U A **B**

Universitat Autònoma
de Barcelona

Dipòsit digital
de documents
de la UAB

Generation of urban data to support climate research: developing Local Climate Zones for Oslo, Norway

Institut de Ciència i Tecnologia Ambientals

Departament de Geografia – Universitat Autònoma de Barcelona

Institut Cartogràfic i Geològic de Catalunya

Author: Laurence Sigler, 5 July 2019

Supervisors: Gara Villalba, ICTA-UAB, Joan Gilabert, UB-ICGC

Integrated System Analysis of Urban Vegetation and Agriculture (URBAG)

Abstract

Meteorological and climate predictions increasingly require more accurate modelling and data. The local climate zone model (LCZ) provides accurate and detailed information on land uses from a climatic perspective, as it integrates information on physical, radiative and metabolic aspects. These data can be used in different configurations and applications, mainly for studies of urban heat island (UHI) and heat waves. This article will demonstrate and explain three different methodologies to generate LCZs for the metropolitan area of Oslo, and the results and comparisons between them will be analysed. Next, the advantages and disadvantages of these methods will be discussed. To finish, an LCZ characterization will be made based on local surface temperature (LST) from a Landsat 8 image pertaining to a heat wave episode. This categorization will help us to know how the temperature is distributed according to land use, indicating which areas are more or less vulnerable to heat effects.

Les prediccions meteorològiques i climàtiques cada vegada necessiten una modelització i dades més precises. El model de zona de clima local (LCZ) proporciona informació precisa y detallada sobre els usos del sòl des d'una perspectiva climàtica, ja que integra informació sobre aspectes físics, radiatius i metabòlics . Aquestes dades es poden utilitzar en diferents configuracions i aplicacions, principalment per a estudis d'illa de calor urbana (UHI) i onades de calor. Aquest article demostrarà i explicarà tres metodologies diferents per generar LCZ per l'àrea metropolitana d'Oslo, i s'analitzaran els resultats i les comparacions entre elles. A continuació, es discutiran els avantatges i els desavantatges d'aquests mètodes. Per finalitzar, es farà una caracterització de les LCZ a partir de la seva temperatura en superfície (LST) mitjançant una imatge Landsat 8 pertanyent a una onada de calor que va afectar durant el juliol i agost de l'any 2018. Aquesta categorització ens ajudarà a saber com es distribueix la temperatura segons cada us del sòl assenyalant quines són les zones més o menys exposades al calor.

Las predicciones meteorológicas y climáticas cada vez necesitan una modelización y datos más precisos. El modelo de zona de clima local (LCZ) proporciona información precisa y detallada sobre los usos del suelo desde una perspectiva climática, ya que integra información sobre aspectos físicos, radiativos y metabólicos. Estos datos se pueden utilizar en diferentes configuraciones y aplicaciones, principalmente para estudios de isla de calor urbana (UHI) y olas de calor. Este artículo demostrará y explicará tres metodologías diferentes para generar LCZ por el área metropolitana de Oslo, y se analizarán los resultados y las comparaciones entre ellas. A continuación, se discutirán las ventajas y las desventajas de estos métodos. Para finalizar, se hará una caracterización de las LCZ a partir de su temperatura en superficie (LST) mediante una imagen Landsat 8 perteneciente a una ola de calor que afectó durante el julio y agosto del año 2018. Esta categorización nos ayudará a saber cómo se distribuye la temperatura según cada uso del suelo señalando cuáles son las zonas más o menos expuestas al calor.

Contents

1. Introduction

1.1 Affiliated institutions and provenance

This study is a Master of Science (MSc – *Màster Universitari*) in Geographical Information Science (*Geoinformació*) thesis (*Treball fin de Màster*), organized jointly by the Autonomous University of Barcelona (*UAB - Universitat Autònoma de Barcelona*) Department of Geography, the Catalan Cartographic and Geological Institute (*ICGC Institut Cartogràfic i Geològic de Catalunya*) and the Institute for Environmental Science and Technologies (*ICTA Institut de Ciència i Tecnologies Ambientals*). This is according to Article 75 of the statutes of the Universitat Autònoma de Barcelona and the authorization of the University President (*Rector*), signed 2019.

1.2 Orientation and scope

This study is part of the Integrated System Analysis of Urban Vegetation and Agriculture (URBAG) project of ICTA. The main objective of URBAG is to evaluate which combinations of urban/peri-urban agriculture and green spaces result in the best performance in terms of local and global environmental impacts (ICTA-UAB (URBAG), 2019). More specifically, URBAG uses two case study cities to develop these methods: the Metropolitan Areas of Barcelona and Oslo. One option to map out the vegetative and urban fractions of urban and peri urban areas is to use the Local Climate Zone (LCZ) classification.

One initial task of the project is to examine the method options available for LCZ generation, and select a way based on commonly available GIS tools and publicly-available data that would be feasible to everyday professional users.

This study generates the LCZ for Oslo, using three different methods, a comparison of the three methods involved in generation of said LCZs, their advantages and disadvantages, and a general interpretation of the correlation of the LCZs to local surface temperatures (LST). The methods are documented and explained, with corresponding results, validations and comparisons. The project is limited to the generation of a basic LCZ for the specified region of interest (ROI).

1.2.1 URBAG

The URBAG (Integrated System Analysis of Urban Vegetation and Agriculture) project is a five-year project starting September 2019, funded by the Horizon 2020 European Research Council (ERC) and hosted by ICTA-UAB.

It aims to determine to what degree green infrastructures can be a source of sustainable food, reduce environmental impacts, and promote a more efficient use of resources in urban regions. The project has two case-study cities, Barcelona and Oslo.

The three main research objectives of the project are:

1. To assess the potential of optimizing urban/peri urban agriculture and vegetation in terms of nutrient, water, and energy, considering urban morphology and weather as well as life cycle impacts.
2. To analyse the direct and indirect effects of vegetation and urban agriculture on the urban and regional atmosphere.
3. To provide urban planners and policy makers with the guidance and tools to create green infrastructure strategies to help meet sustainability targets, avoid unintended environmental consequences, and promote wider and diffused social benefits.

Within the initial tasks of the project is the gathering of accurate data, including LCZ mapping for Oslo, the principal objective of this study, as that for Barcelona was performed previously in 2018 (ICTA-UAB (URBAG), 2019).

1.3 State of the Art

1.3.1 Problems and present-day situation of cities

Cities and climate change

Cities account for 55% of the world's population, which is projected to increase to 68% by 2050, while only occupying 3% of its surface area (Bechtel, et al., 2015), which will increase given present trends of urbanization and economic migration (United Nations, 2018). This concentration of human activity in small areas typically results in an increase in Impervious Surface Areas (ISAs), closely linked to an increase in Surface Urban Heat Island (SUHI) effects and waste generation. In consequence cities are a primary catalyst for increased temperatures and the global warming effect ("climate change") (Li, et al., 2017).

Within the cities themselves the effects of climate change are especially pronounced, as they possess distinct morphologies which produce distinct microclimates and conditions. The high surface impermeability and low albedo of cities renders them vulnerable to temperature increases and weather extremes. Combined with the factors discussed above the climactic effect and pollution increase is hazardous to a city's inhabitants and can become increasingly unmanageable (O'Malley, Piroozfarb, Farr, & Gates, 2014). Cities are sensitive to the rising temperature dynamic in the short and long term (IPCC, 2014). Aggravated hot-spots associated LCZs which artificially increase local temperatures, and prolonged heat waves are occurring more frequently (Li & Bou-Zeid, 2012); (DeJarnett & Pittman, 2017); (Sheridan & Dixon, 2016), and the consequential dangers to health are a major factor of the emerging environmental crisis (Wolf & McGregor, 2013).

Solutions

To undertake any project or research to improve the design of a city or to remediate its environmental footprint, detailed and sound data is needed of the actual conditions in cities, and how their organization and structure affects the local meteorological and climactic patterns. There is a lack of accurate detailed spatial data of cities as the technology to gather it has been time consuming and costly until relatively recently (Stewart, 2011). Contemporary data gathering techniques have been greatly aided by technological advances and permit the detailed mapping of cities to fill the data gap (Tatem, Goetz, & Hay, 2008). During the last decade, the automated classification of urban structures based on improved classifiers and different datasets made considerable progress. Various techniques including SAR, hyperspectral, thermal, and VHR optical data have been used to gather detailed data of urban surfaces and structures (Bechtel & Daneke, 2012).

The advancement in satellite sensor technology has been, logically, a particularly potent boon for the field of Earth observation. The two images (Figures 1) below show the same land use mappings of a town near Barcelona (Cardedeu) using the EU Corine Land Cover classifications from 1990 and 2018 respectively. The data from 1990 used imagery from Landsat 5's LSS/TM sensors gathered between 1986 and 1998. It has an accuracy of within approximately 50m. The 2018 data is from the Sentinel 2 and Landsat 8 satellites, and has an accuracy of within less
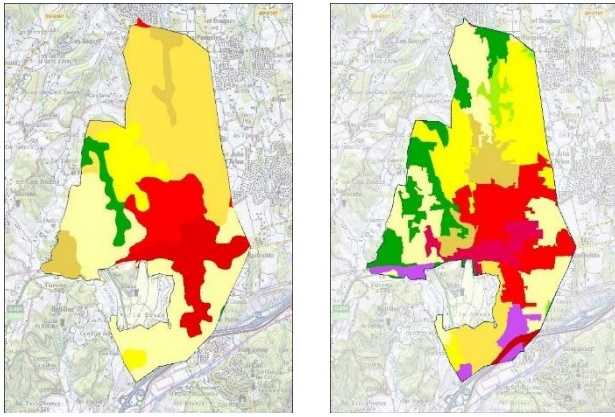
than 10m, and was gathered in less than a year (ESA, 2019). One can clearly note the difference in the level of precision between the two datasets, as the town itself has conserved its surrounding agricultural zones and limited urban sprawl.

Likewise, the technology of GIS software has improved, offering multiple options and avenues for research and practical uses. The evolution of GIS parallels that of IT in general, with thousands of applications offering solutions unimaginable only several years past (Demspey, 2014). Such specialization enables GIS use across domains and increased sharing of knowledge (1000 GIS Applications & Uses – How GIS Is Changing the World, 2019).

Figure 1, Corine Land Cover, Cardedeu, 1990(L), 2018 (R)(ESA, L Sigler)

1.3.2 Local climate zones

There had been lacking a standard methodology to characterize a city or its constituent parts, and to catalogue the properties that affect thermal variations (Stewart, 2011), which hampered a consistent approach to developing an international appraisal of the issue, though the concept of such classifications was not new. Approaching any scientific research would require a base common nomenclature for sharing data and results (i.e. the periodic table, mathematical formulae notation, et cetera.), and the Local Climate Zone classification emerged in response. (Stewart & Oke, 2012).

The local climate zone methodology as used in this study was developed in 2012 by I. Stewart and T. Oke, segregating cities into 17 classes based on their general morphologies such as land cover or building sizes, orientations and uses. Following the rationale of this method, a central business district such as in Figure 2 has an LCZ identifier of 1, and a typical suburban residential area (Figure 3) has an LCZ identifier of 6 (Bechtel, et al., 2015).

Each climate zone has a description of its physical, radiative and metabolic aspects for each land use (Appendix 2: LCZ class characteristics). This standardized information can be further detailed for each city. It signifies a new valuable impulse for meteorological and climatological modelling on the urban scale (Stewart & Oke, 2012).

The LCZ system of classification is designed to describe land cover by its surface properties, and it can be used to

Figure 2, LCZ 1, Dense high-rise. Chicago Loop (downtown)

Figure 3, LCZ 6, Open low-rise. Suburban Orlando

predict potential problem areas of cities such as areas vulnerable to SUHI and to such phenomenon as heat waves which present a risk to public health. It is one of the most useful methods for modelling climatology and meteorology on an urban scale, and it is in this use that LCZs are finding their greatest present use.

The 17 LCZ classifications exist for urbanized zones, forested areas, parks, agricultural zones, industrial estates, open paved areas (such as a port or an airport tarmac, or a parking lot), and so on as seen in Table 1. Also differentiated is the ground composition, being pervious (soil) or impervious (stone or concrete). Each classification has a standard colourized code for use in cartography. The classifications are listed in Table 1 and Appendix 1: LCZ classes and descriptions.

LCZ identifier	Type	Description
1	Compact high-rise	Dense mix of tall buildings (>10 stories). Few or no trees, mostly paved (impervious land cover).
2	Compact mid-rise	Dense mix of mid-rise buildings (3-9 stories). Few or no trees, mostly paved (impervious land cover).
3	Compact low-rise	Dense mix of low-rise buildings (<3 stories). Few or no trees, mostly paved (impervious land cover).
4	Open high-rise	Open mix of tall buildings (>10 stories). Many trees or green areas (pervious land cover).
5	Open mid-rise	Open mix of mid-rise buildings (3-9 stories). Many trees or green areas (pervious land cover).
6	Open low-rise	Open mix of low-rise buildings (<3 stories). Many trees or green areas (pervious land cover).
7	Lightweight low-rise	Dense mix of single-story buildings, of lightweight construction (wood/thatch/scrap metals). Few trees or green areas, but unpaved (pervious land cover).
8	Large low-rise	Open mix of large low-rise buildings (<3 stories). Few or no trees, mostly paved (impervious land cover).
9	Sparsely built	Sparse mix of small or mid-size buildings. Many trees or green areas (pervious land cover).
10	Heavy industry	Low-rise and mid-rise industry (factories/chimneys). Few or no trees, mostly paved or unpaved (pervious or impervious land cover).
A (101)	Dense trees	Heavily wooded, forest, pervious land cover.
B (102)	Scattered trees	Lightly wooded, forest or park, pervious land cover.
C (103)	Bush or scrub	Open mix of bush or short trees. Pervious land cover. Bush or semi-desert.
D (104)	Low plants	Pervious land cover. Grassland, steppe, agriculture, park.
E (105)	Bare rock / paved	Impervious land cover. Desert / rock, or paved (transportation)
F (106)	Bare soil or sand	Pervious land cover. Open soil / sand. Desert, beach or agriculture.
G (107)	Water	Water (Lakes, rivers, reservoirs, other bodies)

Table 1, LCZ classifications

With this system a detailed mapping can be done of an area's morphology, especially important in urban areas, which tend to be very heterogenous in their layouts and structure. The level of detail depends on the data available. With accurate contemporary imagery from Landsat or Sentinel, and accurate heights of buildings from LIDAR sources, an extremely accurate map can be developed.

For areas or climates with distinctive seasonal changes, such as with deciduous tree cover, abundant snowfall, or pronounced wet-dry seasons, further detail can be added, such as snow cover or deciduous tree patterns (Table 2). Thus, one could specify an LCZ of A for dense deciduous forest cover, with a b sub-identifier for this area in winter, or D for an agricultural zone with a d sub-identifier for a dry field in drought, fallow, or a dry season.

This research takes data from 3 July 2018 and does not include variable categorization of LCZs. Fallow fields were considered as category D (low vegetation) and not F (bare soil).

LCZ sub-identifier	Type	Description
b	Bare trees	Leafless deciduous trees. Reduced albedo.
s	Snow cover	Snow >10cm. Increased albedo.
d	Dry ground	Parched soil. Increased albedo.
w	Wet ground	Soaked soil. Reduced albedo.

Table 2, LCZ sub-classifications (variable land covers)

WUDAPT

The WUDAPT (World Urban Database and Access Portal Tools) project seeks to acquire and store urban data using a common framework and to link these data to available methods for climate analysis and for current and what-if scenario development (WUDAPT, 2012). WUDAPT promotes the creation of LCZs for urban areas via the use of a common methodology and provides the results for public use. It is a collaborative, international approach based on three levels of increasing detail: level 0 for a city-wide local climate zone format, level 1 for a neighbourhood scale format, and level 2 for a building scale format. This study develops a level 0 format.

The results from WUDAPT are ideally suited for further research, such as WRF-BEP/BEM and WRF-Chem (Weather Research and Formatting) uses, as is being done in URBAG. There are presently more than ninety cities incorporated in the project database, and the methodology is applicable to any area, based on open-source and data. Thanks to the WUDAPT portal, all the data is freely available.

2. Objectives

2.1 Research objective

The objective of the study is to map the specific local climate zones of the Oslo area for submission to the WUDAPT project. Within the URBAG project the further objectives are to provide a base land use model for further research goals, defined in Task 1.1 of URBAG: Creation of a high-resolution geo-spatial model to delineate present land uses in Oslo such as green and constructed areas, as well as the location of wastewater treatment and solid waste treatment plants (ICTA-UAB (URBAG), 2019).

The objective includes further demonstration of the viability of two more models of LCZ mapping, that of reclassification of EU Urban Atlas data, and of using Landsat data through Google Earth Engine, an online GIS suite.

2.1.1 Region of Interest

The region of interest (ROI) of the study is an area of approximately 3,700km^2 centred on the urban area (Appendix 4: Study Area – ROI within Europe) encompassing the three municipalities of Oslo, Baerum and Nittegard (Appendix 6: Oslo ROI and URBAG target municipalities). This area is larger than would have been habitually selected for an LCZ, including large areas of wilderness, however this is a requirement for the greater URBAG project. This is also indicative of the large area/low urbanization of Norwegian municipalities.

2.2 Further objectives

Other objectives included in this study, in the domain of education pursuant to an MSc, are

- Familiarization with the LCZ concept and rationales behind its creation

- Familiarization with the tools and methodologies of LCZ creation

- Generation of accurate LCZ for Oslo area via SAGA/WUDAPT method

- Generation of accurate LCZ for Oslo area via Google Earth Engine method

- Generation of accurate LCZ for Oslo area via conversion of Urban Atlas data

- Perform validation of generated LCZs

- Generation of LST data for region of interest

- Characterization of LCZ using of LST data

3. Implementation

3.1 Software and tools used

The tools used to generate the study include the following.

‹ Google Earth Pro – Complimentary client software of Google Maps/Earth websites with more advanced features, able to perform basic GIS functions and export datasets, also provides integrated Landsat and Sentinel data (Google, n.d.). Used for general reference, initial identification of urban morphologies, and generation of shapefiles for further use depending on methodology of LCZ creation.

‹ Google Earth Engine – Further iteration of Google Earth software which provides cloud computing availability for Earth imagery processing and historical analysis. Uses a JavaScript coding interface and can store geographic data. Has a large set of readily available data, specifically the same Landsat 8 imagery used throughout this research (Google Earth Engine, 2019).

‹ Google Maps – Online cartographic service, extremely popular and common use, provides both topographic and photographic services available at high resolution. Also used for general reference, navigation and identification (Google, n.d.).

‹ USGS Earth Explorer –USGS Earth observation data portal. Used to obtain Landsat 8 imagery (USGS, n.d.).

‹ SAGA - System for Automated Geoscientific Analyses, GIS software focused on raster analysis. Developed in Germany and purely open-source. Based on GDAL (Geospatial Data Abstraction Library) software, a suite of GIS tools (SAGA, n.d.).

‹ QGIS – Popular open source GIS software with a robust feature capability, also based on GDAL. QGIS functions as geographic information system (GIS) software, allowing users to analyse and edit spatial information, in addition to composing and exporting graphical maps. QGIS supports both raster and vector layers; vector data is stored as either point, line, or polygon features. Multiple formats of raster images are supported, and the software can georeference images (QGIS, n.d.).

   Although QGIS is included in the list of tools, this project is developed with ArcGIS. QGIS is included as it is a completely viable alternative.

‹ ArcGIS – Market leading GIS software, commercial product. Developed by Esri corporation and with QGIS the most common GIS used cotemporally (Esri, n.d.).

‹ Random Forest – Machine learning method for classification and regression. The name is merely coincidental and has nothing to do with climatology nor forestry. It is an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Ho, 1995).

ICGC
Institut
Cartogràfic i Geològic
de Catalunya

UAB Universitat Autònoma
de Barcelona

URBAG Integrated System Analysis of
Urban Vegetation and Agriculture

ICTA

Institut de Ciència
i Tecnologia Ambientals - UAB

màster en
geoinformació
aplicacions i gestió

‹   R and Python – Common programming languages used for reclassification in GIS and for calculating LST/LCZ correlations. Python is a popular high-level, object-oriented language and R is principally used for statistics and graphics generation.

3.2 Data used and study area

3.2.1 Oslo   study area

The capital of Norway is located at 59°55'0 N 10°44'0 E on the northern end of the *Oslofjord*, which gives the city access to the North Sea. The denser urbanization follows the coast in a compact conurbation, confined by hills (Appendix 5: Oslo area). The overall population is lower than most European capitals and spreads over a large area. It is not a particularly old city compared to its European peers, being in a different location previously and much smaller. The density of urbanization is somewhat low and much of the municipality is protected wilderness. The propensity of greenery within the urbanized zones is higher than average (Wikipedia, n.d.) for Europe.

Oslo has a humid continental climate, influenced by the coastal location, compared to other areas at the northern latitudes of the city. The Köppen classification is Dfb and is more moderate than areas further inland (Strahler & Strahler, 1994).

As of 2018, the city proper had 673,469 residents, with the urban area having approximately one million. The administrative Greater Oslo Region (*Stor-Oslo-regionen*) includes a larger area and has over 1,500,000 inhabitants (Statistics Norway, 2019).

Oslo by European standards is wealthy. It has a GDP per capita of 69,061USD (2016) and is the wealthiest area in Norway. It's GDP is 20% of Norway as a whole (OECD, 2018).

3.2.2 Landsat 8

Imagery for sampling the area was taken from Landsat 8. This satellite provides 11 bands of coverage (Figures 4 and 5). It has a 16-day revisit period for any Earth point. Along with ESA's Sentinel satellites, it is the most widely used Earth observation platform.
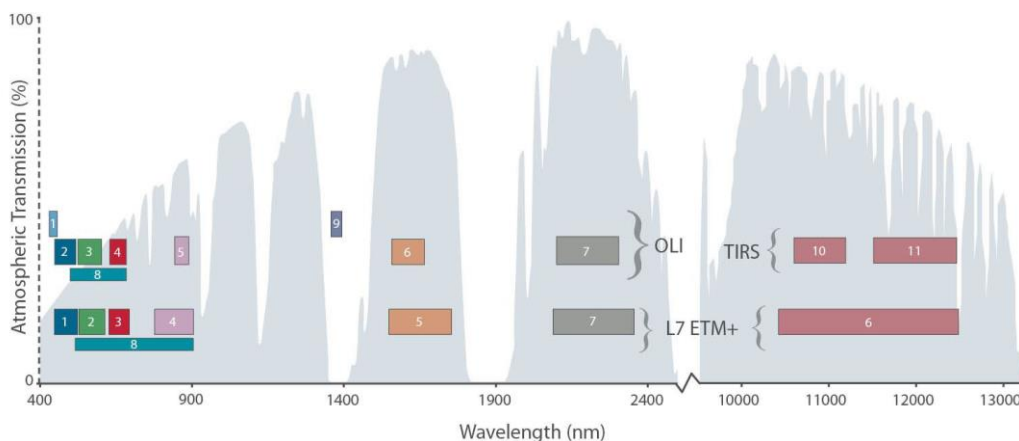

Figure 4, Landsat 8 Spectral Bands: OLI & TIRS coverages (USGS, 2019)

Landsat 8 has two sensors to provide 11 bands of data — the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The sensors have a spatial resolution of 30 meters (visible spectrum, NIR, SWIR); 100 meters (thermal); and 15 meters (panchromatic) (Figure 5).

Landsat 8 was developed between NASA and USGS. NASA led the design, construction, launch, and calibration phases, and USGS took over operations in 2013. USGS leads all post-launch activities, operations, data generation, and data archiving at the Earth Resources Observation and Science (EROS) centre in South Dakota (USA). Landsat has been acquiring around 550 scenes per day. This increases the probability of capturing cloud-free scenes of the surface. The Landsat 8 scene size is 185km wide by 180km long. The altitude is 705 km. Cartographic accuracy of 12 m or better (including compensation for terrain effects) is required for Landsat (NASA, n.d.).

	Bands	Wavelength (micrometers)	Resolution (meters)
Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) Launched February 11, 2013	Band 1 - Coastal aerosol	0.43 - 0.45	30
	Band 2 - Blue	0.45 - 0.51	30
	Band 3 - Green	0.53 - 0.59	30
	Band 4 - Red	0.64 - 0.67	30
	Band 5 - Near Infrared (NIR)	0.85 - 0.88	30
	Band 6 - SWIR 1	1.57 - 1.65	30
	Band 7 - SWIR 2	2.11 - 2.29	30
	Band 8 - Panchromatic	0.50 - 0.68	15
	Band 9 - Cirrus	1.36 - 1.38	30
	Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19	100
	Band 11 - Thermal Infrared (TIRS) 2	11.50 - 12.51	100

Figure 5, Landsat 8 band designations (USGS, 2019)

OLI – Operational Land Imager

As described on the NASA Landsat website, the OLI sensor improves on past Landsat sensors. OLI is a push-broom sensor with a four-mirror telescope and 12-bit quantization. OLI collects data for visible, near infrared, and short wave infrared spectral bands as well as a panchromatic band. It has a five-year design life (NASA, n.d.).

OLI sensor images were used for the training and classification of the Oslo urban morphology in two (WUDAPT/Google Earth Engine) of the three LCZ generation methods.

TIRS – Thermal Infrared Sensor

TIRS collects data for two more narrow spectral bands in the thermal region formerly covered by one wide spectral band on Landsat 4–7. The 100 m TIRS data will be registered to the OLI data to create radiometrically, geometrically, and terrain-corrected 12-bit data products (NASA, n.d.).

Local surface temperature data for Oslo collected by the Landsat 8 TIRS was used in the final part (LCZ/LST correlation) of this study.

Specific imagery used

3 July 2018 – Landsat cloud free image obtained for use in LCZ classification (OLI data).

30 July 2018 – Landsat cloud free image obtained for use in LST calculation (TIRS data, diurnal).

3.2.3 Urban Atlas and CORINE Land Cover

Urban Atlas (2012) – ESA Earth mapping project developed as part of the Copernicus Programme (ESA, n.d.) - provides pan-European comparable land use and land cover data for Functional Urban Areas (FUA). The Urban Atlas is a joint initiative of the European Commission Directorate-General for Regional and Urban Policy and the Directorate-General for Enterprise and Industry in the frame of the EU Copernicus programme with the support of the European Space Agency and the European Environment Agency (ESA, 2012). The land use and cover classes can be seen in Figure 6.

CORINE Land Cover data (2018) - CLC is a land categorization and classification project conducted every 6 years by the EEA, collected and implemented by EU member states (ESA, 2019).
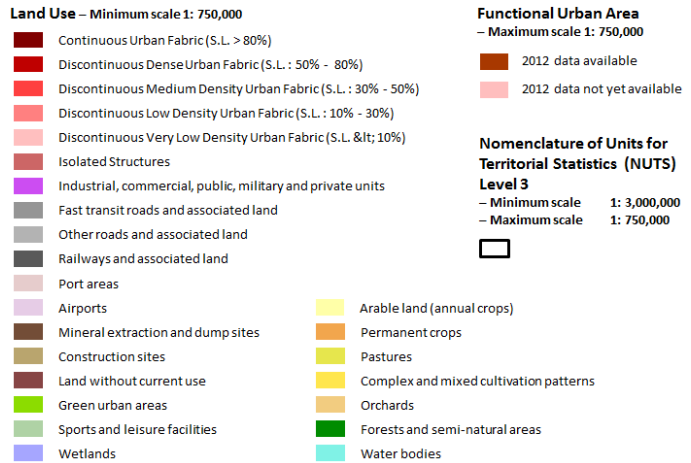
Figure 6, Urban Atlas Land Use Classes (ESA, 2012)

3.2.4 Random Forest Classification

Random decision forests are a machine learning method for classification, regression and similar tasks that construct decision trees via training and can output the classes or predictions (regression) of the individual trees (Ho, 1995).

This method was used in two of the three LCZ generations in this study, WUDAPT and Google Earth Engine.

Decision trees are popular for machine learning, as they mimic many aspects of elemental human behaviour, used in daily life. For example, if one wants to determine how hot or cool it would be, and historical records are available, one would start out with a decision of if it is summer or winter, which would give a historical basis to assume the temperature would be above or below a certain temperature. From this decision one would have a reduced range of temperatures, and an average for the day in question. From this temperature average, one would look at the actual temperature, and again the range is reduced, and from this reduced range one could estimate the temperature coming tomorrow.

> *Initial temperature range -> reduced by season -> reduced by historical average -> reduced by actual -> prediction*

Or to purchase a new automobile:

> *I like this ─ -> is it efficient -> is it in my budget -> is it fast -> buy it*

This prediction, when applied to LCZ generation is obviously much more complex, and in this case alien to human thinking, as humans have much more practical knowledge and information acquired from experience. Whereas a human would look at an area of LCZ 6 and easily identify it, a machine has no relevant knowledge of what a house is nor what LCZ 6 is, nor how it relates to anything else. Identifying it via training areas gives an input from which the machine, via decision trees, can make a relatively accurate estimation of where other examples occur.

For the purposes of regression, or the prediction of the LCZ classes, the simplified process is:

$$R = sum(y - prediction)^2$$

Where R is regression.

For classification, the formula is:

$$C = G = sum(pk * (1 - pk))$$

Where C is classification.

3.3 Methodology

Apart from the principal objective to map a local climate zone scheme of Oslo area for further use in the URBAG project, it is included to explore three possible methodologies for LCZ generation of the same region. These methodologies are the WUDAPT method, the method of reclassification of Urban Atlas data, and the use of Google Earth Engine.

3.3.1 LCZ mapping

Creation of LCZ via WUDAPT method
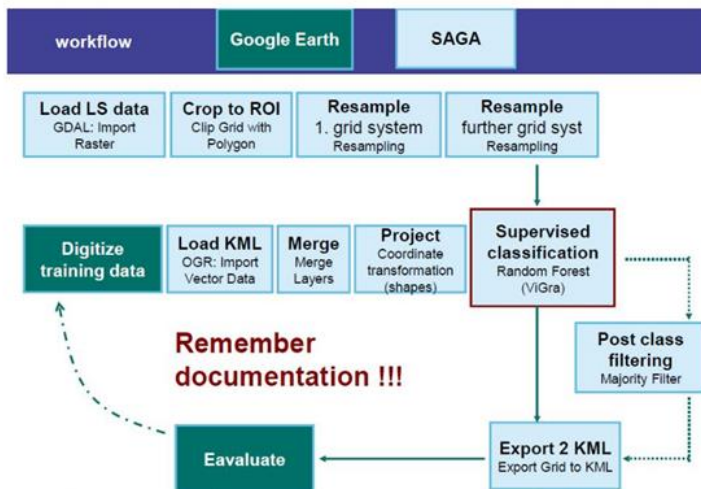
Figure 7, WUDAPT workflow for LCZ creation (WUDAPT, 2012)

The WUDAPT project has developed a standardized approach to generate LCZs for any city using freely available imagery and software.

Their aim is to have cities included on the project website to be freely available for use by researchers. For this their standardized methodology is documented on the WUDAPT website.

The process allows the mapping of LCZs for any area where a Landsat image is available.

Although documented and with examples, a knowledge of geography, information technology and climatology are useful for performing the procedure.

The procedure follows the following workflow (Figure 7):

1. Create ROI in Google Earth Pro according to needs
2. Create TAs in Google Earth Pro, for each LCZ type present
3. Create control (test) TAs in Google Earth Pro, roughly half of the number previously created
4. Obtain Landsat 8 imagery which covers the ROI, with minimal cloud cover
5. Import Landsat 8 imagery (raster) into SAGA
6. Import ROI and TA KML (vector) files into SAGA
7. Adjust spatial references as needed for coherence

8. Clip raster to the ROI
9. Resample raster to 100m^2
10. Execute Random Forest classification
11. Evaluate (repeat if necessary)
12. Validate

First steps

For the process of mapping an LCZ it is first necessary to specify a Region of Interest (ROI), which is the area included in the study and which would be the subject and boundary limits of all activities moving forward. This is done via Google Earth Pro software, as it has global coverage and provides an easy and rapid manner to generate the initial datasets used further on in the process.

Creation of ROI and TAs

In Google Earth Pro, an area including the Greater Oslo area is marked via the polygon drawing tool. This is an area of 3,702 km^2 containing the municipalities of Oslo, Baerum and Nittegard as specified by the URBAG requirements.

Upon creating an ROI (Figure 8) as per WUDAPT guidelines, it is important to include the entire urban area while avoiding unnecessary areas. This is problematic with Oslo, as it is required to include the three municipalities which are included in the URBAG project. In Norway municipalities cover large areas, while the urbanized zones tend to be small and compact. This resulted, as would be a logical assumption, in a preponderance of LCZ zone A (Dense trees) in the results of the study, the most typical unaltered land state of Norway.
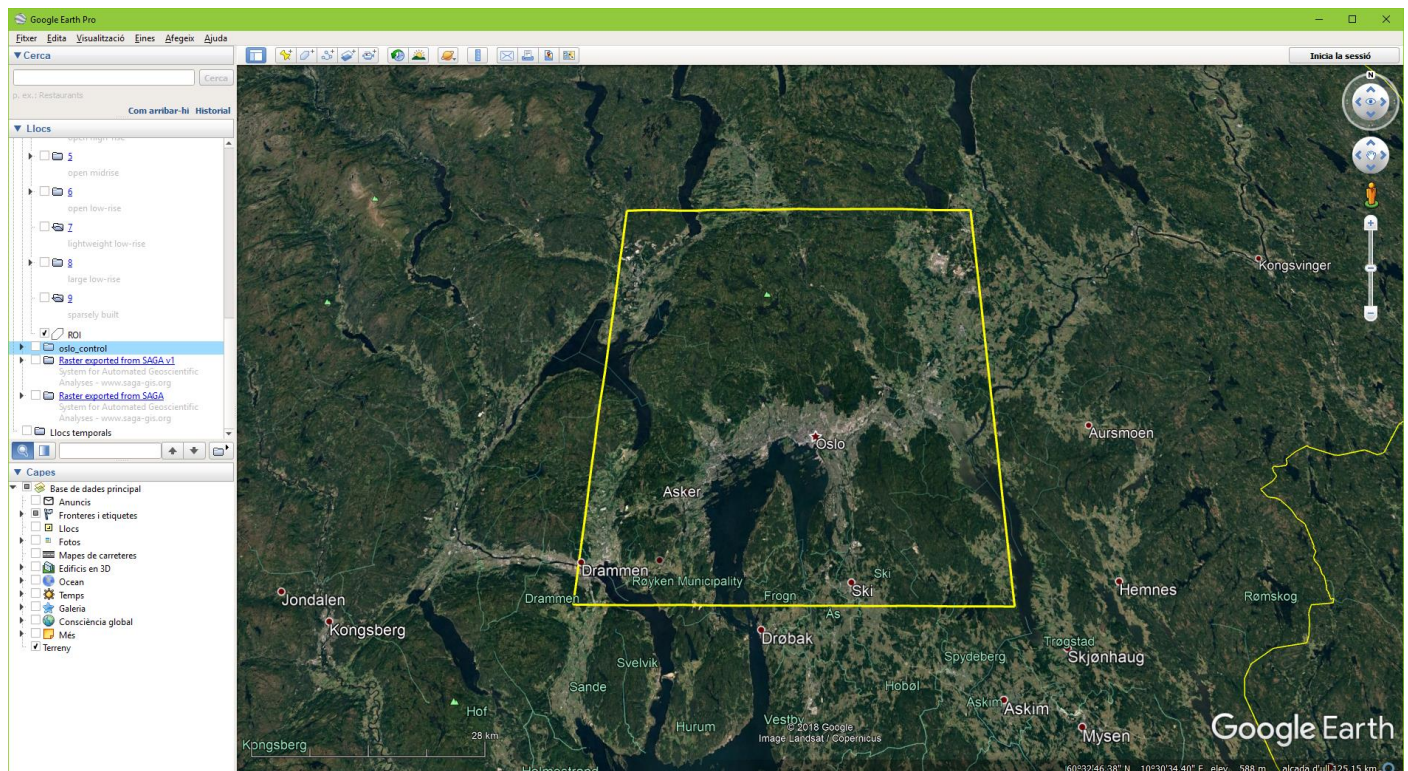

Figure 8, ROI polygon created in Google Earth Pro, with Oslo at centre

Generation of Local Climate Zones for Oslo Metropolitan Area - 16

Creation of training areas (TA)

Once the ROI is designated, it is necessary to provide the training algorithm (Random Forest) with data to analyse. Within the area, samples of each LCZ type are marked (in Google Earth Pro) by creating more polygons. For example, to designate a training area (TA) of type LCZ 1 one must find an area of Oslo with dense construction of buildings more than 25m tall and outline a polygon around it, categorizing it under LCZ 1 in the expandable tree in Google Earth Pro. Training areas should be between 150-1000m wide and not overlap nor abut others. As the Landsat data used is at a resolution of $100m^2$ cells, the TAs are drawn to be larger than this at least. Approximately 15-20 areas are marked for each LCZ type present in the city. Certain areas are not present such as LCZ type 7 (lightweight low-rise, indicative of shanty-town areas), or LCZ type 4 (open high-rise, indicative of large-scale suburban office or residential areas). Nor is present type C, scrub or semi-desert bushland, atypical of Scandinavia.

Avoided are construction sites and other temporal features, a potential issue if the satellite imagery is out of date. As this project uses data from 2018, this is a medium level concern – while the data is recent, many construction works are observable in the city during the specification of the TAs. The heights of buildings were estimated based upon the 3D view available in Google Earth Pro. More detailed data (LIDAR) would be more precise.

Once all the training areas are created they are saved as a keyhole mark-up language (KML) file which contains the geometry of each of the polygons (the ROI and TAs), they are vector files easily imported into other GIS clients.
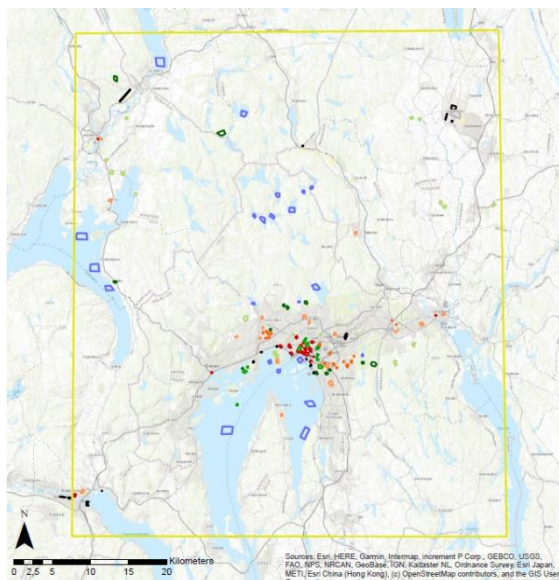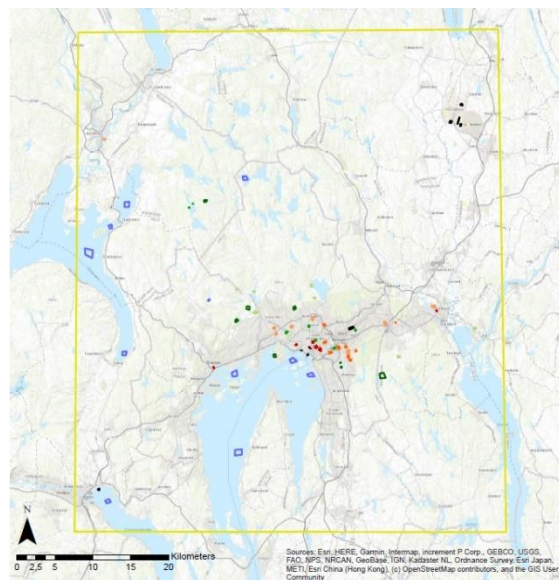

Figure 9, WUDAPT training areas (TA)

Figure 10, WUDAPT control/test TA

Creation of control (testing) TAs

The process of creating TAs is repeated to create another set of analagous TAs, in different locations, and approximately half as many per category. These "control" TAs are not to be used to train a classification with the WUDAPT method, but instead to be used to validate the efficacy of the Random Forest classification (the generation) of the LCZs afterwards.
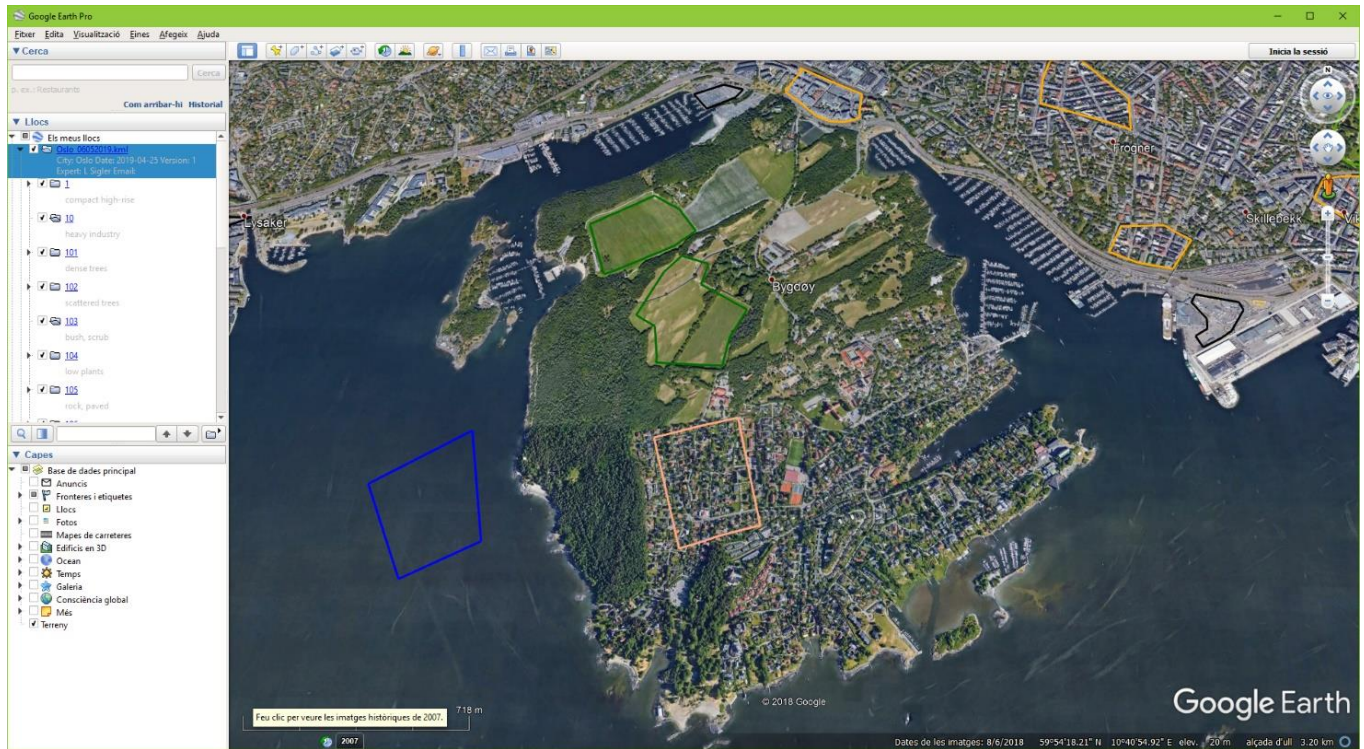
These areas are also saved as KML files.

Figure 11, TAs in Google Earth Pro: LCZ types 2 (orange), 6 (pink), D (green), E (black) and G (blue)

Obtention of Landsat satellite imagery

For the imagery used to classify the designated areas (the previously created TAs within the ROI), images from the above described Landsat 8 satellite are used. Specifically, Landsat 8 level 1 data (including OLI and TIRS) of all 11 bands, with minimal cloud cover, during the summer, was searched, to coincide with an area over and including the ROI for Oslo, as shown in Figure 12, (Appendix 7: Landsat 8 swath, Southern Norway with ROI (3 July 2018)).
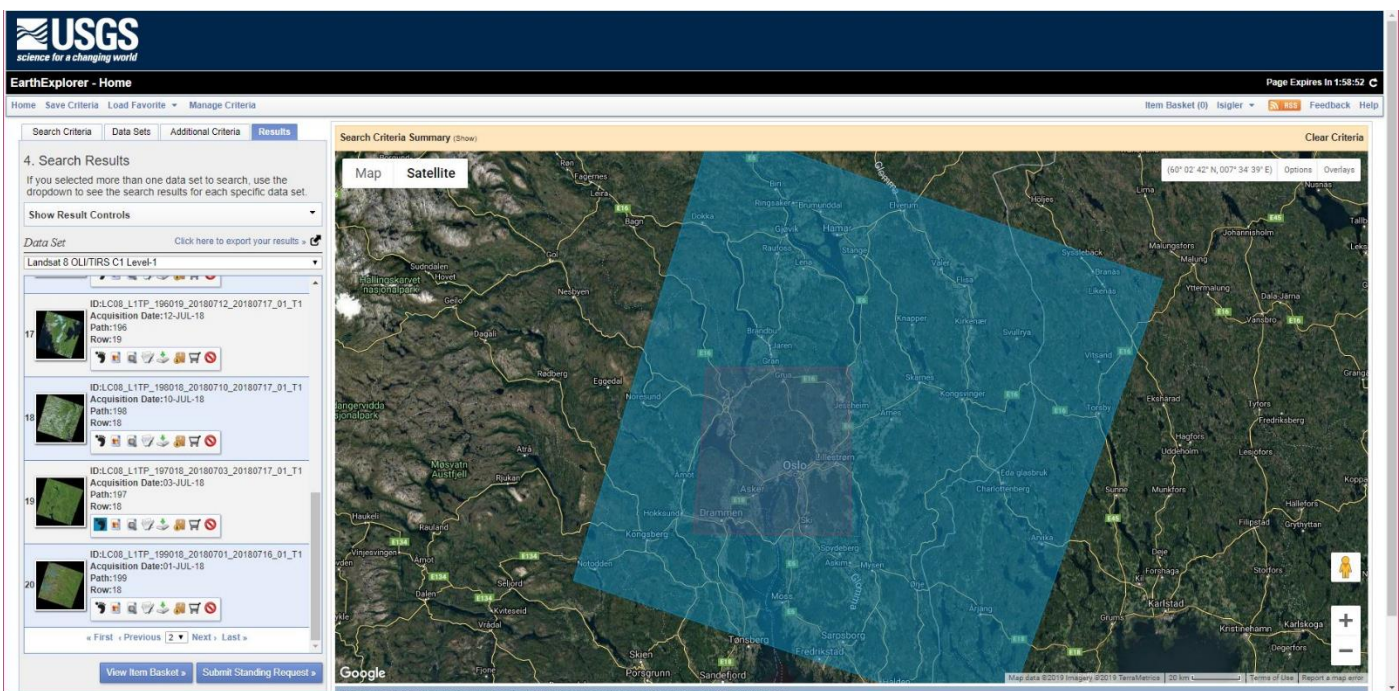
Figure 12, USGS Landsat 8 Earth Explorer site prior to download of imagery (area covered by pass on 3 July 2018 in blue), with ROI for URBAG project requirements (red, inset)

Generation of Local Climate Zones for Oslo Metropolitan Area - 18

These images are freely available from the United States Geological Survey (USGS). The search parameters are customizable, by detail level, sensor and cloud cover. Days with low or no cloud cover were specified. Given the 16-day return period, there was more than one set of imagery to choose from.

An appropriate image was found taken during the satellite's pass over Oslo on 3 July 2018 which completely covers the ROI and has no cloud cover.
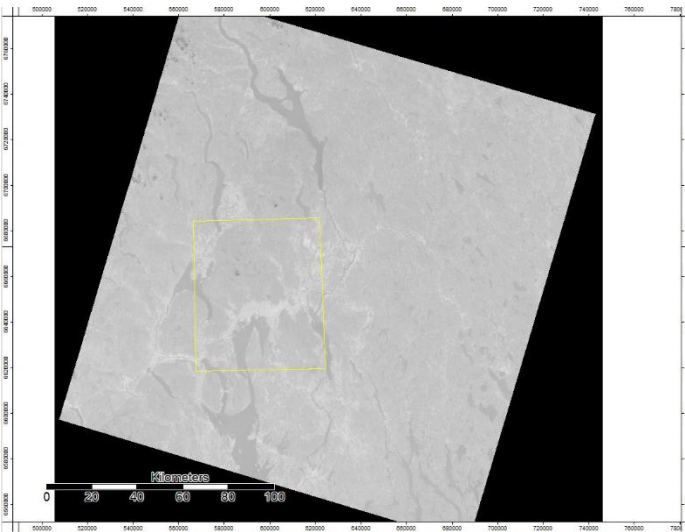
Importation and conversion of ROI, TAs and Landsat data via SAGA GIS

Once created the training areas, and the satellite imagery obtained, the two are combined for analysis using SAGA, an open source GIS client specialized in raster manipulation.

The satellite images are imported via SAGA's raster import functionality, while the ROI and TAs are imported via the program's vector import functionality from the KML files previously created via Google Earth Pro. Both functionalities use the GDAL (Geospatial Data Abstraction Library), an open source library for GIS functions, also found in QGIS.

The imported files are all set to the same spatial reference, EPSG 32632 UTM 32N, though for final display purposes (maps), the reference is set to ETRS89 NTM Zone 10, centred on Oslo for legibility.

Figure 13, Landsat 8 imagery and ROI vector data (yellow) combined in SAGA. Oslo can be seen within the ROI (light colours)

Once both datasets are imported into SAGA, the 17 separate TA files and the ROI (the KML files) are merged into one vector file via the merge layers function of SAGA. Figure 14 shows the merged TAs and ROI without the satellite imagery present. These files are set to the same spatial reference as previously described. Figure 15 shows all the data (ROI and TA vector and satellite imagery raster files) gathered combined in SAGA ready for the following steps.
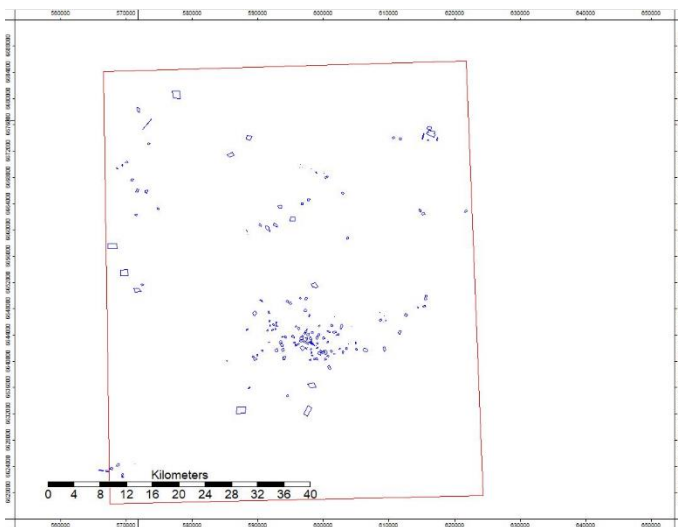

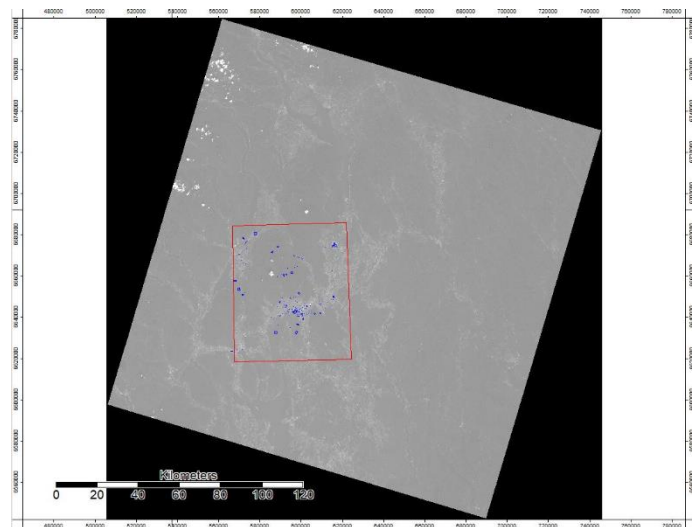*Figure 14, ROI and TA vector files displayed in SAGA*

Figure 15, ROI, TAs and Landsat data displayed in SAGA

Clip raster to ROI and resample

This imagery is then clipped to the ROI, as most of the area covered by the Landsat image is outside the area covered by the project. SAGA applies a-colour palette upon doing this, as seen in Figure 16. These values are not related to the LCZ classification.

After the datasets have been merged and clipped correctly, it is necessary to resample everything to a spatial resolution of 100m^2, as the Landsat scenes vary by sensor used: the bands 11 and 12 (TIRS) are 100m, the band 8 (panchromatic) is 15m, and everything else is 30m. As URBAG requires the spatial resolution to be 100m^2, as the intended use is for use of the layers for modelling and forecasting, the rasters are all resampled as such in SAGA. This generates new files and the old rasters are removed and saved.

At this point the steps are completed to begin the classification/training process, using the Random Forest machine learning method incorporated in SAGA. This is done by selecting the satellite imagery, and by inputting the training areas. This process

Figure 16, Raster clipped to ROI in SAGA

results in the classification of the imagery within the ROI according to the 17 categories from the TA samples. The first attempt at this process is shown in Figure 17.
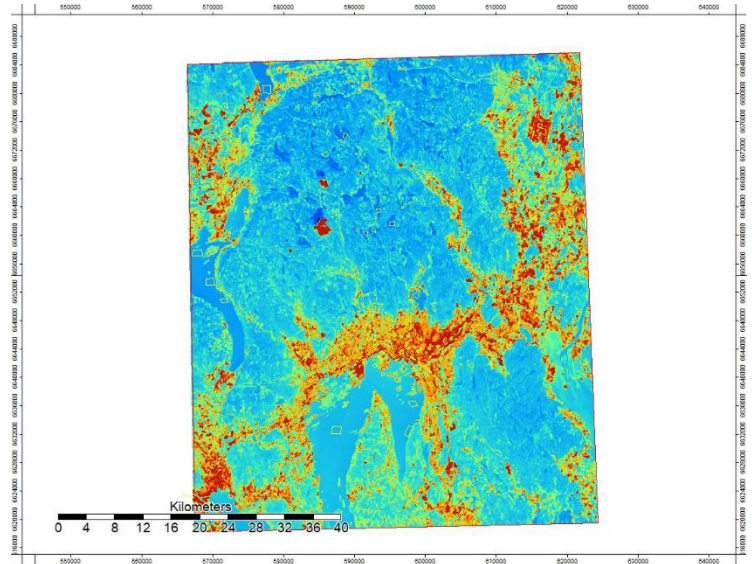
The image in Figure 17 still does not have the correct WUDAPT colour palette applied, and the software applies a random colour scheme. There are errors of misclassification of certain areas. Large sections of water area (LCZ G) are classified incorrectly as soil (LCZ F), and certain areas of soil (LCZ F) are classified as compact mid-rise (LCZ 2). The results are exported and colourized to the WUDAPT palette and can be seen in Figure 18.
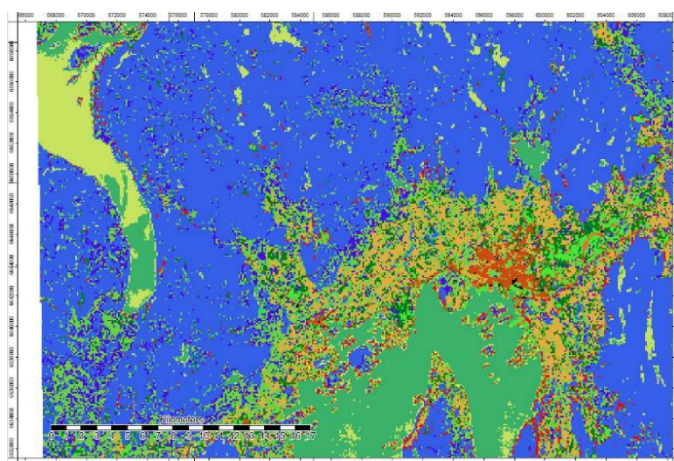

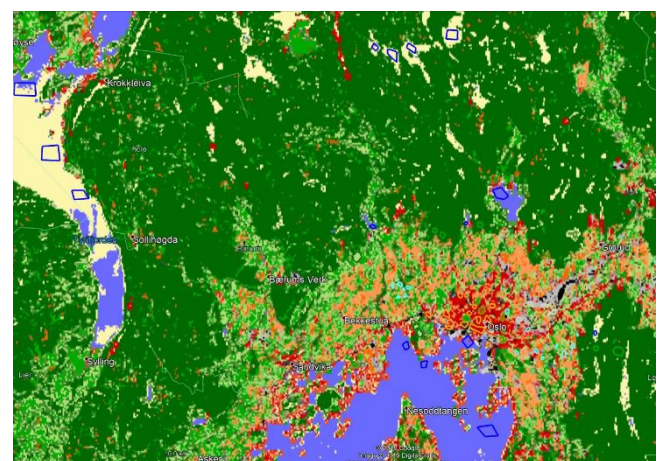Figure 17, Random Forest classification in SAGA, 1st attempt

Figure 18, LCZs, 1st attempt, WUDAPT colour palette applied

The misclassification problem is solved by adding additional training areas, specifically over areas that the Random Forest classification has difficulty with. The water areas in question are shallow or coloured somewhat differently in the original satellite image (silty) or have colouring like other LCZ types (concrete coloured). The second

Generation of Local Climate Zones for Oslo Metropolitan Area - 20

reclassification is shown partially in Figure 19 with all zones more accurately identified, and no large anomalous areas.

The LCZ results can be seen in Appendix 8: Oslo LCZ via WUDAPT method: 100m^2 resolution.
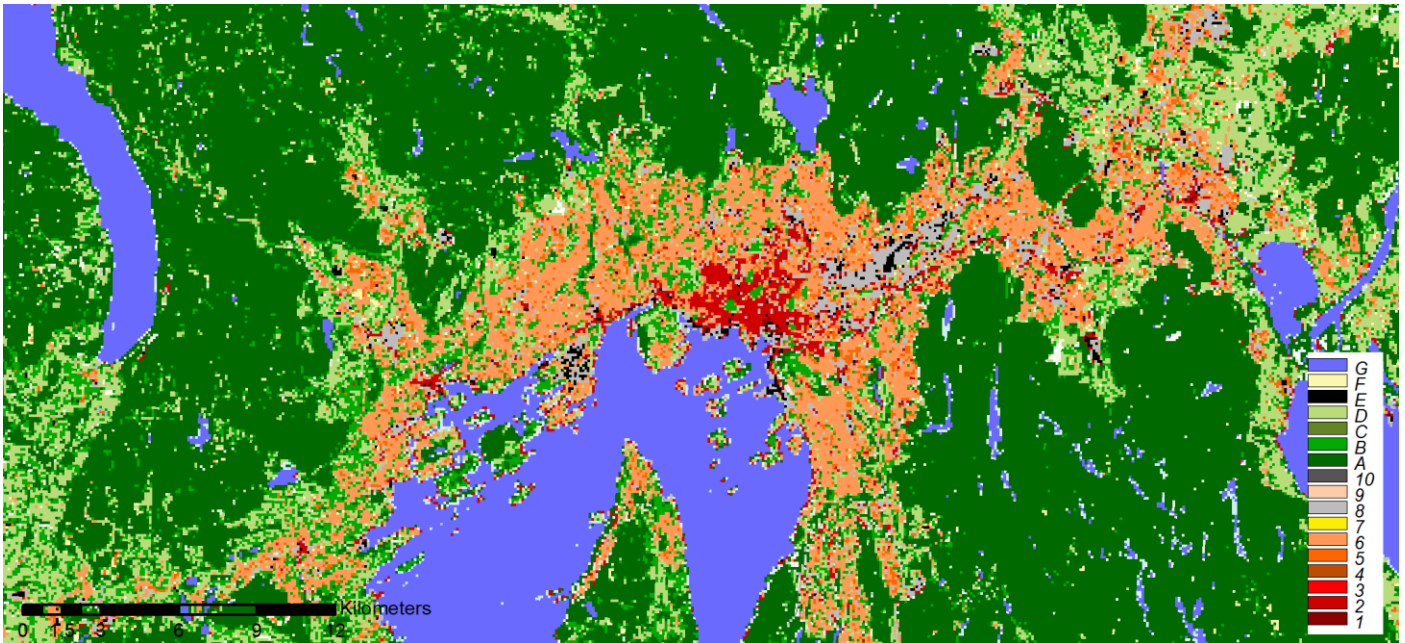

Figure 19, Final WUDAPT LCZ classification, inset (Oslo at centre)

Creation of LCZ via Urban Atlas

For the second method of creating a local climate zone, ESA's Copernicus' Urban Atlas data from 2012 is used. This dataset has available data only from the official area of Metropolitan Oslo (Appendix 9: Urban Atlas (Metropolitan Oslo) with ROI), which does not correspond to the area covered by the ROI. This data is reclassified from its native categories into LCZ categories as seen in Table 3, with a corresponding LCZ type., via a python script in ArcGIS (Appendix 13: Code used). This data does not include height information of the structures; therefore, assumptions and generalizations of classes must be made. This is a limitation of the UA method of LCZ generation.

Urban Atlas Type	LCZ Type	Urban Atlas Code	LCZ Code
Continuous urban fabric	Compact	11100	1, 2 or 3
Discontinuous dense urban fabric	Open	11210	4, 5 or 6
Discontinuous medium density urban fabric	Open	11220	5 or 6
Discontinuous low-density urban fabric	Open	11230	6
Discontinuous very low-density urban fabric	Open	11240	6
Isolated structures	Sparsely built	11300	9
Industrial, commercial, public, military and private units	--	12100	1-6
Fast transit roads and associated land	Bare rock or paved	12210	E (105)
Other roads and associated land		12220	E
Railways and associated land		12230	E
Port areas		12300	E
Airports		12400	E
Mineral extraction and dump sites		13100	E
Construction sites	Not calculated	13300	E
Land without current use	Not calculated	13400	F
Green urban areas	Scattered trees or low plants	14100	B or D
Sports and leisure facilities	Low plants	14200	D
Arable land (annual crops)	Low plants	21000	D
Permanent crops	Low plants	22000	D
Pastures	Low plants	23000	D
Complex and mixed cultivation patterns	Low plants	24000	D
Orchards at the fringe of urban classes	Scattered trees	25000	B
Forest	Dense trees	31000	A (101)
Herbaceous vegetation associations	Scattered trees	32000	B (102)
Open spaces with little or no vegetation	Bush, scrub	33000	C (103)
Wetlands	Scattered trees or low plants	40000	B or D
Water	Water	50000	G (107)
No data (cloud/shadow)		91000	
No data (no image)		92000	

Table 3, Urban Atlas reclassification matrix

Performing a clip of the Urban Atlas shapefile to the ROI gives a result that covers roughly two-thirds of it. The area not included has no data will not enter into the results (Figure 20).

Once reclassified in ArcMap, the new LCZ values can then be symbolized in ArcGIS and show the now familiar LCZ output for Oslo, similar to the WUDAPT results, as seen in Figure 21.
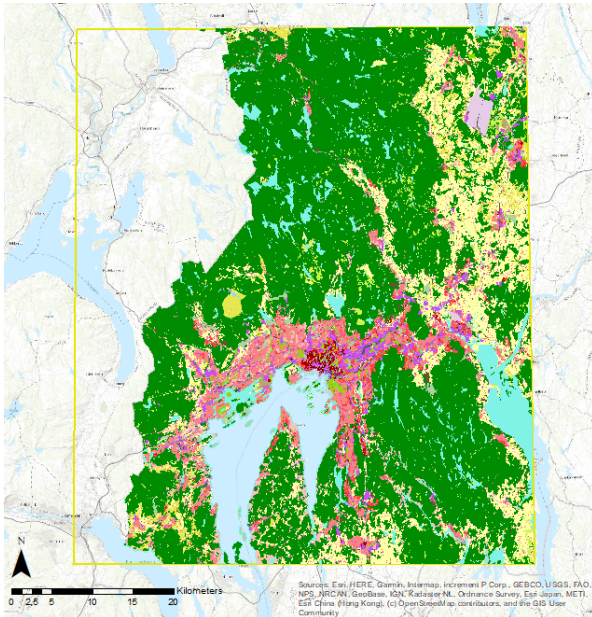
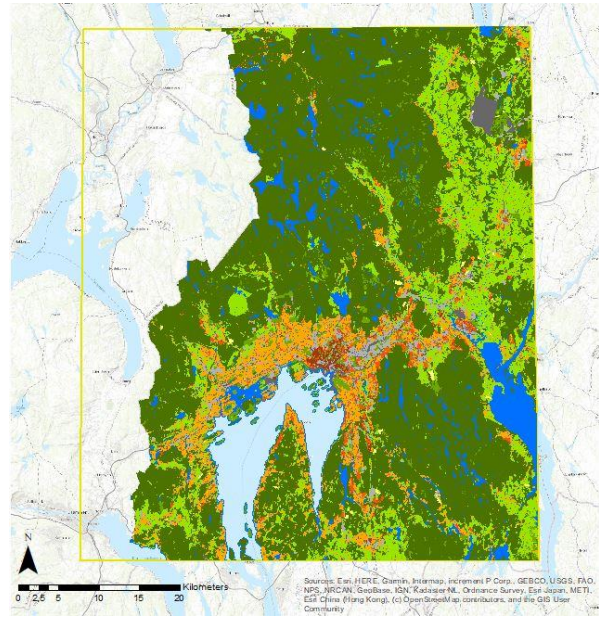*Figure 20, Urban Atlas dataset clipped to ROI with native symbology*

Figure 21, Urban Atlas dataset clipped to ROI with reclassified symbology (LCZ)

The dataset is converted to a raster format and resampled to a 100m² resolution for further use involving LST correlation to local climate zones.

Creation of LCZ via Google Earth Engine

The final method of LCZ creation is performed using Google Earth Engine, all online. To illustrate the functionality screenshots are shown.

The first step is to obtain Landsat 8 imagery, for consistency, the same imagery from the pass of 3 July 2018 is used, a clear day in the summertime. Bands 4, 5 and 6 are used, as they are the most adequate for the application of the

Figure 22, Selection of Landsat 8 characteristics in GEE

Random Forest algorithm (Figure 22). The software only allows the selection of three bands, which results in slightly differing results compared to the WUDAPT method.
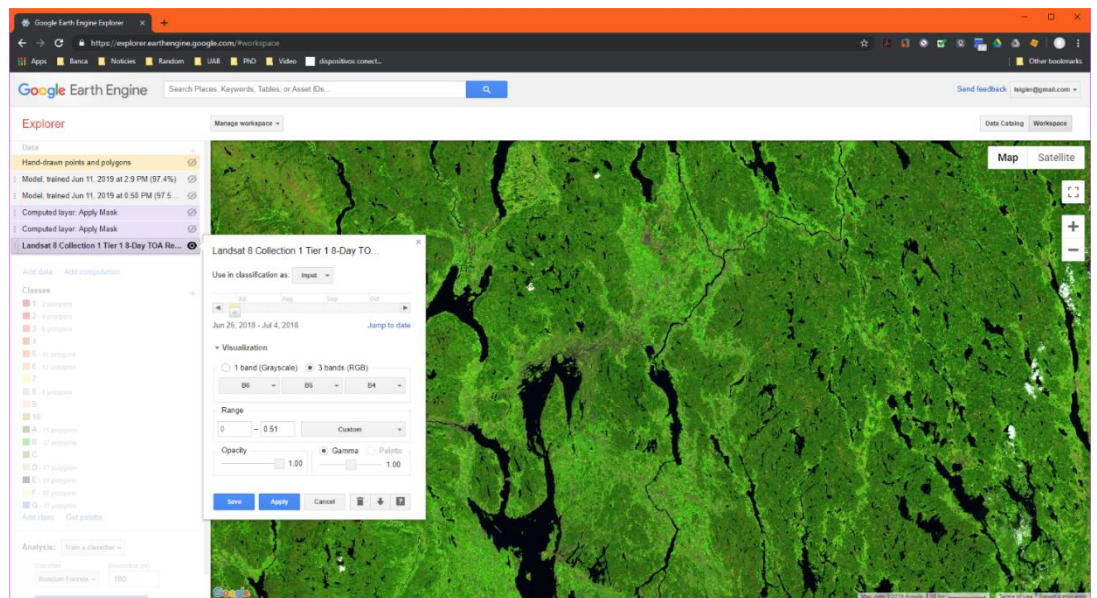
After the satellite imagery has been selected, an area of the ROI is delimited, and applied in the software, by selecting one band (only a clipping action is being performed) and drawing a rectangle over the area in question. This step creates a mask which is then applied to the previously selected satellite imagery (Figures 23, 24).
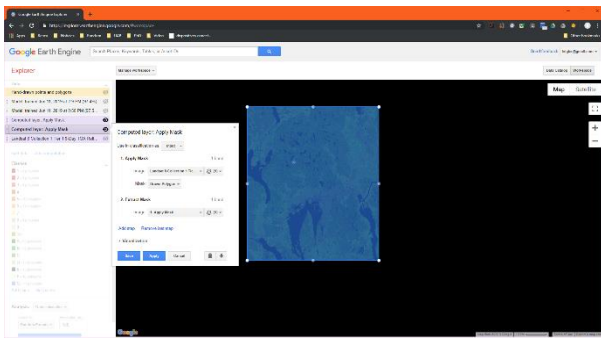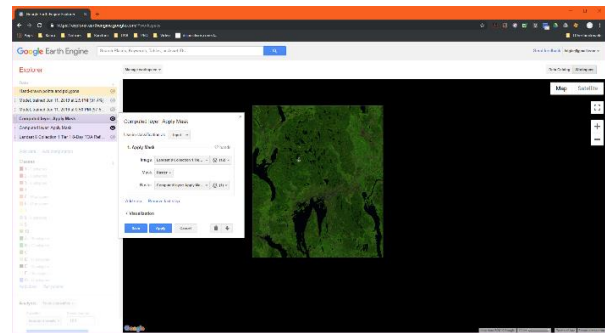

Figure 23, Creation of ROI mask

Figure 24, Application of ROI mask to Landsat image

Once created the ROI, it is necessary to create the training areas. For consistency, the same TAs from the WUDAPT process are used. The first step is to create the classification classes in Google Earth Engine, 17, for each type of LCZ, as seen in Figure 25. The WUDAPT colour palette is used.

The TAs were previously imported from KML and merged in SAGA, and used in the previous operations, so these already exist as a shapefile ready to use.

The shapefile is imported as a Google Fusion Table, which replicates the standard attribute table (*.dbf format) of the shapefile, including all the geometric (geographic) information. It is necessary to alter the data type of the source column (which contains the LCZ type) in the shapefile to numeric, as the software will not recognize and classify on text values. Once imported as a Fusion Table, an alphanumeric identifier is given for the table, which can be referenced in the Earth Engine Explorer, by selecting the appropriate column which has the LCZ class (*source* from the TA shapefile) and assigning it to the corresponding classification previously entered in the software (Figure 25).
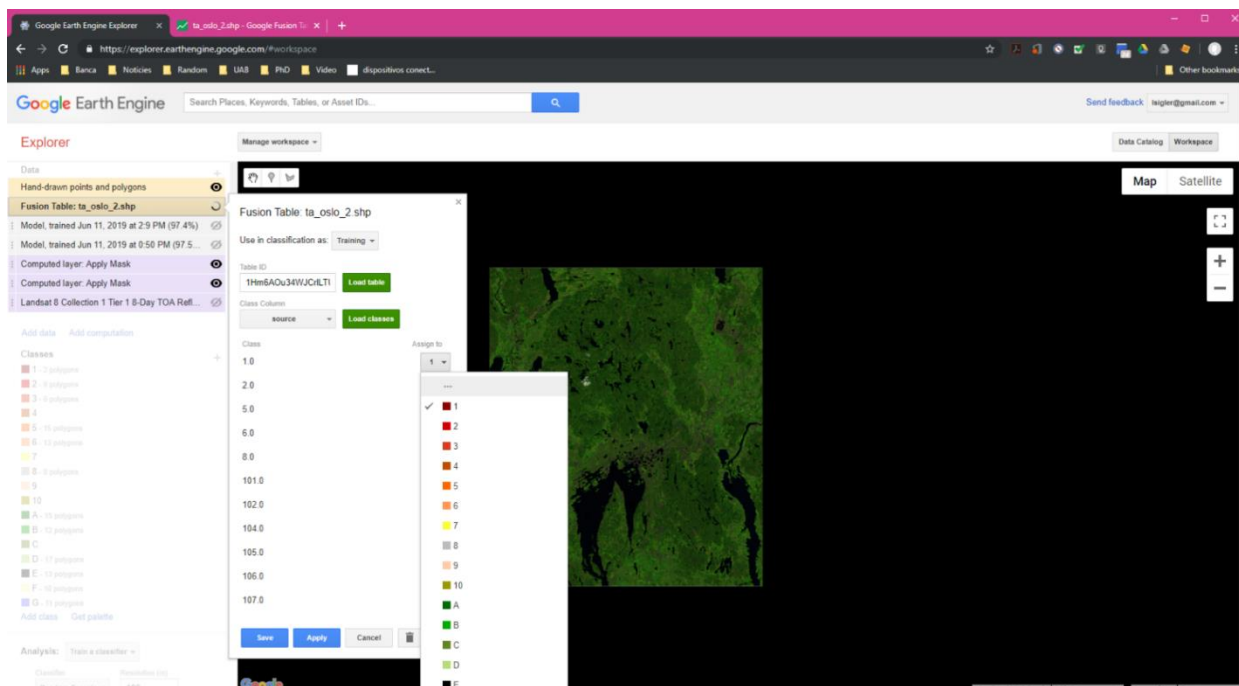

Figure 25, Matching shapefile TA LCZ values to Google Earth Engine classes

The Earth Engine software has the Random Forest functionality included, which is trained on the imported shapefile at a resolution of 100m^2, identical to the WUDAPT data.

The results of the procedure from Earth Engine Explorer can be seen in Figure 26 and (Appendix 11: Oslo LCZ via Google Earth Engine method: 100m^2 resolution).
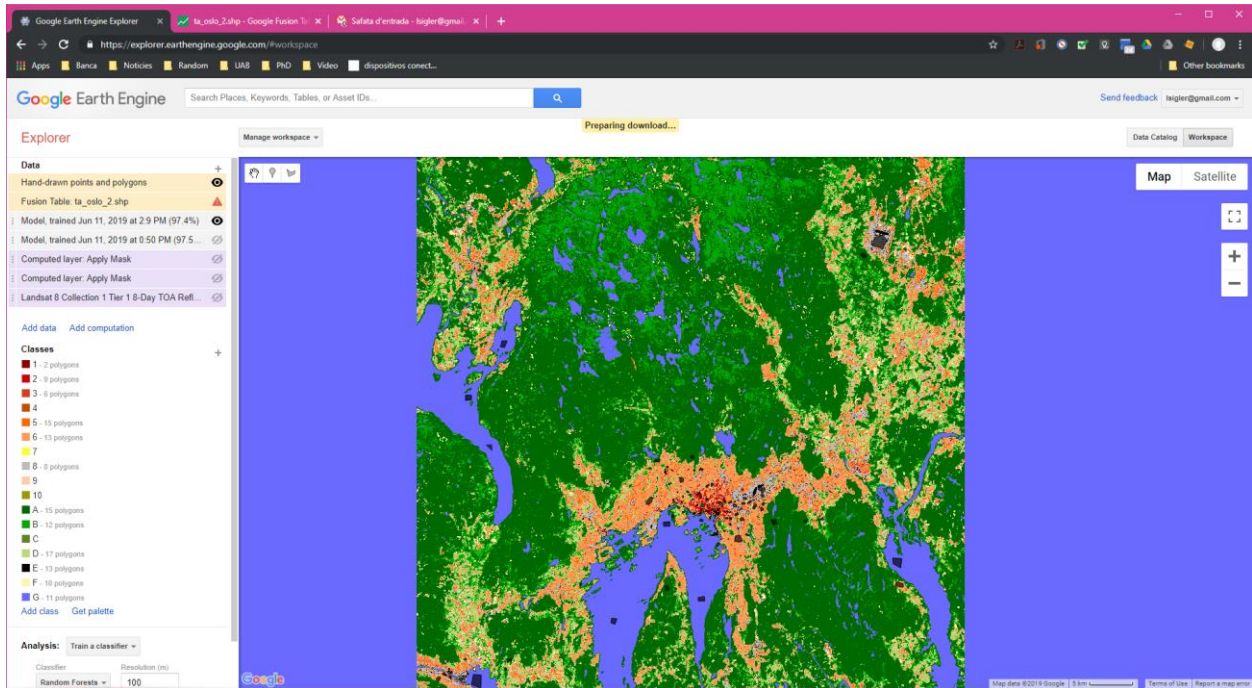

Figure 26, LCZ results for Oslo ROI via Google Earth Engine

3.3.2 Validation

WUDAPT results validation

To perform a validation of the classification of the WUDAPT method the following methodology is followed.

The raster file from the Random Forest classification is converted to a point vector layer, with each point representing a pixel from the raster, or 100m^2 of the ROI.
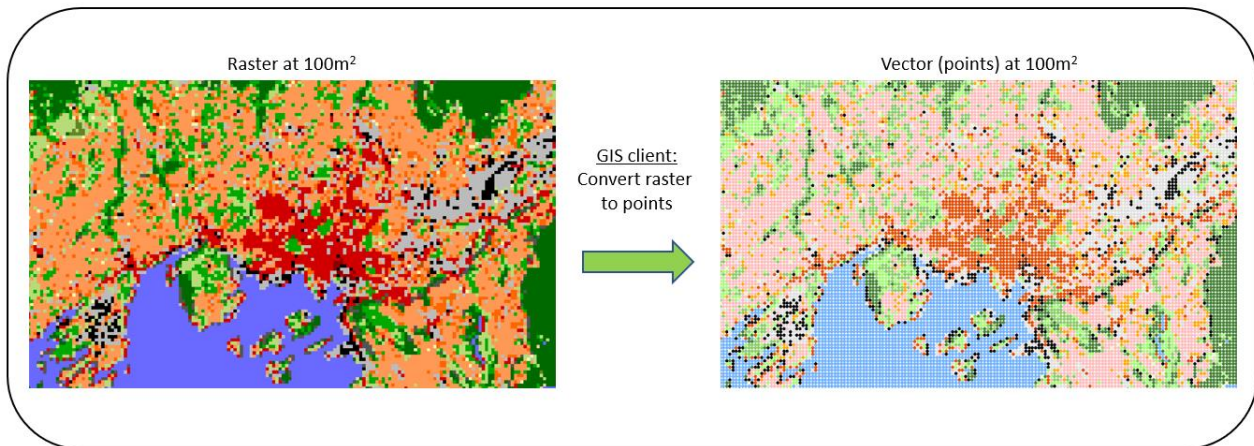

Figure 27, Schematic of raster-point conversion

This new layer is superimposed onto the polygon layer of the training areas and clipped to match each individual TA. It is then a matter of determining the values of the points (the LCZ values) within each training area. The procedure followed is, in the ArcGIS client:

Raster to points -> Clip new point layer to TAs -> Spatial join -> Summary statistics
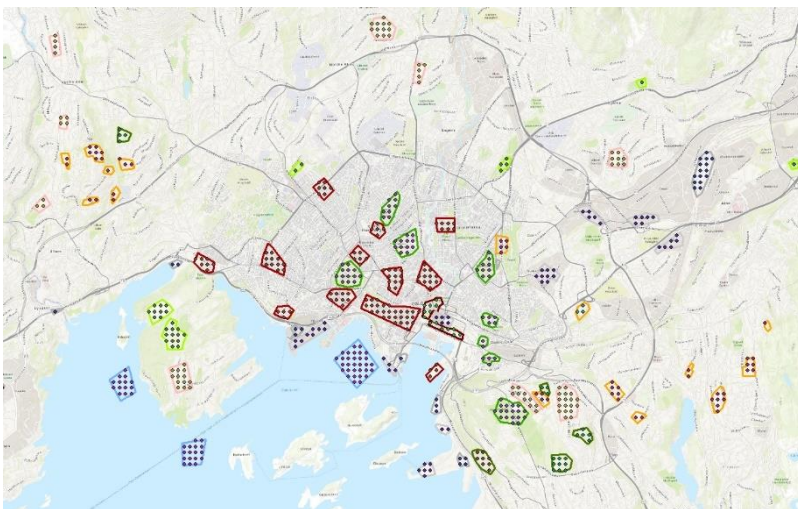

Figure 28, Spatial join of WUDAPT training areas and points (pixel values)

From this operation a combined layer of points and polygons is generated (Figure 28), and an attribute table of which points (corresponding to the original pixels from the Landsat image and from the results) fall in each TA was generated, and from this table the accuracy of the Random Forest training and classification results can be validated.

For this purpose, an Excel spreadsheet is used, with the required formulas inserted to generate the matrix.

measures of validation used are described as:

1. *Kappa Coefficient*, or # ˙M ˙M: measures the percentage of values in the main diagonal of the contingency table and then adjusts these values for agreement that could be expected due to chance alone.

$$k = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

Cohen's kappa measures the agreement between two raters, or classifiers, who each classify *N* items into *C* mutually exclusive categories. Where *Pr(a)* is the relative observed agreement among raters, and *Pr(e)* is the hypothetical probability of agreement, using the observed data to calculate the probabilities of each observer randomly seeing each category (Cohen's Kappa, n.d.).

2. *Overall Accuracy (OA)*: measures the total classification precision.

$$OA = \frac{\sum xi = i'}{2a\sum TA}$$

Where *x(i)* is the number of pixels that has each specific category. The Local Climate Zones (LCZ) are validated by means of training areas (perfect prognostic) and by control/test areas. These test areas are not used to classify the LCZ however are developed with the same criteria of the training areas (TAs), for the sole purpose of testing the classification and training process.

The WUDAPT validation results from the training areas can be seen in Table 4. This exercise resulted in a kappa of 0.985 and an OA of 0.985. A kappa of over 0.6 is considered optimal for use in WUDAPT, therefore the results generated are very accurate.

LCZ	1	2	3	5	6	8	A	B	D	E	F	G	TB	
1	14												14	1,000
2		130	2			1							133	0,977
3			25	3									28	0,893
5				52									52	1,000
6					164							6	170	0,965
8						165							165	1,000
A							157						157	1,000
B								93	19				112	0,830
D									165				165	1,000
E										95			95	1,000
F											35		35	1,000
G												952	952	1,000
TA	14	130	27	55	164	166	157	93	184	95	35	958	2078	0,46102
	1,000	1,000	0,926	0,945	1,000	0,994	1,000	1,000	0,897	1,000	1,000	0,994		

Table 4, Contingency matrix for WUDAPT training areas

The WUDAPT validation results from the control/test areas can be seen in Table 5. This exercise resulted in a kappa of 0.869 and an OA of 0.868.

LCZ	1	2	3	5	6	8	A	B	D	E	F	G	TB	
1	5					1							6	0,833
2	1	20		1		4				8			34	0,588
3			4	2									6	0,667
5				10	8								18	0,556
6		2		23	46		2	1	4				78	0,590
8		3		1		32				11			47	0,681
A				2			98	1			3		104	0,942
B				3			1	13	5	2	2		26	0,500
D							3	4	27		13		47	0,574
E		1				8				17	1		27	0,630
F				2							8		10	0,800
G												397	397	1,000
TA	6	26	4	41	57	45	104	19	36	38	27	397	800	0,496
	0,833333	0,769	0,000	0,561	0,807	0,711	0,942	0,684	0,750	0,447	0,481	1,000		

Table 5, Contingency matrix for WUDAPT test/control areas

The results are positive and acceptable, the classification has slight difficulty with the LCZ zones 2 and 5, and 8 and E, which are superficially similar and easily confused. There is also some overlap along edges, as at 100m^2 a large area is covered by each point/pixel, especially on an urban level.

Urban Atlas results validation

The Urban Atlas data was previously validated by its technicians, as such it is out of the scope of this project, one can only observe results of the classification. Without accurate data of building heights or other observation data (such as LIDAR), and proceeding only with the provided data, the exercise is limited.

The fundamental approach of Urban Atlas is different from that of the other two methods. UA proceeds from a land use classification, whereas LCZ is focused on morphology independent of use, at least on the level of this study. Areas of Oslo categorized as parks (LCZ D) in the other two methods in UA were categorized as governmental, which included LCZ 2 and LCZ E as well.

Regardless of the limitations, the reclassification of the UA layers results in a viable method, though by coincidence the lack of data for the area covered by the ROI limits its use for Oslo. The method would be more complete for a different area that falls entirely within UAs data areas and coupled with height information would produce a reliable model. It provides a very quick generation of LCZs with only a GIS client necessary.

This method could also be performed with CORINE Land Cover data.

Google Earth Engine results validation

The Google Earth Engine software has a built-in validation functionality. From these results one can appreciate a result similar to that of the WUDAPT/SAGA Random Forest classification.

Certain zones with confusion were zones 1 and 2, zones 5 and 6.

	# Points	1	2	3	4	5	6	7	8	9	10	A	B	C	D	E	F	G
1	9	88.89%	11.11%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
2	82	0%	98.78%	0%	0%	0%	1.22%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
3	45	0%	2.22%	91.11%	0%	0%	4.44%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2.22%	0%
4	0	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%
5	167	0%	1.2%	1.2%	0%	92.81%	3.59%	0%	1.2%	0%	0%	0%	0%	0%	0%	0%	0%	0%
6	443	0%	0%	0%	0%	2.03%	96.84%	0%	0.23%	0%	0%	0%	0%	0%	0.45%	0%	0.45%	0%
7	0	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%
8	469	0%	0%	0%	0%	0%	0.85%	0%	98.93%	0%	0%	0%	0%	0%	0.21%	0%	0%	0%
9	0	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%
10	0	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%
A	254	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99.21%	0%	0%	0.79%	0%	0%	0%
B	79	0%	1.27%	0%	0%	0%	0%	0%	0%	0%	0%	3.8%	91.14%	0%	3.8%	0%	0%	0%
C	0	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%	NaN%
D	174	0%	0%	0%	0%	0.57%	3.45%	0%	1.15%	0%	0%	0.57%	0%	93.68%	0%	0.57%	0%
E	182	0%	0.55%	0%	0%	0%	7.14%	0%	0%	0%	0%	0%	0%	0%	92.31%	0%	0%
F	66	0%	0%	0%	0%	0%	4.55%	0%	0%	0%	0%	1.52%	0%	0%	0%	93.94%	0%
G	918	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

Table 6, Contingency matrix for Google Earth Engine. This table is calculated in % b .

The results of the Google Earth Engine method are as accurate as the WUDAPT method, with a kappa of 0.974, on the same TAs. This difference is likely due to the use of only three bands from Landsat and a simpler image, and resampling in the WUDAPT images.

3.3.3 Relation of LCZ to LST and SUHI effects

Figure 29, TIRS LST from Landsat 8 (7/2018)
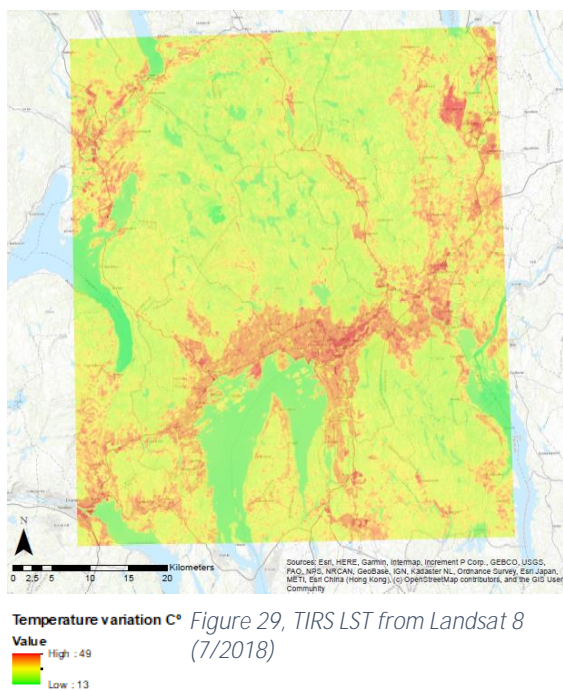
To illustrate the applicability of the LCZ and for further validation, a correlation between the LCZ and local surface temperatures (LST) is determined.

A clear image from the Landsat 8 TIRS during the 2018 summer heat wave was obtained, specifically 30 July 2018, when temperatures reached record highs in Europe, and 34.6°C in Oslo. This is a diurnal mid-day image of surface temperatures.

Methodology

For the LST data obtained from the USGS, their own methodology is followed (USGS, 2019). This omits band 11 data of the TIRS, which is more contaminated with stray visible light than band 10.

To perform the calculations of LST, band 10 digital numbers (DN) are converted to top of atmosphere (TOA) radiance, which is accomplished via the formula:

$$L_\lambda = M_L Q_{CAL} + A_L$$

Where L_λ equals top of atmosphere (TOA) spectral radiance (Watts/m^2*srad*µm), M_L is the band-specific multiplicative rescaling factor from the metadata, Q_{CAL} is the quantized and calibrated standard product pixel values' DN and A_L is the band-specific additive rescaling factor from the metadata.

Once converted to TOA radiance, the at-satellite brightness temperature is determined with the following:

$$BT = \frac{K_2}{Ln\left(\frac{K_1}{L_\lambda} + 1\right)} - 273.15$$

To obtain a more accurate measure of the LST, this value is adjusted for the vegetation presence (obtained from NDVI):

$$NDVI = \frac{NIR\ (B5) - VIS\ (B4)}{NIR\ (B5) + VIS(B4)}$$

$$P_v = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right)^2$$

The LST is then determined by the formula:

$$LST = \frac{\overline{BT}}{1 + w(\frac{\overline{BT}}{p})\,Ln(e)}$$

By generating near table statistics based on the proximity of the points of both layers converted, a correlation between LST and LCZ is generated, for each mapping methodology used in the previous stages of the study. This is then shown in boxplot format generated by a short coding in R (Appendix 13: Code used).

4. Results

4.1 Distribution and occurrence of LCZs in Oslo

To calculate the distribution of LCZs within the ROI, each raster layer is converted to a point layer, and the percentages of points by their LCZ values from the three different layers is shown as seen in Figure 31.

The three methodologies result in similar distributions of zonal types, though it is important to note that the Urban Atlas sample is smaller than the other two, covering approximately two-thirds of the ROI. This makes any direct comparison between UA and the other two results only orientative, with a generalization of climate zone occurrences in the Oslo area.
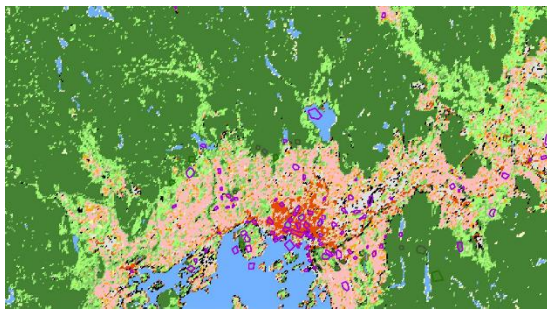
Figure 30, Oslo - detail of LCZ mapping and TAs

The dense trees ($LCZ\ A$) zone is the most common found in all. Given that forest is the natural state for most of the area, and the low population and urbanization, this is not surprising. Some small differences exist in the percentages, with the Urban Atlas having the greatest forest cover.

The LCZ type D (low plants) is the second most common, with the Google Earth method having the lowest occurrence of the three. This zone is of debatable classification, as it can include or omit agricultural zones. Agricultural zones can also be classified as LCZ F (bare soil), and arguments could be made for either route. In this study agricultural fields were classified as D, as that is the ideal state of the area and such an area would likely be vegetated in any case if undisturbed (reforested). The Google Earth method assigns a higher area of type F than the other two. LCZ type D can also include parks, of which there are a significant number in the area.

LCZ type B, scattered trees, can also include parks (which were trained as type D), but generally applies to areas near human habitation or unsuitable for heavy growth. This classification can also occur in areas where resource extraction takes place.

LCZ type C, semi-desert or scrub, does not occur in the region.

LCZ type E includes rocky or paved areas. While there are some rocky areas within the ROI, the principal occurrence of this type is of concrete and paved areas such as parking lots, port areas, and transportation networks. The Urban Atlas classification has a higher occurrence due to its generalization of areas. For example, when training the WUDAPT areas, the airport of Oslo is designated type E for the tarmac and type 8 for the structures. In the Urban Atlas the entire area is classified as airport and converted to LCZ E.

The LCZ areas of type F, sand or soil, are small, confined to beaches (of which there are a few in the area), river plains or artificially cleared areas. Variations between the three methods are small, with the Urban Atlas smallest due to a difference in approach of classification.

LCZ type G indicates water bodies. The WUDAPT and Google Earth methods show an almost identical distribution. The Urban Atlas indicates a much smaller area, as much of the water surface in the *Oslofjord* is not included in the original UA exercise.

For the urban classifications, the most predominant type in all three classifications is type 6, open low-rise, which is the most common in the area. Detached single family homes are the predominant housing type in the Oslo area.

The three methods also correspond in a similar distribution of type 8, large low-rises. This indicates industrial or commercial areas, typical of contemporary developed cities. These areas are also typical of shopping and entertainment uses, with large parking areas for vehicles, or for modern light industries or warehouses.

The remainder of the LCZs are of types 1 through 5. Type 5, the open mid-rise, is the most extensive, due to its large footprint. This is typical of apartment complexes of two to three stories, with large garden areas between buildings.

Zones 1 through 3 are concentrated in the urban cores – types 2 and 3 comprise the Oslo and Drammen central zones. This is the standard European city centre of 19th and 20th century construction. There are small areas of higher buildings (LCZ 1), but this type of development in Oslo is not widespread and is of recent construction.
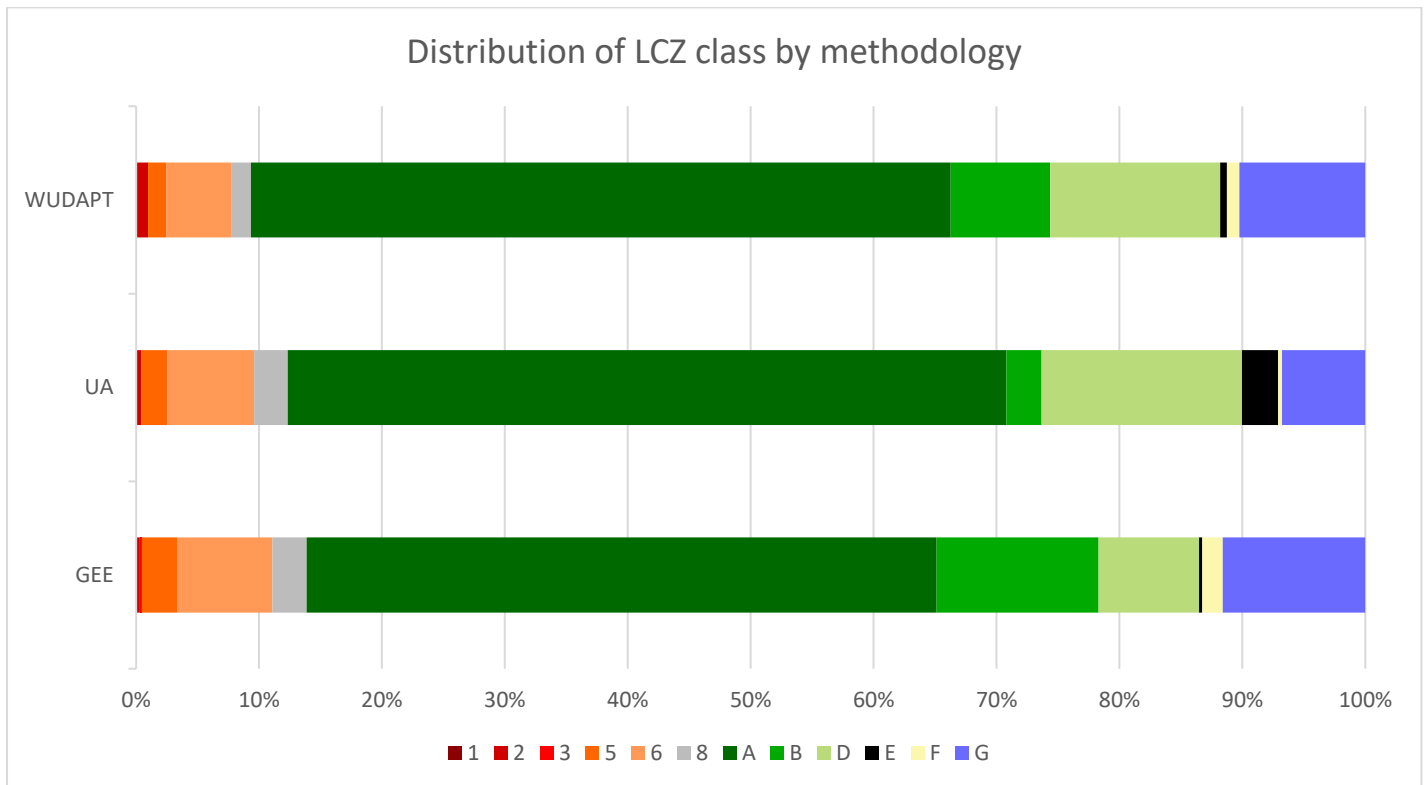
Figure 31, LCZ distribution within ROI by methodology, %

The difference in distribution could have different causes. Firstly, the UA data is generalized and grouped, so the classification is not as precise. The WUDAPT method starts with a full use of the Landsat bands, whereas the GEE method is limited to three. The WUDAPT rasters were converted from 30m for some bands, which might alter the distribution.

A relation between LCZ and surface temperature, surface impermeability, density and vegetation coverage is established. The results show a consistency across the three methods, except for the less detailed Urban Atlas mapping which is examined last. Outliers were all singular in nature.

Results from WUDAPT and Google Earth Engine

Figures 32 and 33, derived from the Random Forest classifications, reflect a correlation between LCZs of types with heavy use of man-made building materials such as concrete, tile, metal and glass, which have a higher surface temperature than areas where colloquially named "natural" materials predominate, such as water, soil or vegetation. Note that we refer only to surface temperature, at midday, in the study.

For LCZ 1, which has a very small footprint in the ROI, the surface temperature is recorded at an average of 37°C in WUDAPT and 35 in Google Earth Engine. There is with WUDAPT an extreme limit of 42°C and with GEE 39°C, or that up to 25% of the points within LCZ 1 reached these temperatures. For this LCZ there are almost no outliers. This is likely due to the small presence of the LCZ.

LCZ 2 is the most common zone found in the central urbanized areas. This area covers a larger area of the ROI, and the lower building mass does not ameliorate the SUHI effect. This LCZ in WUDAPT has a large range of temperature variation, with an extreme high range of 43°C. This is explained by the zone covering a wider area in this classification than in the GEE method. This zone in both WUDAPT and GEE has more outliers due to the increased heterogeneity of the areas included.

LCZ 3 exhibits almost identical behaviour as LCZs 1 and 2, with an even higher surface temperature, up to 41°C in GEE. This LCZ type has lower presence in the ROI and presents a very narrow range of values in the WUDAPT method. In the GEE method the classifier identified more areas of this LCZ, though still a relatively small sample. This zone shows a higher temperature than in the WUDAPT method, though similar in that almost all outliers tend towards the lower temperature increase ranges.

An observation is that in areas where the buildings are particularly dense, and are classified as LCZs 1-3, the potential for SUHI effect is most intense. This effect is localized as Oslo is not highly urbanized, and urban sprawl is more contained than in some cities of similar population elsewhere.

The effect of higher surface temperatures is pronounced and similar in LCZs 5 and 6. These correspond to lower-density areas of cities, indicative of apartment buildings ("garden apartments"), business parks and detached housing. These areas typically have vegetation between structures, but regardless have only slightly lower surface temperatures than the denser urban zones (LCZs 1-3). Both show a narrow temperature range from 35 to 37°C for the majority of the points, with more outliers or exceptions than LCZs 1-3, due to more heterogeneity.

The SUHI effect is most severe in areas where large, open expanses of concrete or stone are present, such as LCZs 8 and E. The lack of shade and high heat capacity of the construction materials increases surface temperature potential.
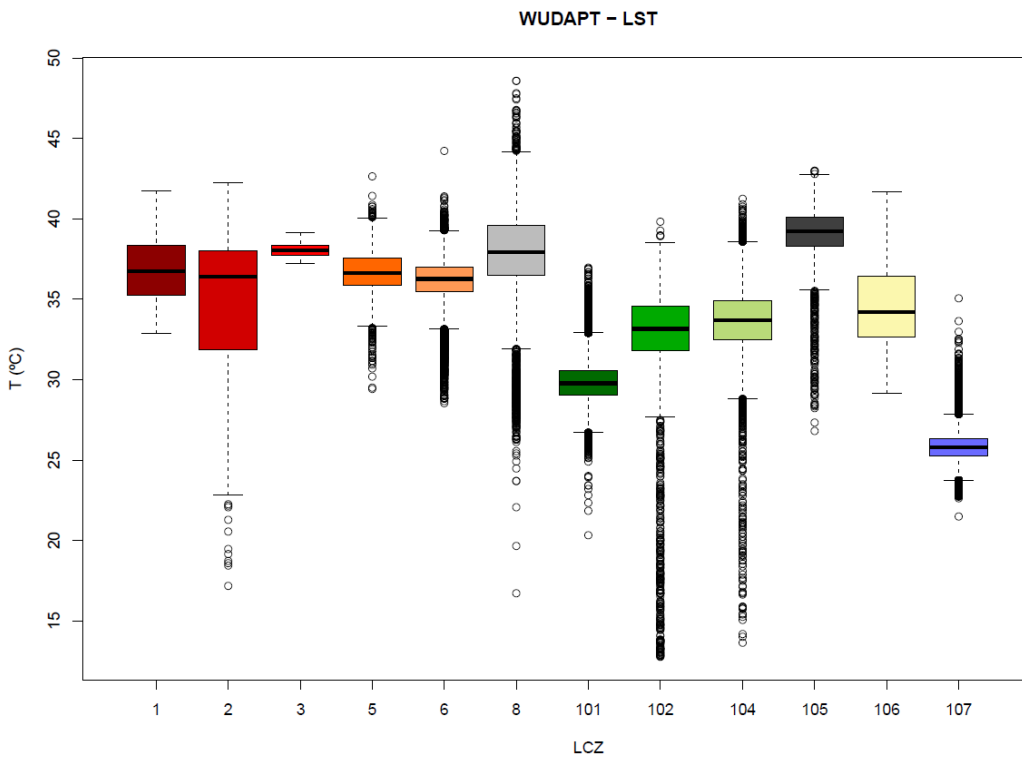
WUDAPT – LST

Figure 32, Boxplot, WUDAPT temperature by LCZ

These 2 LCZ types show a high surface temperature, reaching maximums of almost 50°C in isolated cases, higher than LCZs 1-6, and replicated across all the results. They have a wider range of temperatures, and many more outlying values recorded. This behaviour of wide variation could be attributed to different materials.

The surface temperature is lowest in densely wooded areas (LCZ A, 29°C) and water areas (LCZ G, 26°C), with the severity increasing as the vegetation cover decreases. LCZ A shows a surface temperature average of 30°C,

almost 7°C cooler than the average for LCZ 8, equally in WUDAPT and GEE. The range of extremes is narrower.

LCZ B, with sparser tree cover, exhibits a higher temperature and variation. It also has more outliers trending to the lower ranges, likely increasing with vegetation coverage. These zones had some classification differences between WUDAPT and GEE which can be attributed to the differences in the TOA imagery (fewer bands in GEE).

LCZ D, low plant cover (crops or parks), exhibits a similar temperature as that of LCZ B, with an average of 33 to 34°C. Open areas of permeable ground exhibit lower temperatures compared to impermeable urban areas, as visible from the results of LCZ F, which at an average temperature of 33 to 34°C is much cooler than LCZ E, at 37 to 40°C average.

GEE – LST
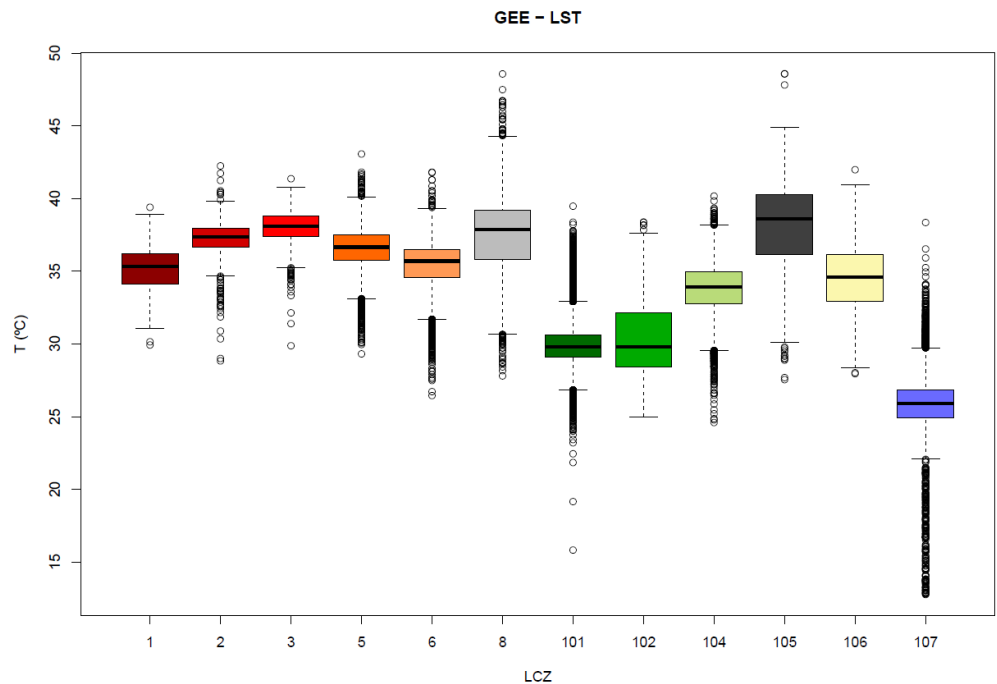
Figure 33, Boxplot, GEE, temperature by LCZ

Generation of Local Climate Zones for Oslo Metropolitan Area - 34

The grouping of LCZs in the Urban Atlas classifications due to lack of height data result in a more generalized result, with wider variations in the temperatures. Areas are surely misclassified, and some are missing. For example, the water areas are largely not included in the original data, and those that are present are in the shallows and would be susceptible to greater surface temperatures during heat episodes.

The discrepancy with the Urban Atlas LCZ E plot is also due to its generalization as a use category to incorporate categories such as government or military installations.

Figure 34, Boxplot, UA method, temperature by LCZ

Regardless of the limitations of the original data, the plot (Figure 34) exhibits identical behaviour to those from the WUDAPT and GEE methods. The permeable and vegetated LCZs have a lower surface temperature, and the LCZs of manufactured materials tend to be higher.

LCZ type 8 demonstrated the highest temperature increase of all the zones in this model – with some extremes approaching 50°C. This category also included more than one reclassification.

5. Conclusions

The Local Climate Zones for Oslo are generated according to three different methodologies: WUDAPT, Urban Atlas, and Google Earth Engine. The LCZ method is a useful method of inventory and identification of urban morphology for climatology and meteorology uses, however it is dependent on accurate data. The possible uses include not only climatology in itself but a range of derived uses such as urban planning, ecological or "green" building priorities, social awareness and emergency planning, and sustainability and resource use and flow optimization.

The three methods of LCZ generation are all valid - the results shown in the three methods' equal placement of LCZs are similar. Each method presents advantages and disadvantages.

The WUDAPT method is the most time consuming. However, this method offers perhaps the most reliable results, with the most options controllable during the process to ensure the outcome of a reliable result due to the complete control over all stages of the process.

The Urban Atlas (UA) method (or for CORINE Land Cover if desired, the processes are almost identical) has less precise results as the data is not intended to be used for mapping morphology, but instead land use. While the two objectives overlap in some respects, the specification of only land use leaves unspecified variables when applied to climate zone mapping. For example, in the UA categories, sections of Oslo were classified under the government and military categories, which in an LCZ mapping could be any Local Climate Zone. For an LCZ mapping as per the scope of this project, the land use was secondary. This is not to say in a more extensive study it would not be, as a use of a high albedo, low permeability zone as an airport would be of interest due to the increased effects on such an LCZ from its use (pollution, heat) factors. The UA method regardless of limitations remains a quick and easy means to generate an LCZ scheme rapidly.

The Google Earth Engine method delivers results very similar to the WUDAPT method, with increased automation and ease of use. The results leave some unanswered questions as to the Random Forest calculations, which, using the identical training areas from the WUDAPT method, should have produced identical results. Likely the differences are due to the limitation of the GEE software to only three bands of the Landsat 8 image.

The relation between the local climate zones and the local surface temperature, in that areas with higher albedo and low permeability and vegetation cover are susceptible to severe temperature increases, is demonstrated by the results of this study.

This strong correlation between the LCZ types and the higher LST areas also shows the high accuracy of the LCZ mapping methodology, by demonstrating the locations of areas of higher LST swings are found in areas of dense urbanization. It is demonstrated that the lesser the vegetation cover, the greater the surface temperature during episodes of heat waves. It is also demonstrated that the materials of the ground and infrastructures in an area can increase the land surface temperature (LST) greatly - specifically concrete, metal and other commonly used materials in cities. This is shown in the three LCZ generation methods

Figure 35, WUDAPT LCZ (L) 3 July 2018, Landsat TIRS Oslo 30 July 2018 (R)

by correlating the areas of higher SUHI effect (temperature increases) with LCZs of urban construction (1-8, E). This is clearly visible not only from the plotted data (Appendix 13: Oslo LST – LCZ Temperature data) but by observation of imagery (Figure 35). Interestingly, the presence of smaller, individual buildings interspersed with small green areas did not diminish the SUHI (Surface Urban Heat Island) effects significantly.
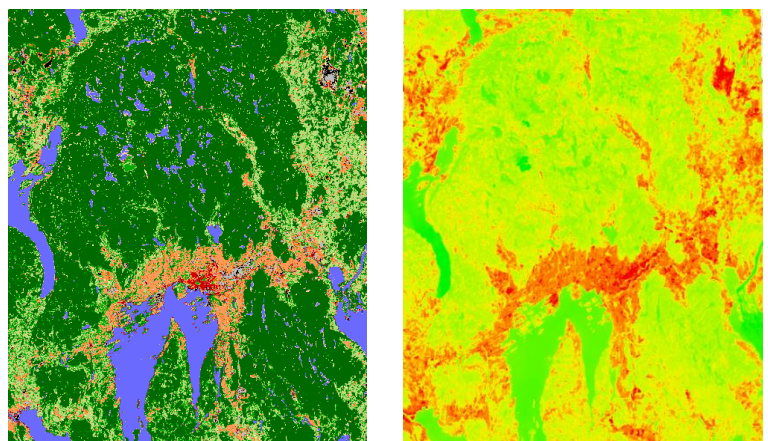
This study could be used as a template for other cities or regions to develop a climate map. It also can help other researchers understand the advantages and limitations of each individual method, and to further refine their own individual results and improve on the methods explored here.

References

1000 GIS Applications & Uses – How GIS Is Changing the World. (2019). Retrieved from GIS Geography: https://gisgeography.com/gis-applications-uses/

Bechtel, B., & Daneke, C. (2012). Classification of Local Climate Zones Based on Multiple Earth Observation Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*(4), 1191 – 1202.

Bechtel, B., Alexander, P., Böhner, J., Ching, J., Conrad, O., Feddema, J., . . . Stewart, I. (2015). Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. *ISPRS Int. J. Geo-Inf, 4*, 199-219.

Ching J, Mills, J., Bechtel, B., See, L., Feddema, J., Wang, X., & Ren, C. (2018). World urban data base and access portal tools (WUDAPT), an urban weather, climate and environmental modeling infrastructure for the Anthropocene. *Bulletin of the American Meteorological Society, 99*, 1907–1924.

Cohen's Kappa. (n.d.). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Cohen%27s_kappa

DeJarnett, N., & Pittman, M. (2017). *Protecting the health and well-being of communities in a changing climate: Proceedings of a workshop—in brief.* Washington: National Academies Press.

Demspey, C. (2014). *GIS Timeline*. Retrieved from GIS Lounge: https://www.gislounge.com/gis-timeline/

ESA. (2012). *Urban Atlas*. Retrieved from Copernicus: Europe's Eyes on Earth: https://land.copernicus.eu/local/urban-atlas

ESA. (2019). *Corine Land Cover*. Retrieved from Copernicus - Europe's Eyes on Earth: https://land.copernicus.eu/pan-european/corine-land-cover

ESA. (n.d.). *Copernicus, Europe's Eyes on Earth*. Retrieved from Copernicus, Europe's Eyes on Earth: https://www.copernicus.eu/en

Esri. (n.d.). *About ArcGIS*. Retrieved from Esri : https://www.esri.com/en-us/arcgis/about-arcgis/overview

Geletic, J., & Lehnert, M. (2016). GIS-based delineation of local climate zones: The case of medium-sized Central European cities. *Moravian Geographical Reports*, 2-12.

Geletic, J., Lehnert, M., & Dobrovolny, P. (2016). GIS-based delineation of local climate zones: The case of medium-sized Central European cities. *Remote Sensing, 8*(10), 2072-4292.

Gilabert, J., Tardà, A., & Corbera, J. (2016). Assessment of local climate zones over metropolitan area of Barcelona and added value of Urban Atlas, Corine Land Cover and Copernicus layers under INSPIRE specifications. *INSPIRE Conference 2016.* Barcelona: ResearchGate.

Google Earth Engine. (2019). Retrieved from Google Earth Engine: https://earthengine.google.com/

Google. (n.d.). *Google Earth*. Retrieved from Google Earth: https://www.google.com/earth/

Google. (n.d.). *Google Maps*. Retrieved from Google Maps: http://maps.google.com

Hammerberg, K., Brousse, O., Martilli, A., & Mahdavi, A. (2018). Implications of employing detailed urban canopy parameters for mesoscale climate modelling: a comparison between WUDAPT and GIS databases over Vienna, Austria. *International Journal of Climatology*, 1241-1257.

Ho, T. K. (1995). Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition* (pp. 278-282). Montreal: ScienceGate.

ICTA-UAB (URBAG). (2019). *URBAG - Integrated System Analysis of Urban Vegetation and Agriculture*. Retrieved from URBAG: https://urbag.eu/

ICTA-UAB. (n.d.). *ICTA-UAB - Institut de Ciència i Tecnologia Ambientals*. Retrieved from ICTA.UAB: https://ictaweb.uab.cat/

IPCC. (2014). *AR5 Synthesis Report: Climate Change 2014*. Retrieved from IPCC: https://www.ipcc.ch/report/ar5/syr/

Li, D., & Bou-Zeid, E. (2012). Synergistic Interactions between Urban Heat Islands and Heat Waves: The Impact in Cities Is Larger than the Sum of Its Parts. *Journal of Applied Meteorology and Climatology*, 2051-2064.

Li, H., Zhou, Y., Li, X., Meng, L., Wang, X., Wu, S., & Sodoudi, S. (2017). A new method to quantify surface urban heat island intensity. *Science of the total Environment, 624*, 262-272.

NASA. (n.d.). *Landsat Science.* Retrieved from NASA: https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/

Nowak, D., Crane, D., & Stevens, J. (2006). Air pollution removal by urban trees and shrubs in the United States. *Urban Forestry and Urban Greening, 4*, 115-123.

OECD. (2018). Retrieved from OECD: https://www.oecd.org/cfe/NORWAY-Regions-and-Cities-2018.pdf

O'Malley, C., Piroozfarb, P., Farr, E., & Gates, J. (2014). An investigation into minimizing urban heat island (UHI) effects: A UK perspective. *6th International Conference on Sustainability in Energy and Buildings, SEB-14.* Cardiff: Elsevier.

QGIS. (n.d.). *QGIS*. Retrieved from QGIS: http://www.gqis.org

SAGA. (n.d.). *SAGA*. Retrieved from SAGA: http://http://www.saga-gis.org/

Sheridan, S., & Dixon, P. (2016, July 19). Spatiotemporal trends in human vulnerability and adaptation to heat. *Anthropocene*.

Skarbit, N., Gál, T., & Unger, J. (2015). Airborne surface temperature differences of the different Local Climate Zones in the urban area of a medium sized city. *Joint Urban Remote Sensing Event (JURSE).* Lausanne: IEEE.

Statistics Norway. (2019). *Statistics Norway*. Retrieved from http://www.ssb.no/

Stewart, I. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature. *International Journal of Climatology*, 200-217.

Stewart, I., & Oke, T. R. (2012). Local Climate Zones for urban temperature studies. *Bulletin of the American Meteorological Society, 93*, 1879-1900.

Stewart, I., Oke, T. R., & Krayenhoff, E. S. (2014). Evolution of the 'local climate zone' scheme using temperature observations and model simulations. *International Journal of Climatology, 34*, 1062-80.

Stewart, I., Oke, T., & Krayenhoff, E. (2014). Evaluation of the 'local climate zone'scheme using temperature observations and model simulations. *International Journal of Climatology* .

Strahler, A., & Strahler, A. (1994). Climate Classification. *Physical Geography*, pp. 147-169.

Tan, J. (2009). The urban heat island and its impact on heat waves and human health in Shanghai. *International Journal of Biometeorology, 54*, 75-84.

Tatem, A., Goetz, S., & Hay, S. (2008). Fifty Years of Earth Observation Satellites. *Am Sci, 96*(5), 390-398.

United Nations. (2018, 5 16). *68% of the world population projected to live in in urban areas by 2050*. Retrieved from United Nations: http://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

USGS. (2019). *https://www.usgs.gov/land-resources/nli/landsat*. Retrieved from USGS Landsat Missions: https://www.usgs.gov/land-resources/nli/landsat

USGS. (n.d.). *USGS Earth Explorer*. Retrieved from USGS: https://earthexplorer.usgs.gov/

Wang, R., Ren, C., Xu, Y., Lau, K., & Shi, Y. (2018). Mapping the local climate zones of urban areas by GIS-based and WUDAPT methods: A case study of Hong Kong. *Urban Climate*, 567-576.

Wikipedia. (2016). *Greater Oslo Region*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Greater_Oslo_Region
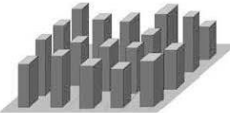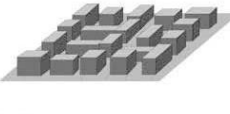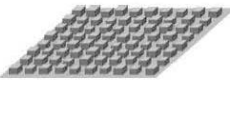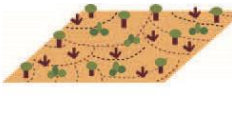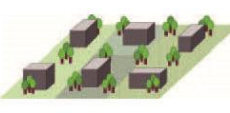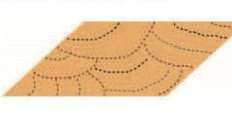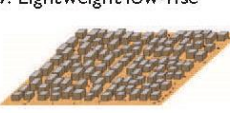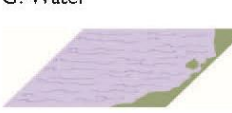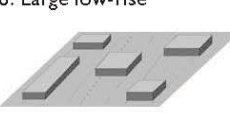
Wikipedia. (n.d.). *Oslo, Norway*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Oslo

Wolf, T., & McGregor, G. (2013). The development of a heat wave vulnerability index for London, United Kingdom. *Weather and Climate Extremes*, 59-68.

WUDAPT. (2012). Retrieved from WUDAPT: http://www.wuapt.org

Appendix

Appendix 1: LCZ classes and descriptions

Built types	Definition	Land cover types	Definition
1. Compact high-rise	Dense mix of tall buildings to tens of stories. Few or no trees. Land cover mostly paved. Concrete, steel, stone, and glass construction materials.	A. Dense trees	Heavily wooded landscape of deciduous and/or evergreen trees. Land cover mostly pervious (low plants). Zone function is natural forest, tree cultivation, or urban park.
2. Compact midrise	Dense mix of midrise buildings (3–9 stories). Few or no trees. Land cover mostly paved. Stone, brick, tile, and concrete construction materials.	B. Scattered trees	Lightly wooded landscape of deciduous and/or evergreen trees. Land cover mostly pervious (low plants). Zone function is natural forest, tree cultivation, or urban park.
3. Compact low-rise	Dense mix of low-rise buildings (1–3 stories). Few or no trees. Land cover mostly paved. Stone, brick, tile, and concrete construction materials.	C. Bush, scrub	Open arrangement of bushes, shrubs, and short, woody trees. Land cover mostly pervious (bare soil or sand). Zone function is natural scrubland or agriculture.
4. Open high-rise	Open arrangement of tall buildings to tens of stories. Abundance of pervious land cover (low plants, scattered trees). Concrete, steel, stone, and glass construction materials.	D. Low plants	Featureless landscape of grass or herbaceous plants/crops. Few or no trees. Zone function is natural grassland, agriculture, or urban park.
5. Open midrise	Open arrangement of midrise buildings (3–9 stories). Abundance of pervious land cover (low plants, scattered trees). Concrete, steel, stone, and glass construction materials.	E. Bare rock or paved	Featureless landscape of rock or paved cover. Few or no trees or plants. Zone function is natural desert (rock) or urban transportation.
6. Open low-rise	Open arrangement of low-rise buildings (1–3 stories). Abundance of pervious land cover (low plants, scattered trees). Wood, brick, stone, tile, and concrete construction materials.	F. Bare soil or sand	Featureless landscape of soil or sand cover. Few or no trees or plants. Zone function is natural desert or agriculture.
7. Lightweight low-rise	Dense mix of single-story buildings. Few or no trees. Land cover mostly hard-packed. Lightweight construction materials (e.g., wood, thatch, corrugated metal).	G. Water	Large, open water bodies such as seas and lakes, or small bodies such as rivers, reservoirs, and lagoons.
8. Large low-rise	Open arrangement of large low-rise buildings (1–3 stories). Few or no trees. Land cover mostly paved. Steel, concrete, metal, and stone construction materials.	**VARIABLE LAND COVER PROPERTIES** Variable or ephemeral land cover properties that change significantly with synoptic weather patterns, agricultural practices, and/or seasonal cycles.	
9. Sparsely built	Sparse arrangement of small or medium-sized buildings in a natural setting. Abundance of pervious land cover (low plants, scattered trees).	b. bare trees	Leafless deciduous trees (e.g., winter). Increased sky view factor. Reduced albedo.
		s. snow cover	Snow cover >10 cm in depth. Low admittance. High albedo.
10. Heavy industry	Low-rise and midrise industrial structures (towers, tanks, stacks). Few or no trees. Land cover mostly paved or hard-packed. Metal, steel, and concrete construction materials.	d. dry ground	Parched soil. Low admittance. Large Bowen ratio. Increased albedo.
		w. wet ground	Waterlogged soil. High admittance. Small Bowen ratio. Reduced albedo.

Appendix 2: LCZ class characteristics

TABLE 3. Values of geometric and surface cover properties for local climate zones. All properties are unitless except height of roughness elements (m).

Local climate zone (LCZ)	Sky view factor[a]	Aspect ratio[b]	Building surface fraction[c]	Impervious surface fraction[d]	Pervious surface fraction[e]	Height of roughness elements[f]	Terrain roughness class[g]
LCZ 1 *Compact high-rise*	0.2–0.4	> 2	40–60	40–60	< 10	> 25	8
LCZ 2 *Compact midrise*	0.3–0.6	0.75–2	40–70	30–50	< 20	10–25	6–7
LCZ 3 *Compact low-rise*	0.2–0.6	0.75–1.5	40–70	20–50	< 30	3–10	6
LCZ 4 *Open high-rise*	0.5–0.7	0.75–1.25	20–40	30–40	30–40	>25	7–8
LCZ 5 *Open midrise*	0.5–0.8	0.3–0.75	20–40	30–50	20–40	10–25	5–6
LCZ 6 *Open low-rise*	0.6–0.9	0.3–0.75	20–40	20–50	30–60	3–10	5–6
LCZ 7 *Lightweight low-rise*	0.2–0.5	1–2	60–90	< 20	<30	2–4	4–5
LCZ 8 *Large low-rise*	>0.7	0.1–0.3	30–50	40–50	<20	3–10	5
LCZ 9 *Sparsely built*	> 0.8	0.1–0.25	10–20	< 20	60–80	3–10	5–6
LCZ 10 *Heavy industry*	0.6–0.9	0.2–0.5	20–30	20–40	40–50	5–15	5–6
LCZ A *Dense trees*	<0.4	>1	<10	<10	>90	3–30	8
LCZ B *Scattered trees*	0.5–0.8	0.25–0.75	<10	<10	>90	3–15	5–6
LCZ C *Bush, scrub*	0.7–0.9	0.25–1.0	<10	<10	>90	<2	4–5
LCZ D *Low plants*	>0.9	<0.1	<10	<10	>90	<1	3–4
LCZ E *Bare rock or paved*	>0.9	<0.1	<10	>90	<10	<0.25	1–2
LCZ F *Bare soil or sand*	>0.9	<0.1	<10	<10	>90	< 0.25	1–2
LCZ G *Water*	>0.9	<0.1	<10	<10	>90	–	1

[a] Ratio of the amount of sky hemisphere visible from ground level to that of an unobstructed hemisphere

[b] Mean height-to-width ratio of street canyons (LCZs 1–7), building spacing (LCZs 8–10), and tree spacing (LCZs A–G)

[c] Ratio of building plan area to total plan area (%)

[d] Ratio of impervious plan area (paved, rock) to total plan area (%)

[e] Ratio of pervious plan area (bare soil, vegetation, water) to total plan area (%)

[f] Geometric average of building heights (LCZs 1–10) and tree/plant heights (LCZs A–F) (m)

[g] Davenport et al.'s (2000) classification of effective terrain roughness (z_0) for city and country landscapes. See Table 5 for class descriptions

TABLE 4. Values of thermal, radiative, and metabolic properties for local climate zones. All values are representative of the local scale.
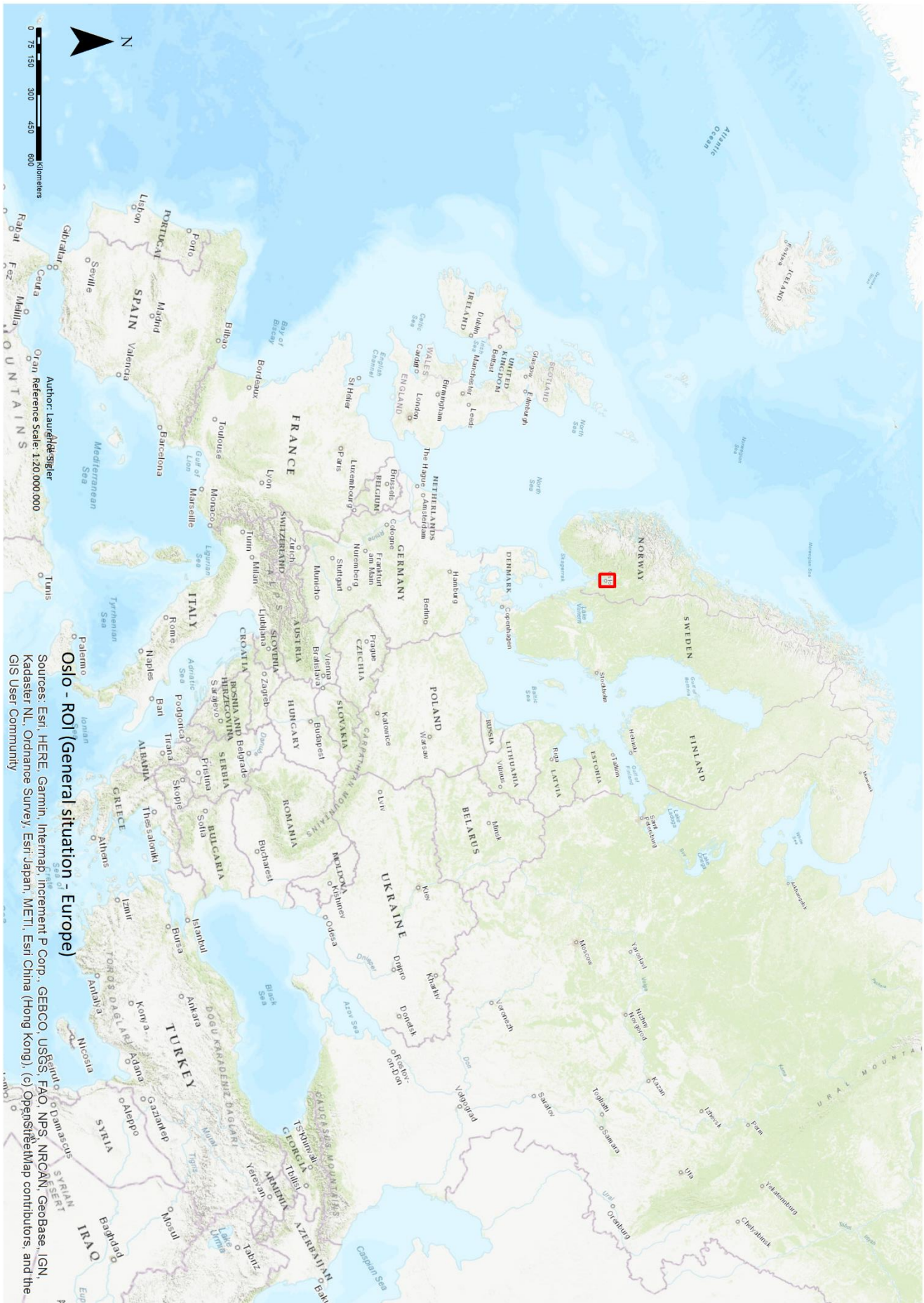
Local climate zone (LCZ)	Surface admittance[a]	Surface albedo[b]	Anthropogenic heat output[c]
LCZ 1	1,500–1,800	0.10–0.20	50–300
Compact high-rise			
LCZ 2	1,500–2,200	0.10–0.20	<75
Compact midrise			
LCZ 3	1,200–1,800	0.10–0.20	<75
Compact low-rise			
LCZ 4	1,400–1,800	0.12–0.25	<50
Open high-rise			
LCZ 5	1,400–2,000	0.12–0.25	<25
Open midrise			
LCZ 6	1,200–1,800	0.12–0.25	<25
Open low-rise			
LCZ 7	800–1,500	0.15–0.35	<35
Lightweight low-rise			
LCZ 8	1,200–1,800	0.15–0.25	<50
Large low-rise			
LCZ 9	1,000–1,800	0.12–0.25	<10
Sparsely built			
LCZ 10	1,000–2,500	0.12–0.20	>300
Heavy industry			
LCZ A	unknown	0.10–0.20	0
Dense trees			
LCZ B	1,000–1,800	0.15–0.25	0
Scattered trees			
LCZ C	700–1,500	0.15–0.30	0
Bush, scrub			
LCZ D	1,200–1,600	0.15–0.25	0
Low plants			
LCZ E	1,200–2,500	0.15–0.30	0
Bare rock or paved			
LCZ F	600–1,400	0.20–0.35	0
Bare soil or sand			
LCZ G	1,500	0.02–0.10	0
Water			

[a] Ability of surface to accept or release heat ($J\ m^{-2}\ s^{-1/2}\ K^{-1}$). Varies with soil wetness and material density. Few estimates of local-scale admittance exist in the literature; values given here are therefore subjective and should be used cautiously. Note that the "surface" in LCZ A is undefined and its admittance unknown.

[b] Ratio of the amount of solar radiation reflected by a surface to the amount received by it. Varies with surface color, wetness, and roughness.

[c] Mean annual heat flux density ($W\ m^{-2}$) from fuel combustion and human activity (transportation, space cooling/heating, industrial processing, human metabolism). Varies significantly with latitude, season, and population density.

Oslo - ROI (General situation - Europe)

Oslo - Urban Area

ETRS 89 NTM Zone 10
Author: Laurence Sigler
Reference Scale: 1:100.000
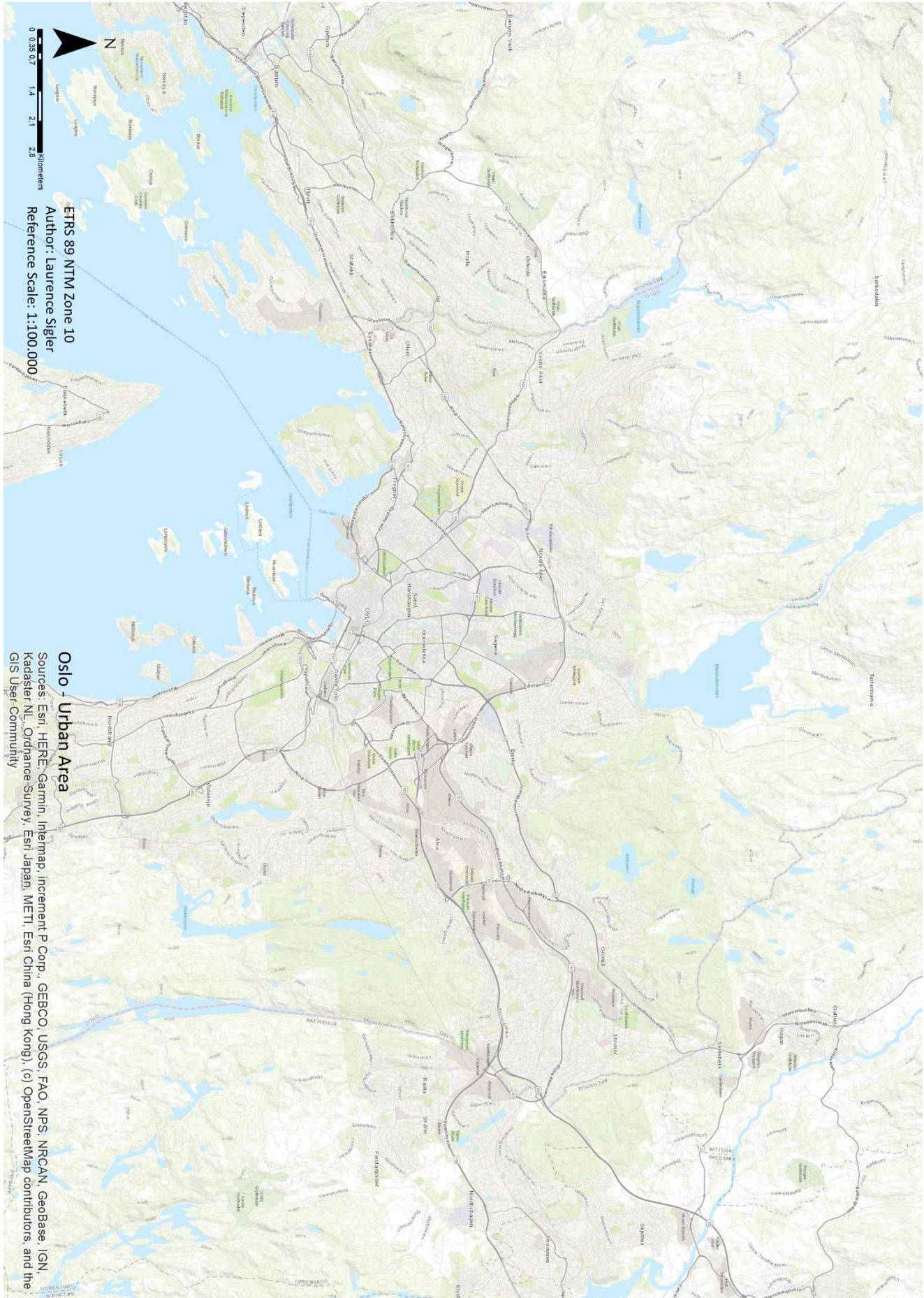
Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community
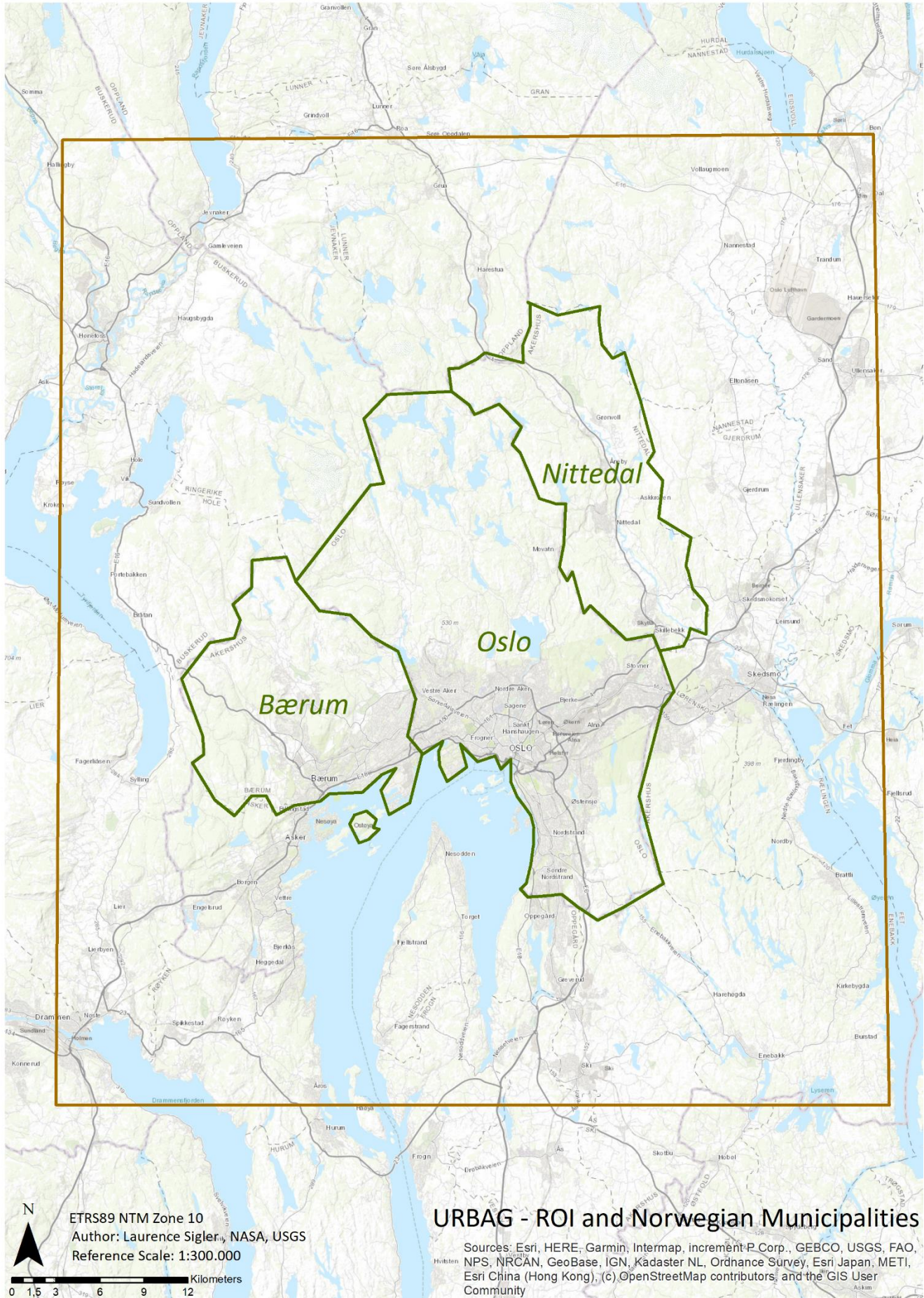
URBAG - ROI and Norwegian Municipalities

Generation of Local Climate Zones for Oslo Metropolitan Area - 46

N

ETRS89 NTM Zone 10
Author: Laurence Sigler, NASA,USGS
Reference Scale: 1:1.300.000

Kilometers
0 5 10   20   30   40

ROI and Landsat 8 (Swath: 3-7-2018)

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

Generation of Local Climate Zones for Oslo Metropolitan Area - 47

LCZ 1 - Compact high-rise
LCZ 2 - Compact mid-rise
LCZ 3 - Compact low-rise
LCZ 5 - Open mid-rise
LCZ 6 - Open low-rise
LCZ 8 - Large low-rise
LCZ A - Dense trees
LCZ B - Scattered trees
LCZ D - Low plants
LCZ E - Rock / paved
LCZ F - Bare soil / sand
LCZ G - Water

ETRS89 NTM Zone 10
Author: Laurence Sigler
Reference Scale: 1:300.000

N

Kilometers
0 1,5 3   6   9   12

LCZ Oslo - WUDAPT 100m^2

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

Generation of Local Climate Zones for Oslo Metropolitan Area - 48

Legend

- 11100: Continuous Urban fabric (S.L. > 80%)
- 11210: Discontinuous Dense Urban Fabric (S.L.: 50% - 80%)
- 11220: Discontinuous Medium Density Urban Fabric (S.L.: 30% - 50%)
- 11230: Discontinuous Low Density Urban Fabric (S.L.: 10% - 30%)
- 11240: Discontinuous very low density urban fabric (S.L. < 10%)
- 11300: Isolated Structures
- 12100: Industrial, commercial, public, military and private units
- 12210: Fast transit roads and associated land
- 12220: Other roads and associated land
- 12230: Railways and associated land
- 12300: Port areas
- 12400: Airports
- 13100: Mineral extraction and dump sites
- 13300: Construction sites
- 13400: Land without current use
- 14100: Green urban areas
- 14200: Sports and leisure facilities
- 21000: Arable land (annual crops)
- 22000: Permanent crops
- 23000: Pastures
- 24000: Complex and mixed cultivation patterns
- 25000: Orchards
- 31000: Forests
- 32000: Herbaceous vegetation associations
- 33000: Open spaces with little or no vegetations
- 40000: Wetlands
- 50000: Water

ETRS89 NTM Zone 10
Author: Laurence Sigler, ESA, EU
Reference Scale: 1:500.000

Kilometers
0 2,5 5   10   15   20

URBAG - ROI and Urban Atlas (2012)

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

Generation of Local Climate Zones for Oslo Metropolitan Area - 49

LCZ Oslo - Urban Atlas 100m²

LCZ 8 - Large low-rise
LCZ 6 - Open low-rise
LCZ 5 - Open mid-rise
LCZ 2 - Compact mid-rise
LCZ G - Water
LCZ F - Bare soil / sand
LCZ E - Rock / paved
LCZ D - Low plants
LCZ B - Scattered trees
LCZ A - Dense trees

ETRS89 NTM Zone 10
Author: Laurence Sigler
Reference Scale: 1:300.000

Kilometers
0 1,5 3   6   9   12

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

■	LCZ 1 - Compact high-rise	■	LCZ 6 - Open low-rise	■	LCZ D - Low plants
■	LCZ 2 - Compact mid-rise	■	LCZ 8 - Large low-rise	■	LCZ E - Rock / paved
■	LCZ 3 - Compact low-rise	■	LCZ A - Dense trees	■	LCZ F - Bare soil / sand
■	LCZ 5 - Open mid-rise	■	LCZ B - Scattered trees	■	LCZ G - Water

N

ETRS 89 NTM Zone 10
Author: Laurence Sigler
Reference Scale: 1:300.000

Kilometers
0   1,5   3   6   9   12

LCZ Oslo - Google Earth Engine 100m²

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

Generation of Local Climate Zones for Oslo Metropolitan Area - 51

Temperature variation Cº

Value

High : 49

Low : 13

Nittedal

Oslo

Bærum

ETRS89 NTM Zone 10
Author: Laurence Sigler, NASA, USGS
Reference Scale: 1:300.000

N

Kilometers
0 1.5 3   6   9   12

ROI: Land Surface Temperature (30 July 2018)

Sources: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community

Generation of Local Climate Zones for Oslo Metropolitan Area - 52

Appendix 12: Glossary

CLC – CORINE Land Cover

DN – Digital Numbers

ESA – European Space Agency

GEE – Google Earth Engine

ICTA- Institut de Ciència i Tecnologies Ambiental

KML – Keyhole Mark-up Language

LCZ – Local Climate Zone

LST – Land Surface Temperature

NASA – National Aeronautics and Space Administration

OLI – Operational Land Imager

ROI – Region of Interest

SUHI – Surface Urban Heat Island

TA – Training Area

TIRS – Thermal Infrared Sensor

UA – Urban Atlas

UAB – Universitat Autònoma de Barcelona

URBAG – Integrated System Analysis of Urban Vegetation and Agriculture

USGS – United States Geological Survey

WUDAPT – World Urban Database and Access Portal Tools

Appendix 13: Code used

Python code for reclassification of Urban Atlas data:

```python
def Reclass(CODE2012):

    if CODE2012 == "11100":
      return 2
    elif CODE2012 == "11210":
      return 2
    elif CODE2012 == "11220":
      return 5
    elif CODE2012 == "11300":
      return 5
    elif CODE2012 == "11230":
      return 6
    elif CODE2012 == "11240":
      return 9
    elif CODE2012 == "12000":
      return 8
    elif CODE2012 == "12100":
      return 2
    elif CODE2012 == "31000":
      return 101
    elif CODE2012 == "14100":
      return 102
    elif CODE2012 == "14200":
        return 102
    elif CODE2012 == "25000":
        return 102
    elif CODE2012 == "32000":
      return 102
    elif CODE2012 == "21000":
      return 104
    elif CODE2012 == "22000":
      return 104
    elif CODE2012 == "23000":
      return 104
    elif CODE2012 == "24000":
      return 104
    elif CODE2012 == "33000":
      return 104
    elif CODE2012 == "12210":
      return 105
    elif CODE2012 == "12220":
      return 105
    elif CODE2012 == "12230":
      return 105
    elif CODE2012 == "12300":
      return 105
    elif CODE2012 == "12400":
      return 105
    elif CODE2012 == "13300":
      return 105
    elif CODE2012 == "13100":
      return 106
    elif CODE2012 == "13400":
      return 106
    elif CODE2012 == "40000":
      return 107
    elif CODE2012 == "50000":
      return 107
    else:
      return 0


Reclass(!CODE2012!)
```

R code and steps for boxplot generation：

1. Convert dbf table to plain text (*.txt)

2. Load this text file in R, setting the variables as needed (filename, directory):

```
FILE <- read.delim2("~/directory/FILE.txt")
>   View(FILE)
```

3. Generate and export boxplot, setting the variables as needed (column names, title, axis names):

```
boxplot(FILE$COLUMN1~FILE$COLUMN2, main= "TÍTLE", xlab= "axisname_x", ylab = "axisname_y")
```

Appendix 14: Urban Atlas Land Use Classes

- 11100: Continuous Urban fabric (S.L. > 80%)
- 11210: Discontinuous Dense Urban Fabric (S.L.: 50% - 80%)
- 11220: Discontinuous Medium Density Urban Fabric (S.L.: 30% - 50%)
- 11230: Discontinuous Low Density Urban Fabric (S.L.: 10% - 30%)
- 11240: Discontinuous very low density urban fabric (S.L. < 10%)
- 11300: Isolated Structures
- 12100: Industrial, commercial, public, military and private units
- 12210: Fast transit roads and associated land
- 12220: Other roads and associated land
- 12230: Railways and associated land
- 12300: Port areas
- 12400: Airports
- 13100: Mineral extraction and dump sites
- 13300: Construction sites
- 13400: Land without current use
- 14100: Green urban areas
- 14200: Sports and leisure facilities
- 21000: Arable land (annual crops)
- 22000: Permanent crops
- 23000: Pastures
- 24000: Complex and mixed cultivation patterns
- 25000: Orchads
- 31000: Forests
- 32000: Herbaceous vegetation associations
- 33000: Open spaces with little or no vegetations
- 40000: Wetlands
- 50000: Water