297.330

3.3.4 Accuracy Assessment of Urban Green Areas Product

The validation of the Green Area mapping results is done in a similar way as the validation for the Land Use Land Cover product. The necessary amount of sampling points is calculated according to the formula of Goodchild et al. (1994), which is given in Table 8.

Table 8: Calculation of the minimum number of samples according Goodchild et al. (1994).

Variables	Values
p	0.85
q	0.15
E	0.05
z	1.96
$n = \frac{p * q}{(E / z)^2}$	**196**
n with 10% oversampling	**215**
with:	
p = required accuracy of the data	
q = 1-p	
E = Level of acceptable (allowable) sample error	
Z = value from table (for the given level of significance)	

The calculated number of 215 sample points was randomly distributed among the entire map and overlaid on the VHR data of each epoch. The following Figure (see Figure 4) shows the mapping result with the overlaid sample points.

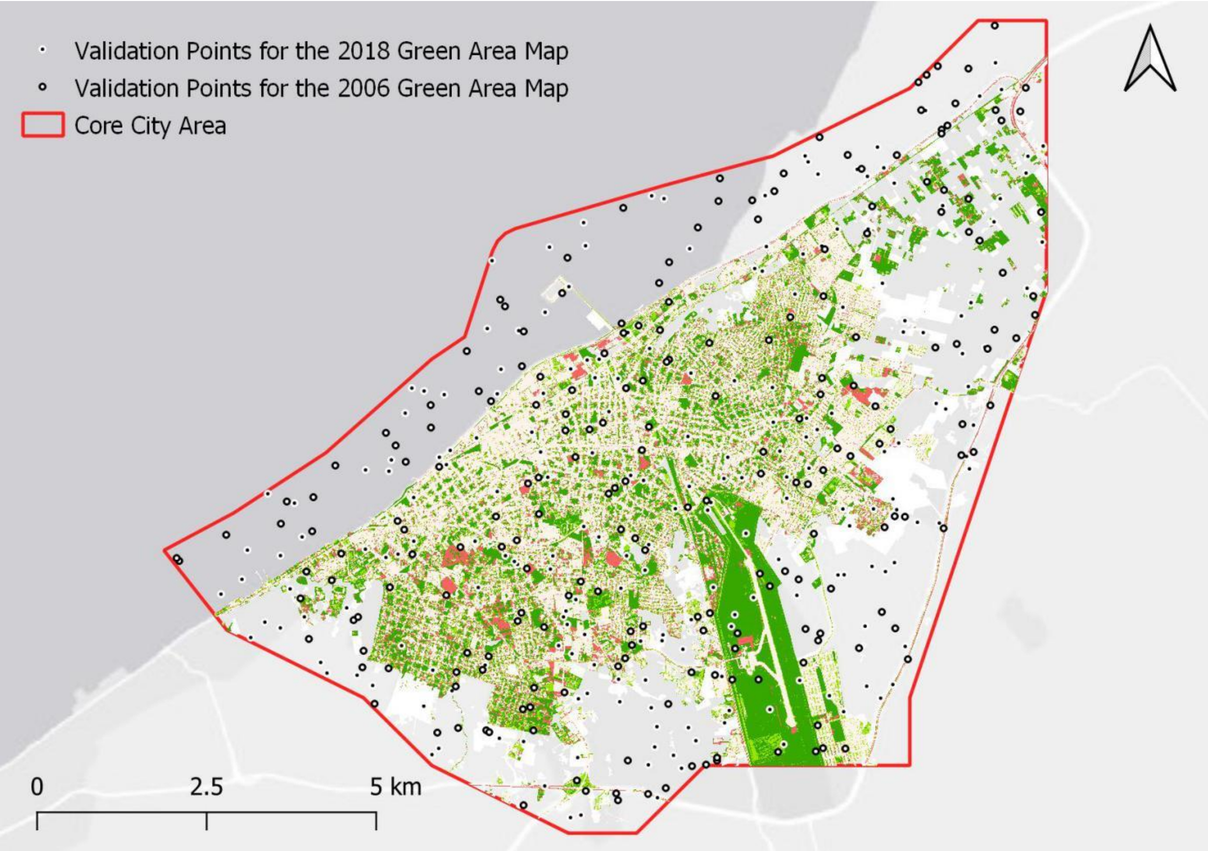

Figure 4: Result of the Urban Green Area mapping in Campeche (change product) with sampling points used for product validation.

At each sample point location the reference data was collected by visual interpretation of the VHR data. The size of the area to be observed had to be related to the Minimum Mapping Unit (MMU) of the map product to be assessed. Finally, visual interpreted land cover type was compared with the mapping results and the numbers of correctly and not-correctly classified observations were recorded. From this information the specific error matrices and statistics were computed.

The confusion matrices show the mapping error for each relevant class. For each class the number of samples which are correctly and not correctly classified are listed in the Tables below. They allow the calculation of the user and producer accuracies for each class as well as the confidence interval at 95% confidence levels based on the formulae above. The results of the Accuracy Assessment are listed in Table 9 and Table 10 below, for 2006 and 2018 respectively.

Table 9: Results of the Accuracy Assessment of Urban Green Areas in Campeche, 2006.

Overall Accuracy: 97.7 %.

Urban Green 2006	Reference Data		Totals
	0 - Non-Urban Green Area	1 - Urban Green Area	
0 - Non-Urban Green Area	168	2	170
1 - Urban Green Area	3	42	45
Totals	171	44	**215**

Table 10: Results of the Accuracy Assessment of Urban Green Areas in Campeche, 2018.

Overall Accuracy : 97.21 %.

Urban Green 2018	**Reference Data**		**Totals**
	0 - Non-Urban Green Area	1 - Urban Green Area	
0 - Non-Urban Green Area	158	3	161
1 - Urban Green Area	3	51	54
Totals	161	54	**215**

The confusion matrices are additionally provided within the Quality Control documentation in Annex 2 and showing the mapping error for each relevant class. For each class the number of samples, which are correctly and not correctly classified, are listed, which allows the calculation of the user and producer accuracies for each class as well as the confidence interval at 95% confidence levels.

3.4 Quality Control/Assurance

A detailed Quality Control and Quality Assurance (QC/QA) system has been developed which records and documents all quality relevant processes ranging from the agreed product requirements, the different types of input data and their quality as well as the subsequent processing and accuracy assessment steps. The main goal of the QC/QA procedures was the verification of the completeness, logical consistency, geometric and thematic accuracy and that metadata are following ISO standards on geographic data quality and INSPIRE data specifications. These assessments were recorded in Data Quality Sheets which are provided in Annex 2. The QC/QA procedures were based on an assessment of a series of relevant data elements and processing steps which are part of the categories listed below:

- Product requirements;
- Specifications of input data: EO data, in-situ data, ancillary data;
- Data quality checks: EO data quality, in-situ data quality, ancillary data quality;
- Geometric correction & accuracy, data fusion (if applicable), data processing;
- Thematic processing: classification, plausibility checks;
- Accuracy: thematic accuracy, error matrices
- Delivery checks: completeness, compliancy with requirements

After each intermediate processing step a QC/QA was performed to evaluate products appropriateness for the subsequent processing (see Figure 5).

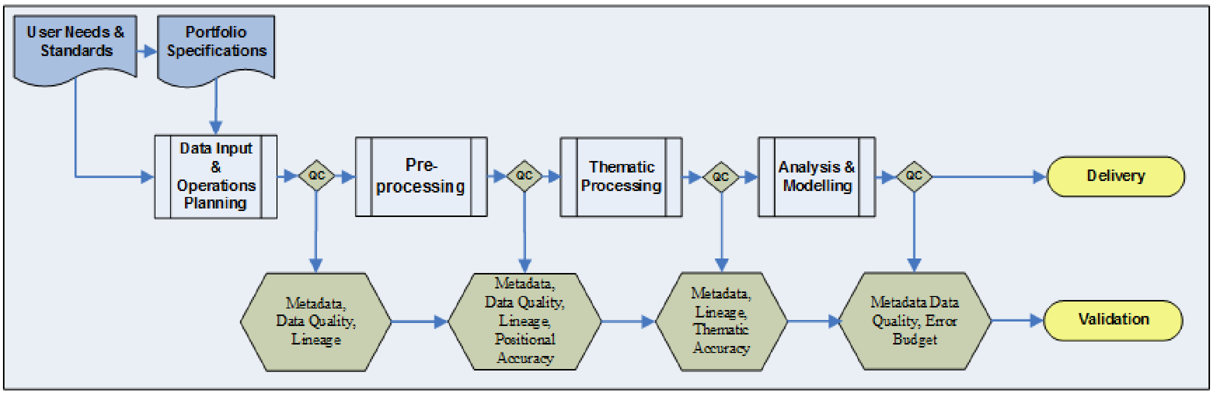

Figure 5: Quality Control process for EO4SD-Urban product generation. At each intermediate processing step output properties are compared against pre-defined requirements.

After the initial definition of the product specifications (output) necessary input data were defined and acquired. Input data include all satellite data and reference data e.g. in-situ data, reference maps, topographic data, relevant studies, existing standards and specifications, statistics. These input data were the baseline for the subsequent processing and therefore all input data had to be checked for **completeness**, **accuracy** and **consistency**. The evaluation of the quality of input data provides confidence of their suitability for further use (e.g. comparison with actual data) in the subsequent processing line. Data processing towards the end-product required multiple intermediate processing steps. To guarantee a traceable and quality assured map production the QC/QA assessment was performed and documented by personnel responsible for the Quality Control/Assurance. The results of all relevant steps provided information of the acceptance status of a dataset/product.

The documentation is furthermore important to provide a comprehensive and transparent summary of each production step and the changes made to the input data. With this information the user will be able to evaluate the provided services and products. Especially the accuracy assessment of map products and the related error matrices are highly important to rate the quality and compare map products from different service providers.

3.5 Metadata

Metadata provides additional information about the delivered products to enable it to be better understood. In the current project a harmonised approach to provide metadata in a standardised format applicable to all products and end-users was adopted. Metadata are provided as XML files, compliant to the ISO standard 19115 "Metadata" and ISO 19139 "XML Scheme Implementation". The metadata files have been created and validated by the GIS/IP-operator for each map product with the Infrastructure for Spatial Information in Europe (INSPIRE) Metadata Editor available at: http://inspire-geoportal.ec.europa.eu/editor/.

The European Community enacted a Directive in 2007 for the creation of a common geo-data infrastructure to provide a consistent metadata scheme for geospatial services and products that could be used not only in Europe but globally. The geospatial infrastructure called INSPIRE was built in a close relation to existing International Organization for Standardization (ISO) standards. These are ISO 19115, ISO 19119 and ISO 15836. The primary incentive of INSPIRE is to facilitate the use and sharing of spatial information by providing key elements and guidelines for the creation of metadata for geospatial products and services.

The INSPIRE Metadata provides a core set of metadata elements which are part of all the delivered geo-spatial products to the users. Furthermore, the metadata elements provide elements that are necessary to perform queries, store and relocate data in an efficient manner. The minimum required information is specified in the Commission Regulation (EC) No 1205/2008 of 3 December 2008 and contains 10 elements:

- Information on overall Product in terms of: Point of contact for product generation, date of creation
- Identification of Product: Resource title, Abstract (a short description of product) and Locator
- Classification of Spatial Data
- Keywords (that define the product)
- Geographic information: Area Coverage of the Product
- Temporal Reference: Temporal extent; date of publication; date of last revision; date of creation
- Quality and Validity: Lineage, spatial resolution
- Conformity: degree of conformance to specifications
- Data access constraints or Limitations
- Responsible party: contact details and role of contact group/person

These elements (not exhaustive) constitute the core information that has to be provided to meet the minimum requirements for Metadata compliancy. Each element and its sub-categories or elements have specific definitions; for example, in the element “Quality” there is a component called “Lineage” which has a specific definition as follows: “a statement on process history and/or overall quality of the spatial data set. Where appropriate it may include a statement whether the data set has been validated or quality assured, whether it is the official version (if multiple versions exist), and whether it has legal validity. The value domain of this element is free text,” (INSPIRE Metadata Technical Guidelines, 2013). The detailed information on the Metadata elements and their definitions can be found in the “INSPIRE Metadata Implementing Rules: Technical Guidelines,” (2013). Each of the EO4SD-Urban products will be accompanied by such a descriptive metadata file. It should be noted that the internal use of metadata in these institutions might not be established at an operational level, but the file format (*.xml) and the web accessibility of data viewers enable for the full utility of the metadata.

4 Analysis of Mapping Results

This Chapter presents and assesses all results that were produced within the framework of the current project, and provides the results of some standard analytics undertaken with these products including the following:

- Settlement Extent – Developments from 1985, 1990, 1995, 2005, 2010 to 2015
- Land Use / Land Cover - Status and Trends between 2006 and 2018
- Urban Green Areas - Status and Change between 2006 and 2018

These analytics provide information on general trends and developments in the AOIs which can then be further interpreted and used by Urban planners and the City Authorities for city planning.

It should be noted that all digital data sets for these products are provided with this City Report, along with all the related metadata and Quality Control documentation.

4.1 Settlement Extent – Developments 1985, 1990, 1995, 2000, 2005, 2010 and 2015

The Urban or World Settlement Extent (WSE) product in the EO4SD-Urban project is provided by the German Aerospace Centre (DLR) for 7 distinct years. This product and its accuracy were described in Sections 2.5 and 3.3.2. In the current project, the Urban Extent product for Campeche was first used to assess historical developments from 1985-2015 (see Figure 6 and Figure 7). Further analysis, aiming at assessing urbanization extent patterns based on administrative units, can be achieved by overlaying administrative boundaries.

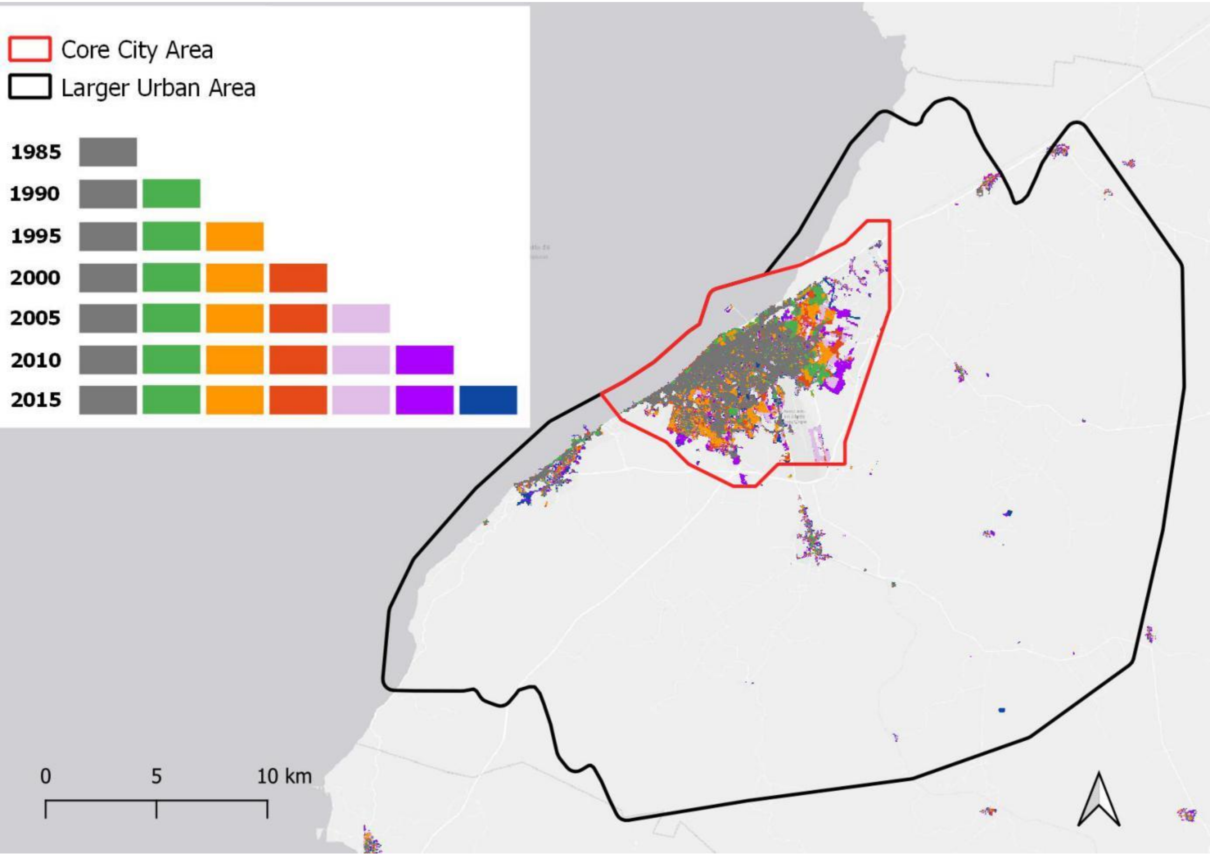

Figure 6: Settlement Extent developments in the epochs 1985 to 1990, 1990 to 1995, 1995 to 2000, 2000 to 2005, 2005 to 2010 and 2010 to 2015 in Campeche and surrounding region.

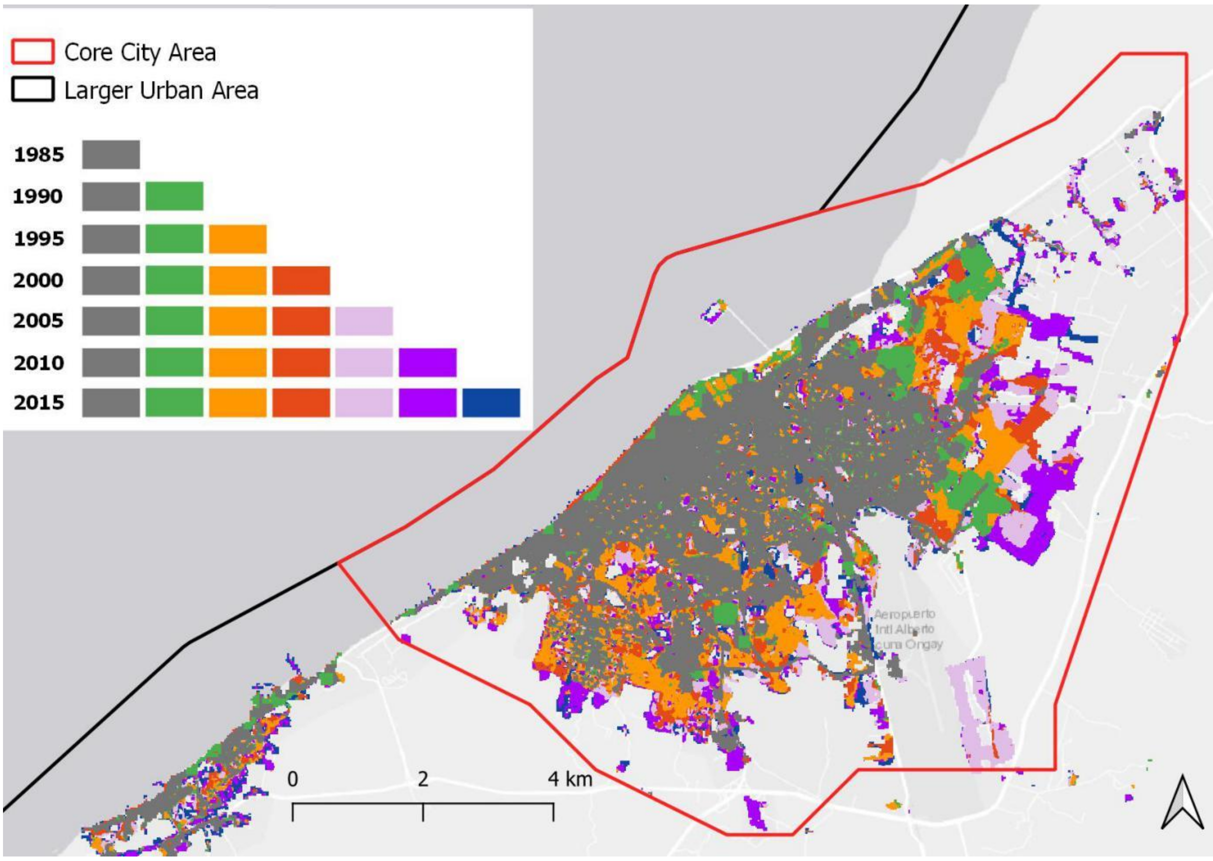

Figure 7: Settlement Extent developments in the epochs 1985 to 1990, 1990 to 1995, 1995 to 2000, 2000 to 2005, 2005 to 2010 and 2010 to 2015 in Campeche within the High Density Area.

4.2 Land Use / Land Cover for 2006 and 2018

This Section presents the results of the LU/LC mapping for 2006 and 2018 as well the statistical information on the changes between these two epochs.

4.2.1 LU/LC Mapping for Core City Area

The LU/LC map generated for 2018 reference year over the Core City Area is depicted in Figure 8, and a more detailed vue is provided through Figure 9. A cartographic version of the map layout is provided as a PDF file in addition to the geo-spatial product.

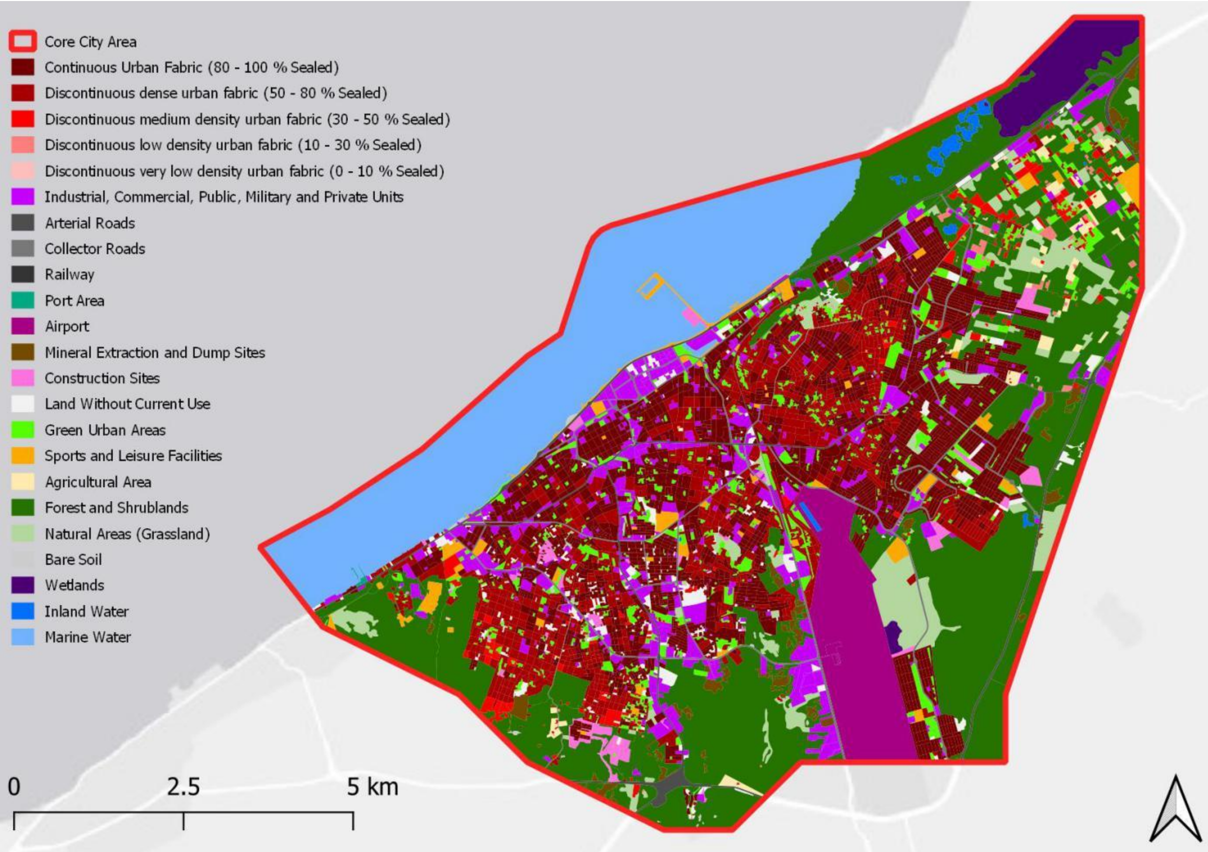

Figure 8: Core City Area - Detailed Land Use Land Cover 2018 over Campeche.

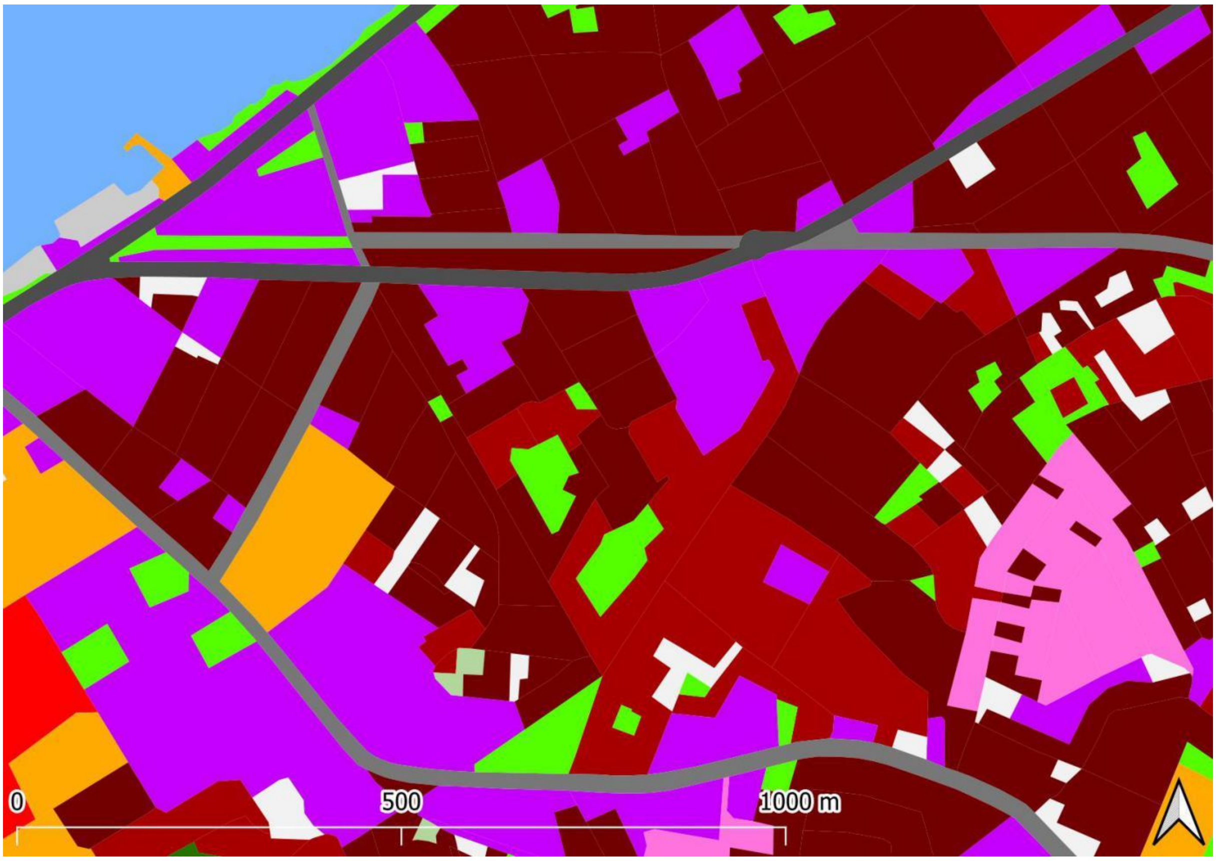

Figure 9: Core City Area - Insight on the detailed Land Use Land Cover 2018 inside the city.

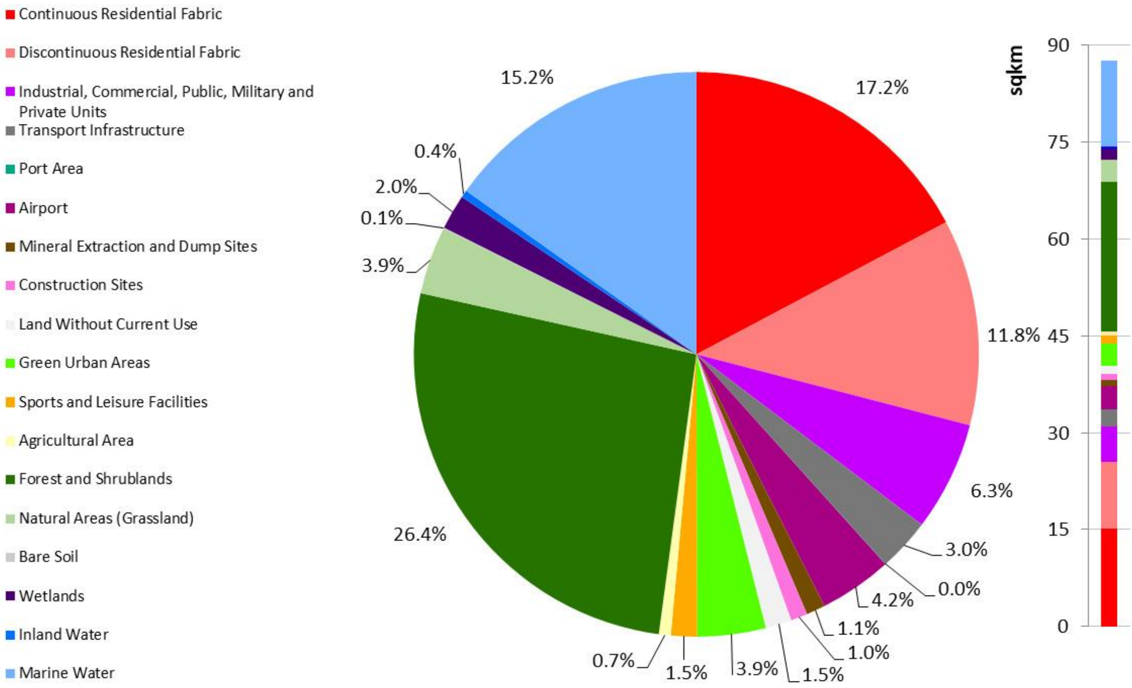

Figure 10: Core City Area - Detailed Land Use Land Cover 2018 structure presented in % (left) and km² (right).

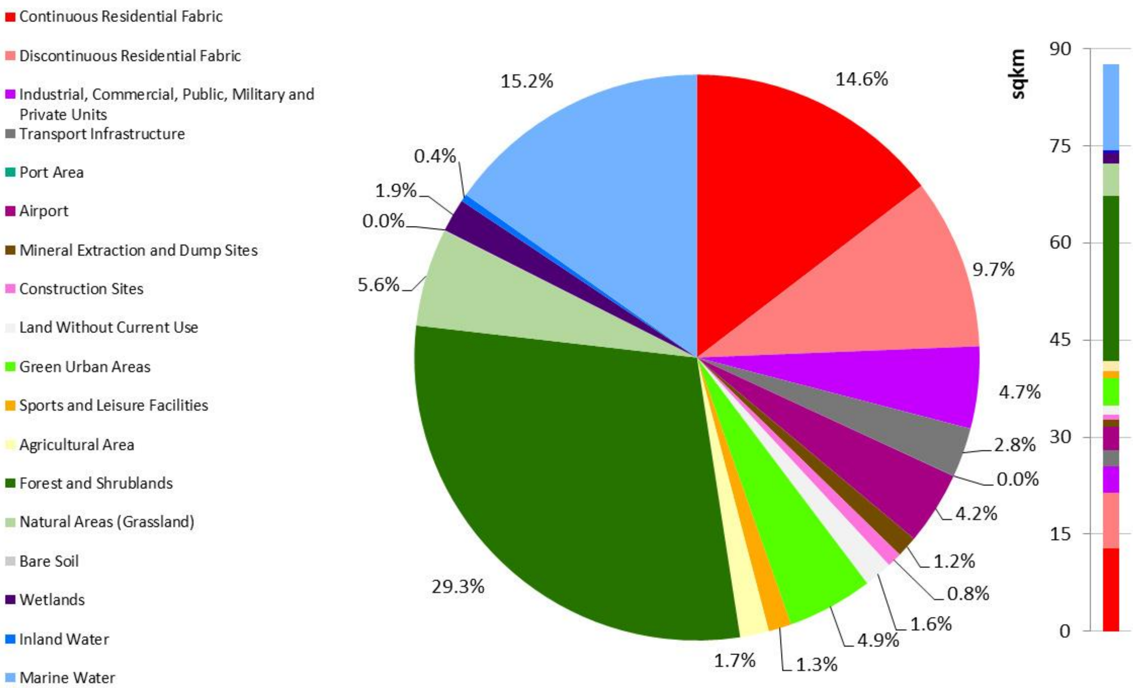

Figure 11: Core City Area - Detailed Land Use Land Cover 2006 structure presented in % (left) and km² (right).

Figure 10 and Figure 11 provide detailed information on the class disaggregation and area coverage for the two epochs, respectively 2006 and 2018. Even within the Core City, forests and shrublands are a very important class in terms of spatial extent, with a total of 26% of the surface in 2018. It represented about 30% in 2006, and mainly disappeared in favor of the expansion of the residential fabric (24% in 2006, 29% of the territory in 2018). The Urban Greeneries represented about 5% in 2006, but due to the densification inside the City these areas lost 1% between 2006 and 2018. The Ingeniero Alberto Acuña Ongay International Airport consumes just over of 4% of the Core City territory.

Table 11: Core City Area - Detailed information on area and percentage of total area for each class for 2006 and 2018 as well as the changes.

LU/LC Classes	2018		2006		Change		Change per Year	
	sqkm	% of total	sqkm	% of total	sqkm	%	sqkm	%
1110 - Continuous Urban Fabric (80 - 100 % Sealed)	15.11	17.25%	12.81	14.62%	2.30	18.0%	0.19	1.5%
1121 - Discontinuous dense urban fabric (50 - 80 % Sealed)	8.90	10.16%	7.56	8.63%	1.34	17.8%	0.11	1.5%
1122 - Discontinuous medium density urban fabric (30 - 50 % Sealed)	1.13	1.29%	0.79	0.90%	0.34	43.1%	0.03	3.6%
1123 - Discontinuous low density urban fabric (10 - 30 % Sealed)	0.30	0.34%	0.18	0.21%	0.11	63.1%	0.01	5.3%
1124 - Discontinuous very low density urban fabric (0 - 10 % Sealed)	0.00	0.00%	0.01	0.01%	0.00	-17.3%	0.00	-1.4%
1210 - Industrial, Commercial, Public, Military and Private Units	5.49	6.26%	4.10	4.69%	1.38	33.7%	0.12	2.8%
1221 - Arterial Roads	1.21	1.38%	1.17	1.34%	0.03	2.8%	0.00	0.2%
1222 - Collector Roads	1.32	1.51%	1.19	1.36%	0.13	11.1%	0.01	0.9%
1223 - Railway	0.11	0.13%	0.11	0.13%	0.00	0.1%	0.00	0.0%
1230 - Port Area	0.02	0.02%	0.01	0.02%	0.00	16.9%	0.00	1.4%
1240 - Airport	3.66	4.18%	3.66	4.18%	0.00	0.0%	0.00	0.0%
1310 - Mineral Extraction and Dump Sites	0.95	1.09%	1.06	1.20%	-0.10	-9.9%	-0.01	-0.8%
1330 - Construction Sites	0.84	0.96%	0.73	0.84%	0.11	14.5%	0.01	1.2%
1340 - Land Without Current Use	1.31	1.49%	1.40	1.60%	-0.10	-6.9%	-0.01	-0.6%
1410 - Green Urban Areas	3.43	3.92%	4.29	4.89%	-0.85	-19.9%	-0.07	-1.7%
1420 - Sports and Leisure Facilities	1.30	1.48%	1.14	1.31%	0.16	13.6%	0.01	1.1%
2000 - Agricultural Area	0.60	0.68%	1.45	1.65%	-0.85	-58.7%	-0.07	-4.9%
3100 - Forest and Shrublands	23.09	26.35%	25.63	29.26%	-2.54	-9.9%	-0.21	-0.8%
3200 - Natural Areas (Grassland)	3.40	3.88%	4.89	5.58%	-1.49	-30.5%	-0.12	-2.5%
3300 - Bare Soil	0.05	0.05%	0.04	0.05%	0.00	6.7%	0.00	0.6%
4000 - Wetlands	1.75	2.00%	1.67	1.90%	0.08	5.1%	0.01	0.4%
5100 - Inland Water	0.37	0.42%	0.37	0.42%	0.00	0.2%	0.00	0.0%
5200 - Marine Water	13.29	15.17%	13.35	15.23%	-0.06	-0.5%	-0.01	0.0%
Total	**87.61**	**100%**	**87.61**	**100%**	-	-	-	-

In addition to the overall LU/LC classification for the two epochs it is interesting to further assess the different trends between classes over the 12-year time gap. The quantitative figures for each class are first provided in Table 11 as an overview. The next Section provides more detailed analysis of the LU/LC changes between the two epochs.

The classes with the total areas that increased the most are the residential urban fabric, mainly with constructions and expansion in south-east and south-west of the city. As a very probable consequence, the Natural areas registered a loss of 30% of their surfaces, and so did the urban greeneries (loss of 20%). Agricultural surfaces on the east of the Core City are also used for the expansion of the urban area, going from 1.45 sqkm to 0.60 sqkm.

4.2.2 Spatial Distribution of Main LU/LC Change Categories for Core City Area

In order to better analyze the growth trend and the spatial distribution of changes, meaningful aggregations of the LU/LC classes in both epochs were used. The following categories were chosen:

- Urban Residential Expansion: all changes from Non-Residential classes to a Residential one;
- Other Urban Land Use Expansion: all changes from Non-Urban or Urban Residential classes to Other Urban Land Use classes;
- Urban to Agricultural or Natural/Semi-Natural Areas: all changes from artificial classes returning to non-urban classes (agricultural areas, forest, wetlands or water)
- Natural or Semi-Natural to Agricultural Areas: all changes from non-urban to agricultural areas
- Agricultural to Natural or Semi-Natural Areas: all changes from agricultural to natural or semi-natural classes
- Changes within Natural and Semi-Natural Areas: all changes in between the natural and semi-natural classes (e.g. Forest into Agriculture).

Overall analysis of these aggregated categories for both epochs, 2006 and 2018, is depicted in Figure 12.

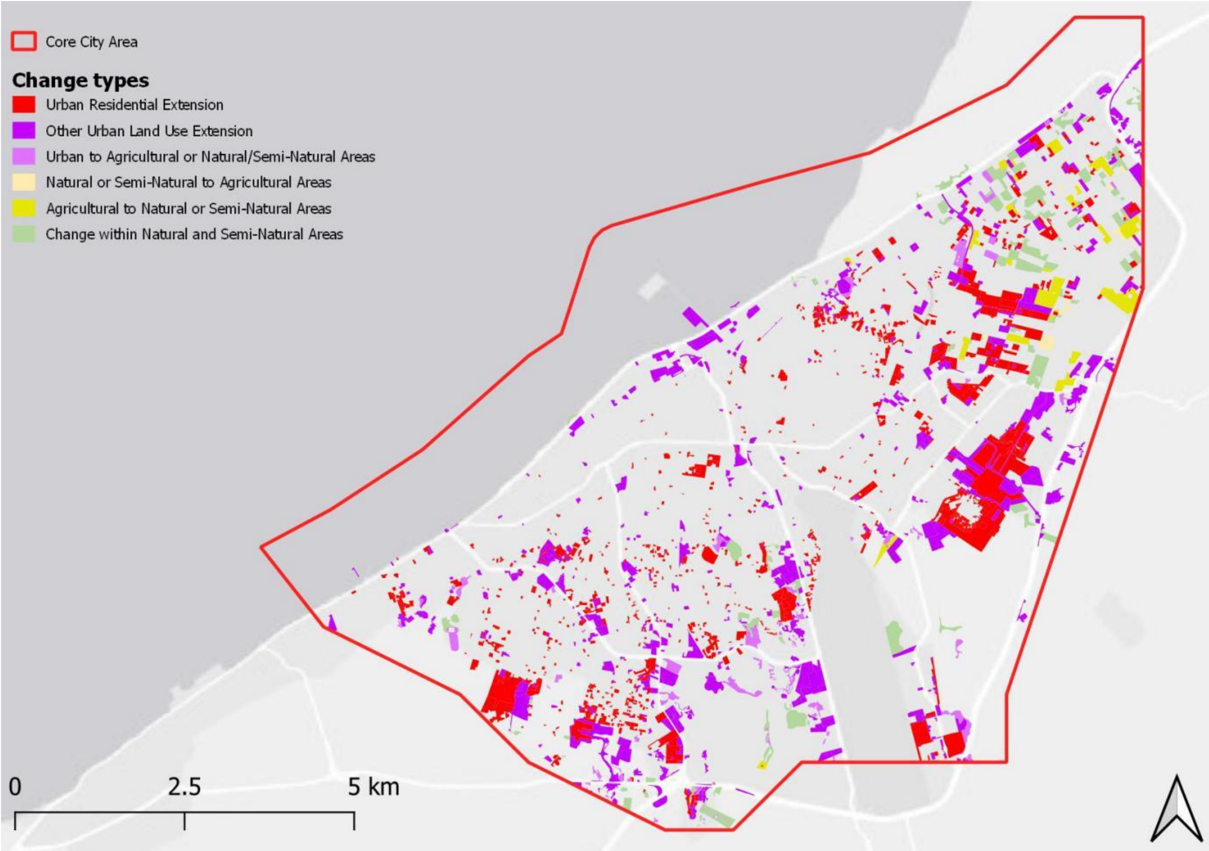

Figure 12: Core City Area - Land Use Land Cover change types and spatial distribution.

The spatial distribution of the change types confirms the analysis in the way that urban densification is almost non-existent and urban residential expansion quite important. It can be summarized as important units scattered throughout this very urban territory, especially on the south-east and the south-west of the Core City territory. Otherwise, this change mapping highlights the location and extent of the land where agriculture or natural areas have given way to human activities other than residential ones, mainly around the International Airport and the proximity of the coast. Important changes can also be noted on the eastern part, inside the natural and semi-naturals areas, with agricultural areas used as orchards or tree nurseries abandoned and returning to forest or simple shrublands.

The statistics of the Change categories are presented in Figure 13 and Table 12.

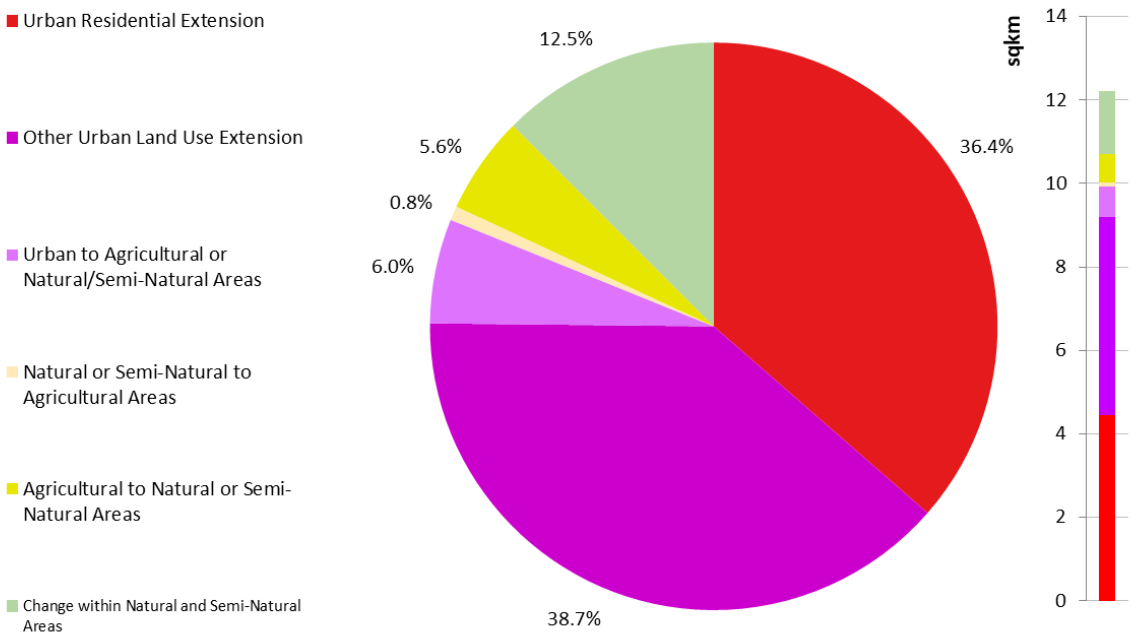

Figure 13: Core City Area - Land Use Land Cover Change Types 2006-2018 presented in % (left) and km² (right).

The quantitative results also reveal the limited extent of the changes that occurred during the period, since they represent only 12 km² over an area of interest of 88 sqkm (13%). Most of the change areas are related to the expansion of the city (75%, 9.2 sqkm, with half of it for the residential expansion), ant up to 10% of the changes concern evolution inside Natural and Semi-Natural areas.
Because of the very small surfaces inside the Core City, the evolutions concerning the Agricultural surfaces do not exceed 13% of the total land use changes.

Table 12: Core City Area - Overall Main LU/LC Changes Statistics.

Change Classes	Change Overall	
	sqkm	%
Urban Residential Extension	4.45	36.4%
Other Urban Land Use Extension	4.74	38.7%
Urban to Agricultural or Natural/Semi-Natural Areas	0.73	6.0%
Natural or Semi-Natural to Agricultural Areas	0.10	0.8%
Agricultural to Natural or Semi-Natural Areas	0.68	5.6%
Change within Natural and Semi-Natural Areas	1.53	12.5%
Total	**12.22**	**100%**

4.2.3 LU/LC Mapping for Larger Urban Area

The LU/LC map generated for 2018 reference year over the Larger Urban Area is depicted in Figure 14 and insight provided through Figure 15. A cartographic version of the map layout is provided as a PDF file in addition to the geo-spatial product.

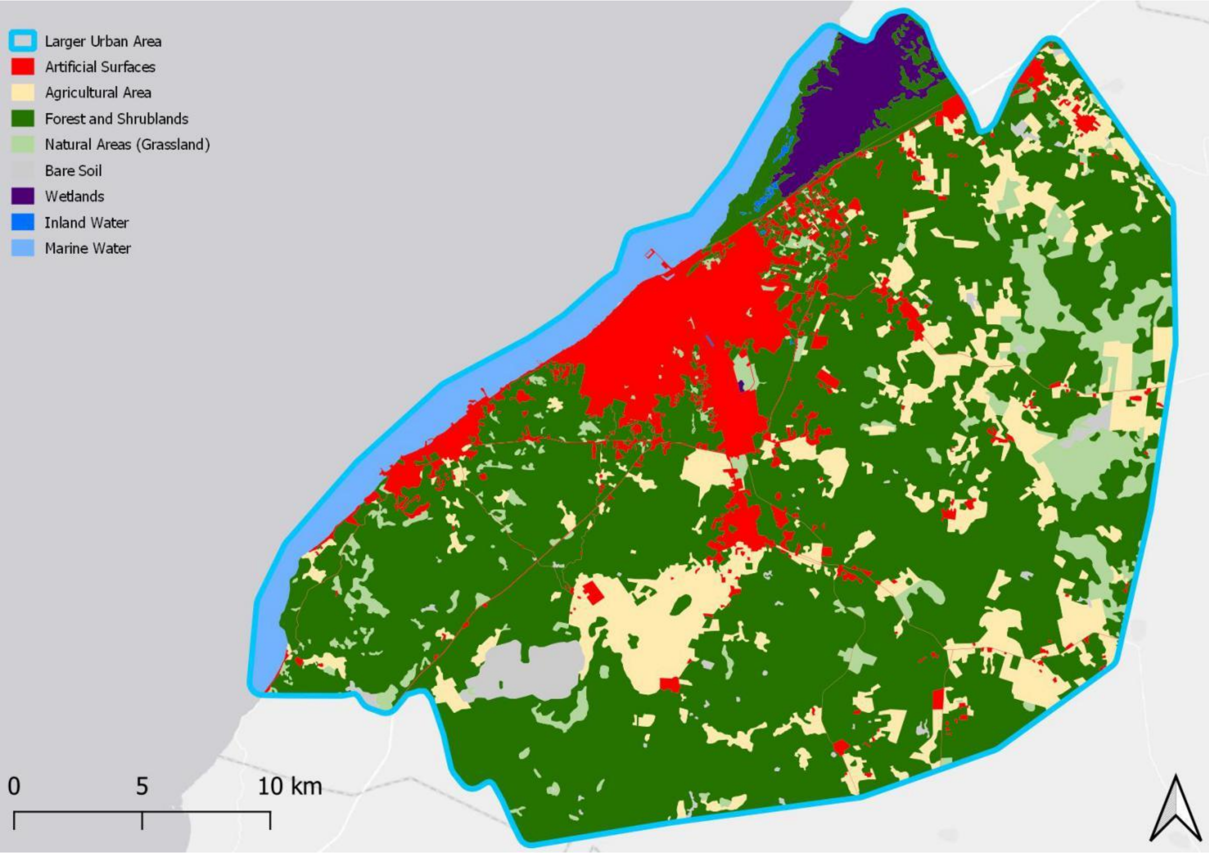

Figure 14: Larger Urban Area - Land Use Land Cover 2018 over Campeche.

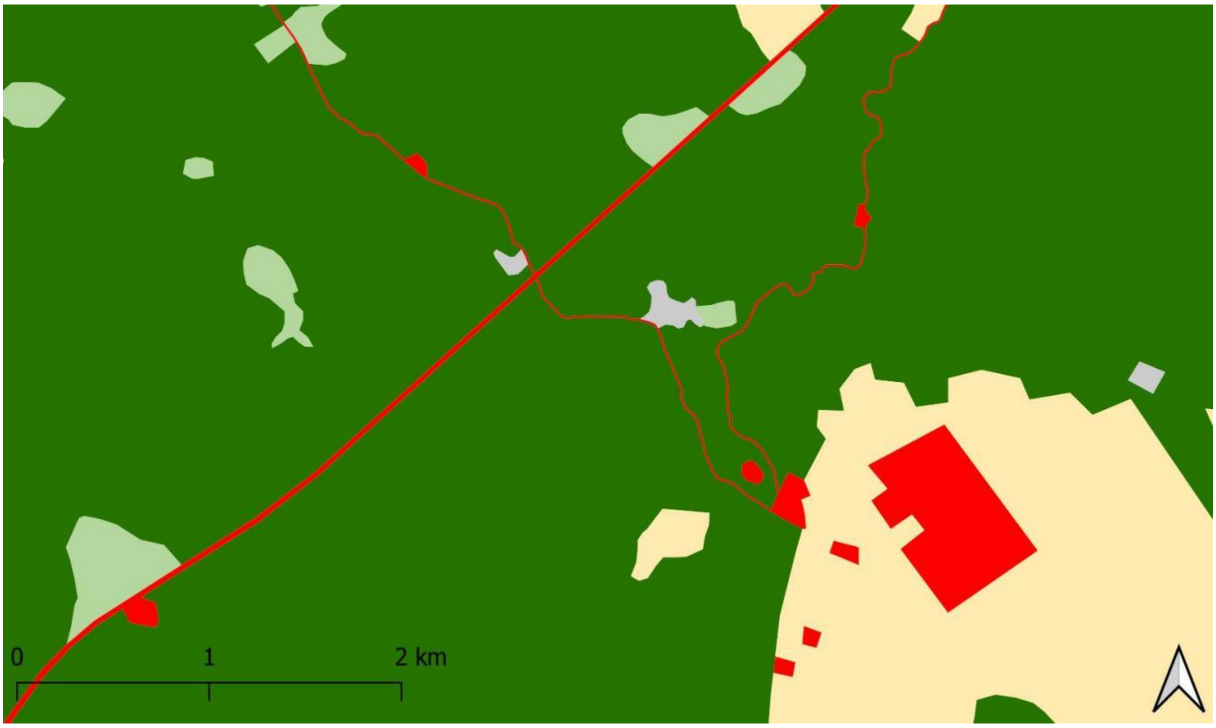

Figure 15: Larger Urban Area - Insight on the Land Use Land Cover 2018 outside the city.

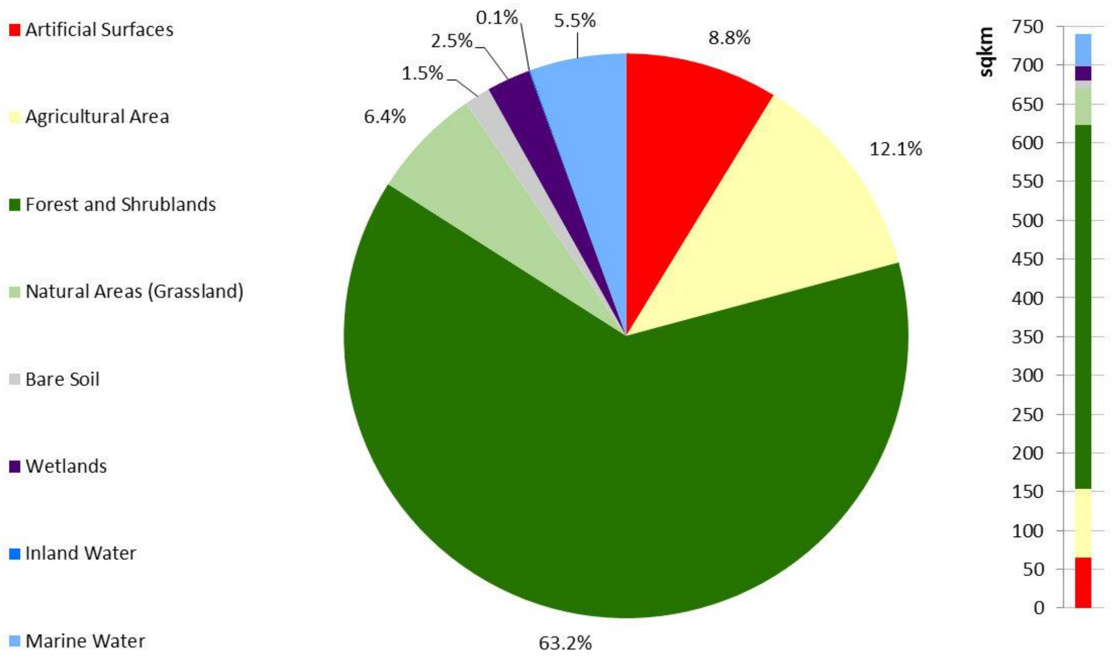

Figure 16: Larger Urban Area - Land Use Land Cover 2006 structure presented in % (left) and km² (right).

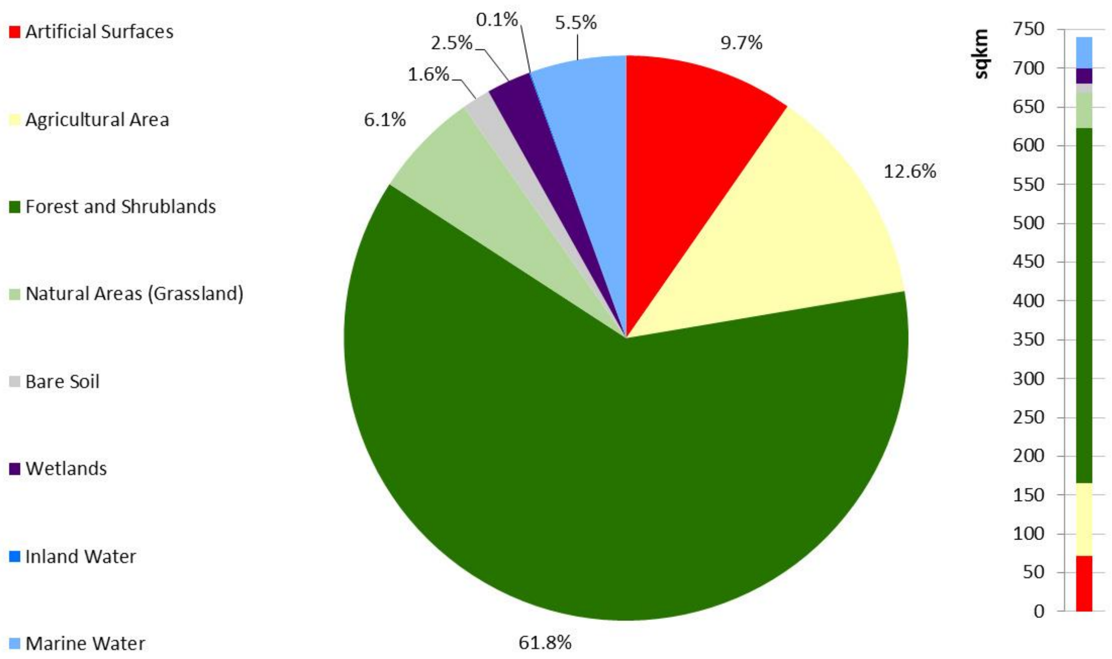

Figure 17: Larger Urban Area - Land Use Land Cover 2018 structure presented in % (left) and km^2 (right).

Figure 16 and Figure 17 provide detailed information on the class disaggregation and area coverage for the epochs 2006 and 2018 respectively. The Larger Urban Area of Campeche is largely dominated by the presence of forests and shrublands, that represent more than 60% of the surface. On the contrary, the artificial surfaces represent less than 10% of the territory. The agricultural areas are also represented within the territory (around 12%), with a small augmentation between 2006 and 2018.

Table 13: Larger Urban Area - Detailed information on area and percentage of total area for each class for 2006 and 2018 as well as the changes.

LU/LC Classes	2018		2006		Change		Change per Year	
	sqkm	% of total	sqkm	% of total	sqkm	%	sqkm	%
1000 - Artificial Surfaces	71.78	9.7%	64.82	8.8%	6.97	10.7%	0.58	0.9%
2000 - Agricultural Area	93.66	12.6%	89.32	12.1%	4.34	4.9%	0.36	0.4%
3100 - Forest and Shrublands	457.89	61.8%	468.15	63.2%	-10.25	-2.2%	-0.85	-0.2%
3200 - Natural Areas (Grassland)	44.96	6.1%	47.15	6.4%	-2.19	-4.6%	-0.18	-0.4%
3300 - Bare Soil	12.16	1.6%	11.08	1.5%	1.08	9.8%	0.09	0.8%
4000 - Wetlands	18.77	2.5%	18.57	2.5%	0.20	1.1%	0.02	0.1%
5100 - Inland Water	0.55	0.1%	0.58	0.1%	-0.03	-4.6%	0.00	-0.4%
5200 - Marine Water	40.76	5.5%	40.88	5.5%	-0.12	-0.3%	-0.01	0.0%
Total	**740.54**	**100%**	**740.54**	**100%**	-	-	-	-

In addition to the overall LU/LC classification for the two epochs it is interesting to further assess the different trends between classes over the 12-year time. The quantitative figures for each class are first provided in Table 13 to get an overview. The next Section will highlight the LU/LC change information between the two epochs in more detail.

During the considered period, the artificial surface gained 11% in terms of surface, going from 65 sqkm to 72 sqkm.

Even if the forest was consumed for the artificial expansion, the diminution of its surfaces is around 2% when analysed at the scale of the Larger Urban Area. The agricultural areas progressed between the two dates, with an increase of 5%, small contributor to the loss of Forest/Shrublands and Natural Areas.

4.2.4 Spatial Distribution of Main LU/LC Change Categories for Larger Urban Area

In order to better analyze the growth trends and the spatial distribution of changes, meaningful aggregations of the LU/LC classes in both epochs were made. The following categories were used:

- Urban expansion: all changes from Non-Artificial classes to Artificial class;
- Agriculture development: all changes conducting to development of Agricultural Areas;
- Natural or Semi-Natural development: changes form Artificial to Non-artificial areas;
- Natural and Semi-Natural Areas internal changes: all changes between the natural and semi-natural classes (e.g. Forest/shrubland, Natural/Semi-Natural, Wetlands).

Overlay analysis of these aggregated categories for the 2006 and 2018 epochs is depicted in Figure 18.

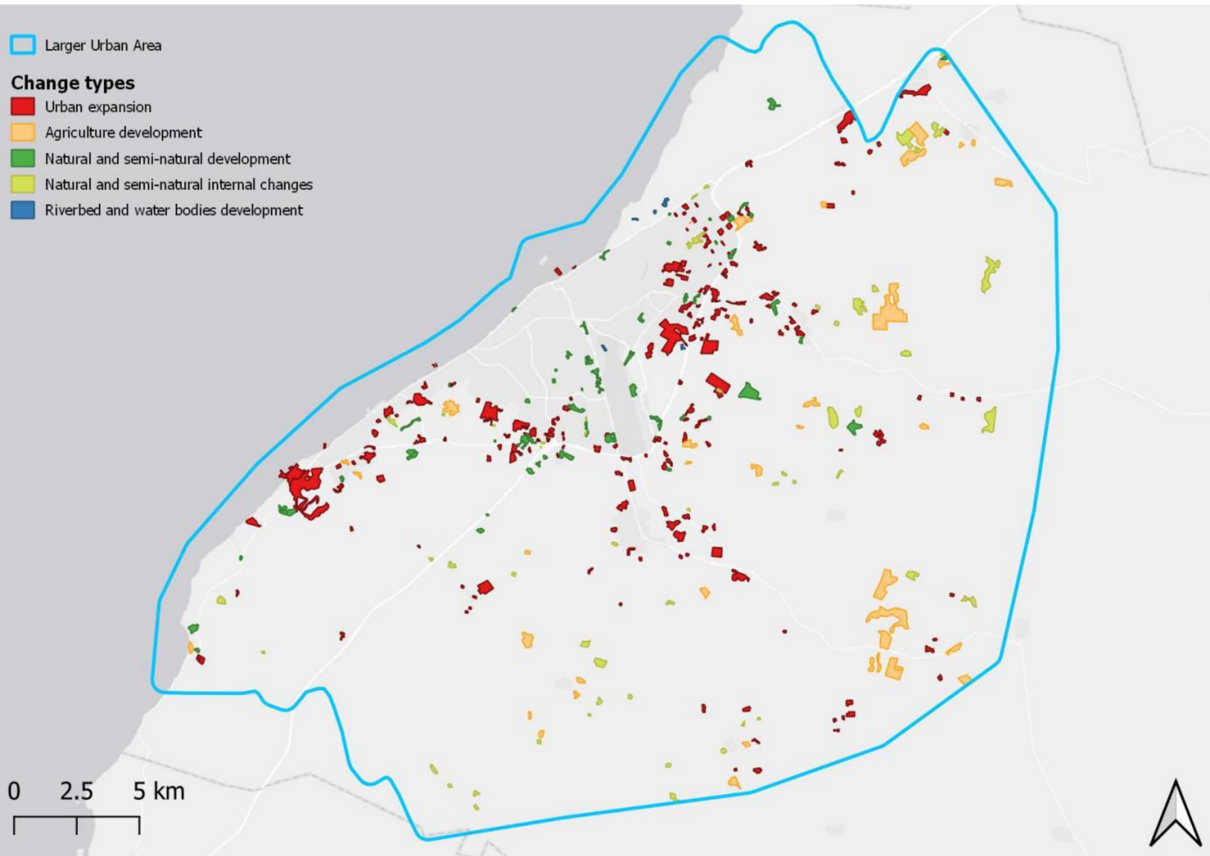

Figure 18: Larger Urban Area - Land Use Land Cover change types and spatial distribution.

The spatial distribution of the change types clearly shows the Urban Expansion around the existing city of Campeche, but also on the west with the construction of the Campeche Country Club and its surroundings.
The development of Agricultural Areas is also an important trend in the south and south-east of the territory.

The statistics of the Change categories are presented in Figure 19 and Table 14.

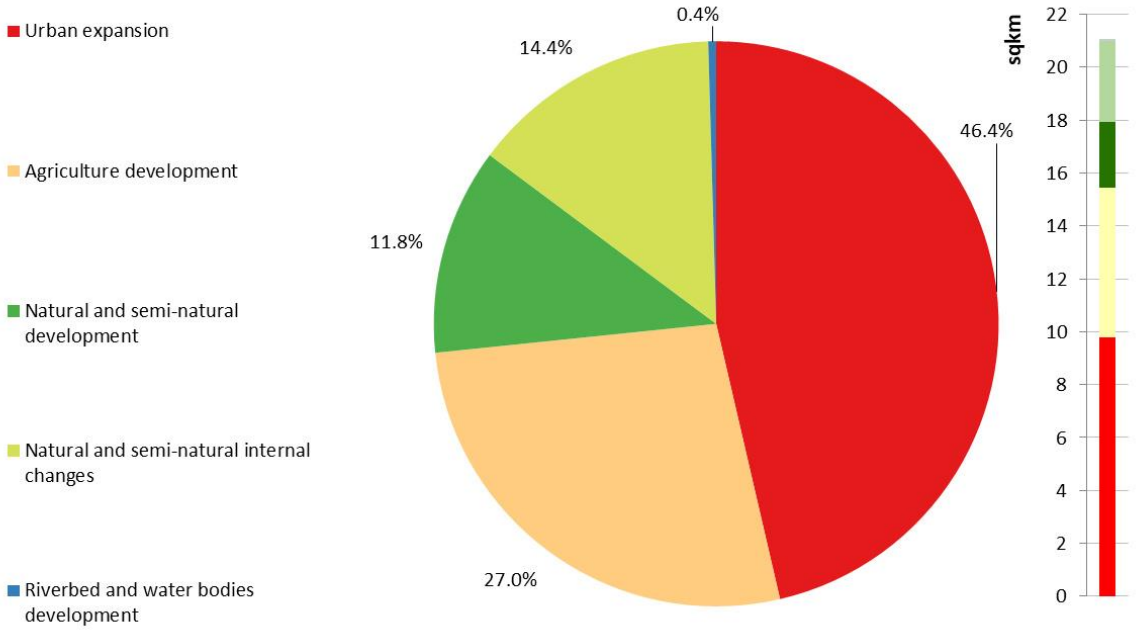

Figure 19: Larger Urban Area - Land Use Land Cover Change Types 2006-2018 presented in % (left) and km² (right).

The quantitative results reveal the important expansion (above 45%, around 10km²) of the Urban areas, but also of the development of the agriculture (27%).
The Natural and semi-natural development, which represent 12% of the changes for 2.5 sqkm, is particularly due to abandoned quarries, where vegetation has developed again.

Table 14: Larger Urban Area - Overall Main LU/LC Changes Statistics.

Change Classes	Change Overall	
	sqkm	%
Urban expansion	9.77	46.4%
Agriculture development	5.69	27.0%
Natural and semi-natural development	2.49	11.8%
Natural and semi-natural internal changes	3.02	14.4%
Riverbed and water bodies development	0.09	0.4%
Total	**21.07**	**100%**

4.3 Urban Green Areas

The location and extent of green areas are determined within the product of urban land use/ land cover at Level I. Urban green areas refer to land within and on the edges of a city that is partly or completely covered with grass, trees, shrubs, or other vegetation. The product delivered provides accurate information (1 m resolution) on the spatial location and extent of green areas located within the Urban Extent (Level I class: 1000) derived from the baseline LULC information product. Detecting and monitoring urban green coverage requires very high resolution optical satellite images, that explains the product generation over the Core Urban Area of Campeche only.

The Urban Green Areas change map generated over the AOI is depicted in Figure 20. A cartographic version of the map layout is provided as a pdf file in addition to the geo-spatial product.

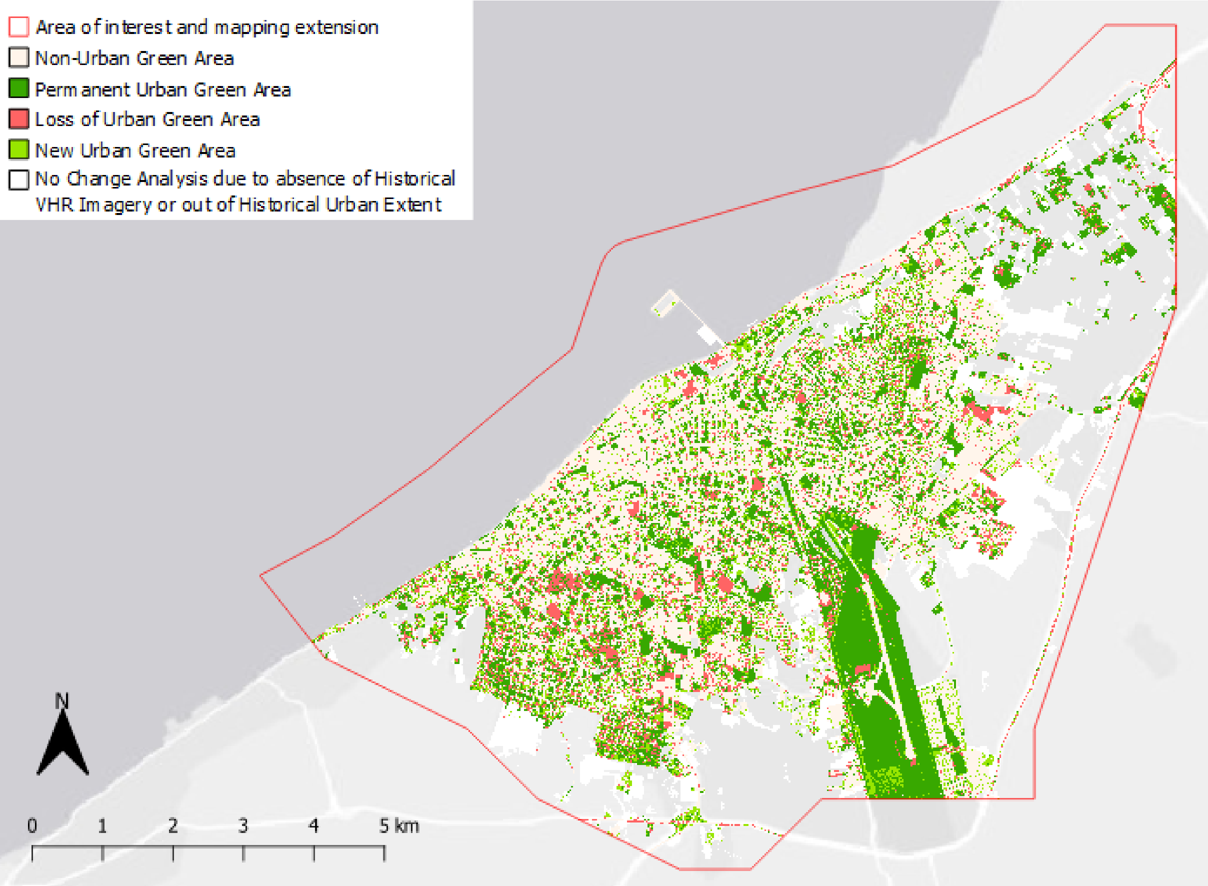

Figure 20: Urban Green Areas changes and spatial distribution.

The urban extent over the AOI is limited to the areas along the coastline. Green areas seem to be quite homogeneously distributed within this urban extent with an equilibrated balance of decrease and increase over time.

The quantitative results are shown in Figure 21 and Figure 22. The permanent green spaces over the period represent 23.8% of the entire area, or 10.9 km². The loss in green areas represents 9.3% and the gain 9.7%. This means that the total green area does not change much over time.

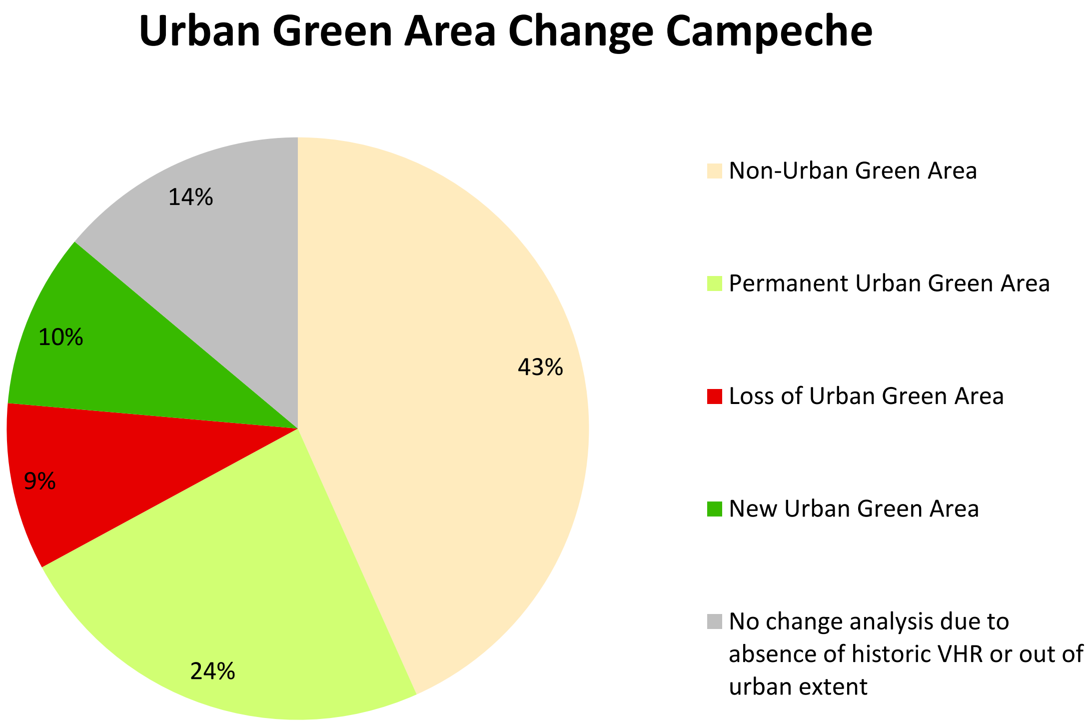

Figure 21: Status and change of urban green areas in-between 2006 and 2018 expressed in %.

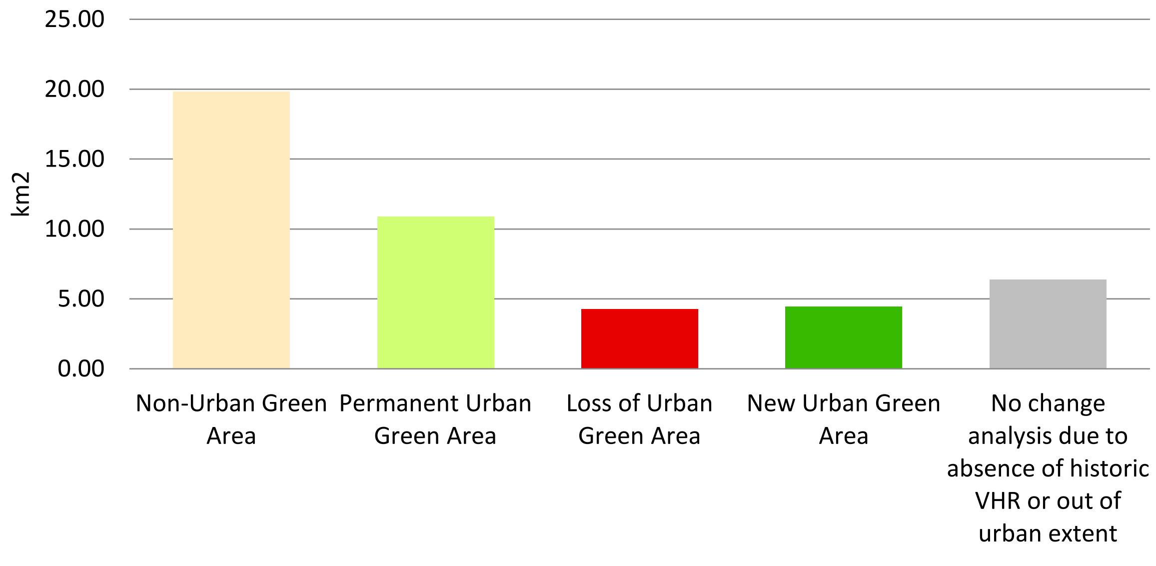

Figure 22: Status and change of urban green areas in-between 2006 and 2018 expressed in area.

4.4 Sustainable Development Goal 11 Indicators

A main objective of the EO4SD-Urban Product Portfolio is to support the reporting requirements of Urban Development Policies and Strategies. One of the most important policy frameworks that countries are trying to implement are the UN Sustainable Development Goals (SDGs). Seventeen SDGs were developed with a focus on “ending extreme poverty; fighting inequality & injustice; and addressing climate change,” by 2030. To achieve the 17 goals there are 169 targets and for each target, indicators will be used to assess the level of achievement of the countries.

The SDG Goal 11 “Make cities and human settlements inclusive, safe, resilient and sustainable” is specifically dedicated to Sustainable Urban Development. A list of Urban Sustainability Indicators specific to the SDG Goal 11, have been defined in March 2016 by the UN and are described in the UN-Habitat “SDG Goal 11 Monitoring Framework Report (UN, 2016a)”.

The EO4SD-Urban project supports seven GPSC cities, namely Bhopal and Vijayawada in India, Campeche in Mexico, Saint-Louis and Dakar in Senegal, Abidjan in Ivory Coast and Lima in Peru. For these seven cities, the indicators for which the needed input data is available were calculated and are described in the following subsections. The EO4SD-Urban products can be fully or partly used for the calculation of four SDG 11 indicators (see Table 15).

Table 15: SDG 11 indicators measurable with the support of EO4SD-Urban products.

TARGETS	INDICATORS
Target 11.1: By 2030, ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums	11.1.1: Proportion of urban population living in slums, informal settlements or inadequate housing
Target 11.2: By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons	11.2.1: Proportion of the population that has convenient access to public transport by sex, age and persons with disabilities
Target 11.3: By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries	11.3.1: Ratio of land consumption rate to population growth rate
Target 11.7: By 2030, provide universal access to safe, inclusive and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities	11.7.1: Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities

A short description of the calculation as well as the needed input data and the achieved outputs are described in the next sections for the indicators 11.2.1, 11.3.1 and 11.7.1. For Campeche, it is not possible to calculate the Indicator 11.1.1, as the needed input data is not available.

More information including the exact calculation steps of each indicator are described in the UN-Habitat Methodological Guidance document to monitor and report on the SDG Goal 11 indicators (UN-Habitat, 2016).

4.4.1 SDG 11 Indicator 11.2.1

The 11.2.1 Indicator calculates the ***Proportion of the population that has convenient access to public transport by sex, age and persons with disabilities*** and describes the Target 11.2: "By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, and children, persons with disabilities and older persons."

The indicator aims to monitor the use and access of public transportation system and move towards reaching a convenient access for all. According to UN-Habitat and described in the Methodological Guidance document (UN-Habitat, 2016) the access to public transport is considered convenient when an officially recognised stop is accessible within a distance of 0.5 km from a reference point such as home, school, workplace, market, etc.

The indicator is calculated by using the following formula:

$$\text{\% with access to public transport} = \frac{100x\text{ (population with convenient access to public transport)}}{\text{city population}}$$

At a diagnosis phase, this indicator helps urban planners in identifying areas that are underserved and to be put as a priority in the Master Plans for the localisation of transport stations and addition of new transport lines (bus, metro, tramway, train).

Calculating this indicator considering parameters such as sex, age and persons with disabilities would require additional census data, as not available through EO data. However, the indicator can be calculated over the Larger Urban Area using the Global Human Settlement Population Layer and the OpenStreetMap (OSM) transportation features (bus and subway stations and stops, railway stations, ferry terminals), both available for the reference year 2015. It provides a first good estimate of the proportion of the population that has convenient access to public transport.

The results are presented in Figure 23 below. For comparative reasons the graphic shows the indicator results for all GPSC cities, but Bhopal and Vijayawada. The proportion of the population that has convenient access to public transport is estimated around 20% of the total population of Campeche and 30% of Abidjan and Saint-Louis, while this indicator is higher for Dakar and Lima with a value close to 50%.

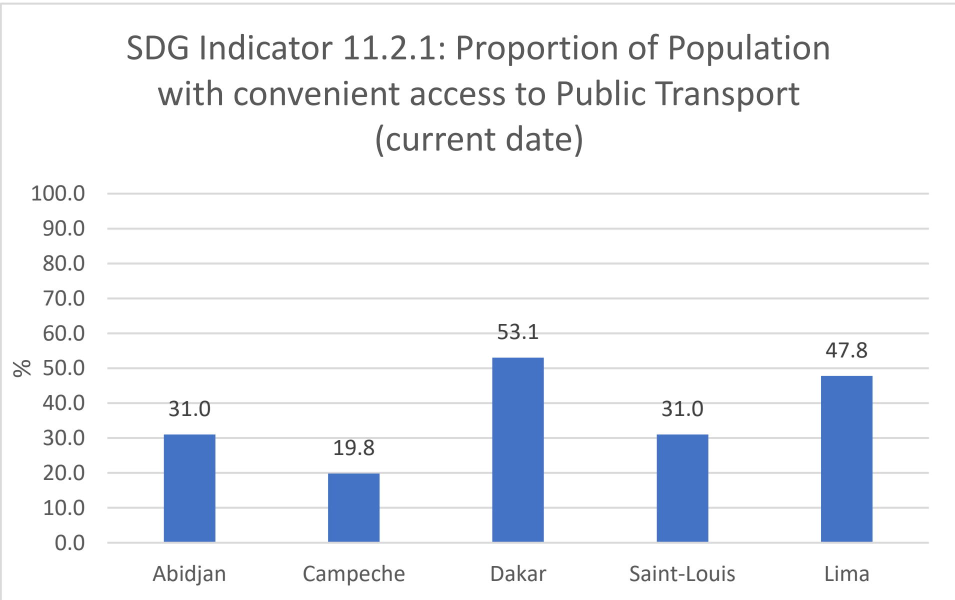

Figure 23: Proportion of population with convenient access to public transport.

4.4.2 SDG 11 Indicator 11.3.1

The 11.3.1 Indicator calculates the ***Ratio of land consumption rate to population growth rate*** and describes the Target 11.3: "By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated and sustainable human settlement planning and management in all countries."

The indicator needs the definition of the two components population growth and land consumption rate. According to the UN-Habitat Methodological Guidance document (UN-Habitat, 2016) the population growth rate (PGR) is the increase of population in a country during a specific period, usually one year. The PGR is expressed as a percentage of the population at the start of that period.

Further, the land consumption rate includes a) the expansion of build-up area that can be directly measured and b) the absolute extent of land that is subject to exploitation by agriculture, forestry or other economic activities and c) the over-intensive exploitation of land that is used for agriculture and forestry.

The indicator is calculated by using following formula:

$$\text{Ratio of land consumption rate to population growth rate (LCRPGR)} = \frac{\text{Land consumption rate}}{\text{Annual population growth rate}}$$

The ratio of land consumption rate to population growth rate is an indicator for measuring land use efficiently and is intended to answer the questions of whether the remaining undeveloped urban land is being developed at a rate that is less than or greater than the prevailing rate of population growth. As the ratio of land consumption rate to population growth rate is dimensionless and not straightforward in its interpretation, several countries report the urban expansion and the population growth rate in terms of percentage change instead of using the ratio values (Nicolau et. al., 2018).
In the following, the ratio (see Figure 24) and the percentage change values (see Figure 25) for all GPSC cities were calculated. For the calculation of the population growth rate the Global Human Settlement Population Layer available for the years 2000 and 2015 were used. For the calculation of the land consumption rate the built-up area extracted from the Larger Urban Area LU/LC classification is taken by dissolving all artificial classes.

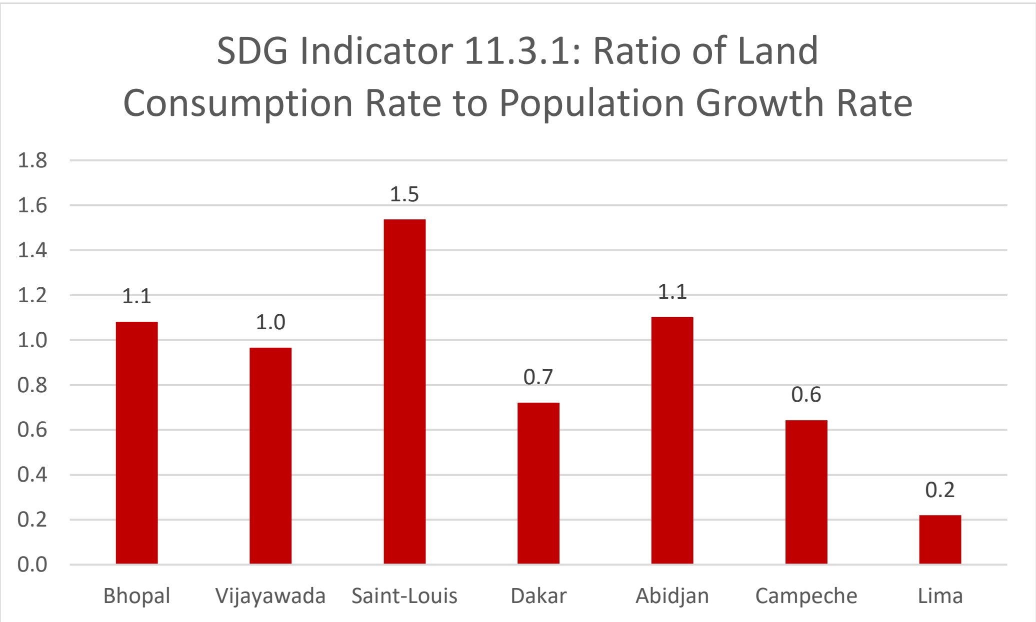

Figure 24: Ratio of land consumption rate to population growth rate between 2005 and 2015.

Figure 24 shows the ratio of land consumption rate to population growth rate for all GPSC cities. All GPSC cities are visualised in the bar chart for comparative reasons. Campeche, as well as Dakar and Lima, has a value significantly below one, which means that the population growth rate is higher than the land consumption rate and let assume that the land is efficiently used.
Cities with values close to one have a population growth rate similar to the land consumption rate. This indicates that the land is efficiently used too.
On the contrary, cities with values significantly above one have a higher land consumption rate than a population growth rate. This indicates that the land is not as efficiently used as for example in Campeche.
European countries, for comparison very often have values below zero. This means that either the population or the land consumption shows a decrease.

Looking at the percentage change values of population and land consumption between 2000/2006 and 2015/2018 all cities have a growing population and a growing urban extent, which is typical for cities in developing countries. Campeche's population grew by 21.9% between 2000 and 2015. Its land consumption grew by 10.7% between 2006 and 2018. This means that the population grew faster than the built-up area of the city and indicates also that the city seems to grow in a compact way.

In Saint-Louis for example, it is the other way around. Here the land consumption grew by 21% while the population grew by only 17%, indicating that the city had a less compact growth in the last years.

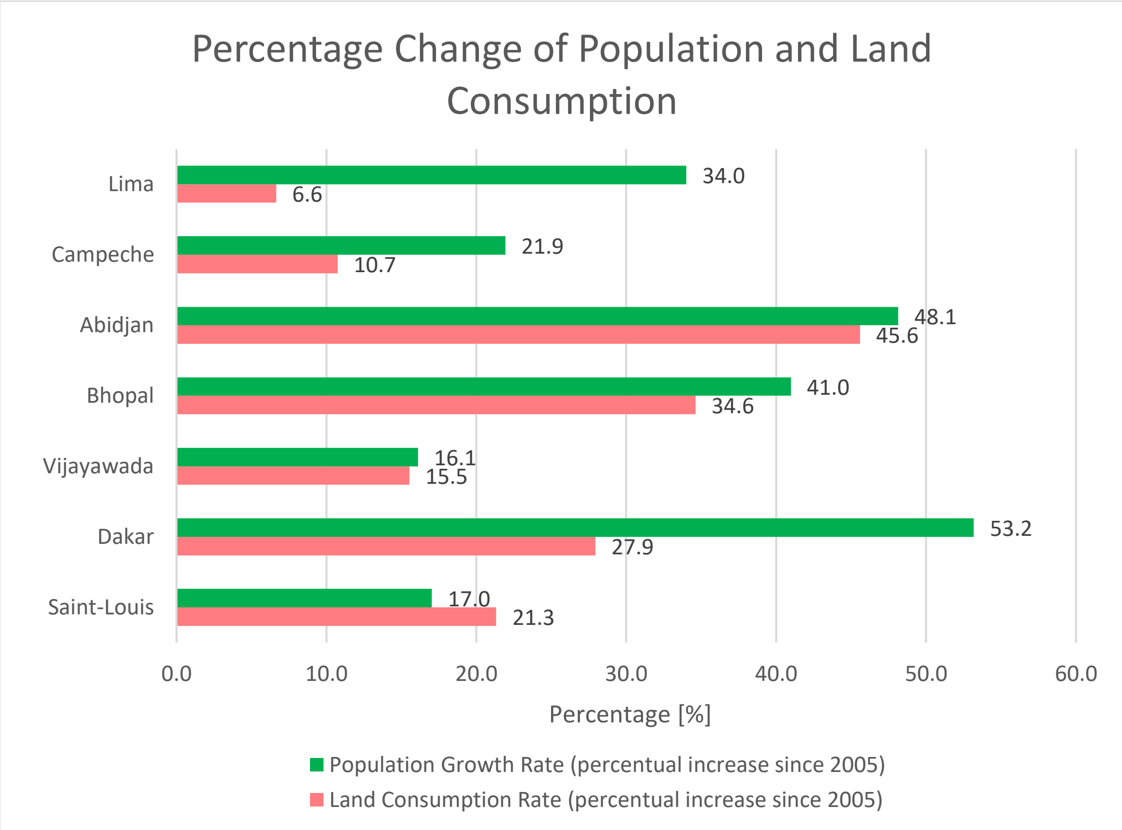

Figure 25: Percentage change of population and land consumption between 2005 and 2015.

A significant limitation of this indicator is that the approach captures only the urban extent change, not the internal city dynamics.

4.4.3 SDG 11 Indicator 11.7.1

The SDG 11 Indicator 11.7.1 ***"Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities"*** refers to the Target 11.7.: By 2030, provide universal access to safe, inclusive and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities.

The indicator aims to monitor the amount of land that is dedicated by cities for public space. According to the UN-Habitat Methodological Guidance document (UN-Habitat, 2016) public space includes open spaces and streets and should be accessible by all.

The indicator is calculated by using following formula:

$$\text{\% of land that is dedicated by cities for public space (open spaces and streets)} = \frac{(\text{Total surface of open public space} + \text{Total surface of land allocated to streets})}{\text{Total surface of built up area of the urban agglomeration}}$$

The share of land in public open spaces cannot be obtained directly from the use of high-resolution satellite imagery, because it is not possible to determine the ownership or use of open spaces by remote sensing. Additional metadata that helps to describe the land use patterns in the locale is additionally required to map out land that is for public and non-public use.
As this information is not available, the LU/LC classes *Urban Green Areas* and *Sports and Leisure Facilities*, which are available in the Core City Area LU/LC classification, were taken with the assumption that these places are public places and accessible by all.
To calculate the total surface of land allocated to streets, the road network was used. Different buffers were applied for three different road types (6m for Arterial Roads, 5m for Collector Roads and 3m for Local Roads) to assess the total surface of streets. The total surface of built-up area of the urban agglomeration is extracted from the LU/LC classification by summarising all artificial classes of the Core City Area.

The results are presented in Figure 26 below. For comparative reasons the graphic shows the indicator results for all GPSC cities, but Lima.
The indicator was calculated for two points in time i.e. around 2005 and 2015. In Campeche, the average share of built-up area that is open space for public use decreases from 28.1% in 2006 to 24.5% in 2018. Bhopal, Vijayawada and Saint-Louis also show a decrease, while Abidjan and Dakar show an increase in open spaces.

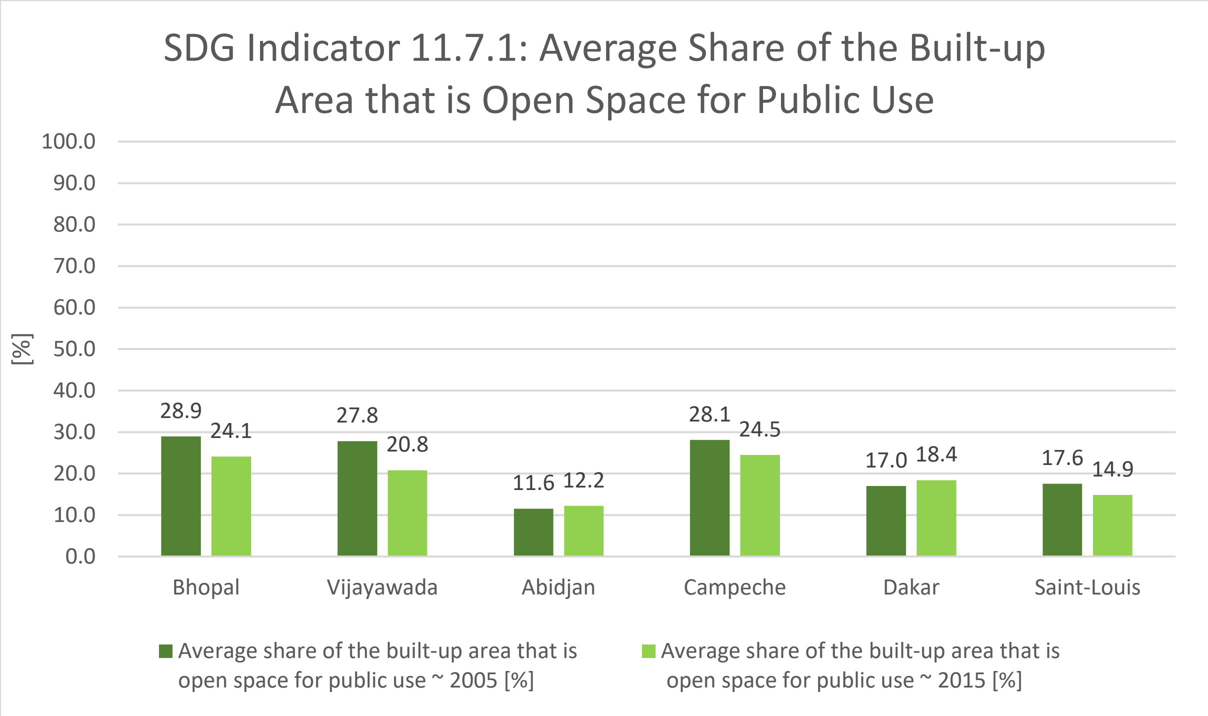

Figure 26: Average share of the built-up area that is open space for public use.

4.5 Concluding Points

This Chapter 4 presented a summary and an overview of what is possible in terms of analytics with the geo-spatial datasets provided for Campeche in the current project. This Report is a living document and will be complemented with further analysis during the project when relevant. Important would be to further analyse the EO4SD Urban datasets with the main stakeholders, i.e. the IADB and the INFOCAM, in order to enhance the latter for planning purposes.

5 References

Czaplewski, R. L. (2003). Chapter 5: Accuracy assessment of maps of forest condition: statistical design and methodological considerations, pp. 115–140. In Michael A. Wulder, & Steven E. Franklin (Eds.), Remote sensing of forest environments: concepts and case studies. Boston: Kluwer Academic Publishers (515 pp.).

European Union (2011). Mapping Guide for a European Urban Atlas, Version 11.0

Goodchild, M., Chih-Chang, L. and Leung, Y. (1994): Visualizing fuzzy maps, pp. 158-67. In Heamshaw, H.H. and Unwin, D.J. (Eds.), Visualization in geographical information systems. Chichester: Wiley.

Olofsson, P., Foody, G. M., Stehman, S. V., & Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment, 129, 122–131. doi:10.1016/j.rse.2012.10.031

Selkowitz, D. J., & Stehman, S. V. (2011). Thematic accuracy of the National Land Cover Database (NLCD) 2001 land cover for Alaska. Remote Sensing of Environment, 115(6), 1401–1407. doi:10.1016/j.rse.2011.01.020.

Internet

Road classification, European Commission, 2017, https://ec.europa.eu/transport/road_safety/specialist-/knowledge/road/designing_for_road_function/road_classification_en, last accessed 2017.08.17

This page is intentionally left blank!

Annex 1 – AOI Calculation based on the DG Regio approach

This page is intentionally left blank!

AOI Calculation Methodology based on the DG Regio Approach

So far, no internationally accepted definition for the term "Urban Area" and the related Peri-Urban area exists. Different initiatives are currently trying to address a standardised approach for defining the term "Urban Area". During discussions with the GPSC Co-ordinator it was considered important to use an uniform definition for the GPSC cities in order for the cities to exchange information and share products/experiences and conduct potential comparative studies.

In this context, it was decided to use an international approach for the demarcation of the Areas of Interest (AOI) for mapping the GPSC cities in terms of Core Urban area and Peri-Urban area. Thus, the approach is based on the European Union's Directorate-General for Regional and Urban Policy (DG REGIO) method and the definitions are described in the Regional Working Paper 2014 from the European Commission on "A harmonised definition of cities and rural areas: the new degree of urbanisation" (European Commission, 2014). Following the naming of the DG Regio approach, the Urban Core is named as "High Density Core" and the Peri-Urban area is termed "Urban Cluster". Within the DG REGIO approach, the High Density Core is defined as contiguous grid cells of 1 km^2 with a density of at least 1 500 inhabitants per km^2 and a minimum population of 50 000. The Urban Cluster is defined as clusters of contiguous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 5 000.

The DG REGIO methodology used in the EO4SD-Urban project was slightly adjusted to Non-European countries. For the first two GPSC cities (namely Bhopal and Vijayawada) produced within the project the Global Human Settlement Population (GHSP) grid with a spatial resolution of 1 km were used for the classification into "High Density Core" and "Urban Cluster". The raster dataset is available for the years 1975, 1990, 2000, 2015. This dataset depicts the distribution and density of population, expressed as the number of people per cell. The data can be downloaded under following link http://data.jrc.ec.europa.eu/dataset/jrc-ghsl-ghs_pop_gpw4_globe_r2015a.

In 2019, a higher resolution population layer (spatial resolution of 10 m) produced by the German Aerospace Centre (DLR) became available. The AOIs for the remaining GPSC cities (namely Melaka, Abidjan, Dakar and Campeche) were produced based on the DLR population layer.

In the following, a more detailed description of the calculation methodology for the High-Density Core and the Urban Cluster follows. The calculation is exemplary described on the AOI generation for the city of Melaka.

To start with, Figure 27 shows the city of Melaka and the surrounding area. Figure 28 shows the population distribution grid over Melaka produced by the European Commission. Figure 29 shows the DLR population grid with 10 meter spatial resolution for Melaka while Figure 30 illustrates the aggregated DLR population grid with 1 km spatial resolution.

Figure 27: Satellite image showing Melaka and the surrounding area.

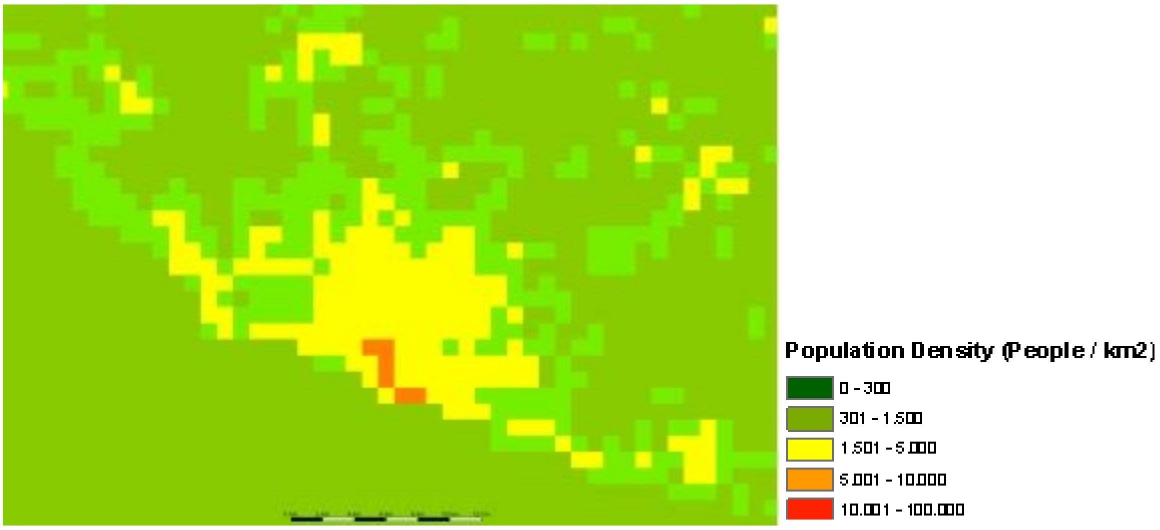

Figure 28: Global Human Settlement Population Layer (spatial resolution of 1 km).

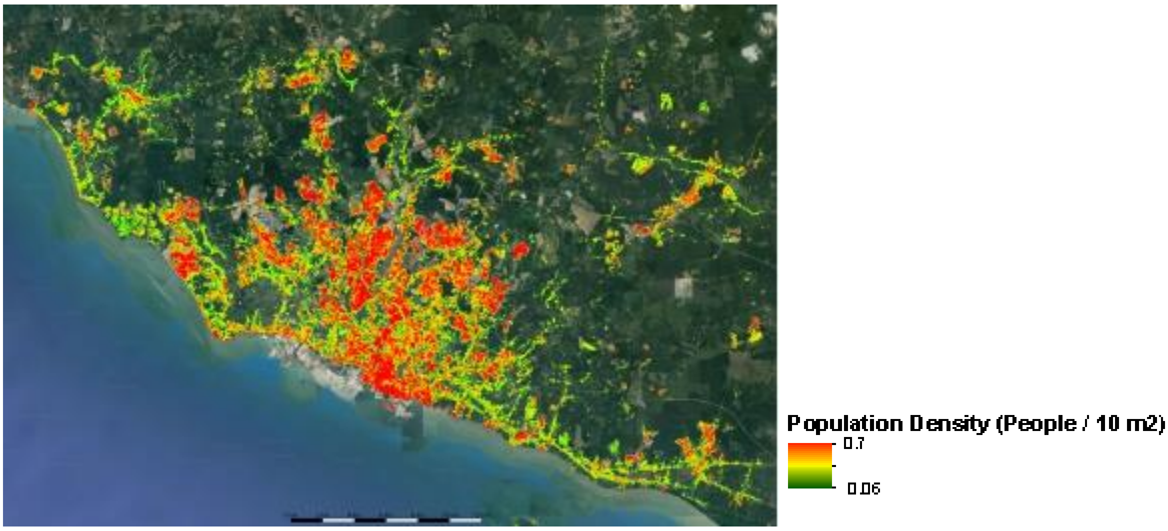

Figure 29: DLR population layer (spatial resolution of 10 m).

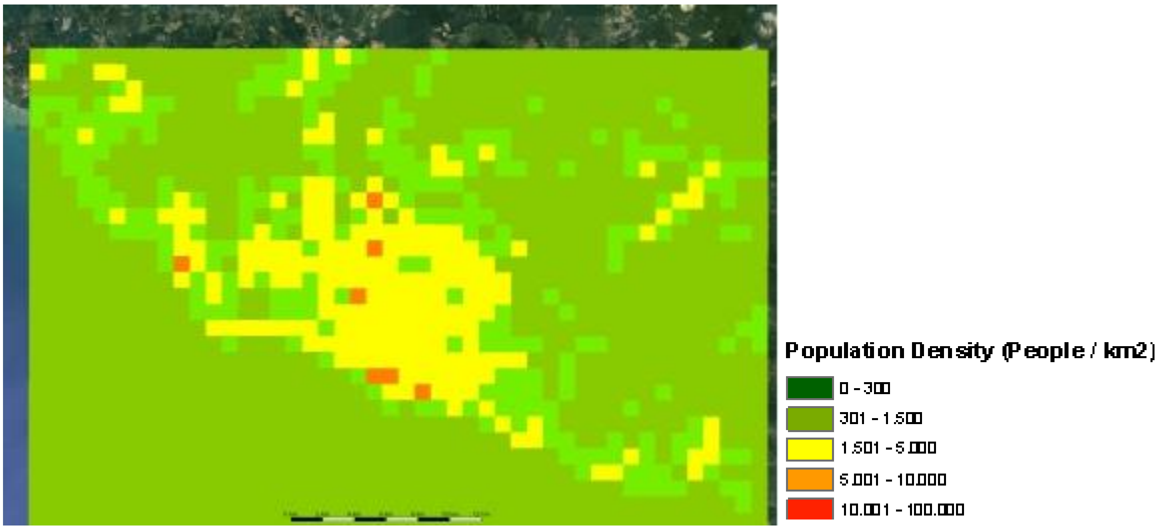

Figure 30: Aggregated DLR population layer (spatial resolution of 1 km).

The High Density Core AOI for a city is created by merging the contiguous grid cells of 1 km^2 with a density of at least 1500 inhabitants per km^2 and a minimum population of 50 000. In the definition of the High Density Core the contiguity is only allowed via a vertical or horizontal connection. In a next step, gaps are filled. Due to the coarse resolution of the population grid cells additional grid cells were in a last step added for under estimated settlement areas. The same was done for over estimations, here grid cells were removed.

Figure 31 shows the High Density Core AOI (red line) overlaid on the DLR population layer (left) and on a RGB satellite image (right).

Figure 31: High Density Core area of Melaka calculated based on the aggregated DLR population layer.

The image on the left shows the AOI overlaid on the DLR population layer. On the right, the AOI is overlaid on a RGB satellite image.

The Urban Cluster is created very similar to the High Density Core. Continuous grid cells of 1 km^2 with a density of at least 300 inhabitants per km^2 and a minimum population of 500 are merged together to form the Urban Cluster. The contiguity within the Urban Cluster can also be diagonal. After gaps are filled, areas, which were over or under estimated by the population grid were removed or added to the AOI. Figure 32 shows the Urban Cluster AOI (magenta line) overlaid on the DLR population layer (left) and on a RGB satellite image (right).

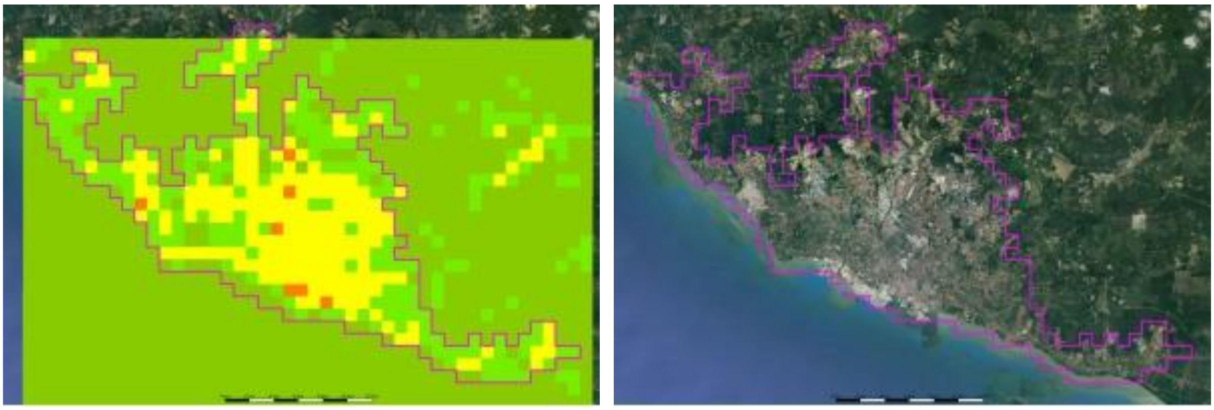

Figure 32: Urban Cluster area of Melaka calculated based on the aggregated DLR population layer.

In some cases, the city counterparts requested that the AOIs for the High Density Core and the Urban Cluster follow the municipal or administrative boundary of the city. In this case, the municipal/administrative boundary was used but enlarged in areas where the AOI created according to the adjusted DG Regio approach was bigger.

This page is intentionally left blank!

Annex 2 – Processing Methods for EO4SD-Urban Products

This page is intentionally left blank!

Summary of Processing Methods

Urban Land Use/Land Cover and Change

The input includes Very High Spatial Resolution (VHR) imagery from different sensors acquired at different time. The data is pre-processed to ensure a high level of geometric and radiometric quality (ortho-rectification, radiometric calibration, pan-sharpening).

The complexity when dealing with VHR images comes from the internal variability of the information for a single land-use. For instance, an urban area is represented by a high number of heterogeneous pixel values hampering the use of automated pixel-based classification techniques.

For these VHR images, it is possible to identify textures (or pattern) inside an entity such as an agricultural parcel or an urban lot. In other words, whereas pixel-based techniques focus on the local information of each single pixel (including intensity / DN value), texture analysis provides global information in a group of neighbouring pixels (including distribution of a group intensity / DN values but also spatial arrangement of these values). Texture and spectral information are combined with a segmentation algorithm in an Object Based Image Analysis (OBIA) approach to reach a high degree of automation for most of the peri-urban rural classes. However, within urban land, land use information is often difficult to obtain from the imagery alone and ancillary/in situ data needs to be used. The heterogeneity and format of these data mean that another information extraction method based on Computer Aided Photo-Interpretation techniques (CAPI) need to be used to fully characterise the LULC classes in urban areas. Therefore, a mix of automated (OBIA) and CAPI are used to optimise the cost/quality ratio for the production of the LULC/LUCC product. The output format is typically in vector form which makes it easier for integration in a GIS and for subsequent analysis.

Level 4 of the nomenclature can be obtained based on additional information. These can be generated by more detailed CAPI (e.g. identification of waste sites) or by an automated approach based on derived/additional products. An example is illustration by categorising the density of the urban fabric which is related to population density and can then subsequently used for disaggregating population data.

Information on urban fabric density can be obtained through several manners with increasing level of complexity. The Imperviousness Degree (IMD) or Soil Sealing (SL) layer (see separate product) can be produced relatively easily based on the urban extent derived from the LULC product and a linear model between imperviousness areas and vegetation vigour that can be obtained from Sentinel 2 or equivalent NDVI time series. This additional layer can be used to identify continuous and discontinuous urban fabric classes. Four urban fabric classes can be extracted based on a fully automated procedure:

- Continuous Dense Urban Fabric (Sealing Layer-S.L. > 80%)
- Discontinuous Dense Urban Fabric (S.L. 50% - 80%)
- Discontinuous Medium Density Urban Fabric (S.L. 30% - 50%)
- Discontinuous Low Density Urban Fabric (S.L. 10% - 30%)
- Discontinuous Very Low Density Urban Fabric (S.L. < 10%)

Manual enhancement is the final post-processing step of the production framework. It will aim to validate the detected classes and adjust classes' polygon geometry if necessary to ensure that the correct MMU is applied. Finally, a thorough completeness and logical consistency check is applied to ensure the topological integrity and coherence of the product.

Change detection: Four important aspects have to be considered to monitor land use/land cover change effectively with remote sensing images: (1) detecting that changes have occurred, (2) identifying the nature of the change, (3) characterising the areal extent of the change and (4) assessing the spatial pattern of the change.

The change detection layer can be derived based on an image-to-image approach provided the same sensor is used. An original and efficient image processing chain is promoted to compare two dates' images and provide multi-labelled changes. The approach mainly relies on texture analysis, which has the benefits to deal easily with heterogeneous data and VHR images. The applied change mapping approach is based on spectral information of both dates' images and more accurate than a map-to-map comparison.

Summary of Processing Methods

World Settlement Extent

The rationale of the adopted methodology is that given a series of radar/optical satellite images for the investigated AOI, the temporal dynamics of human settlements are sensibly different than those of all other land-cover classes.

While addressing settlement-extent mapping for the period 2014-2015 multitemporal S1 IW GRDH and Landsat-8 data acquired at 10 and 30m spatial resolution were taken into account. Concerning radar data, each S1 scene is pre-processed by means of the SNAP software available from ESA; specifically, this task includes: orbit correction, thermal noise removal, radiometric calibration, Range-Doppler terrain correction and conversion to dB values. Scenes acquired with ascending and descending pass are processed separately due to the strong influence of the viewing angle in the backscattering of built-up areas. As a means for characterizing the behaviour over time, the backscattering temporal maximum, minimum, mean, standard deviation and mean slope are derived for each pixel. Texture information is also extracted to ease the identification of lower-density residential areas. As regards optical data, only Landsat-8 scenes with cloud cover lower than 60% are taken into consideration (indeed, further rising this threshold often results in accounting for images with non-negligible misregistration error). Data are calibrated and atmospherically corrected using the LEDAPS tool available from USGS and the CFMASK software is applied for removing pixels affected by cloud-cover and cloud-shadow. Next, a series of 6 spectral indices suitable for an effective delineation of settlements (identified through extensive experimental analysis) are extracted; these include – among others – the Normalized Difference Built-Up Index (NDBI), the Modified Normalized Difference Water Index (MNDWI) and the Normalized Difference Vegetation Index (NDVI). For all of them, the same set of 5 key temporal statistics used in the case of S1 data are generated for each pixel in the AOI. Moreover, to improve the detection of suburban areas, for each of the 6 temporal mean indices also here texture information is computed. For matching the spatial resolution of Sentinel data, the whole stack of Landsat-based features is finally resampled to 10m spatial resolution.

To identify reliable training points for the settlement and non-settlement class, a strategy has been designed which jointly exploits the temporal statistics computed for both S1 and Landsat data, along with additional ancillary information. In the case of optical data, in general the most of settlement pixels can be effectively outlined by properly jointly thresholding the corresponding NDBI, NDVI, and MNDWI temporal mean; likewise, this holds also for non-settlement pixels. Regarding radar data, it generally occurs that the temporal mean backscattering of most settlement samples is sensibly higher than that of all other non-settlement classes. Nevertheless, in complex topography regions: i) radar data show high backscattering comparable to that of urban areas; and ii) bare rocks are present, which often exhibit a behaviour similar to that of settlements in the Landsat-based temporal statistics. Accordingly, to exclude these from the analysis, all pixels are masked whose slope - computed based on SRTM 30m DEM for latitudes between -60° and +60° and the ASTER DEM elsewhere - is higher than 10 degrees.

Support Vector Machines (SVM) are used in the classification process. However, as the criteria defined above for outlining training samples might results in a high number of candidate points, for AOIs up to a size of ~10000 km² the most effective choice proved extracting 1000 samples for both the settlement and non-settlement class. Nonetheless, since results might vary depending on the specific selected training points, as a means for further improving the final performances and obtain more robust classification maps, 20 different training sets are randomly generated and given as input to an ensemble of as many SVM classifiers; then, a majority voting is applied. Afterwards, the stacks of Landsat-8-based and S1-based temporal features are classified separately as this proved more effective than performing a single classification on their merger. In both cases, a grid search with a 5-fold cross validation approach is employed to identify for each training set the optimal values for the learning. Here, those resulting in the highest cross-validation overall accuracy are then selected and used for classifying the corresponding study region.

A final post-classification phase is dedicated to properly combining the Landsat- and S1-based classification maps and automatically identifying and deleting potential false alarms. To this purpose, an advanced post-editing object-based approach has been specifically designed.

The above-described methodology has been further adapted for outlining the settlement extent in the past solely based on Landsat-5/7 imagery available since 1984; indeed, no long-term SAR data archive at comparable spatial resolution is freely accessible for the same timeframe (e.g., ESA ERS-1/2 data are available from 1991 without systematic world coverage and often proved too complicated to pre-process). In particular, for the given target period and AOI, all available Landsat imagery with cloud cover lower than 60% is pre-processed in the same fashion as described in the previous paragraphs and the same set of temporal statistics and texture features are extracted. Based on the hypothesis that settlement growth occurred over time (meaning that a pixel cannot be marked as settlement at an earlier time if it has been defined as non-settlement at a later time), all pixels categorized as non-settlement in the 2014-2015 extent map are excluded from the analysis. Then, training samples are derived by thresholding the temporal mean NDBI, MNDWI and NDVI; specifically, a dedicated strategy has been implemented for automatically determining the thresholds for the 3 indices by comparing their cumulative distribution function (CDF) for the target period with that exhibited for the period 2014-2015. Also in this case, an ensemble of 20 SVMs is used, each one trained on a different subset of 2000 samples (i.e., 1000 for the settlement and 1000 for the non-settlement class) and majority voting is then employed for generating the final map. It is worth noting that, when deriving the past settlement extent for multiple times, both the masking and threshold adaptation are performed on the basis of the results derived for the next target period.

Summary of Processing Methods

Percentage Impervious Surface

Imperviousness product is intended to represent the impervious surfaces because of urban development, layers of completely or partly impermeable artificial material (asphalt, concrete, etc.) and infrastructure construction. Therefore, the Imperviousness Degree (IMD) or Soil Sealing (SL) information can be produced relatively easily based on the Urban Extent derived from the baseline LULC information product and the linear model between impervious areas and vegetation presence that can be determined and characterized from Landsat or Sentinel-2 NDVI time series.

More precisely, the raster product is generated at 10m - 30m spatial resolution by properly exploiting Landsat-4/5/7/8 or Sentinel-2 multitemporal imagery acquired over the study area within a given time interval of interest in which no relevant changes are expected to occur (typically a time period of 1-2 years allows to get very accurate results). Each acquired EO data is pre-processed (ortho-rectification, radiometric calibration, pansharpening, cloud-masking). Then, the Normalized Difference Vegetation Index (NDVI) is extracted for each image within the urban mask (corresponding to Urban Extent product). NDVI is inversely correlated with the amount of impervious areas, i.e. the higher the NDVI is, the higher the expected presence of vegetation, hence the lower the corresponding imperviousness degree. The core idea is to compute per each pixel its temporal maximum which depicts the status at the peak of the phenological cycle. It is worth noting that for different pixels in the study area, different number of scenes might be available.

However, in the hypothesis of sufficient minimum number of acquisitions unavailable for computing consistent statistics, this does not represent an issue. Indeed, in this framework, it is also possible to get spatially consistent datasets useful for the desired analyses, even when investigating large territories. Areas associated with different levels of impervious surfaces are then extracted by visual interpretation from data sources with higher spatial resolution (e.g. VHR imagery, Google Earth imagery). OSM layers or information derived from in-situ campaigns are other auxiliary data sources which can also be used for this purpose. At the end, reference data are extracted in various parts of the study region and then rasterized and aggregated at the spatial resolution of input EO data.

A support vector regression SVR module is then used for properly correlating the resulting training information with the temporal maximum NDVI to finally derive the Percentage of Impervious Surface (PIS) or Imperviousness Degree (IMD) for the entire AOI. Specifically, 8bit integer values from the raster product range from 0 (no impervious surface in the given pixel) to 100 (completely impervious surface in the given pixel).

Summary of Processing Methods

Urban Green Areas

The location and extent of green areas are determined within the product of urban land use/ land cover at Level I. Urban green areas refer to land within and on the edges of a city that is partly or completely covered with grass, trees, shrubs, or other vegetation. This includes public parks, private gardens, cemeteries, forested areas as well as trees, river alignments, hedges etc. The product delivered within EO4SD-Urban project thus provides accurate information (1 m resolution) on the spatial location and extent of the green areas located within the Urban Extent (Level I class: 1000) derived from the baseline LULC information product.

Detecting and monitoring urban green coverage needs very high resolution optical satellite images, which explains the product generation over the Core Urban Area of AOI only. The same images have been logically used for generating the LULC information product. Consequently, the usual preliminary quality check and pre-processing tasks were already implemented.

Urban Green Areas have been detected using automated supervised classification method. More precisely, each single multispectral VHR scene has been classified by specifying the most appropriate algorithm and class number. Then, pixel units from the classes considered as representing green areas have been combined into 1 single class. From this operation results the required binary raster product. At this stage, it only remains necessary to apply some post-processing steps:

- Morphological filter is applied to fill small gaps within the green areas (caused by shadow)
- Resampling of the data to the provided spatial resolution of 1m
- Removing small pixel groups under the minimum mapping unit.
- Integrating the information provided by the LULC product (e.g. class Urban Parks, Cemeteries).
- Validation of Mapping results

Furthermore, using archive very high resolution images, current and historic extent of urban green areas are compared to identify their temporal evolution – extent growth or reduction. Quality control and accuracy assessment tasks are performed by means of visual interpretation considering also the LULC dataset.

Annex 3 – Filled Quality Control Sheets

Quality Control Sheets for the following products are provided in the form of independent documents:

- Urban Land Use / Land Cover
- Urban Green Areas

This page is intentionally left blank!