FACULTÉ
DES SCIENCES

UNIVERSITÉ LIBRE DE BRUXELLES

Mapping 20 Years of Urbanization in Sub-Saharan Africa from Space

An approach based on multi-sensor satellite imagery and volunteered geographic information

Thesis submitted by Yann FORGET
in fulfilment of the requirements of the PhD Degree in Agronomy and Bioengineering ("Docteur en Sciences")
Academic year 2019-2020

Supervisor: Professor Marius GILBERT

Spatial Epidemiology Lab (SpELL)

Thesis jury:

Pierre COHEUR (Université Libre de Bruxelles, Chair)
Moritz LENNERT (Université Libre de Bruxelles, Secretary)
Marius GILBERT (Université Libre de Bruxelles, Supervisor)
Tim VANDEVOORDE (Vrije Universiteit Brussel)
Anton VAN ROMPAEY (Katholieke Universiteit Leuven)

Mapping 20 Years of Urbanization in Sub-Saharan Africa from Space

An approach based on multi-sensor satellite imagery and volunteered geographic information

Yann Forget

Contents

Acknowledgments

The present thesis has been funded by the STEREO III program (https://eo.belspo.be) of the Belgian Science Policy Office (BELSPO) in the context of the MAUPP project (https://maupp.ulb.ac.be). I would like to thank the members of the STEREO team and the members of the project's steering committee for their support and their valuable inputs: Prof. Paolo Gamba (University of Pavia), Prof. Richard Sliuzas (University of Twente), Sergio Freire (EU Joint Research Centre), and Dr. Thomas Esch (German Aerospace Center).

I would like to thank my supervisor, Prof. Marius Gilbert, for his precious, comprehensive and continuous support since the beginning of my PhD in 2014. His trust, his scientific guidance and the free & creative working environment he built has been essential to my work throughout the years. I would also like to thank Prof. Catherine Linard (University of Namur) and Dr. Michal Shimoni (Royal Military Academy of Brussels) for their support during the whole PhD. Their scientific advices and inputs are everywhere in this thesis.

My thanks also go out to the other members of the MAUPP project–whithout whom none of this work would have been possible–quite particularly Prof. Eleonore Wolff and Dr. Taïs Grippa from the Geography Department of the Université Libre de Bruxelles. My appreciations also goes to my colleagues at the Spatial Epidemiology Lab, especially Jean Artois who started its PhD with me in 2014!

And finally to Méline, who shared every single moments of this PhD and whithout whom I would not have the courage to start and finish this journey.

Abstract

Between 2015 and 2050, half of the net increase in the world's urban population is expected to take place in Sub-Saharan Africa (SSA), driving drastic land cover changes and challenging the spatial organization of human societies. Understanding past and present dynamics of this urbanization process is critical to achieve a sustainable pattern of urban development, yet is limited by the lack of accurate and multi-temporal spatial data on urban expansion. Since the 2000s, the rise of satellite-based Earth Observation (EO) enabled the production of several global urban maps, thereby mitigating the issue of data scarcity. But SSA is still characterized by lower accuracies in satellite-based maps because of various issues, such as: a lower satellite imagery availability, a lack of reference datasets, and a high heterogeneity across the urban areas of the region.

In this thesis, I propose to leverage open-access satellite catalogs along with volunteered geographic information to improve large-scaled and automated mapping of the built environment in SSA. The proposed approach makes use of OpenStreetMap to support model training and calibration, thereby bypassing the need for reference datasets or manual digitization campaigns. This method was assessed in 10 urban areas of SSA, reaching classification performances similar to manual approaches.

Moreover, the combined use of multispectral and synthetic-aperture radar (SAR) imagery was explored. In 11 out of 12 case studies in SSA, multi-sensor classification schemes outperformed single-sensor approaches. More specifically, multi-sensor classification dramatically increased built-up detection rates in arid and semi-

arid regions—where bare soil and buildings may share a similar spectral signature.

These findings were implemented to map the built environment of 46 urban areas at five different dates from 1995 to 2015, with an average F1-score of 0.93. The statistical interpretation of the produced dataset revealed the high heterogeneity that characterizes urban areas in SSA, and confirmed that the spatial patterns of urbanization highly depends on demographic and economic factors.

Overall, the present thesis provides promising insights for large-scaled and automated mapping of the built environment in data-scarce regions. Several issues are still affecting the mapping accuracies, such as: multi-temporal inconsistencies caused by the use of imagery from 7 different sensors, low availability of historical imagery in SSA, or missing data in OpenStreetMap. Still, with the growing availability of open-access EO catalogs and the increasingly completeness of OpenStreetMap, the proposed approach is expected to become even more relevant in the near future.

List of Figures

List of Tables

List of Abbreviations

AOI	Area of Interest
API	Application Programming Interface
AVHRR	Advanced Very-High-Resolution Radiometer
CAGR	Compound Annual Growth Rates
CASR	Compound Annual Sprawl Rates
CBERS	China-Brazil Earth Resources Satellite
CV	Cross Validation
DBI	Dry Built-up Index
DBSI	Dry Bare-Soil Index
DEM	Digital Elevation Model
ERS	European Remote Sensing
ETM	Enhanced Thematic Mapper
EO	Earth Observation
GDP	Gross Domestic Product
GE	Google Earth
GHSL	Global Human Settlement Layer

GLCM Grey-Level Co-Occurence Matrix

GRD Ground Range Detected

GRUMP Global Rural-Urban Mapping Project

GUF Global Urban Footprint

HBASE Global Human Built-up And Settlement Extent

IW Interferometric Wide

LaSRC Landsat Surface Reflectance Code

LULC Land Use and Land Cover

MAUPP Modelling and forecasting African Urban Population Patterns

MERIS MEdium Resolution Imaging Spectrometer

MODIS Moderate Resolution Imaging Spectroradiometer

MNDWI Modified Normalized Difference Water Index

NBLI Normalized Bare Land Index

NDBal Normalized Difference Bareness Index

NDBI Normalized Difference Built-up Index

NDSV Normalized Difference Spectral Vector

NDVI Normalized Difference Vegetation Index

OLI Operational Land Imager

OSM OpenStreetMap

PCA Principal Component Analysis

RF Random Forest

SAR Synthetic Aperture Radar

SNAP Sentinel Application Platform

SRTM	Shuttle Radar Topographic Mission
SSA	Sub-Saharan Africa
TM	Thematic Mapper
UI	Urban Index
UN	United Nations
USGS	United States Geological Survey
VGI	Volunteered Geographic Information
VHSR	Very High Spatial Resolution

Chapter 1

Introduction

The world's population is expected to increase from 7.4 billion in 2015 to 9.8 billion in 2050 (United Nations, 2017), driving an urbanization process that transforms the Earth's surface and the spatial organization of human societies. Cities emerge and grow everywhere across the globe, eroding natural habitats at their borders, and reaching far beyond for food, material and energy inputs (Seto, Guneralp and Hutyra, 2012). In low-income countries, public services struggle to keep up with the expansion rates. The resulting lack of infrastructure threatens human health and well-being: unsafe housing, air pollution, water contamination, vulnerability to climate change and extreme weather events (McGranahan *et al.*, 2009; United Nations, 2015). "Inclusive, safe, resilient and sustainable cities" are the 11 sustainable development goal of the United Nations. But how to tackle urbanization issues when we know so little about it?

Traditionally, urbanization was thought as a byproduct of the capitalist development, thereby correlated with industrialization and economic growth. Yet, Sub-Saharan Africa (SSA) experienced the highest urbanization rates in the world since the 1980s, while its economies were stagnating. For a long time, policy makers view the "*urbanization without growth*" occurring in poor countries as a dangerous anomaly that should be restrained—for instance by discouraging rural-to-urban migrations (Jedwab and Vollrath, 2015). But as low- and middle-income countries

achieve their urban transitions, the anomaly is becoming the norm. This switch led to a redefinition of urbanization as a global historical process driven by food availability and mortality decline, with or without economic development (Fox, 2012). In other words, the expansion of urban areas cannot be restrained. Moreover, the social and environmental issues associated with urbanization cannot always be thought as a direct consequence of urban expansion. Above all, urban poverty and pressure on natural resources in SSA result from decades of region-wide economic crisis and structural adjustment programs (Becker and Morrison, 1997).

Still, in the near future, the welfare of millions of people will depend on the social ability to achieve a beneficial pattern of urban development. Tackling urbanization problems requires a sound understanding of the *urbanization process* and, to that end, spatially accurate and up-to-date data are needed to better understand, anticipate and predict those urban dynamics. Just a decade ago, we had little idea of the total surface covered by built-up areas: somewhere between 0.3 and 3.5 million km^2 (Seto *et al.*, 2011). Today, the growing availability of Earth Observation (EO) imagery is an opportunity to improve our understanding of the urbanization dynamics.

The following chapter focuses on the specifics of urbanization in SSA, on the need for a spatial analysis of the process and how EO can support urban mapping. More specifically, starting from the observation that SSA is characterized by lower accuracies in global urban maps, I identify three main challenges: data availability, spectral confusion and urban heterogeneity. To address these issues, I investigate the opportunities brought by open-access EO data and Volunteered Geographic Information (VGI), and integrate them into a proposed mapping approach.

1.1 Urbanization in Sub-Saharan Africa

1.1.1 Past, current and future trends

Urban settlements in SSA have existed for thousands of years. Yet, they were relatively small and dispersed before the 20^{th} century. The early independence period marks the real beginning of the urban transition. The development of better health

facilities and the successive vaccination campaigns enabled a steady decline of mortality rates. Rural-to-urban migrations were driven by the increasing employment opportunities (the *Africanization* of the civil administrations) and the elimination of the residence restrictions established by the colonial powers (Fox, 2012). Those factors led to exceptionally high rates of urbanization between 1960 and 1975. After the region-wide economic crisis in the late 1970s, the 1980s and the 1990s were characterized by a decline of the urbanization rates. Trade liberalization and structural adjustment programs caused a rollback of public services, lower wages and high unemployment rates (Bocquier, 2003). From this period, the poor living conditions in urban areas held back the rural-to-urban migrations. In the 2000s, two thirds of the urban population growth was caused by simple natural increase (Martine *et al.*, 2008) and the share of rural-to-urban migrations in the growth of the urban population became minor in most of the Sub-Saharan African countries.

As a result, the *urbanization process*—defined as the increase in the proportion of people living in urban areas—is not particularly rapid in SSA. According to the data provided by the World Urbanization Prospects (United Nations, 2017), the share of the urban population increased from 28% to 39% between 1990 and 2015, that is a compound annual growth rate of 1.4%. In fact, today's urbanization growth rate of SSA is comparable to the ones observed in South America or Eastern Europe in the 1960s (1.5%), while their urbanization levels were considerably higher (more than 50%).

SSA is not urbanizing at particularly high rates. What is true, however, is that the absolute numbers are massive. By 2050, the global population is expected to increase by 2.4 billion, and more than half of that increase would take place in SSA (United Nations, 2017). Driven by the population growth rates, the urban population of SSA will increase by 235% over the same period. This means that urban areas in SSA will have to absorb a net increase of 0.9 billion of new dwellers by 2050. As shown in Figure 1.1, most of this growth will occur in Western and Eastern Africa—Southern Africa being more advanced in its urban transition. Many have concluded that SSA's urban dwellers will soon be living in mega-cities. If by 2020, Africa will host 11 mega-cities with more than 10 million people (Mundia

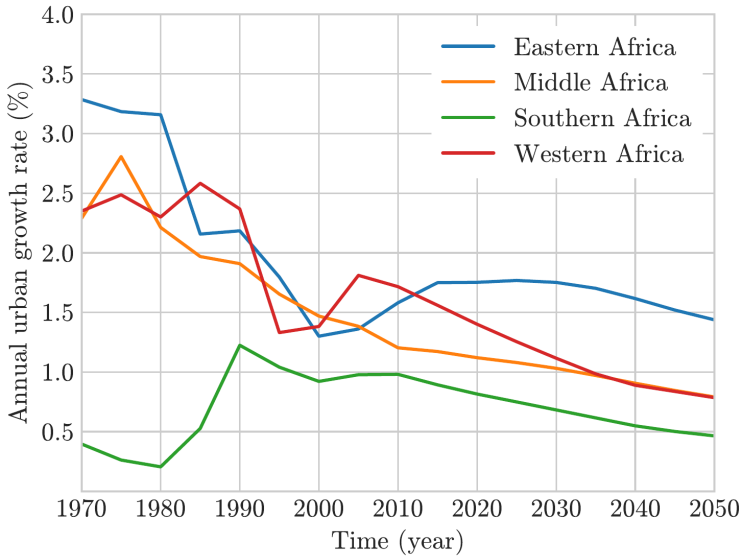

Figure 1.1: Annual urban population growth rates in SSA between 1970 and 2050 according to the World Urbanization Prospects (United Nations, 2017).

and Murayama, 2010), the majority of the growth will actually occur in smaller urban areas with less than 500,000 people (Cohen, 2004).

1.1.2 Social and environmental impacts

Such a rapid growth—especially in a context of urban poverty—has consequences on human health and the environment. Thanks to a better access to health services, urban areas are characterized by lower childhood mortality rates and, on average, a better health than in rural areas (Hay *et al.*, 2005). Likewise, the risks associated with malaria exposure appear to be reduced in urban settings (Tatem, 2004). However, high inequalities are observed depending on the income level (Dye, 2008). In 2015, 64% of urban dwellers in SSA did not have access to a safely managed drinking water service accessible on premises and free from contamination (WHO, 2017). According to the WHO, urban poor are also affected by a large range of environmental factors, including outdoor air pollution, indoor smoke from solid fuels, or physical hazards (car accidents, flooding, fire). Urban areas in SSA are also among

the most affected by the global environmental changes. Sea-level rise increases the risk of flooding for the 60 million people that live in low-elevated coastal lands across Africa (Martine *et al.*, 2008). Salt intrusion into bodies of fresh water and ecosystems degradation threatens the supply of critical resources (water, food, energy, building materials) to the surrounding urban areas. Likewise, the degradation of the agro-climatic conditions in the region may increase the rural-to-urban migrations (Henderson, Storeygard and Deichmann, 2017)—thereby accelerating the growth of the nearby urban areas.

The *urbanization process* is not directly responsible for all the aforementioned issues. Since the 1960s, urban areas in SSA had to absorb a tenfold increase of the region's urban population in a context of (1) inefficient and exclusionary structures inherited from the racist and capitalist-oriented colonial economies (Mabogunje, 1990), and (2) a region-wide economic crisis along with unfavorable terms on trade on the international market and structural adjustment programs (Becker and Morrison, 1997). During this period, the urban population developed its own resilience through the creation of informal economies and settlements, and many have stated their legal integration as a requirement for urban sustainability (Anyamba, 2005; Mutisya and Yarime, 2011; UN-Habitat, 2015). In parallel, local and regional policy making are limited by the poor knowledge of urban dynamics.

1.2 Mapping urbanization from space

1.2.1 From census- to EO-based urban monitoring

To study the *urbanization process*, scientists and policy-makers mainly rely on census-based data provided by the territorial administrations. As a matter of fact, the estimates provided by the United Nations (UN) in their "*World Urbanization Prospects*" are a compilation of data from various national authorities that do not share the same definitions of urban and rural populations. Criteria to distinguish between urban and rural areas include administrative designations, population sizes and densities, or socio-economic variables (United Nations, 2018b). The number of people above which a settlement is considered to be urban ranges from 200 to 50,000 inhab-

itants. In SSA, those thresholds are generally low—2,000 in Ethiopia and Liberia, 5,000 in Zambia and Sudan—thereby potentially leading to an over-estimation of the urbanization levels. These significant temporal and spatial variations of what "urban" means are a major obstacle for local and regional urban planning, as they limit comparative analyses spanning multiple countries. Furthermore, the reliability of census-based data depends on the quality of the censuses and their frequency over time, which can be lower in SSA than in other parts of the world.

In short, demographic data in low-income countries is not detailed enough to support the monitoring of urban dynamics. In that context, EO satellites can provide a monitoring system that is consistent both in space and time. Compared to census-based statistics, urban data extracted from satellite images provide standardized and spatially accurate information on the urban extents. Satellite imagery cannot replace socio-economic and highly detailed census data, but it allows the measurement of one of the most important change caused by urbanization: the expansion of the urban land.

1.2.2 Global urban maps

In the 2000s, the launch of second-generation coarse-resolution EO satellites—such as MODIS, Medium Resolution Imaging Spectrometer (MERIS) and Système pour l'Observation de la Terre (SPOT)—allowed the production of the first global urban maps. As shown in Table 1.1, the open-access policy adopted by the Landsat mission in 2008 led to the release of several medium-resolution global urban—or land cover—maps: the Global Human Settlement Layer (GHSL), the Global Human Built-up And Settlement Extent (HBASE) and GlobeLand30. More recently, the Global Urban Footprint (GUF) made use of high-resolution radar satellites to map the world's built environment at an unprecedented spatial resolution of 12 m, thereby enabling a new range of intra-urban studies (Esch *et al.*, 2018).

Table 1.1: Global maps of human settlements.

Product name	**Data source**	**Resolution**
Global Rural-Urban Mapping Project (CIESIN, 2011)	NOAA's Nighttime Lights	1 km
MODIS Urban Land Cover 500m (Schneider, Friedl and Potere, 2009)	MODIS	500 m
GlobCover v2 (Arino *et al.*, 2007)	MERIS	300 m
Global Human Settlement Layer (Martino Pesaresi and Daniele Ehrlich *et al.*, 2016)	Landsat	30 m
Global Human Built-up & Settlement Extent (Wang *et al.*, 2017)	Landsat	30 m
GlobeLand30 (Chen *et al.*, 2015)	Landsat, HJ-1	30 m
Global Urban Footprint (Esch *et al.*, 2017)	TanDEM-X, TerraSAR-X	12 m

Compared to census-based statistics, urban data extracted from satellite images provide consistent and spatially accurate information on the urban extents. Still, the definition problem was not solved at all. What are the objective conditions for a given surface (a pixel) to be considered as *urban*, or even as *built* ? If the condition refers to the presence of built structures, how much of the surface must be covered ? What kind of structures are considered (elevated buildings, roads) ? In early global maps, divergences on the urban definition caused massive differences regarding the world's total urban extent: 3,727,000 km^2 in Global Rural-Urban Mapping Project (GRUMP), 727,000 km^2 in MODIS-500, and 308,000 km^2 in GlobCover (Potere and Schneider, 2007). Thanks to an increasing accuracy, the difference is lower between newer products but still significant: 774,000 km^2 in the GHSL (Pesaresi *et al.*, 2017), and 834,000 km^2 in the GUF (Esch *et al.*, 2018).

Furthermore, the *urbanization process* is fundamentally dynamic. This means that, to be fully understood, accurate and consistent multi-temporal maps are required.

However, most of the global products provide urban extents for only one or two dates. As of today, the GHSL is the only initiative providing maps for four different dates between 1990 and 2015. Still, going back in time implies relying more on lower-quality sensors and reference data sets. As a result, the accuracy is generally lower for earlier dates, thereby limiting multi-temporal analyses of the urban dynamics. As stated by Singh (1989), the minimum accuracy of a multi-temporal change detection based on multiple classified images is equal to the product of the accuracies of each individual classification. For instance, four classified images with 85% accuracy might have a joint accuracy of only $0.85^4 \approx 0.52$. In other words, the accuracies of today's global products is not high enough to compute reliable urbanization statistics in the temporal dimension.

1.2.3 Urban and built-up definitions

A major source of confusion is the distinction between urban as a *land use* and urban as a *land cover*. *Land use* refers to the function of a given surface for the human societies, whereas *land cover* is about the bio-physical coverage of the land. In practice, urban land use may include a wide range of elements such as parks, private gardens, residential roads, cemeteries, industrial dumps or golf courses. Urban land cover is more difficult to characterize and no consensus has been reached to this day (Potere and Schneider, 2007)—mainly due to the various geographic objects from which urban areas are made up. That is why most remote sensing studies focus instead on the *built environment* or *built-up areas*. For instance, Potere *et al.* (2009) and Mertes *et al.* (2015) defined urban areas as places dominated by the built environment, where the "built environment" includes all human-constructed elements (roads, buildings, industrial facilities, etc.) and excludes vegetated urban surfaces such as parks or green spaces. In the GHSL, the term of *"built-up"* is preferred and is expressed as a continuous value that corresponds to the proportion of building footprints in a given area—thereby excluding roads from the definition (Pesaresi *et al.*, 2017).

In the present thesis, we refer to *urban* and *rural* areas as geographic objects identified by various demographical and socio-economical variables such as population

densities, economic activities or public infrastructure. However, we focus on the detection of *built-up areas*—as opposed to *not built-up areas*—which can be found in both *urban* and *rural* areas. As stated previously, built-up pixels have been defined as a surface covered by at least 50% of human-constructed elements (Potere *et al.*, 2009; Mertes *et al.*, 2015). In this thesis, we extend the definition by lowering the threshold to 25% to better account for mixed land cover in urban areas (gardens, bare soil, non-asphalted roads, isolated buildings, etc.). Validation metrics are computed according to the aforementioned definition, as well as the comparative assessments with other available urban maps.

Likewise, we refer to *urbanization* as the increasing share of the urban population, and to *built-up growth* as the increasing surface covered by buildings—regardless of the demographics.

1.3 Urban remote sensing in SSA

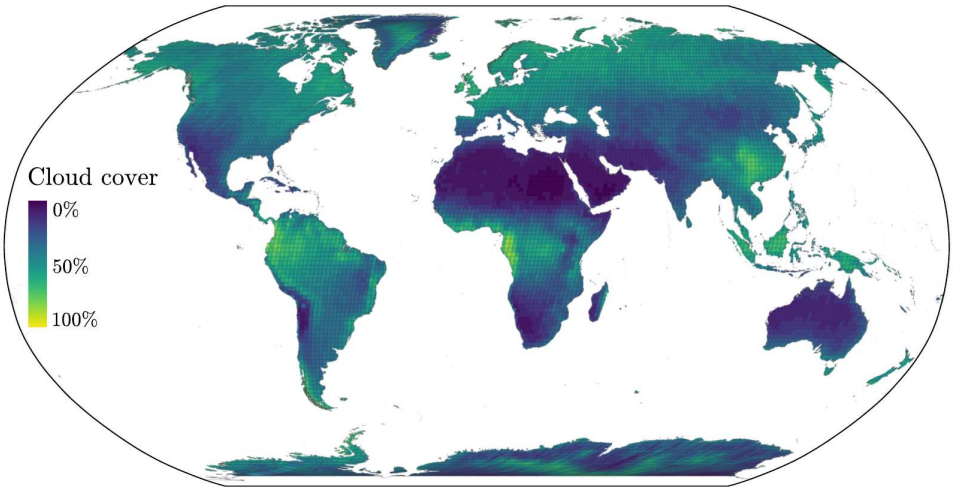

Figure 1.2: Average cloud cover of Landsat 5, Landsat 7 and Landsat 8 scenes.

As previously stated, SSA concentrates most of the stakes associated with urbanization and land cover changes. Yet, the reliability of global maps is generally lower than in other regions. Potere *et al.* (2009) compared the accuracy of eight global urban maps in different world regions and found that misclassifications were more fre-

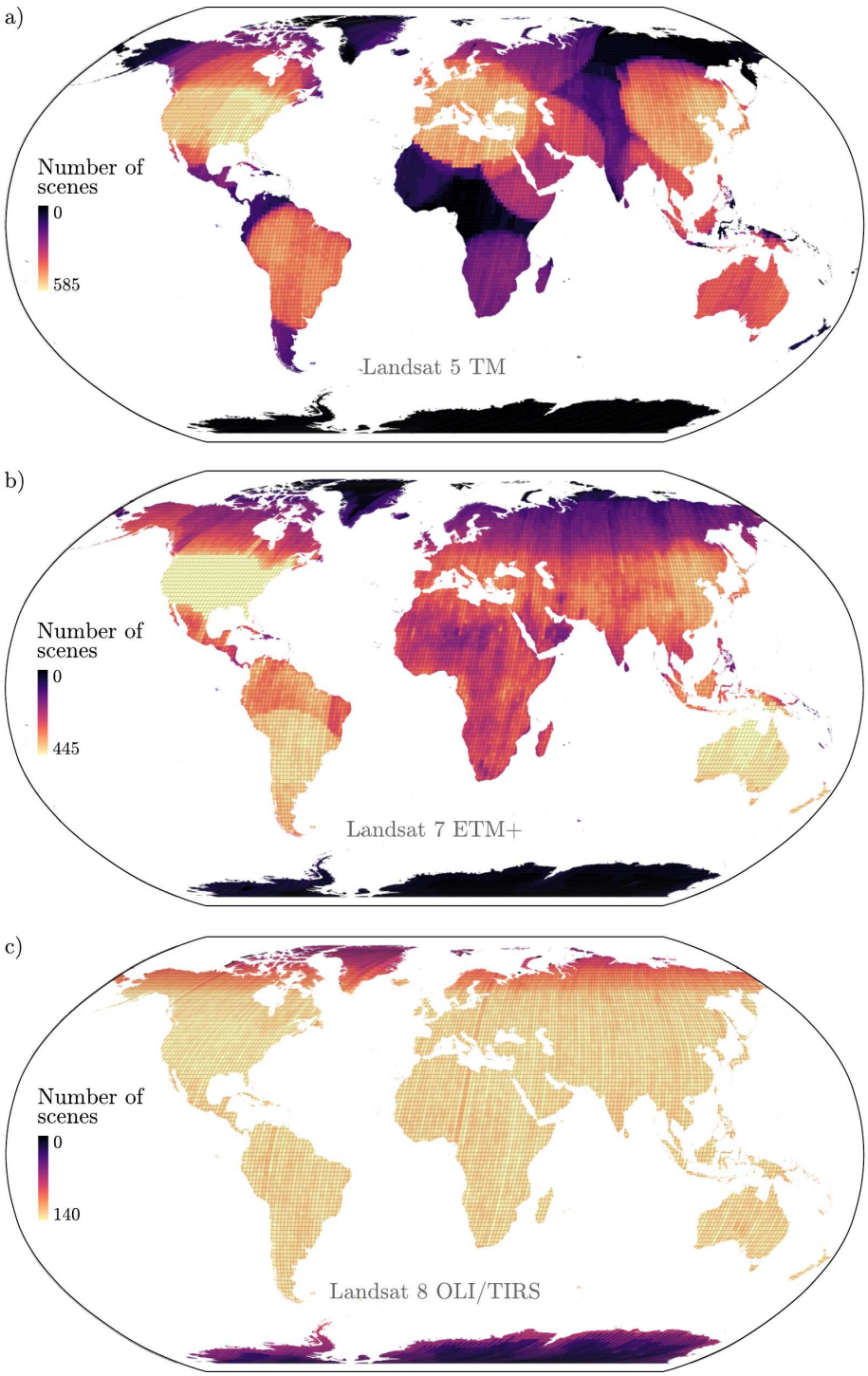

Figure 1.3: Number of available scenes in the Landsat catalog for **a)** Landsat 5 TM, **b)** Landsat 7 ETM+, and **c)** Landsat 8 OLI/TIRS.

quent in Africa (average Kappa value of 0.46) than in industrialized countries (0.57). Likewise, Esch *et al.* (2017) assessed four global products (GUF, GHSL, GlobeLand30 and MODIS-500) in twelve urban areas. In the four products, the average accuracy of the six urban areas located in SSA (Addis Ababa, Dar Es Salaam, Kampala, Kigali, Lagos and Nairobi) was significantly lower than in the other cities (Athens, Beijing, Mexico City, Milan, New York and Perth). The regional gaps in map accuracies may be explained by two main factors: (1) low data availability, and (2) the heterogeneous nature of the urban environment.

1.3.1 Data availability

The accuracy of any land cover classification based on satellite remote sensing depends on the quality and quantity of the input data at three different levels: the *image level*, the *training level* and the *testing level*. The *image level* refers to the images acquired by the satellite sensors. It is mainly sensitive to the quality of the satellite acquisition—the atmospheric conditions at the time of acquisition—, the quality of the pre-processing routines that has been applied—orthorecitification, atmospheric correction, cloud detection—, and the spatial resolution of the sensor. Figure 1.2 shows the average cloud cover of Landsat scenes across the globe. If the African continent is not particularly cloudy, some of the most populated regions of SSA are susceptible to poor atmospheric conditions, such as the West African littoral and the African Great Lakes region. In practice, this implies a lower quantity of data to work with, or the inability to take advantage of the seasonal information to discriminate the built environment. For instance, only five Landsat scenes with less than 10% of cloud cover are available over Kinshasa, D.R. Congo—and none between 1995 and 2005. Furthermore, until recently, many satellite acquisitions over SSA have not been held in the Landsat archive. As shown in Figure 1.3, SSA suffers from a very low availability of Landsat 5 TM scenes compared to the rest of the world—comparable to Siberia, Antarctica or Alaska. This is significantly prejudicial to map urbanization dynamics in SSA since the satellite covers a period of 28 years from 1984 to 2012. However, there is a greater proportion of African acquisitions in the Landsat 7 catalog and, thanks to a more systematic approach for

global acquisition (Roy *et al.*, 2010), the situation is much better with Landsat 8.

The *training* and *testing* levels depend on the availability of ground truth data, that is used (1) to support the training of the classification models in *machine learning* approaches, and (2) to validate the classification outputs. In satellite remote sensing, ground truth data is traditionally collected through on-site measurements, but authors tend to rely more on external datasets when working on large study areas. In the field of urban mapping, these datasets are mainly provided by governmental services, cartographic centers or commercial agencies. The scarcity of such datasets in poor countries affects the classification accuracies in multiple ways. First, less training data implies a decreased representativeness of the target geographic object. Urban areas are fundamentally heterogeneous and made up of a wide range of elements (roofs, asphalt, trees, green spaces, pools, bare soil, etc.) composed of different materials and laid out in varying proportions and shapes (Herold *et al.*, 2004). Classification models must be able to learn all those possible forms of the built environment from the training data, failing which entire human settlements could stay undetected. The lower availability of reference datasets to validate the classification models may lead to miss such misclassifications and, more importantly, communicate erroneous metrics to the end-users. Overall, this means that classification models are either (1) trained with less data, thereby leading to higher rates of misclassification, or (2) calibrated in other data-rich regions and suffering from lower accuracies when applied in SSA.

1.3.2 Urban heterogeneity and spectral confusion

> *"The most consistent characteristic of the urban mosaic is spectral heterogeneity."* (Small, 2005)

The other issues that may be encountered when mapping urban areas in SSA through satellite remote sensing are caused by the heterogeneity that characterizes both the natural and the urban environment of the region. Some of the most populated urban areas in SSA are located in arid and tropical regions, where built-up areas may be harder to detect. Mapping built-up areas in arid regions is a notorious issue in urban remote sensing (Zhang, Chen and Lu, 2015; Rasul

et al., 2018). From a remote sensing point of view, there is no clear boundary between bare soil and built-up areas in arid regions when using optical sensors. The spectral signatures of both land cover types are similar, leading to *spectral confusion* and misclassifications. In urban areas where buildings are made up of materials that belong to the surrounding natural environment, the problem may even be worse—as shown in Figure 1.4 for Gao, Mali. Likewise, the urban mosaics located in tropical regions may be so densely vegetated that their spectral signature becomes similar to the ones of other vegetated land cover types.

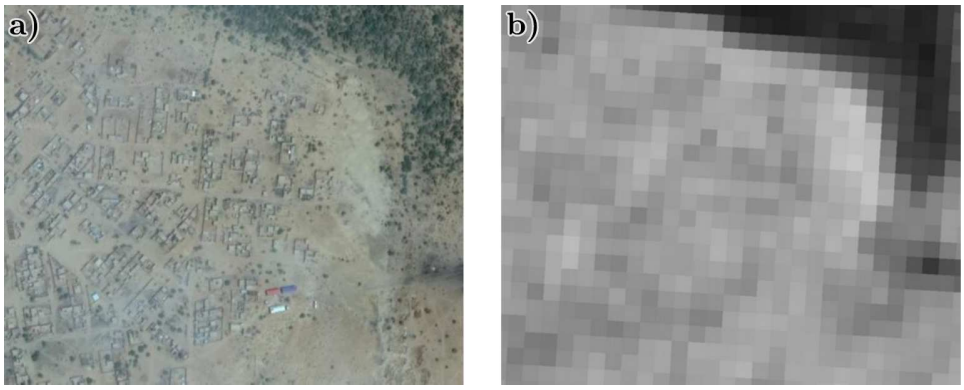

Figure 1.4: Spectral confusion between bare soil and built-up areas in Gao, Mali. **a)** Satellite image of the area of interest, courtesy of Google, and **b)** Near-infrared Landsat band.

Figure 1.5: Inter-urban heterogeneity in **a)** Gao, Mali, **b)** Johannesburg, South Africa, and **c)** Katsina, Nigeria. Satellite imagery courtesy of Google.

More generally, the spectral confusion caused by the combination of different spectral objects in a given pixel is called the "mixed pixel" effect. The issue is predominant in urban remote sensing, since medium-resolution (from 15 to 50 m) urban pix-

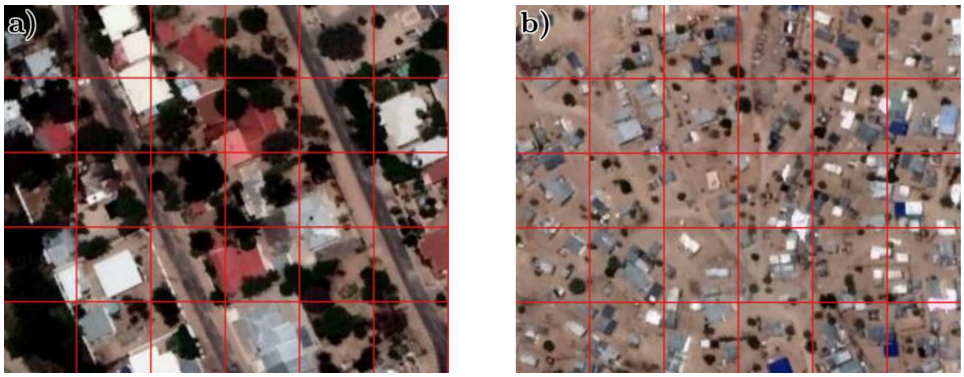

Figure 1.6: Intra-urban heterogeneity in Windhoek, Namibia. Each red square has a surface of 900 m, which corresponds to a Landsat pixel. Satellite imagery courtesy of Google.

els will always be made up of multiple objects with various spectral properties (roofs of various materials, asphalted surfaces of different age, trees, gardens, patches of bare soil, etc.). That is why Small (2005) states that "*the most consistent characteristic of the urban mosaic is its spectral heterogeneity*". The spectral variations of the urban mosaic are both *inter-urban* and *intra-urban*. *Inter-urban* heterogeneity (among different cities) is the result of socio-economic, cultural, historical or environmental differences among the cities of the world, as illustrated in Figure 1.5. In the context of urban remote sensing, this means that a model built for a given urban area cannot be applied on another: the targeted spectral object has changed. As illustrated in Figure 1.6, variations may also be *intra-urban*—among different parts of a given city. In an approach based on machine learning, this means that the representativeness of the training data will be determinant in a model's ability to detect the heterogeneous set of built-up areas in a given city.

1.4 Open data opportunities

1.4.1 Open-access satellite imagery

Coarse-resolution satellite imagery (250–1,500 m) has been free and open for many years—the most famous examples being the Advanced Very-High-Resolution

Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS). But, traditionally, medium-resolution satellite imagery (10–40 m) had to be paid for. The open-access policy adopted by the Landsat program in 2008 has changed that and enabled a dramatic increase in the use of satellite remote sensing for a wide range of applications (Woodcock *et al.*, 2008; Turner *et al.*, 2015), by making freely available more than 25 years of medium-resolution imagery. This positive impact has been illustrated by the significant increase in the distribution of images by the USGS: 25,000 Landsat images were provided to users in 2001 when each individual scene was charged $600, against 2.5 million in 2010 (Wulder *et al.*, 2012). As shown in Table 1.2, many space agencies followed the lead, such as the European Space Agency with the launch of the Sentinel program in 2014, or the China-Brazil Earth Resources Satellite (CBERS) program. As of today, the archives of more than ten medium-resolution satellite imagery datasets have been made freely available. This era of "Big Earth Data" provides opportunities as well as challenges for the EO community (Guo, Wang and Liang, 2016). Provided the computing capacity, open "Big Earth Data" enables land cover mapping at regional- and global-scales at relatively low costs (see, for instance, section 1.2 on global urban maps), and allows a new range of methods that make use of dense timeseries.

Table 1.2: Main open-access satellite imagery datasets.

Sensor	**Type**	**Period**	**Resolution**
Landsat 4/5 MSS	Optical	1982–2013	60 m
Landsat 4/5 TM	Optical	1982–2012	30 m
ERS-1	SAR	1991–2000	25 m
JERS-1	Both	1992–1998	12.5 m
ERS-2	SAR	1995–	25 m
MODIS	Optical	1999–	250 m
Landsat 7 ETM+	Optical	1999–	30 m
CBERS-1/2/3/4	Optical	1999–	20 m
Envisat MERIS	Optical	2002–2012	300 m

Sensor	Type	Period	Resolution
Envisat ASAR	SAR	2002–2012	30 m
Landsat 8 OLI/TIRS	Optical	2013–	30 m
Sentinel-1	SAR	2014–	10 m
Sentinel-2	Optical	2015–	10 m

Another opportunity lies in the availability of multiple sensors with different characteristics. In the context of urban mapping, the theoretical complementarity of multispectral and SAR sensors has been stated many times (Dell'Acqua, 2009). Furthermore, the use of sensors which are not sensitive to clouds—such as SAR—becomes essential in tropical and cloudy regions (see section 1.3.1). Still, combining image data acquired with different sensors and in varying atmospheric conditions remains a challenge.

1.4.2 Volunteered Geographic Information

The increasing availability of spatial data is not limited to satellite imagery. EO satellites provide a systematic and continuous spatial information on the Earth's surface. On the contrary, geographic information databases provide a labeled and discrete information—for instance a road network. For centuries, this function has been reserved to commercial or governmental agencies. But since the 2000s, user-generated mapping has been enabled by the availability of low-cost GPS receivers along with the rise of Internet and home computing. This collaborative web-based effort to collect and share spatial data is called "Volunteered Geographic Information" (VGI) (Goodchild, 2007). The leading example of VGI is OSM, that follows the collaborative model of Wikipedia to create a map of the world whose data is free to use or edit (Haklay and Weber, 2008). Funded in 2004, OpenStreetMap (OSM) reached 50,000 registered users by 2008. In addition to the users contributions, local authorities and commercial navigation information providers started to donate geographic information to the project, thereby making it more and more complete. Today, OSM data cover the entire world and a wide range of spatial data—from road networks to building footprints, parks, points of interest, water bodies, etc.

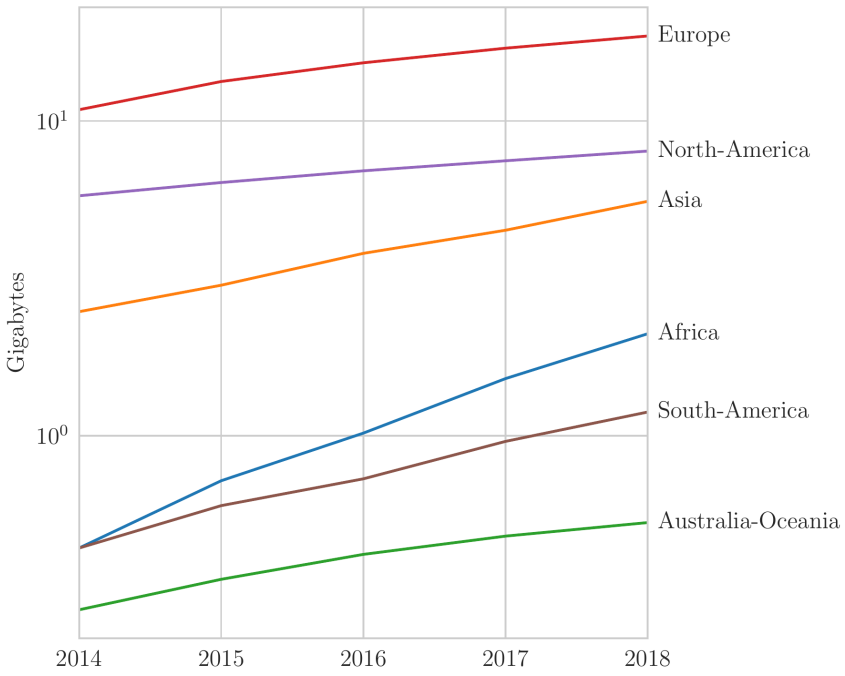

Figure 1.7: Gigabytes of information in the OSM database for each continent between 2014 and 2018 on a logarithmic scale.

Still, the reliability of VGI can be low depending on the region of interest. As a matter of fact, Europe and North America were the main focus of the project until recently. Figure 1.7 shows the gigabytes of information available in the OSM database for each continent between 2014 and 2018. Because of what Goodchild (2007) called the "Digital Divide", OSM data has always been more scarce in the African, South-American and Asian continents. Before 2014, data located in Africa accounted for only 2% of the OSM database 440 MB. But since then, the uneven spatial distribution of OSM data has been partially mitigated: between 2014 and 2018, 11% of the contributed data was located in Africa, for a total increase of 380%. The evolution of the share of Europe and North America in the database reveals the same trend: 81% in 2014, against 73% in 2018. This increasing amount of geographic data available for free is an opportunity for scientists and policy-makers working on developing regions, for which geographic information databases were relatively scarce.

1.4.3 Thesis input data

The open-access datasets supporting this thesis can be divided into two categories: (1) satellite imagery, and (2) volunteered geographic information. Table 1.3 lists these datasets together with the period they cover in the context of the study. Satellite imagery catalogs have been selected according to three criteria: access policies, spatial resolution (at least 30 m), and automation opportunities. Furthermore, because of the multi-temporal dimension of urbanization, consistency between sensors must be taken into account. This is why Sentinel-2 multispectral imagery have been excluded from the study, despite its higher spatial resolution.

Table 1.3: Main datasets used in this thesis and the period they cover.

Dataset	**Type**	**Period Covered**
OpenStreetMap	Volunteered Geographic Information	2015–
Landsat 5	Multispectral Satellite Imagery	1995–2010
Landsat 7	Multispectral Satellite Imagery	2000
Landsat 8	Multispectral Satellite Imagery	2015

Dataset	Type	Period Covered
ERS-1	SAR Satellite Imagery	1995–2000
ERS-2	SAR Satellite Imagery	1995–2010
Sentinel-1	SAR Satellite Imagery	2015

1.5 Thesis organization

1.5.1 Objectives and research hypotheses

The objectives of the present thesis are two-fold. First, we aim to provide insights regarding the improvement of regional- and global-scale maps of the built environment in the context of SSA. To that end, we present a novel approach to map built-up areas at medium-resolution (30 m) that is based on multi-sensor satellite data and VGI. The second objective is to create a dataset of urban expansion in 46 urban areas of SSA for the 1995–2015 period, in order to produce knowledge regarding the relationship between the population dynamics and the expansion of the built environment.

From a methodological point of view, the present work relies on two main research hypotheses.

1.5.1.1 1st hypothesis

> *In the data-scarce context of SSA, VGI can enable large-scale mapping of built-up areas by supporting the training of the classification models without reducing their performances.*

In *supervised learning*, the main factor that determines the performance of a classification model is the quality of the input training data. Traditionally, training samples are collected by conducting in situ measurements or digitization of very high resolution imagery. Both processes are manual and time-consuming, especially when working on regional or global scales. In that case, many rely on external geographic databases maintained by governmental or commercial agencies. As pre-

viously stated in section 1.3.1, the data-scarce context of SSA makes it difficult to rely on such datasets. One of the solution is to invest more time is samples collection ; another is to train the classification models in other data-rich regions. In the end, this implies sacrificing either classification performance or time. In the present thesis, an alternative approach based on VGI (see section 1.4.2) is proposed. More specifically, the use of OSM to collect training data is investigated. Several challenges are expected. First, OSM is the result of a collaborative effort, and thereby may lack completeness and reliability. In some regions, this could lead to a lack of representativeness in the training data—thereby causing lower classification performances. Secondly, the OSM database is primarily focused on the urban environment. This means that samples corresponding to the built environment could be easier to collect than information related to the surrounding natural environment. Chapter 2 proposes an implementation of this approach and discusses its limitations. The proposed approach is assessed in the context of ten different urban areas located in SSA.

1.5.1.2 2nd hypothesis

> *Using SAR sensors (ERS, Envisat, Sentinel-1) in combination with multispectral data (Landsat) can lead to a more consistent detection of the built environment, especially in arid regions.*

Across the world, urban mosaics are always characterized by a high spectral heterogeneity. At medium-resolution (10–40 m), a built-up pixel may contain a wide range of geographic objects such as buildings, roads, bare soil, gardens or pools. As stated in section 1.3.2, the issue is exacerbated when a given built-up pixel is partially made up of elements similar to the surrounding natural environment, such as dense urban vegetation in a tropical environment. Spectral confusion is particularly severe in arid regions, where built-up and bare land may share a similar spectral signature. Contrary to optical sensors, SAR is mainly sensitive to the ground texture, and thereby consistently able to discriminate between flat bare land and elevated built-up structures. In that sense, optical and SAR sensors complement one another. Still, data fusion is not straightforward. For instance, SAR images are char-

acterized by the presence of a granular "speckle noise"—making them less suited to pixel-level classifications. Moreover, because of the sensitivity of SAR to ground texture, densely vegetated areas can be confused with elevated built-up structures. In chapter 3, an implementation based on both Landsat 8 and Sentinel-1 imagery is proposed. The approach is assessed and discussed in the context of 12 urban areas of SSA.

Finally, the methodological contributions presented in chapter 2 and 3 are leveraged to create an urban expansion dataset for 46 case studies located in SSA for the 1995–2015 period. Chapter 4 presents an implementation based on seven open-access EO datasets: Landsat 5 TM, Landsat 7 ETM+, Landsat 8, ERS-1, ERS-2, Envisat and Sentinel-1. Here, spatio-temporal consistency is expected to be the main challenge. Given that OSM may lack spatial homogeneity, is the approach equally suited in all urban areas, regardless of population size and data availability ? Considering the use of seven sensors, are the results sufficiently consistent in time to compute urbanization statistics ? In this chapter, the proposed approach is assessed and discussed. Furthermore, the results obtained are compared to existing global urban maps, and are used to analyze the relationship between built-up expansion and population growth.

1.5.2 A note on reproducibility

> "*An article about computational results is advertising, not scholarship. The actual scholarship is the full software environment, code and data, that produced the result.*" (Buckheit and Donoho, 1995)

Reproducibility refers to an ensemble of practices that enable reuse, redistribution and reproduction of the research outputs of a study (Sandve *et al.*, 2013; Stodden and Miguez, 2014). In practice, this implies sharing the code that supports a given analysis along with the full software environment and the input data. In the field of EO, two main challenges can be identified. First, a wide range of software solutions are used—for instance for image preprocessing and classification. These software packages can be written in various programming languages (Python, R, C/C++, FORTRAN) and are not always available on all operating systems. This can make

the reproduction of the software environment difficult, thereby limiting the ability of the reader to actually run the shared code. Second, EO data can be huge—especially when working on regional and global scales. This makes it impossible to share the raw input data in conventional data repositories. In that case, two different approaches can be adopted: sharing only the output data, or sharing the scripts that have been used for data acquisition. Overall, a reproducible approach in EO studies may includes up to five dimensions:

1. Allowing the reproduction of the software environment. This means providing a way to easily recreate a full software environment, that can be made up of various software packages in specific versions and from different programming languages. Package management systems such as Conda[1] or container-based technologies such as Docker[2] can be used to solve the problem.

2. Enabling the reproduction of the research outputs. This includes sharing the code that generates any intermediary result supporting methodological decisions and output data. It also implies providing a way to easily run the code without having to manually modify the source—for instance by providing Python modules and scripts instead of isolated chunks of code. More importantly, code must be shared along with its documentation and, optionally, with its version control history—for instance through online code sharing platforms such as Gitlab[3] or Github[4].

3. Sharing interactive analyses. This includes sharing exploratory analyses, unpublished figures or failed experiments—in other words, anything that supported decision making. It can also refers to in-depth data stories that are not suited to the format of a journal article or a thesis dissertation. Interactive notebook environments such as RMarkdown[5] or Jupyter[6] are perfectly suited for that type of task.

4. Data sharing. This means making available any input data required by the provided code. It also includes providing output data to make it accessible without

[1] https://conda.io
[2] https://docker.com
[3] https://gitlab.com
[4] https://github.com
[5] https://rmarkdown.rstudio.com
[6] https://jupyter.org

having to run the code again. To that end, one can rely on research data repositories such as Dataverse[7] or Zenodo[8]. In EO, raw data may be too large to be uploaded on the aforementioned platforms. In that case, an alternative is to provide the code for data acquisition.

The present thesis follows the proposed approach and uses Conda, Github, Jupyter and Zenodo to ensure its reproducibility. Code and data are provided on a per-chapter basis:

Chapter 2
Data: https://zenodo.org/record/1291961
Code: https://github.com/yannforget/builtup-classification-osm

Chapter 3
Data: https://zenodo.org/record/1450932
Code: https://github.com/yannforget/landsat-sentinel-fusion

Chapter 4
Data: https://zenodo.org/record/3234908
Code: https://github.com/yannforget/maupp

Additionally, the dataset produced in the fourth chapter can be downloaded or visualized in interactive maps on the website of the project (https://maupp.ulb.ac.be/page/wp1results/).

References

Anyamba, T. J. C. (2005) 'Nairobi's Informal Modernism', in *6th N-AERUS Workshop, Promoting Social Inclusion in Urban Areas: Policies and Practice*. Lund, Sweden.

Arino, O. *et al.* (2007) 'Globcover-a global land cover service with MERIS', in *Proceedings of the ENVISAT Symposium*, pp. 23–27.

Becker, C. M. and Morrison, A. R. (1997) 'Beyond Urban Bias in Africa: Urbanization in an Era of Structural Adjustment', *The International Journal of African Historical Studies*, 29(3), p. 597. doi: 10.2307/221370.

[7] https://dataverse.org
[8] https://zenodo.org

Bocquier, P. (2003) 'Analyzing Urbanization in Sub-Saharan Africa', in *New Forms of Urbanization*. London: Routledge, pp. 133–150.

Buckheit, J. B. and Donoho, D. L. (1995) 'WaveLab and Reproducible Research', in Bickel, P. et al. (eds) *Wavelets and Statistics*. New York, NY: Springer New York, pp. 55–81. doi: 10.1007/978-1-4612-2544-7_5.

Chen, J. *et al.* (2015) 'Global land cover mapping at 30m resolution: A POK-based operational approach', *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp. 7–27. doi: 10.1016/j.isprsjprs.2014.09.002.

CIESIN (2011) 'Global Rural-Urban Mapping Project, Version 1 (GRUMPv1)'. doi: 10.7927/H4GH9FVG.

Cohen, B. (2004) 'Urban Growth in Developing Countries: A Review of Current Trends and a Caution Regarding Existing Forecasts', *World Development*, 32(1), pp. 23–51. doi: 10.1016/j.worlddev.2003.04.008.

Dell'Acqua, F. (2009) 'The Role of SAR Sensors', in *Global mapping of human settlement: Experiences, datasets, and prospects*. Boca Raton: CRC Press (Taylor & Francis series in remote sensing applications), pp. 309–319.

Dye, C. (2008) 'Health and Urban Living', *Science*, 319(5864), pp. 766–769. doi: 10.1126/science.1150198.

Esch, T. *et al.* (2018) 'Where We LiveA Summary of the Achievements and Planned Evolution of the Global Urban Footprint', *Remote Sensing*, 10(6), p. 895. doi: 10.3390/rs10060895.

Esch, T. *et al.* (2017) 'Breaking new ground in mapping human settlements from space The Global Urban Footprint', *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, pp. 30–42. doi: 10.1016/j.isprsjprs.2017.10.012.

Fox, S. (2012) 'Urbanization as a Global Historical Process: Theory and Evidence from sub-Saharan Africa', *Population and Development Review*, 38(2), pp. 285–310. doi: 10.1111/j.1728-4457.2012.00493.x.

Goodchild, M. F. (2007) 'Citizens as sensors: The world of volunteered geography', *GeoJournal*, 69(4), pp. 211–221. doi: 10.1007/s10708-007-9111-y.

Guo, H., Wang, L. and Liang, D. (2016) 'Big Earth Data from space: A new engine for Earth science', *Science Bulletin*, 61(7), pp. 505–513. doi: 10.1007/s11434-016-1041-y.

Haklay, M. and Weber, P. (2008) 'OpenStreetMap: User-Generated Street Maps', *IEEE Pervasive Computing*, 7(4), pp. 12–18. doi: 10.1109/MPRV.2008.80.

Hay, S. I. *et al.* (2005) 'Urbanization, malaria transmission and disease burden in Africa: Opinion tropical infectious diseases', *Nature Reviews Microbiology*, 3(1), pp. 81–90. doi: 10.1038/nrmicro1069.

Henderson, J. V., Storeygard, A. and Deichmann, U. (2017) 'Has climate change driven urbanization in Africa?', *Journal of Development Economics*, 124, pp. 60–82. doi: 10.1016/j.jdeveco.2016.09.001.

Herold, M. *et al.* (2004) 'Spectrometry for urban area remote sensingDevelopment and analysis of a spectral library from 350 to 2400 nm', *Remote Sensing of Environment*, 91(3-4), pp. 304–319. doi:

10.1016/j.rse.2004.02.013.

Jedwab, R. and Vollrath, D. (2015) 'Urbanization without growth in historical perspective', *Explorations in Economic History*, 58, pp. 1–21. doi: 10.1016/j.eeh.2015.09.002.

Mabogunje, A. L. (1990) 'Urban Planning and the Post-Colonial State in Africa: A Research Overview', *African Studies Review*, 33(2), p. 121. doi: 10.2307/524471.

Martine, G. *et al.* (2008) *The new global frontier: Urbanization, poverty and environment in the 21st century*. London ; Sterling, VA: Earthscan.

McGranahan, G. *et al.* (2009) *Africa's urban transition and the role of regional collaboration*. 5. London: International Institute for Environment; Development (IIED).

Mertes, C. M. *et al.* (2015) 'Detecting change in urban areas at continental scales with MODIS data', *Remote Sensing of Environment*, 158, pp. 331–347. doi: 10.1016/j.rse.2014.09.023.

Mundia, C. N. and Murayama, Y. (2010) 'Modeling Spatial Processes of Urban Growth in African Cities: A Case Study of Nairobi City', *Urban Geography*, 31(2), pp. 259–272. doi: 10.2747/0272-3638.31.2.259.

Mutisya, E. and Yarime, M. (2011) 'Understanding the grassroots dynamics of slums in Nairobi: The dilemma of Kibera informal settlements', *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*, 2(2), pp. 197–213.

Pesaresi, M. *et al.* (2016) *Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014*. Joint Research Centre, European Commission.

Pesaresi, M. *et al.* (2017) *Atlas of the Human Planet 2016: Mapping Human Presence on Earth with the Global Human Settlement Layer*. Joint Research Centre, European Commission.

Potere, D. and Schneider, A. (2007) 'A critical look at representations of urban areas in global maps', *GeoJournal*, 69(1), pp. 55–80. doi: 10.1007/s10708-007-9102-z.

Potere, D. *et al.* (2009) 'Mapping urban areas on a global scale: Which of the eight maps now available is more accurate?', *International Journal of Remote Sensing*, 30(24), pp. 6531–6558. doi: 10.1080/01431160903121134.

Rasul, A. *et al.* (2018) 'Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates', *Land*, 7(3), p. 81. doi: 10.3390/land7030081.

Roy, D. P. *et al.* (2010) 'Accessing free Landsat data via the Internet: Africa's challenge', *Remote Sensing Letters*, 1(2), pp. 111–117. doi: 10.1080/01431160903486693.

Sandve, G. K. *et al.* (2013) 'Ten Simple Rules for Reproducible Computational Research', *PLoS Computational Biology*. Edited by P. E. Bourne, 9(10), p. e1003285. doi: 10.1371/journal.pcbi.1003285.

Schneider, A., Friedl, M. A. and Potere, D. (2009) 'A new map of global urban extent from MODIS satellite data', *Environmental Research Letters*, 4(4), p. 044003. doi: 10.1088/1748-9326/4/4/044003.

Seto, K. C. *et al.* (2011) 'A Meta-Analysis of Global Urban Land Expansion', *PLoS ONE*. Edited by J. A. Añel, 6(8), p. e23777. doi: 10.1371/journal.pone.0023777.

Seto, K. C., Guneralp, B. and Hutyra, L. R. (2012) 'Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools', *Proceedings of the National Academy of Sciences*, 109(40), pp. 16083–16088. doi: 10.1073/pnas.1211658109.

Singh, A. (1989) 'Review Article Digital change detection techniques using remotely-sensed data', *International Journal of Remote Sensing*, 10(6), pp. 989–1003. doi: 10.1080/01431168908903939.

Small, C. (2005) 'A global analysis of urban reflectance', *International Journal of Remote Sensing*, 26(4), pp. 661–681. doi: 10.1080/01431160310001654950.

Stodden, V. and Miguez, S. (2014) 'Best Practices for Computational Science: Software Infrastructure and Environments for Reproducible and Extensible Research', *Journal of Open Research Software*, 2(1). doi: 10.5334/jors.ay.

Tatem, A. J. (2004) 'Measuring Urbanization Pattern and Extent for Malaria Research: A Review of Remote Sensing Approaches', *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 81(3), pp. 363–376. doi: 10.1093/jurban/jth124.

Turner, W. *et al.* (2015) 'Free and open-access satellite data are key to biodiversity conservation', *Biological Conservation*, 182, pp. 173–176. doi: 10.1016/j.biocon.2014.11.048.

UN-Habitat (2015) *Sustainable Urban Development in Africa*. Nairobi, Kenya: UN-Habitat.

United Nations (2015) 'World Urbanization Prospects: The 2014 Revision'.

United Nations (2017) *World Population Prospects: The 2017 Revision*. New York.

United Nations (2018) *World Urbanization Prospects: The 2018 Revision, Methodology*. Working Paper No. ESA/P/WP.252. New York.

Wang, P. *et al.* (2017) 'Human Built-up And Settlement Extent (HBASE) Dataset From Landsat'. NASA Socioeconomic Data; Applications Center (SEDAC).

WHO (2017) *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. Geneva.

Woodcock, C. E. *et al.* (2008) 'Free Access to Landsat Imagery', *Science*, 320(5879), pp. 1011a–1011a. doi: 10.1126/science.320.5879.1011a.

Wulder, M. A. *et al.* (2012) 'Opening the archive: How free data has enabled the science and monitoring promise of Landsat', *Remote Sensing of Environment*, 122, pp. 2–10. doi: 10.1016/j.rse.2012.01.010.

Zhang, C., Chen, Y. and Lu, D. (2015) 'Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery', *International Journal of Remote Sensing*, 36(17), pp. 4483–4500. doi: 10.1080/01431161.2015.1084552.

Chapter 2

Automated supervised learning based on OpenStreetMap

This chapter is based on the following publication:

Forget Y., Linard C., & Gilbert M. (2018). "Supervised Classification of Built-Up Areas in Sub-Saharan African Cities Using Landsat Imagery and OpenStreetMap". *Remote Sensing*, 10 (7). 10.3390/rs10071145

Abstract. The Landsat archives have been made freely available in 2008, allowing the production of high resolution built-up maps at the regional or global scale. In this context, most of the classification algorithms rely on supervised learning to tackle the heterogeneity of the urban environments. However, at a large scale, the process of collecting training samples becomes a huge project in itself. This leads to a growing interest from the remote sensing community toward Volunteered Geographic Information (VGI) projects such as OpenStreetMap (OSM). Despite the spatial heterogeneity of its contribution patterns, OSM provides an increasing amount of information on the earth's surface. More interestingly, the community has moved beyond street mapping to collect a wider range of spatial data such as building footprints, land use, or points of interest. In this paper, we propose a classification method that makes use of OSM to automatically collect training samples

for supervised learning of built-up areas. To take into account a wide range of potential issues, the approach is assessed in ten Sub-Saharan African urban areas from various demographic profiles and climates. The obtained results are compared with: (1) existing high resolution global urban maps such as the Global Human Settlement Layer (GHSL) or the Human Built-up and Settlements Extent (HBASE); and (2) a supervised classification based on manually digitized training samples. The results suggest that automated supervised classifications based on OSM can provide performances similar to manual approaches, provided that OSM training samples are sufficiently available and correctly pre-processed. Moreover, the proposed method could reach better results in the near future, given the increasing amount and variety of information in the OSM database.

2.1 Introduction

The population of Africa is predicted to double by 2050 (UN-Habitat, 2014), alongside a rapidly growing urbanization. In this context, reliable information on the distribution and the spatial extent of human settlements is crucial to understand and monitor a large set of associated issues, such as the impacts on both environmental systems and human health (Grimm *et al.*, 2008; Dye, 2008; Wentz *et al.*, 2014). In the 2000s, the remote sensing community took advantage of the availability of coarse-resolution satellite imagery based on the MERIS or the MODIS programs to produce global land cover maps such as GlobCover (Arino *et al.*, 2007) or the MODIS 500m Map of Global Urban Extent (MOD-500) (Schneider, Friedl and Potere, 2009). Subsequently, Landsat data have been made freely available in 2008 and dramatically reduced the operative cost of high resolution satellite imagery acquisition and processing (Wulder *et al.*, 2008). This enabled the production of high resolution global land cover maps based on the Landsat catalog, such as the GHSL (Martino Pesaresi and Daniele Ehrlich *et al.*, 2016), Global Land Cover (GLC) (Chen *et al.*, 2015) or the HBASE (Wang *et al.*, 2017).

However, land cover classification in urban areas remains a challenge because of the inherent complexity of the urban environment which is characterized by both in-

traurban and interurban heterogeneity (Herold *et al.*, 2004; Small, 2005). Because of the complexity associated with the urban mosaic at high resolution, supervised classification methods have been shown to provide the best results (Gamba and Herold, 2009; Li, 2014; Belgiu and Drăguț, 2016). However, such an approach requires a large amount of training samples to grasp the heterogeneity of the urban environment. As a result, the process of collecting training samples for large-scale supervised classification becomes an unaffordable task. Additionally, studies have shown that global urban maps suffer from higher rates of misclassifications in developing regions such as Sub-Saharan Africa or South Asia (Potere *et al.*, 2009) because of the lack of reference data in both quantity and quality.

The training samples collection step can be automated by using existing land cover information in ancillary datasets. Coarser resolution global maps such as GlobCover or MOD-500 have been widely used to identify training sites (Potere *et al.*, 2009; Trianni *et al.*, 2015). However, integrating such datasets leads to the introduction of noisy samples which have been shown to decrease the performance of the classifiers. In this context, the increasing availability of VGI brings new opportunities. Defined as the spatial dimension of the web phenomenon of user-generated content (Goodchild, 2007), VGI drives a new way of collecting geographic information that relies on the crowd rather than official or commercial organizations. Founded in 2004, OSM is the most famous of the VGI projects. Initially, the objective was to provide free user-generated street maps (Haklay and Weber, 2008). In the following years, OSM became a collaborative effort to create a free and editable map of the whole world which is not limited to the road network (Mooney and Minghini, 2017). OSM uses a data model based on three object types: nodes (points), ways (polylines or polygons), and relations (logical connections of ways, e.g., a closed way that forms a polygon). Each object is described by at least one key/value pair (a *"tag"*) (Haklay and Weber, 2008; Mooney and Minghini, 2017). This simple data model allows the mapping of a large range of spatial features such as building footprints, points of interest, natural elements, or land use. As a result, and as the database grows, OSM is increasingly used for Land Use and Land Cover (LULC) classification (Estima and Painho, 2013, 2015; Jokar Arsanjani *et*

al., 2013; Fonte *et al.*, 2017). Furthermore, the increasing data availability and quality enable the use of OSM data to support automated supervised classifications of remote sensing imagery. Schultz *et al.* (2017) successively used OSM objects to fill the data gaps in the global Open Land Cover product based on Landsat imagery. Similarly, Yang *et al.* (2017) used Landsat imagery and training points from OSM to map land use in Southeastern United States with an overall accuracy of 75%. This demonstrates that OSM becomes increasingly relevant to support the training of large-scale LULC classifications.

However, the use of OSM data also brings new issues to consider, including: (1) its non-exhaustive nature; and (2) the spatial heterogeneity of the contribution patterns across the regions. Indeed, users are more likely to contribute information where they live. According to Coleman, Georgiadou and Labonte (2009), economical interest and the *"Pride of Place"* are among the main factors that encourage people to contribute. Additionally, Juhász and Hochmair (2018) demonstrated that users are more likely to contribute in specific places such as natural areas and city centers. Furthermore, because of the *"Digital Divide"* (Goodchild, 2007) caused by inequalities in access to education and Internet, developing and developed countries does not benefit from the same amount of contributions. As of March 2018, the amount of information (in bytes) in the OSM database was ten times bigger for the European continent than for Africa.[1] Another example of such heterogeneity is that Germany contained two times more bytes of information than Sub-Saharan Africa. However,OSM data availability in developing regions is rapidly increasing over the last few years, thanks to local contributors and initiatives such as Humanitarian OpenStreetMap Team[2] or Missing Maps.[3] In fact, Africa is the continent where contributions are increasing at the highest rate since 2014. This makes OSM increasingly relevant in developing regions such as Sub-Saharan Africa.

This paper focuses on the use of OSM to collect training samples for the classification of built-up and non-built-up areas in Landsat scenes of ten Sub-Saharan

[1] http://download.geofabrik.de
[2] https://www.hotosm.org
[3] https://www.missingmaps.org

African urban agglomerations. To support global urban mapping, our research aimed to answer the following questions: What information can be extracted from the OSM database to collect built-up and non-built-up training samples? What post-processing must be applied? What performance loss can we expect compared to a strategy based on a manually digitized dataset? Finally, what is the practicability of this approach in the context of a developing region such as Sub-Saharan Africa?

2.2 Materials and methods

2.2.1 Case studies

As stated previously, the spectral profiles of urban areas are characterized by high interurban variations caused by environmental, historical, or socioeconomic differences (Small, 2005). This makes the selection of case studies a crucial step when seeking to ensure the generalization abilities of a method. Our set of case studies is comprised of ten Sub-Saharan African cities described in Table 2.1. Climate is one of the most important sources of variation among the urban areas of the world because it determines the abundance and the nature of the vegetation in the urban mosaic and at its borders. Urban areas located in tropical or subtropical climates (Antananarivo, Johannesburg, Chimoio, and Kampala) can be spectrally confused with vegetated areas because of the presence of dense vegetation in the urban mosaic. This can lead to a mixed-pixel problem and result in misclassifications. On the contrary, cities located in arid or semi-arid climates (Dakar, Gao, Katsina, Saint-Louis, and Windhoek) are characterized by a low amount of vegetation. Bare soil being spectrally similar to built-up, the separation between built-up and non-built-up classes can be more difficult in such areas (Zhang, Chen and Lu, 2015; Li *et al.*, 2017), especially when construction materials are made up from nearby natural resources. Population of an urban area impacts both the distribution of built-up and the data availability. In the context of our study, the population is mainly used as a proxy of the spatial contribution patterns of OSM. Highly populated urban areas (Dakar, Johannesburg, Nairobi, and Kampala) are more likely to benefit from a

high density of information in OSM. On the other hand, smaller cities (Chimoio, Gao, Saint-Louis, and Windhoek) can suffer from a lack of OSM contributions (Coleman, Georgiadou and Labonte, 2009).

Table 2.1: Environmental and demographic characteristics of the case studies. Climate zones are identified according to the Köppen–Geiger classification (Kottek *et al.*, 2006). Population numbers are estimated according to the AfriPop/WorldPop dataset (Linard *et al.*, 2012) for the AOI of each case study.

City	**Country**	**Climate**	**Population**
Antananarivo	Madagascar	Subtropical highland	2,452,000
Chimoio	Mozambique	Humid subtropical	462,000
Dakar	Senegal	Hot semi-arid	3,348,000
Gao	Mali	Hot desert	163,000
Johannesburg	South Africa	Subtropical highland	4,728,000
Kampala	Uganda	Tropical rainforest	3,511,000
Katsina	Nigeria	Hot semi-arid	1,032,000
Nairobi	Kenya	Temperate oceanic	5,080,000
Saint-Louis	Senegal	Hot desert	305,000
Windhoek	Namibia	Hot semi-arid	384,000

2.2.2 Satellite imagery

The Landsat 8 imagery is provided by the U.S. Geological Survey (USGS) through the Earth Explorer portal. The scenes are acquired as Level-1 data products, therefore radiometrically calibrated and orthorectified. The product identifiers and the acquisition dates of each scene are shown in Table 2.2. Calibrated digital numbers are converted to surface reflectance values using the Landsat Surface Reflectance Code (LaSRC) (Vermote *et al.*, 2016) made available by the USGS. Clouds, cloud shadows and water bodies are detected using the Function of Mask (FMASK) algorithm (Zhu and Woodcock, 2012; Zhu, Wang and Woodcock, 2015). The acquisition dates range from August 2015 to October 2016, because of availability issues caused by cloud cover. To reduce the processing cost of the analysis, the satellite

images are masked according to an area of interest (AOI), which is defined as a 20 km rectangular buffer around the city center (as provided by OSM). As a result, all AOI have a surface of 40×40 km^2, regardless of the actual city footprints.

Table 2.2: Product identifiers and acquisition dates of each Landsat scene.

City	Landsat Product Identifier	Acquisition Date
Antananarivo	LC08_L1TP_159073_20150919_20170404_01_T1	2015–09–19
Chimoio	LC08_L1TP_168073_20150529_20170408_01_T1	2015–05–29
Dakar	LC08_L1TP_206050_20151217_20170331_01_T1	2015–12–07
Gao	LC08_L1TP_194049_20160114_20170405_01_T1	2016–01–14
Johannesburg	LC08_L1TP_170078_20150831_20170404_01_T1	2015–08–31
Kampala	LC08_L1TP_171060_20160129_20170330_01_T1	2016–01–29
Katsina	LC08_L1TP_189051_20160111_20170405_01_T1	2016–01–11
Nairobi	LC08_L1TP_168061_20160124_20170330_01_T1	2016–01–24
Saint-Louis	LC08_L1TP_205049_20161009_20170320_01_T1	2016–10–09
Windhoek	LC08_L1TP_178076_20160114_20170405_01_T1	2016–01–14

2.2.3 Reference dataset

Reference samples for four land cover classes are collected using very high spatial resolution (VHSR) imagery from Google Earth (GE): built-up, bare soil, low vegetation (sparse vegetation, farms) and high vegetation (forests). The history slider of the GE interface has been used to ensure that the acquisition dates of the images are in a one-year range of the corresponding Landsat scenes. Even if our classification problem is binary (built-up vs. non-built-up), the collection of reference samples for specific land covers was preferred to ensure the spectral representativeness of the non-built-up landscape. As shown in Figure 2.1, samples were collected as polygons to include the inherent spectral heterogeneity of urban land covers. Reference built-up areas deliberately included mixed pixels provided that they contain at least 20% of built-up.

In total, over all the case studies, more than 2,400 polygons were digitized, cor-

responding to more than 180,000 pixels after rasterization. In the context of our case study, reference samples were used for: (1) assessing the quality of the training samples extracted from OSM; (2) assessing the performance of the built-up classifications; and (3) producing a reference classification for comparison purposes.

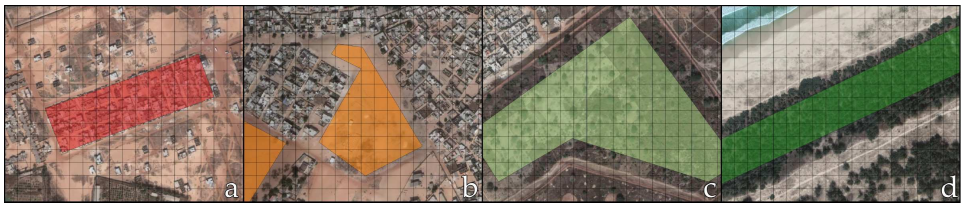

Figure 2.1: Examples of digitized samples in Dakar, Senegal: (**a**) Built-up; (**b**) Bare soil; (**c**) Low vegetation; and (**d**) High vegetation. The grid corresponds to the 30 meters Landsat pixels. Satellite imagery courtesy of Google Earth.

2.2.4 OpenStreetMap

OSM data were acquired in January 2018 using the Overpass API.[4] Four different objects were collected from the database: (1) `highway` polylines (the road network); (2) `building` polygons (the building footprints); (3) the `landuse`, `leisure`, and `natural` polygons (potentially non-built-up objects); and (4) `natural=water` polygons (the water bodies). Complex geometries such as polygons with holes were not considered to simplify the processing.

As previously stated, spatial contribution patterns of OSM are not homogeneous. The evolution of OSM data availability for each type of object for each case study is shown in Figure 2.2. The trends observed at the continental scale are confirmed in the context of our case studies. As suggested by its name, OSM was initially focused on street mapping. Street mapping appears as a continuous effort that leads to a regular increase of the number of roads in the database. Later, contributors started to integrate building footprints, points of interest, or land use and land cover features. As a result, the number of building footprints, natural and land use polygons more than doubled between 2016 and 2018. These trends suggest that

[4] http://overpass-api.de

OSM can now support large-scale supervised classification in developing regions. They also reveal that an increasing amount of data will be available in the near future.

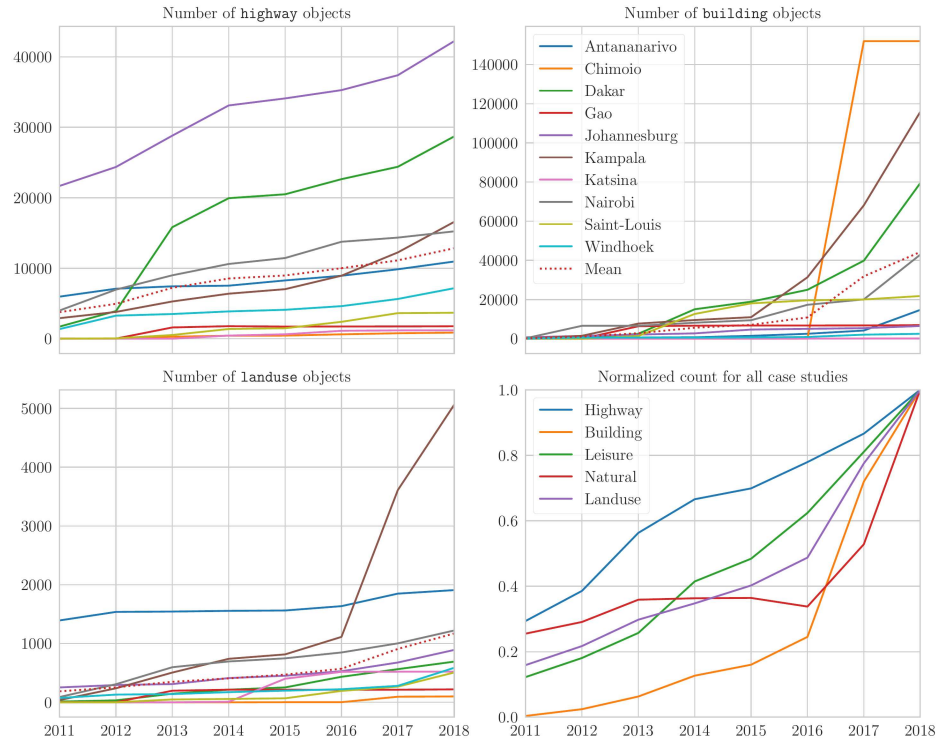

Figure 2.2: Evolution of OSM data availability in our case studies between 2011 and 2018.

2.2.5 Training samples

2.2.5.1 Built-up training samples

In the OSM database, the `building` key is used to mark an area as a building. When they are available, the building footprints are the perfect candidates for built-up training samples collection thanks to their unambiguous spatial definition. However, as shown in Figure 2.3, they are not consistently available among the cities. Highly populated urban areas such as Nairobi, Dakar or Johannesburg contain more than 1,000 ha of building footprints, whereas smaller cities such as Katsina and Gao

only contain a few hectares, thus reducing the representativeness of the full built-up spectral signature. Such data availability issue implies that additional samples must be collected from another data source. Figure 2.3 also reveals that the typical building footprint does not cover more than 15% of the surface of a Landsat pixel. It means that, when going from the vector space to the 30 m raster space of our analysis, the geographic object is not the building footprint anymore but the percentage of the pixel which is effectively covered by any footprint. As a result, the decision to include or exclude a pixel from the built-up training samples relies on a binary threshold, under the assumption that, the higher the threshold, the lower the risk is to include mixed pixels.

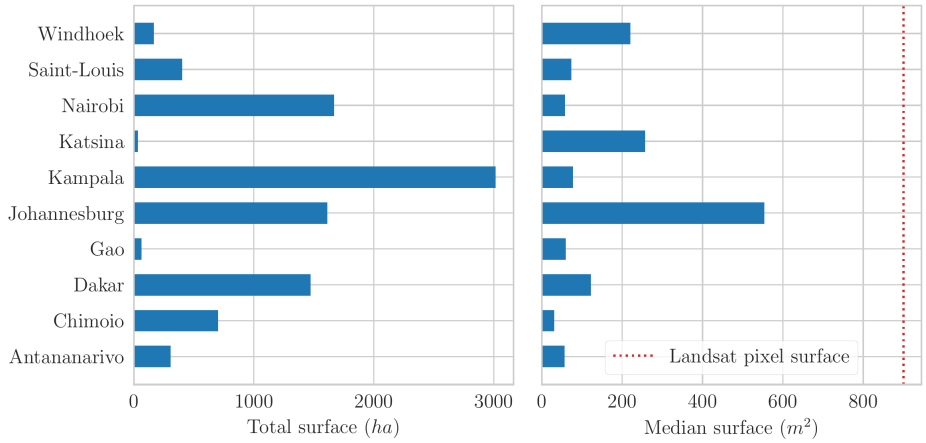

Figure 2.3: Availability and median surface of building footprints in each case study.

As previously stated, OSM building footprints are not a sufficient data source to collect built-up training samples because of inconsistencies in data availability among cities. The road network remains the most exhaustive feature in OSM: even the smallest cities among our case studies contain hundreds of kilometers of roads, and new streets are being mapped each month. As illustrated in Figure 2.4, built-up information can be derived from these road networks using the concept of urban blocks, under the assumption that the smallest areas surrounded by streets are likely to contain buildings. Here, urban blocks are defined as the polygons shaped by the intersection of the roads. To focus on residential blocks, only roads tagged as `resi-`

dential, tertiary, living_street, unclassified or with the generalist road value were used. Major roads such as highways, express ways or national roads were avoided, as well as service roads, tracks, and paths. In the case of Katsina and Gao, for which the building footprints did not provide a sufficient amount of built-up training samples, the process resulted in the availability of more than 1,000 blocks. One assumption can be made regarding the reliability of such geographic objects to collect more built-up training samples: large blocks have a higher probability of containing mixed pixels or non-built-up areas.

Figure 2.4: Urban blocks extracted from the OSM road network in Windhoek, Namibia (transparent: surface greater than 10 ha; red: surface greater than 1 ha; green: surface lower than 1 ha). Satellite imagery courtesy of Google.

2.2.5.2 Non-built-up training samples

Because of the focus of the OSM database on the urban objects, the extraction of non-built-up samples was less straightforward than the extraction of built-up samples. The OSM database includes information on the physical materials at the surface of the earth according to: (1) the description of various bio-physical landscape features such as grasslands or forests with the natural key; (2) the primary usage for an area of land such as farms, or managed forests and grasslands with the landuse key; and (3) the mapping of specific leisure features such as parks or nature reserves with the leisure key. Such objects are not ensured to be non-built-up and must be filtered according to their assigned value. From the 105 available values in our case studies, the following 20 values were selected: sand, farmland, wetland, wood, park, forest, nature_reserve, golf_course, cemetery, sand, quarry, pitch, scree,

`meadow`, `orchard`, `grass`, `recreation_ground`, `grassland`, `garden`, `heath`, `bare_rock`, `beach` and `greenfield`.

However, the availability of non-built objects was not consistent among the case studies. Antananarivo, Johannesburg, Kampala, or Nairobi contained more than 1,000 non-built-up objects according to the previously stated definition. On the contrary, smaller cities such as Chimoio or Katsina benefited from less than 50 objects. Given the spectral and spatial heterogeneity of the non-built landscape which may consist of different types of soil and vegetation, a low amount of non-built-up objects may induce a lack of representativeness in the training dataset. However, a large amount of urban information is available through the road network or the digitized building footprints. In the case of low OSM data availability, this information allows for the discrimination of areas with a low probability of being built-up. The underlying assumption is that the areas which are distant from any urban object, such as roads or buildings, have a low probability of being built-up, thereby making potential candidates for being used as non-built-up training samples. Under the previous assumptions, we define the urban distance as the distance from any road or building:

$$d_{urban} = \min(d_{roads}, d_{buildings}) \tag{2.1}$$

2.2.5.3 Quality assessment of training samples

To assess the quality of the training samples extracted from the OSM database, we measured the distance between their spectral signatures and those of the reference land cover polygons. The spectral signature of an object is the variation of its reflectance values according to the wavelength. In the six non-thermal Landsat bands, the spectral signature S of an object can be defined as:

$$S = (\bar{x}_1, \ \dots \ , \ \bar{x}_n \ , \ \dots \ , \ \bar{x}_6) \tag{2.2}$$

with $\bar{x}_n$ being the mean pixel value of the object for the band n. Therefore, the

euclidean spectral distance d between two objects x and y can be defined as:

$$d(x, y) = \sqrt{\sum_{i=1}^{n=6} (\bar{x}_i - \bar{y}_i)^2} \tag{2.3}$$

More specifically, the optimal value of four parameters was investigated:

1. The minimum coverage threshold for the building footprints, *i.e.* the minimum share of the building footprints inside a given pixel, under the assumption that a lower surface implies more mixed pixels;
2. The maximum surface threshold for the urban blocks, under the assumption that larger blocks are more likely to contain non-built-up and mixed pixels;
3. The OSM tags for the extraction of non-built-up objects;
4. The minimum distance threshold for the random selection of supplementary non-built-up samples from the urban distance raster, under the assumption that the probability to encounter buildings decreases as the urban distance increases.

2.2.6 Classification

Relying on crowd-sourced geographic information to automatically generate a training dataset implies that the resulting sample will be more noisy compared to a manual sampling strategy. Therefore, a larger amount of samples may be required to compensate the mislabeled points and the lack of representativeness. Consequently, the binary classification task (built-up vs. non-built) was performed using the Random Forest (RF) classifier, which has been shown to be computationally efficient and relatively robust to outliers and noisy training data (Rodriguez-Galiano *et al.*, 2012; Mellor *et al.*, 2015). A different classification model was trained for each case study. The implementation was based on a set of Python libraries, including: NumPy (Oliphant, 2015) and SciPy (Jones *et al.*, 2001) for scientific computing, Rasterio (Gillies, 2013) for raster processing, Shapely (Gillies, 2007) and Geopandas for vector analysis, and Scikit-learn (Pedregosa *et al.*, 2011) for machine learn-

ing. The code used to support the study is available on Github.[5]

To remove errors and ambiguities caused by variations in acquisition conditions, eight Landsat bands (blue, green, red, NIR, SWIR1, SWIR2, TIR1, and TIR2) were transformed to a Normalized Difference Spectral Vector (NDSV) (Angiuli and Trianni, 2014) before the classification. The NDSV is a combination of all normalized spectral indices, as defined in Equation 2.4 and Equation 2.5. In the case of Landsat, this leads to a vector of 28 normalized spectral indices.

$$f(b_i, b_j) = \frac{b_i - b_j}{b_i + b_j} \tag{2.4}$$

$$NDSV = \begin{bmatrix} f(b_1, b_2) \\ \dots \\ f(b_i, b_j) \\ \dots \\ f(b_7, b_8) \end{bmatrix} \tag{2.5}$$

To assess the ability of OSM for training supervised built-up classification, a comparative approach was adopted. Three distinct classifications were carried out using different training datasets, as described in Table 2.3. A reference classification (REF) was performed using the reference land cover polygons as training samples to assess the relative performance of OSM-based approaches. In this case, reference polygons were randomly split between a training and a testing dataset of equal sizes. The procedure was repeated twenty times and the validation metrics were averaged. Training samples of the two other classifications (OSM_a and OSM_b) were exclusively extracted from OSM. The first one used first-order features from OSM: building footprints as built-up samples, and land use, natural and leisure polygons as non-built-up samples. The second one was designed to tackle the OSM data availability issue which may be encountered in less populated urban areas. It used second-order features derived from first-order objects such as urban blocks and ur-

[5] https://github.com/yannforget/builtup-classification-osm

ban distance.

Table 2.3: Training samples data sources for each classification scheme.

	Built-Up	**Non-Built-Up**
REF	Reference built-up polygons	Reference non-built polygons
OSM_a	Building footprints	Non-built features
OSM_b	Building footprints & urban blocks	Non-built features & urban distance

In all three cases, RF parameters were set according to the recommendations of the literature (Rodriguez-Galiano *et al.*, 2012; Mellor *et al.*, 2015). RF decision tree ensembles were constructed with 100 trees, and the maximum number of features per split was set to the square root of the total number of features. Imbalance issues between built-up and non-built-up training datasets sizes were tackled by over-sampling the minority class (Lemaître, Nogueira and Aridas, 2017). Additionally, fixed random seeds were set to ensure the reproducibility of the analysis.

2.2.7 Validation

Classification performances were assessed using the manually digitized reference land cover polygons as a testing dataset. Three validation metrics were computed: F1-score, precision and recall. The metrics were computed for the three classifications, as well as for two existing Landsat-based urban maps: the GHSL and the HBASE datasets.

2.3 Results and discussion

2.3.1 Built-up training samples

The extraction of built-up training samples from the OSM building footprints required the selection of a minimum coverage threshold. The impact of this threshold has been assessed by measuring the spectral distance of the resulting samples to the reference built-up samples. As shown in Figure 2.5, the assumption that increas-

ing the threshold would minimize the spectral distance to the reference built-up is not verified. Indeed, the highest spectral distances are reached when only fully covered pixels are selected. As shown in Figure 2.5, the optimal threshold appears to be reached when a minimum of 20,000 samples are available. This reveals the importance of maximizing the representativeness of the sample by ensuring that a sufficient amount of samples is available. Furthermore, given the non-exhaustive nature of the OSM database, a pixel that contains a building footprint of any size have a high probability to contain additional unmapped built-up structures. However, low threshold values (between 0 and 0.2) appear to effectively increase the spectral similarity with the reference built-up samples by eliminating pixels covered by small and isolated buildings. Overall, a minimum coverage threshold of 0.2 appears to maximize both samples quality and quantity.

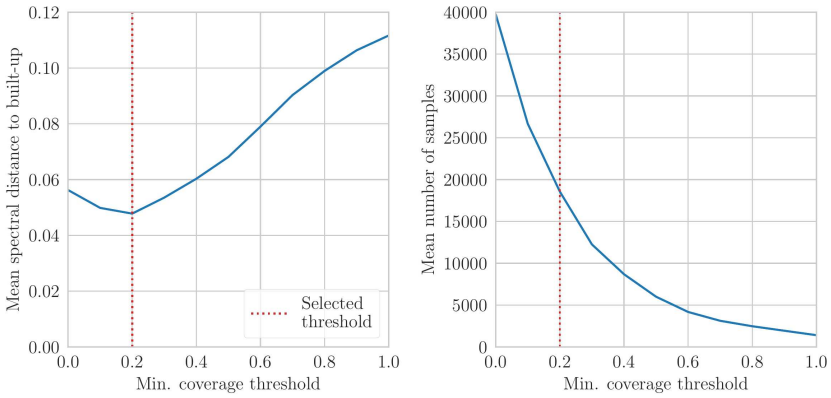

Figure 2.5: Quality and quantity of built-up training samples extracted from OSM building footprints according to the minimum coverage threshold in the 10 case studies: (**a**) mean spectral distance to the reference built-up samples; and (**b**) mean number of samples (in pixels).

Urban blocks enabled the collection of built-up training samples where buildings footprints were lacking. Figure 2.6 shows the impact of the maximum surface threshold on both samples quality and quantity. As expected, excluding large blocks increases the spectral similarity with the reference built-up samples by avoiding highly mixed pixels and bare lands. The highest similarity is reached when only

including the blocks with a surface lower than 1 ha. However, this conservative threshold dramatically reduces the sample size in small urban agglomerations such as Katsina, Gao, or Saint-Louis. Therefore, a maximum surface threshold of 3 ha was selected to ensure a sufficient amount of samples while minimizing the spectral distance to the reference built-up samples.

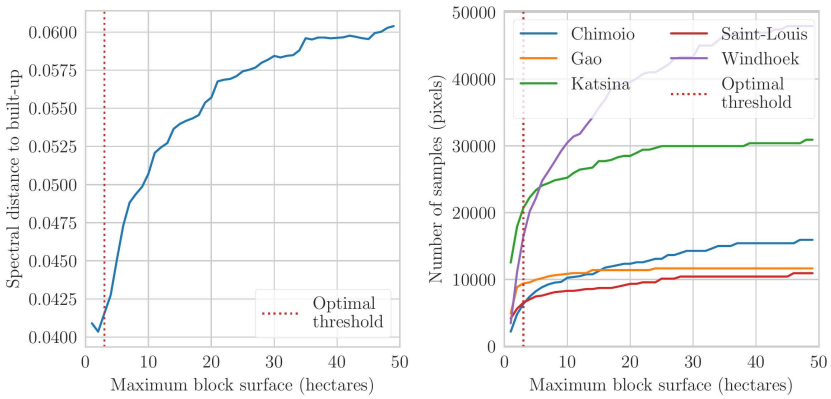

Figure 2.6: Quality and quantity of built-up training samples extracted from OSM urban blocks according to maximum surface threshold in the 10 case studies: (**a**) mean spectral distance to the reference built-up samples; and (**b**) number of samples (in pixels) in the five case studies with the lowest data availability.

2.3.2 Non-built-up training samples

The non-built-up landscape is spectrally complex due to its irregular spatial patterns and the variations of soils and vegetation types. Figure 2.7 shows the most similar land cover of each OSM non-built-up tag in terms of spectral distance. The analysis reveals the spectral variability of OSM non-built-up objects across the case studies. Urban features such as `garden`, `recreation_ground`, `pitch` or `park` can have a spectral signature closer to built-up than to bare soil or lowly vegetated areas. The small surface covered by these features can lead to a high proportion of mixed pixels. Additionally, their urban nature makes highly probable the presence of human-constructed elements. On the contrary, natural features such as `orchard`, `meadow`, `forest` or `wood` are more consistently close to the spectral signature of vegetated

areas. Generally, most of the features providing bare soil samples may be confused with built-up areas because of their urban nature (`pitch`) or their spectral similarity (`beach`). However, the decision boundary between built-up and bare soil pixels being the most prone to errors in urban areas, we choose to not exclude them in order to maximize the representativeness. Overall, these inconsistencies also highlight the fact that a supervised multi-class land cover classification based on OSM would be difficult to set up as of today.

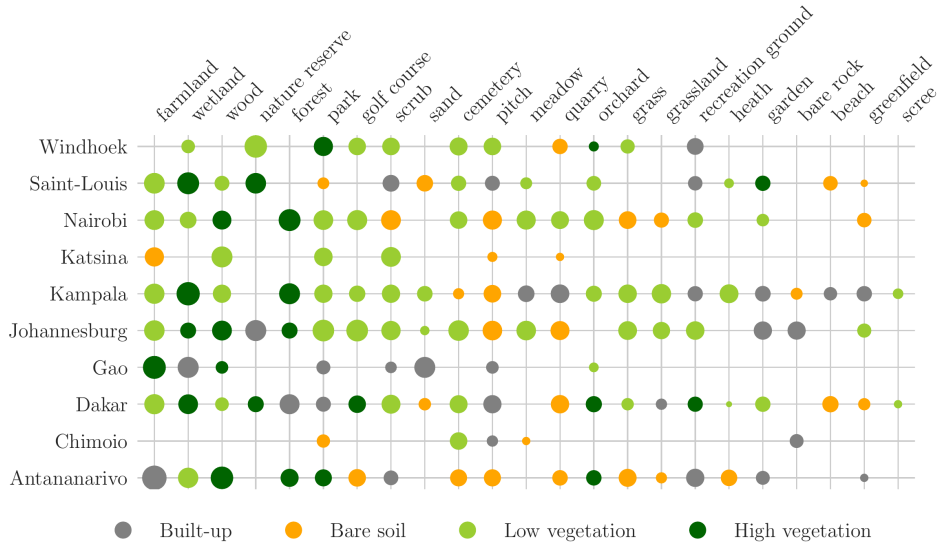

Figure 2.7: Most similar land cover of each OSM non-built-up object according to its tag. Circles are logarithmically proportional to the number of pixels available.

In case studies where OSM non-built-up objects were not sufficiently available, an urban distance raster was used to randomly collect supplementary training samples in remote areas. Figure 2.8 shows the relationship between the remoteness and the spectral distance to the reference built-up samples. As expected, the spectral distance increases with the urban distance. However, the spectral variations become inconsistent and are mainly caused by changes in the non-built-up landscape (e.g., forests, mountains, or bare lands). In highly urbanized agglomerations such as Johannesburg or Dakar, the road network covers the whole area of interest, leading to a very low amount of remote pixels. Consequently, a minimum distance

threshold of 250 m was used.

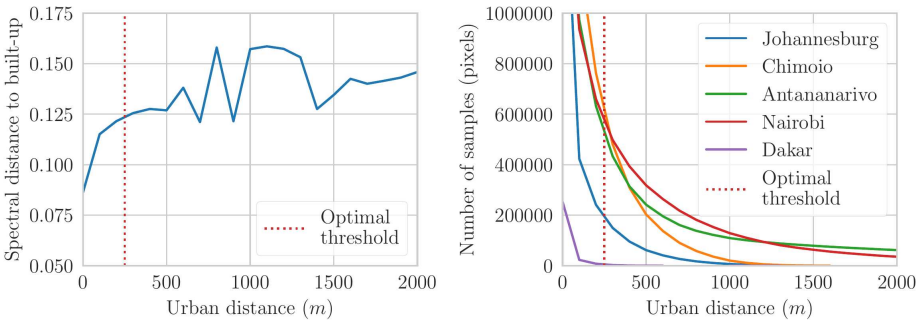

Figure 2.8: Quality and quantity of non-built-up training samples extracted from the OSM-based urban distance raster: (**a**) mean spectral distance to the reference built-up samples according to the urban distance; and (**b**) number of samples (in pixels) in the five case studies with the lowest sample availability.

2.3.3 GHSL and HBASE assessment

The assessment metrics for the GHSL and HBASE datasets in the context of our case studies are shown in Table 2.4. They are provided as an indication of their relevance in the context of our case studies and our definition of a built-up area, since they use validation samples manually digitized according to the built-up definitions of the present study. They also reveal which case studies may be problematic for an automated built-up mapping method. For example, the arid urban area of Gao suffers from low recall scores because of the spectral confusion that occurs between the buildings materials and the bare surroundings areas. This leads to the misclassification of large built-up areas as bare lands. To a lesser extent, the semi-arid urban areas of Saint-Louis, Windhoek, and Katsina present the same issue. On the contrary, subtropical urban areas such as Antananarivo or Chimoio are characterized by an abundant vegetation in the urban mosaic. Thus, high rates of misclassifications are observed in the peripheral areas where built-up is less dense. A similar phenomenon is also noticed in the richest residential districts of Johannesburg. Overall, both datasets reach a mean F1-score of 0.82 when excluding Gao.

Table 2.4: Assessment metrics of the GHSL and HBASE datasets.

Case study	.	GHSL	.	.	HBASE	.
	F1-Score	**Precision**	**Recall**	**F1-Score**	**Precision**	**Recall**
Antananarivo	0.83	0.82	0.83	0.79	0.67	0.96
Chimoio	0.47	0.98	0.31	0.82	0.94	0.73
Dakar	0.85	0.74	0.99	0.81	0.69	0.98
Gao	0.35	0.98	0.21	0.72	0.94	0.59
Johannesburg	0.92	0.86	0.99	0.90	0.82	0.99
Kampala	0.96	0.95	0.96	0.95	0.93	0.97
Katsina	0.90	0.92	0.88	0.64	0.76	0.56
Nairobi	0.84	0.96	0.75	0.88	0.81	0.97
Saint Louis	0.76	0.95	0.63	0.81	0.97	0.70
Windhoek	0.81	0.92	0.73	0.78	0.65	0.99
Mean	0.77	0.91	0.73	0.81	0.82	0.85
Standard dev.	0.20	0.08	0.28	0.09	0.12	0.18

2.3.4 Classification results

Assessment metrics of the three classification schemes are presented in Table 2.5. The reference classification, which has been trained with manually digitized samples, reached a mean F1-score of 0.92 and a minimum of 0.84 in Gao. Such results suggest that high classification performances can be achieved in most of the case studies provided that the training dataset is sufficiently large and representative. The first OSM-based classification scheme (OSM_a) made use of first-order OSM objects: buildings footprints and objects associated with a non-built up tag. Therefore, a limited availability in either of the aforementioned objects was highly detrimental to the classification performance. Katsina, Windhoek, and Johannesburg suffered from a low availability in building footprints with, respectively, 110, 2,636, and 6,724 objects. This led to an unrepresentative built-up training sample consisting mainly of large administrative structures or isolated settlements. In Chimoio, more than 150,000 building footprints were available. However, only 12 non-built-up

polygons have been extracted from the OSM database, all related to forested areas. As a result, the lack of information regarding the spectral characteristics of the heterogeneous non-built-up landscape did not enable the separation between built-up and bare areas. A similar issue was also encountered in Antananarivo, where most of the non-built-up training samples were located in natural reserves and forests.

The second OSM-based classification scheme (OSM_b) was designed to tackle the aforementioned issues by deriving second-order features from the road network. The addition of built-up and non-built-up training samples collected from urban blocks and remote areas solved the data availability and representativeness issues in all the case studies, leading to better scores in nine out of ten cases. Overall, OSM_b reached scores that were comparable to those of the reference classification. More specifically, OSM_b had the highest recall scores, suggesting that the model was more successful in the detection of isolated, informal or peripheral settlements. Additionally, the use of larger training datasets (from 30,000 to 500,000 samples per case study) led to higher consistencies in the classification performance with a standard deviation of 0.02.

Table 2.5: Assessment metrics (F1-score, precision, and recall) for the three classification schemes.

	OSM_a			OSM_b			REF		
	F1	**Pre.**	**Rec.**	**F1**	**Pre.**	**Rec.**	**F1**	**Pre.**	**Rec.**
Antananarivo	0.78	0.99	0.65	0.93	0.91	0.96	0.92	0.97	0.87
Chimoio	0.77	0.63	0.97	0.92	0.90	0.95	0.85	0.93	0.79
Dakar	0.95	0.98	0.92	0.96	0.94	0.98	0.94	0.98	0.90
Gao	0.81	0.96	0.69	0.90	0.94	0.86	0.84	0.84	0.86
Johannesburg	0.60	0.98	0.43	0.92	0.99	0.86	0.96	0.98	0.94
Kampala	0.98	1.00	0.97	0.98	0.99	0.96	0.98	0.99	0.96
Katsina	0.20	0.84	0.11	0.91	0.95	0.87	0.94	0.99	0.90
Nairobi	0.91	0.94	0.89	0.94	0.97	0.92	0.93	0.97	0.89
Saint-Louis	0.95	0.98	0.93	0.94	0.92	0.96	0.92	0.98	0.88
Windhoek	0.68	0.98	0.52	0.95	0.93	0.98	0.93	0.96	0.90

	OSM_a			OSM_b			REF		
Mean	0.76	0.93	0.71	0.94	0.95	0.93	0.92	0.96	0.89
Standard dev.	0.23	0.11	0.29	0.02	0.03	0.05	0.04	0.05	0.05

With OSM_b, three case studies still had recall scores lower than 0.9: Gao, Johannesburg and Katsina. This suggests that the model did not effectively detect built-up in some areas. Figure 2.9 shows some examples of such areas. In Katsina, higher rates of misclassifications were observed in the northeast part of the city, where urban vegetation was denser that in other parts of the agglomeration. Furthermore, because of a less dense road network and the unavailability of building footprints, no training samples were available in this area. Likewise, the richest neighborhoods in Johannesburg are characterized by isolated buildings in a denser urban vegetation, leading to a higher rate of misclassification. In Gao, errors were mainly caused by the spectral confusion which occurred between built-up and bare soil areas. The phenomenon was exacerbated by the arid climate and the buildings materials made off nearby natural resources.

Figure 2.9: Areas with high rates of misclassifications in: (**a**) Katsina; (**b**) Johannesburg; (**c**) Gao; and (**d**) Dakar. Satellite imagery courtesy of Google Earth

Generally, as shown in Figure 2.10, the classification scores increased with the number of training samples. Because of the introduction of noise and mislabeled samples inherent to automated approaches, large training datasets were required to make sense of the heterogeneous spectral characteristics of the urban environment. Figure 2.10 suggests that between 10,000 and 20,000 samples are necessary to fit the classification model depending on the spectral complexity of the urban mosaic.

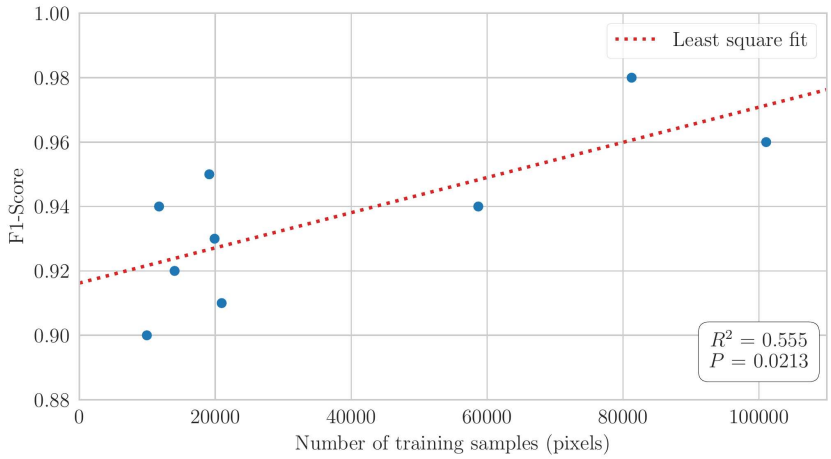

Figure 2.10: Relationship between the number of training samples and the classification F1-score (the outlier Johannesburg is excluded from the graph).

2.4 Conclusion

This study provided important insights regarding the automatic collection of training samples to support large-scale or rapid supervised classification of built-up areas. The proposed method made use of the growing amount of information in the OSM database to automatically extract both built-up and non-built-up training samples. This automated approach can reach classification performances similar to manual sampling strategies, provided that a relevant set of pre-processing routines are applied. In some less populated urban areas, first-order urban features—such as building footprints—can be too scarce. The issue of data scarcity can be tackled by a spatial analysis of the road network to derive second-order features such as urban blocks or urban distance. The proposed approach reached a mean F1-score of 0.93 across our case studies, while the manual approach reached 0.92. Case studies located in arid climates suffer from higher misclassification rates because of the spectral confusion that occurs between the building materials and the bare soil. The issue could be addressed by using a higher resolution imagery such as Sentinel-2. Likewise, the use of Synthetic Aperture Radar (SAR) to extract textural features should lead to a better separation between built-up and bare soil.

Automatically generated training datasets contain more noise than manually collected samples. Additionally, the reliance on crowd-sourced geographic information introduces its own share of errors and inconsistencies. Previous studies have shown that the RF classifier can handle up to 20% noise in the training dataset, provided that the sample size is large enough (Mellor *et al.*, 2015). Overall, these previous findings are confirmed by our results. In our case studies, maximizing the size of the training dataset was more advantageous than minimizing the noise.

Land use and land cover mapping in the OSM community is a relatively new phenomenon, especially in developing regions. In fact, 77% of the building footprints and 50% of the land use polygons used in this study have been mapped after January 2016. The growing amount of available data suggests that the proposed approach will provide better results in following years. More importantly, the mapping of land use and natural elements could enable multi-class supervised land cover classifications in the near future.

References

Angiuli, E. and Trianni, G. (2014) 'Urban Mapping in Landsat Images Based on Normalized Difference Spectral Vector', *IEEE Geoscience and Remote Sensing Letters*, 11(3), pp. 661–665. doi: 10.1109/LGRS.2013.2274327.

Arino, O. *et al.* (2007) 'Globcover-a global land cover service with MERIS', in *Proceedings of the ENVISAT Symposium*, pp. 23–27.

Belgiu, M. and Drăguț, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, pp. 24–31. doi: 10.1016/j.isprsjprs.2016.01.011.

Chen, J. *et al.* (2015) 'Global land cover mapping at 30m resolution: A POK-based operational approach', *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp. 7–27. doi: 10.1016/j.isprsjprs.2014.09.002.

Coleman, D. J., Georgiadou, Y. and Labonte, J. (2009) 'Volunteered Geographic Information: The nature and motivation of produsers', *International Journal of Spatial Data Infrastructures Research*, 4, p. 27. doi: 10.2902/1725-0463.2009.04.art16.

Dye, C. (2008) 'Health and Urban Living', *Science*, 319(5864), pp. 766–769. doi: 10.1126/science.1150198.

Estima, J. and Painho, M. (2013) 'Exploratory analysis of OpenStreetMap for land use classification', in *Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowd-*

sourced and Volunteered Geographic Information. New York: ACM Press, pp. 39–46. doi: 10.1145/2534732.2534734.

Estima, J. and Painho, M. (2015) 'Investigating the Potential of OpenStreetMap for Land Use/Land Cover Production: A Case Study for Continental Portugal', in Jokar Arsanjani, J. et al. (eds) *OpenStreetMap in GIScience*. Cham: Springer International Publishing, pp. 273–293.

Fonte, C. *et al.* (2017) 'Generating Up-to-Date and Detailed Land Use and Land Cover Maps Using OpenStreetMap and GlobeLand30', *ISPRS International Journal of Geo-Information*, 6(4), p. 125. doi: 10.3390/ijgi6040125.

Gamba, P. and Herold, M. (2009) *Global mapping of human settlement: Experiences, datasets, and prospects*. Boca Raton: CRC Press (Taylor & Francis series in remote sensing applications).

Gillies, S. (2007) 'Shapely: Manipulation and analysis of geometric objects'. toblerity.org.

Gillies, S. (2013) 'Rasterio: Geospatial raster I/O for Python programmers'. Mapbox.

Goodchild, M. F. (2007) 'Citizens as sensors: The world of volunteered geography', *GeoJournal*, 69(4), pp. 211–221. doi: 10.1007/s10708-007-9111-y.

Grimm, N. B. *et al.* (2008) 'Global Change and the Ecology of Cities', *Science*, 319(5864), pp. 756–760. doi: 10.1126/science.1150195.

Haklay, M. and Weber, P. (2008) 'OpenStreetMap: User-Generated Street Maps', *IEEE Pervasive Computing*, 7(4), pp. 12–18. doi: 10.1109/MPRV.2008.80.

Herold, M. *et al.* (2004) 'Spectrometry for urban area remote sensingDevelopment and analysis of a spectral library from 350 to 2400 nm', *Remote Sensing of Environment*, 91(3-4), pp. 304–319. doi: 10.1016/j.rse.2004.02.013.

Jokar Arsanjani, J. *et al.* (2013) 'Toward mapping land-use patterns from volunteered geographic information', *International Journal of Geographical Information Science*, 27(12), pp. 2264–2278. doi: 10.1080/13658816.2013.800871.

Jones, E. *et al.* (2001) 'SciPy: Open Source Scientific Tools for Python'.

Juhász, L. and Hochmair, H. (2018) 'OSM Data Import as an Outreach Tool to Trigger Community Growth? A Case Study in Miami', *ISPRS International Journal of Geo-Information*, 7(3), p. 113. doi: 10.3390/ijgi7030113.

Kottek, M. *et al.* (2006) 'World Map of the Köppen-Geiger climate classification updated', *Meteorologische Zeitschrift*, 15(3), pp. 259–263. doi: 10.1127/0941-2948/2006/0130.

Lemaître, G., Nogueira, F. and Aridas, C. K. (2017) 'Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning', *Journal of Machine Learning Research*, 18(17), pp. 1–5.

Li, H. *et al.* (2017) 'Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index', *Remote Sensing*, 9(12), p. 249. doi: 10.3390/rs9030249.

Li, M. (2014) 'A Review of Remote Sensing Image Classification Techniques: The Role of Spatio-contextual Information', *European Journal of Remote Sensing*, pp. 389–411. doi: 10.5721/EuJRS20144723.

Linard, C. *et al.* (2012) 'Population Distribution, Settlement Patterns and Accessibility across Africa in 2010', *PLoS ONE*, 7(2). doi: 10.1371/journal.pone.0031743.

Mellor, A. *et al.* (2015) 'Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin', *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, pp. 155–168. doi: 10.1016/j.isprsjprs.2015.03.014.

Mooney, P. and Minghini, M. (2017) 'A Review of OpenStreetMap Data', in *Mapping and the Citizen Sensor*. London: Ubiquity Press, pp. 37–59. doi: 10.5334/bbf.c.

Oliphant, T. E. (2015) *Guide to NumPy*. Austin, Tex.: Continuum Press.

Pedregosa, F. *et al.* (2011) 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research*, 12(Oct), pp. 2825–2830.

Pesaresi, M. *et al.* (2016) *Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014*. Joint Research Centre, European Commission.

Potere, D. *et al.* (2009) 'Mapping urban areas on a global scale: Which of the eight maps now available is more accurate?', *International Journal of Remote Sensing*, 30(24), pp. 6531–6558. doi: 10.1080/01431160903121134.

Rodriguez-Galiano, V. F. *et al.* (2012) 'An assessment of the effectiveness of a random forest classifier for land-cover classification', *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, pp. 93–104. doi: 10.1016/j.isprsjprs.2011.11.002.

Schneider, A., Friedl, M. A. and Potere, D. (2009) 'A new map of global urban extent from MODIS satellite data', *Environmental Research Letters*, 4(4), p. 044003. doi: 10.1088/1748-9326/4/4/044003.

Schultz, M. *et al.* (2017) 'Open land cover from OpenStreetMap and remote sensing', *International Journal of Applied Earth Observation and Geoinformation*, 63, pp. 206–213. doi: 10.1016/j.jag.2017.07.014.

Small, C. (2005) 'A global analysis of urban reflectance', *International Journal of Remote Sensing*, 26(4), pp. 661–681. doi: 10.1080/01431160310001654950.

Trianni, G. *et al.* (2015) 'Scaling up to National/Regional Urban Extent Mapping Using Landsat Data', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7), pp. 3710–3719. doi: 10.1109/JSTARS.2015.2398032.

UN-Habitat (2014) *The state of African cities, 2014: Re-imagining sustainable urban transitions*. Nairobi, Kenya.

Vermote, E. *et al.* (2016) 'Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product', *Remote Sensing of Environment*, 185, pp. 46–56. doi: 10.1016/j.rse.2016.04.008.

Wang, P. *et al.* (2017) 'Human Built-up And Settlement Extent (HBASE) Dataset From Landsat'. NASA Socioeconomic Data; Applications Center (SEDAC).

Wentz, E. *et al.* (2014) 'Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing', *Remote Sensing*, 6(5), pp. 3879–3905. doi:

10.3390/rs6053879.

Wulder, M. A. *et al.* (2008) 'Landsat continuity: Issues and opportunities for land cover monitoring', *Remote Sensing of Environment*, 112(3), pp. 955–969. doi: 10.1016/j.rse.2007.07.004.

Yang, D. *et al.* (2017) 'Open land-use map: A regional land-use mapping strategy for incorporating OpenStreetMap with earth observations', *Geo-spatial Information Science*, 20(3), pp. 269–281. doi: 10.1080/10095020.2017.1371385.

Zhang, C., Chen, Y. and Lu, D. (2015) 'Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery', *International Journal of Remote Sensing*, 36(17), pp. 4483–4500. doi: 10.1080/01431161.2015.1084552.

Zhu, Z., Wang, S. and Woodcock, C. E. (2015) 'Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 47, 8, and Sentinel 2 images', *Remote Sensing of Environment*, 159, pp. 269–277. doi: 10.1016/j.rse.2014.12.014.

Zhu, Z. and Woodcock, C. E. (2012) 'Object-based cloud and cloud shadow detection in Landsat imagery', *Remote Sensing of Environment*, 118, pp. 83–94. doi: 10.1016/j.rse.2011.10.028.

Chapter 3

Complementarity between optical and SAR sensors

This chapter is based on the following publication:

Forget Y., Shimoni M., Gilbert M. & Linard C. (2018). "Complementarity Between Sentinel-1 and Landsat 8 Imagery to Map Built-Up Areas in Sub-Saharan Africa". *Preprints*. 10.20944/preprints201810.0695.v1

Abstract. The rapid urbanization that takes place in developing regions such as Sub-Saharan Africa is associated with a large range of environmental and social issues. In this context, remote sensing is essential to provide accurate and up-to-date spatial information to support risk assessment and decision making. However, mapping urban areas remains a challenge because of their heterogeneity, especially in developing regions where the highest rates of misclassification are observed. Nevertheless, urban areas located in arid climates—which are among the most vulnerable to anthropogenic impacts, suffer from the spectral confusion occurring between built-up and bare soil areas when using optical imagery. Today, the increasing availability of satellite imagery from multiple sensors allow to tackle the aforementioned issues by combining optical data with Synthetic Aperture Radar (SAR). In this paper, we assess the complementarity of the Landsat 8 and Sentinel-1 sensors to map built-

up areas in twelve Sub-Saharan African urban areas, using a pixel-level supervised classification based on the Random Forest classifier. We make use of textural information extracted from SAR backscattering data in order to reduce the speckle noise and to introduce contextual information at the pixel level. Results suggest that combining both optical and SAR features consistently improves classification performances, mainly by enhancing the differentiation between built-up and bare lands. However, the fusion was less beneficial in mountainous case studies, suggesting that including features derived from a Digital Elevation Model (DEM) could improve the reliability of the proposed approach. As suggested by previous studies, combining features computed from both VV and VH polarizations consistently led to better classification performances. On the contrary, introducing textures computed from different spatial scales did not improve the classification performances.

3.1 Introduction

Urbanization is a worldwide process associated with a wide range of environmental and human health issues (Grimm *et al.*, 2008; Dye, 2008). In Africa, the urban population is predicted to triple between 2010 and 2050, threatening both social and environmental sustainability (UN-Habitat, 2015). Monitoring built-up areas in developing regions such as Sub-Saharan Africa is therefore crucial to understand, predict, and mitigate the risks associated with such a rapid urbanization (Linard, Tatem and Gilbert, 2013). In this context, remote sensing plays a major role by providing accurate spatial information on built-up areas at a relatively low cost (Gamba and Herold, 2009; Wentz *et al.*, 2014). However, because of the heterogeneity of urban areas in terms of spatial structure and materials, mapping built-up with medium resolution optical imagery remains challenging. In medium spatial resolution imagery (10–50 m), urban pixels are made of a combination of several elements—such as buildings, roads, trees or bare soil. Those mixed pixels can make spectral-based classifications difficult. Furthermore, the spectral characteristics and the spatial distribution of these objects differ across a given urban area. High differences are also observed among the cities of the world because of socioeconomic, cultural, historical and environmental variations (Forster, 1993; Small, 2001, 2005). Developing

regions are the most concerned with the risks associated with urbanization: urban poverty, air and water pollution, lack of sanitation structures, vector-borne diseases, floods and fires (Moore, Gould and Keary, 2003; McGranahan *et al.*, 2009; Bai *et al.*, 2012; Brockmann and Helbing, 2013). Still, previous studies have shown that they also suffer from lower accuracies in global built-up maps (Potere *et al.*, 2009), because of a high urban heterogeneity coupled with a lack of reference datasets to support the training and the validation of the classification models.

Likewise, despite the fact that about one third of the global land surface is characterized by an arid or semi-arid climate according to the Köppen-Geiger classification (Kottek *et al.*, 2006; Rubel *et al.*, 2017), they also suffer from low accuracies when it comes to built-up mapping. Due to their overlapping spectral signatures, the differentiation between bare land and built-up in arid and semi-arid environments has proven to be one of the main challenges associated with optical sensors in urban remote sensing. Previous studies have shown that conventional spectral indices—such as the normalized difference built-up index (NDBI), the normalized difference bareness index (NDBal), or the urban index (UI), are not reliable to differentiate built-up areas from bare land in arid regions (Qian, Zhou and Hou, 2007; Rasul *et al.*, 2018). As a result, new approaches based on object-oriented classification or linear spectral mixture analysis have been proposed (Qian, Zhou and Hou, 2007; Zhang, Chen and Lu, 2015). New spectral indices have also been specifically developed to tackle the issue, such as the normalized bare land index (NBLI) (Li *et al.*, 2017). Likewise, the dry built-up index (DBI) and the dry bare-soil index (DBSI) provide a better separation between bare soil and built-up in arid regions by making use of the blue and thermal bands of Landsat 8 (Rasul *et al.*, 2018). Approaches based on the thresholding of spectral indices do not require any training dataset and benefit from a low computational cost. However, as stated by their authors, their reliability highly depends on the landscape and the climate of the study area. For instance, the DBSI is not considered suitable in humid regions or in urban areas surrounded by vegetation (Rasul *et al.*, 2018).

Because of the aforementioned issues, the idea of combining optical data with complementary sensors such as Synthetic Aperture Radar (SAR) recently gained mo-

mentum. SAR has the advantages of providing high resolution imagery independently from daylight, clouds, or weather conditions. The C-band of the European Remote-Sensing Satellite 1 and 2 (ERS-1/2) has been widely used to monitor urban areas (Weng, 2014). In contrast to optical sensors, SAR is sensitive to the roughness of the terrain—and thus is able to better differentiate between bare soil and built-up (Soergel, 2010). Previous studies have shown that the combined use of optical and SAR data can significantly improve the accuracy of a land cover classification (Waske and van der Linden, 2008; Zhang, Zhang and Lin, 2012; Joshi *et al.*, 2016). However, classifying data provided by different sensors is not straightforward and there is no consensus among the remote sensing community regarding the best fusion approach. Conventional parametric classifiers which concatenate signals from different sensors into one vector have been shown to be inefficient in modeling multi-sensor data distributions, therefore most of the methods rely on machine learning classifiers that do not make any assumption regarding the data distribution (Tupin, 2010). Fusion can occur at four different levels: (1) at the pixel level—by concatenating data from multiple sensors into one stacked vector (Griffiths *et al.*, 2010; Zhu *et al.*, 2012; Zhang, Zhang and Lin, 2014; Braun and Hochschild, 2015), (2) at the feature level in the context of an object-based classification that makes use of image segmentation techniques (Clerici, Valbuena Calderón and Posada, 2017), or (3) at the decision level, for instance by merging several single-source classifiers using neural networks or support vector machines (Benediktsson and Kanellopoulos, 1999; Fauvel, Chanussot and Benediktsson, 2006; Waske and van der Linden, 2008; Shao *et al.*, 2016).

Previous studies have reported that pixel-level fusion approaches are inappropriate at high spatial resolutions because of the lack of information about the spatial context of a given pixel and the speckle noise inherent to SAR data (Tupin, 2010; Gamba, 2014; Zhang, Zhang and Lin, 2014). The extraction of textural features from SAR backscattering partially solves the aforementioned issues (Zhang, Zhang and Lin, 2014; Braun and Hochschild, 2015), for instance by computing the grey level co-occurrence matrix (GLCM) texture features (Haralick, Shanmugam and Dinstein, 1973; Gotlieb and Kreyszig, 1990). However, there is no consensus on

which features are the most relevant in the context of a combined use with optical data, or on the optimal size of the moving window used to compute the GLCM.

In this paper, we investigate the combined use of Landsat 8 and Sentinel-1 imagery to detect built-up in twelve Sub-Saharan African case studies characterized by various climates, landscapes and population patterns ; including both medium-sized and large urban areas. We assess the complementarity of optical and SAR data in the context of a pixel-level supervised classification based on the extraction of 18 GLCM texture features, with several window sizes and from the two polarizations available with Sentinel-1—VV and VH.

3.2 Materials and methods

3.2.1 Case studies

Compared to natural land covers, built-up areas are highly heterogeneous at both the interurban and the intraurban scales. As a result, a method developed in the context of an European urban area has no guarantee to be reliable in an urban agglomeration of Sub-Saharan Africa. This is why a diverse set of case studies is crucial when seeking to maximize the generalization potential of a method. In the context of built-up mapping using both optical and SAR data, the reliability of each sensor is expected to be highly dependent on landscape and climate variables. The selected case studies for the present analysis are presented in Table 3.1. The area of interest for each case study corresponds to the rectangular 20 km buffer around the city center.

The set contains urban areas with various climates, landscapes and population characteristics. In the context of built-up mapping from both optical and SAR data, optical sensors are expected to perform well in tropical, subtropical and temperate climates (Antananarivo, Bukavu, Chimoio, Kampala or Johannesburg) because of their ability to differentiate between the spectral signatures of built-up and vegetation. However, densely vegetated urban mosaics could cause some confusion. On the contrary, SAR sensor is expected to perform better in dry climates (Dakar, Gao, Katsina, Saint-Louis, Ouagadougou, Windhoek), and lower in mountainous urban

areas surrounded with dense vegetation and steep slopes (Antananarivo, Bukavu). More generally, the various climates and population sizes ensure that multiple urban morphologies will be encountered.

Table 3.1: Climate, topography and population for each case study. Values are aggregated for the area of interest. Climate data is derived from the Koppen-Geiger classification (Kottek *et al.*, 2006; Rubel *et al.*, 2017). Mean slope and elevation are computed from the Shuttle Radar Topographic Mission (SRTM) 30 m (JPL, 2013). Population is estimated using the AfriPop/WorldPop dataset (Linard *et al.*, 2012; Worldpop, 2016).

City (Country)	Climate	Slope	Elevation	Population
Antananarivo (MDG)	Subtropical highland	8°	1,319 m	2,436,196
Bukavu (COD)	Subtropical highland	13°	1,836 m	1,041,703
Chimoio (MOZ)	Humid subtropical	4°	612 m	455,612
Dakar (SEN)	Hot semi-arid	2°	14 m	3,332,985
Gao (MLI)	Hot desert	3°	273 m	161,172
Johannesburg (ZAF)	Subtropical highland	4°	1,608 m	4,668,844
Kampala (UGA)	Tropical rainforest	5°	1,177 m	3,498,376
Katsina (NGA)	Hot semi-arid	2°	495 m	1,027,729
Nairobi (KEN)	Temperate oceanic	4°	1,692 m	5,064,548
Ouagadougou (BFA)	Hot semi-arid	2°	308 m	2,256,479
Saint-Louis (SEN)	Hot desert	2°	7 m	300,518
Windhoek (NAM)	Hot desert	9°	1,811 m	383,503

3.2.2 Data acquisition and preprocessing

Sentinel-1 and Landsat 8 product types and acquisition dates are presented in Table 3.2. Landsat 8 imagery was acquired through the Earth Explorer portal of the U.S. Geological Survey (USGS) with the `landsatxplore` software (Forget, 2018b), using cloud cover as the main criterion. Additionally, monthly NDVI values from the MODIS-based MOD13C2 dataset (Didan, 2015) were used to favor the most vegetated periods, which are different depending on the case study. Scenes were ac-

quired as Level-1 data products, thus radiometrically calibrated and orthorectified. Calibrated digital numbers were converted to surface reflectance values using the Landsat Surface Reflectance Code (LaSRC) (Vermote *et al.*, 2016) made available by the USGS. Cloudy pixels were masked using the Function of Mask (FMASK) algorithm (Zhu and Woodcock, 2012; Zhu, Wang and Woodcock, 2015).

Sentinel-1A images were acquired through the Copernicus Open Access Hub using the `sentinelsat` software (Clauss *et al.*, 2018). Scenes belonging to the dry season were favored based on the monthly NDVI values provided by the MODIS-based MOD13C2 dataset. Additionally, the CPC Global Unified Precipitation dataset (Chen *et al.*, 2008), provided by the NOAA Climate Prediction Center, was used to require at least two days without any precipitation before the acquisition date. The last criterion for the final scene selection was the temporal proximity to the selected Landsat 8 scene. The scenes were acquired as Ground Range Detected (GRD) Level-1 products in the Interferometric Wide (IW) swath mode, therefore multi-looked and projected to ground range using an Earth ellipsoid model. Pre-processing was performed using the Sentinel Application Platform (SNAP) (ESA, 2018). Firstly, orbit state vectors were refined with precise orbit files and GRD data was calibrated to β nought.

SAR images are also characterized by the presence of a granular noise, known as *speckle noise*, which can reduce the effectiveness of image classification. Simple local mean or median filters can be applied to reduce the noise. However, filters specifically developed for speckle noise reduction (Frost, Lee, or Sigma) are better at preserving details, edges and linear features (Lee *et al.*, 1994; Kupidura, 2016). Therefore, a 3x3 Refined Lee filter (Lee *et al.*, 1994) was applied to the SAR images in order to maximize the preservation of the textural information. Finally, terrain flattening (Small, 2011) and Range-Doppler terrain correction (Small and Shubert, 2008) were applied based on the SRTM 1-sec Digital Elevation Model (DEM).

Table 3.2: Sentinel-1 and Landsat 8 product types and acquisition dates.

	Sentinel-1	**Landsat 8**
Antananarivo	S1A_IW_GRDH 2015-10-07	LC08_L1TP 2015-06-15
Bukavu	S1A_IW_GRDH 2016-06-10	LC08_L1TP 2015-09-21
Chimoio	S1A_IW_GRDH 2015-04-27	LC08_L1TP 2016-03-28
Dakar	S1A_IW_GRDH 2016-05-12	LC08_L1TP 2015-12-17
Gao	S1A_IW_GRDH 2016-06-13	LC08_L1TP 2016-07-08
Johannesburg	S1A_IW_GRDH 2015-10-20	LC08_L1TP 2015-12-21
Kampala	S1A_IW_GRDH 2016-07-04	LC08_L1TP 2016-01-29
Katsina	S1A_IW_GRDH 2016-05-12	LC08_L1TP 2015-10-23
Nairobi	S1A_IW_GRDH 2016-10-27	LC08_L1TP 2016-01-24
Ouagadougou	S1A_IW_GRDH 2015-04-13	LC08_L1TP 2016-12-22
Saint-Louis	S1A_IW_GRDH 2016-06-05	LC08_L1TP 2016-10-09
Windhoek	S1B_IW_GRDH 2016-10-09	LC08_L1TP 2016-01-14

3.2.3 Feature extraction

Grey Level Co-Occurence Matrix (GLCM) textures were computed with an inter-pixel distance of 1 and 32 levels of quantization using the Orfeo Toolbox (Grizonnet *et al.*, 2017), after a 2% histogram cutting on the source SAR data. GLCMs were constructed for the direction angles 0, 45, 90, and 135 degrees ; but only the average value was considered. The GLCMs were computed independently for each polarization (VV and VH) and with four different window sizes (5×5, 7×7, 9×9, and 11×11) commonly used for land cover classification in urban environments (Braun and Hochschild, 2015). A set of 18 textures was extracted: energy, entropy, correlation, inertia, cluster shade, cluster prominence, Harralick correlation, mean, variance, dissimilarity, sum average, sum variance, sum entropy, difference of entropies, difference of variances, and two information measures of correlation (IC1 and IC2). This resulted in the extraction of 72 texture features for each polarization, that is, 144 in total.

Several GLCMs texture features can be highly correlated, such as energy and entropy, or inertia and the inverse different moment. In order to reduce the dimensionality of the dataset, a Principal Component Analysis (PCA) was performed on each combination of polarization and window size. Only the first six components of each PCA—which consistently explained more than 95% of the variance, were retained. This reduced the number of features from 144 to 48.

SAR features were reprojected to Universal Transverse Mercator (UTM) coordinate system only after the computation of GLCM textures in order to minimize the destruction of textural information. All Landsat 8 bands—including thermal bands, were retained without further processing except of co-registration to the spatial resolution of SAR products (i.e. about 10 m).

3.2.4 Classification

The binary classification task—built-up vs. non-built-up, was performed using the Random Forest (RF) classifier, which has been shown to be relatively effective in the context of multisource and multimodal data classification (Pal, 2005; Gislason, Benediktsson and Sveinsson, 2006; Belgiu and Drăguț, 2016). The implementation was based on Python and a set of libraries, including: NumPy (Oliphant, 2007), SciPy (Oliphant, 2007), and Rasterio (Gillies, 2013) for raster processing, Shapely (Gillies, 2007) and Geopandas for vector processing and Scikit-learn (Pedregosa *et al.*, 2011) for machine learning. The Python code that supported the present study is available on Github[1], and the associated datasets can be acquired through Zenodo[2].

In order to assess the optimal combination of features in the context of a pixel-based supervised classification, 12 different classifications were performed using different input features. Table 3.3 lists the features used for each classification scheme. The schemes #1 to #9 were single-source classifications, based either on optical or SAR data. Previous studies suggested that combining several window sizes could improve classification accuracies by including spatial information from multiple scales

[1] https://github.com/yannforget/landsat-sentinel-fusion
[2] https://zenodo.org/record/1450932

(Puissant, Hirsch and Weber, 2005). The schemes #2 to #9 were designed to identify the optimal window size and polarization combination to classify built-up areas using Sentinel-1. Finally, the schemes #10 to #12 included both optical and SAR data with different window sizes for the computation of the GLCMs.

In the 12 classification schemes, the RF ensemble was constructed with 50 trees and a maximum number of features per tree equal to the square root of the total number of features—as suggested by previous studies (Gislason, Benediktsson and Sveinsson, 2006). Imbalance issues in the training dataset between the built-up and the non-built-up classes were overcome by a random over-sampling of the minority class (Lemaître, Nogueira and Aridas, 2017). Additionally, in order to ensure the reproducibility of the results, fixed random seeds were used.

Table 3.3: Label, input features and number of dimensions of the 12 classification schemes.

##	Scheme label	SAR features	Optical features	Dims.
1	`optical`	None	Landsat bands	8
2	`vv_5×5`	PCA GLCM 5×5 VV	None	6
3	`vh_5×5`	PCA GLCM 5×5 VH	None	6
4	`vv_vh_5×5`	PCA GLCM 5×5 [VV, VH]	None	12
5	`vv_vh_7×7`	PCA GLCM 7×7 [VV, VH]	None	12
6	`vv_vh_9×9`	PCA GLCM 9×9 [VV, VH]	None	12
7	`vv_vh_11×11`	PCA GLCM 11×11 [VV, VH]	None	12
8	`vv_vh_5×5_9×9`	PCA GLCM [5×5, 9×9] [VV, VH]	None	24
9	`vv_vh_5×5_11×11`	PCA GLCM [5×5, 11×11] [VV, VH]	None	24
10	`fusion_5×5`	PCA GLCM 5×5 [VV, VH]	Landsat bands	18
11	`fusion_9×9`	PCA GLCM 9×9 [VV, VH]	Landsat bands	18
12	`fusion_11×11`	PCA GLCM 11×11 [VV, VH]	Landsat bands	18

3.2.5 Validation

Reference polygons were digitized from very high spatial resolution imagery through Google Earth to support both the training and the validation of the classification models. Four land cover classes were collected: built-up, bare soil, low vegetation (sparse or small vegetation), and high vegetation (dense and tall

vegetation). As stated in previous studies (Potere and Schneider, 2007; Mertes *et al.*, 2015), no consensus has been reached regarding the definition of a built-up area. In the context of this study, a pixel is considered a built-up area if its surface is covered by at least 25% of elevated constructions. Reference polygons were digitized according to this definition through visual interpretation. All non-built-up samples (bare soil, low vegetation and high vegetation) were concatenated to build the binary (built-up vs. non-built-up) training and validation datasets. However, specific land cover samples were also used to assess the performance of the models in specific areas. Table 3.4 shows the number of polygons collected for each land cover and each case study, together with the amount of resulting samples (in pixels) after rasterization.

Table 3.4: Number of reference pixels for each case study and land cover. Enclosed in brackets: the number of polygons before rasterization.

	Built-up	**Bare Soil**	**Low Vegetation**	**High Vegetation**
Antananarivo	42,596 (110)	31,769 (67)	60,423 (53)	22,338 (50)
Bukavu	30,762 (54)	5,956 (20)	19,196 (21)	7,308 (22)
Chimoio	29,040 (79)	17,405 (59)	11,891 (63)	11,347 (50)
Dakar	123,386 (76)	14,367 (41)	60,993 (53)	29,739 (33)
Gao	18,998 (74)	46,834 (45)	805 (25)	1,348 (25)
Johannesburg	570,282 (260)	69,106 (91)	97,100 (112)	26,315 (37)
Kampala	41,528 (89)	5,049 (34)	21,033 (44)	9,376 (22)
Katsina	37,507 (95)	11,710 (55)	4,411 (31)	2,107 (28)
Nairobi	60,371 (103)	15,030 (46)	18,947 (41)	12,666 (23)
Ouagadougou	83,540 (62)	22,477 (24)	66,624 (15)	26,078 (7)
Saint-Louis	13,154 (64)	24,162 (47)	25,701 (40)	10,388 (22)
Windhoek	62,464 (60)	50,247 (79)	26,032 (48)	14,655 (28)

In order to ensure the independence of the training and validation datasets, the reference samples were randomly split at the polygon level—the samples inside a given polygon being characterized by a high spatial autocorrelation. Each classification is performed ten times with different random splits, then assessment metrics (F1-score (Equation 3.1), land cover accuracies) and classifier characteristics (feature importances, probabilities) were averaged for statistical and visual interpretation.

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{3.1}$$

3.3 Results and discussion

Table 3.5 presents the F1-score obtained with each classification scheme in each case study. In 10 case studies out of 12, optical-based schemes reached a higher F1-score than SAR-based schemes. The two case studies where SAR-based schemes appear more accurate are Gao (+7.6 points) and Katsina (+3.4 points): two small urban areas located in an arid climate characterized by a domination of bare land in the landscape. However, the observation is not confirmed by the low scores obtained by SAR-based schemes in cities such as Ouagadougou and Saint-Louis, which present similar characteristics. These results suggest that optical-based schemes are superior to SAR-based schemes in the context of pixel-based classifications from a single sensor. This can be explained by the speckle noise inherent to SAR data and by the loss of spatial resolution that occurred during the computation of the GLCM textures. However, in 11 case studies out of 12, the multi-sensor classification schemes reached the highest F1-scores—the only exception being Gao, where the SAR-based scheme performed better. In most of the case studies, complementing optical data with SAR features improved the classification performances, sometimes dramatically (+6.7 points in Chimoio, +6.6 in Saint-Louis, +5.8 in Gao, +4 in Bukavu and Katsina). This suggests that combining optical and SAR-based GLCM textures in the context of a pixel-based classification can be a reliable and robust strategy.

Table 3.5: F1-score obtained by each classification scheme in each case study (see Table 3.3 for the characteristics of each scheme).

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
Antananarivo	0.92	0.77	0.61	0.79	0.83	0.86	0.88	0.86	0.88	0.93	0.93	0.93
Bukavu	0.92	0.83	0.74	0.85	0.89	0.91	0.93	0.91	0.93	0.94	0.96	0.96
Chimoio	0.83	0.32	0.10	0.37	0.45	0.54	0.58	0.50	0.55	0.90	0.90	0.90
Dakar	0.95	0.68	0.65	0.74	0.76	0.78	0.80	0.78	0.80	0.95	0.96	0.95
Gao	0.76	0.83	0.71	0.81	0.82	0.83	0.82	0.83	0.83	0.81	0.81	0.80
Johannesburg	0.96	0.88	0.88	0.90	0.91	0.92	0.93	0.92	0.93	0.97	0.98	0.98

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
Kampala	0.98	0.92	0.77	0.93	0.95	0.96	0.97	0.96	0.97	0.98	0.98	0.98
Katsina	0.93	0.94	0.92	0.94	0.95	0.96	0.96	0.96	0.96	0.96	0.97	0.97
Nairobi	0.94	0.79	0.76	0.83	0.86	0.88	0.90	0.87	0.90	0.96	0.96	0.96
Ouagadougou	0.98	0.65	0.43	0.67	0.70	0.72	0.74	0.71	0.73	0.98	0.99	0.99
Saint-Louis	0.89	0.56	0.06	0.55	0.61	0.64	0.67	0.63	0.66	0.96	0.95	0.94
Windhoek	0.97	0.69	0.64	0.76	0.81	0.85	0.88	0.85	0.87	0.97	0.97	0.97

Figure 3.1 summarizes the F1-scores obtained by each classification scheme across all the case studies. The box plot confirms the previously observed trends. The differences between the various SAR-based classification schemes are also highlighted. Classifications based on the VV polarization reached higher scores compared to the ones based on the VH polarization. However, combining texture features derived from both polarizations appears to increase the reliability of the classification models, leading to higher scores and a lower standard deviation. Likewise, the window size used for the computation of the GLCM did influence the classification scores. In our case studies, larger window sizes led to higher scores and reduced the variability across the case studies. We previously stated the hypothesis that combining several window sizes could increase the classification performance by including spatial information from multiple scales. The results obtained tend to refute this hypothesis, as the schemes combining textures from two window sizes did not perform better. Generally, the fusion schemes consistently obtained the highest scores. However, contrary to the trend observed in single-sensor SAR-based schemes, combining optical data with larger window sizes textures does not seem to increase the classification performance. Fusion schemes that include textures computed with smaller window sizes (5×5 and 9×9) benefited from slightly less variability across the case studies.

In classifiers based on a forest of decision trees such as RF, the relative contribution of each feature can be evaluated through the feature importance measure. The value ranges from 0.0 to 1.0, where 0.0 would indicate a feature that does not contribute to the classification, and 1.0 a feature that alone classifies all samples. Figure 3.2 shows the distribution of the importance measure across the input features used in

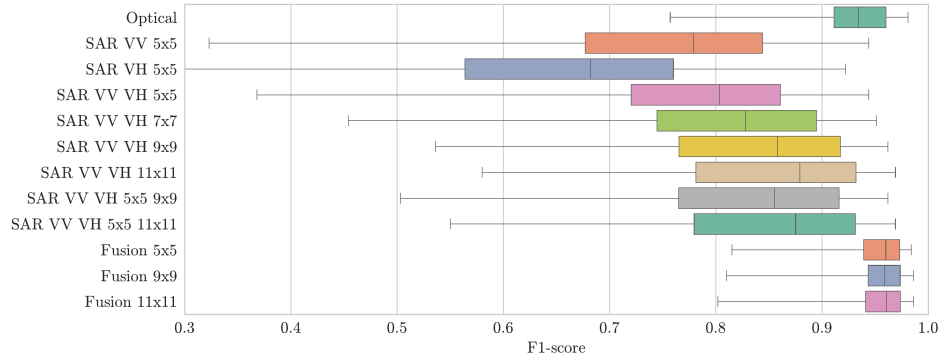

Figure 3.1: Box plot of the F1-score obtained with each classification scheme for each case study.

the fusion classification scheme, grouped by data source: VV or VH polarization for SAR data, and multispectral or thermal for optical data. Generally, texture features derived from the VV polarization and multispectral bands from Landsat 8 were the features contributing the most to the construction of the decision trees. Considering the lower scores obtained in single-sensor classification schemes with SAR data, lower importances for SAR features could be expected. However, the grouped importance of SAR features was superior to the importance of optical features in 5 case studies, and exceeded 40% of the contribution in 10 case studies. Furthermore, the relative importance of optical and SAR features did not appear correlated with their respective scores in the context of a single-sensor classification. For instance, in Antananarivo and Nairobi, the best SAR-based classification scheme reached a F1-score lower than the optical-based scheme by, respectively, 4.2 and 3.9 points. However, in the context of the fusion scheme, the contribution of SAR features was superior to the contribution of optical features. This suggests that the combination of both optical and SAR features adds information to the classifier that cannot be modeled in the context of a single-sensor classification.

As previously stated, combining SAR-based textural information and optical imagery is expected to improve the classification performance in bare lands. Figure 3.3 shows the mean accuracy of each classification scheme in three non-built-up land

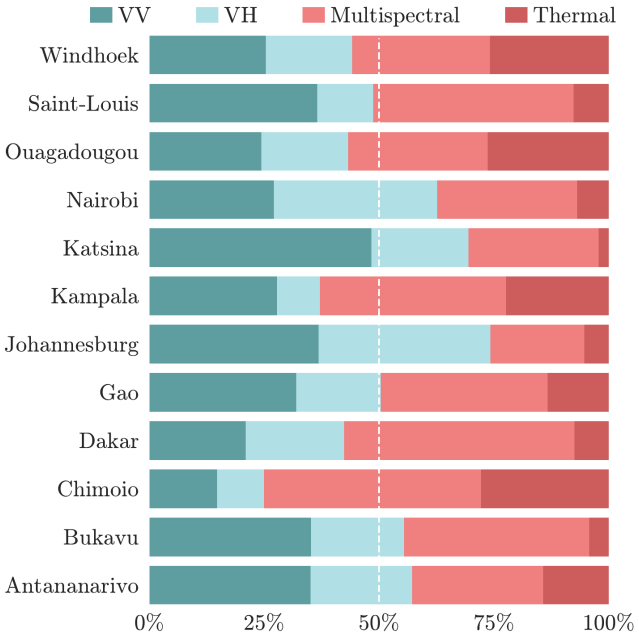

Figure 3.2: Grouped Random Forest feature importances for the fusion scheme in each case study. VV and VH groups correspond to the SAR features derived from a given polarization. Multispectral and thermal groups correspond to the Landsat 8 features.

covers: bare soil, low vegetation and high vegetation. As expected, the performance of the classification model in bare soil areas was superior in SAR-based schemes than in optical-based schemes. On the contrary, SAR-based classification schemes, especially the ones based on the VH polarization, suffered from low accuracies in densely vegetated areas. Finally, fusion schemes based on both SAR and optical data benefited from the complementarity between both sensors and present high accuracies in the three land covers.

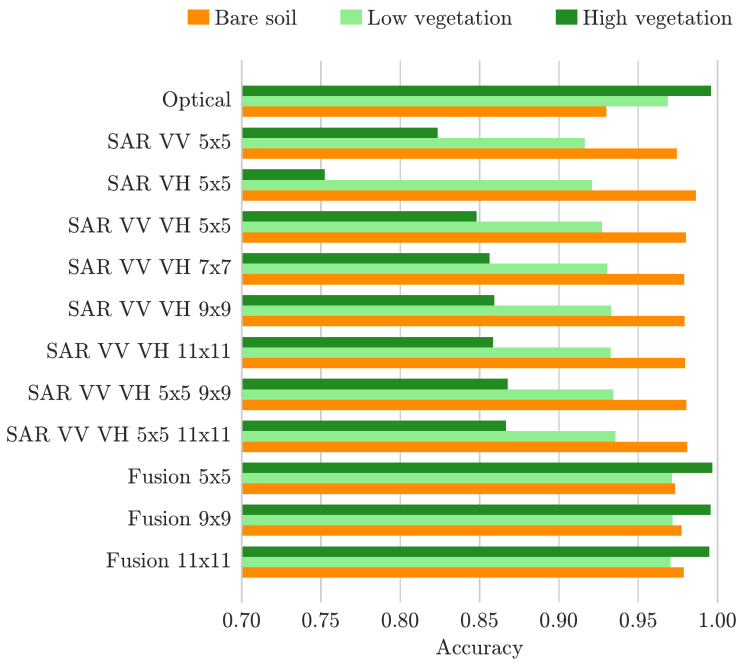

Figure 3.3: Classification accuracy in specific land covers for each scheme.

Figure 3.4 shows the probabilistic output of three different classifiers in Katsina, Nigeria: one based only on optical data, and two based on both optical and SAR data with different GLCM window sizes. Visually, the fusion classifiers appears to better distinguish between the built-up areas and the surrounding bare lands. This leads to lower rates of misclassification after thresholding of the probabilities, as previously shown by the assessment metrics. A side effect of the data fusion is the disappearance of the road network from the built-up class. Indeed, the classifier

highly relies on the textural information from the SAR features to discriminate between built-up and bare land, and roads are therefore excluded from the built-up class.

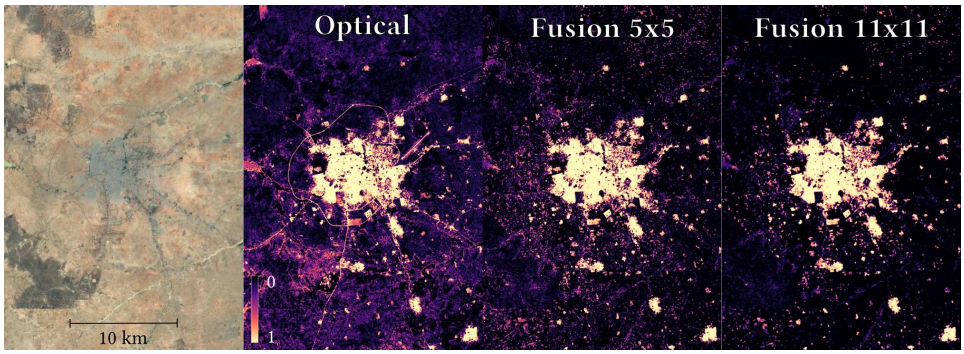

Figure 3.4: Random Forest class probabilities in Katsina, Nigeria. **a**) Aerial view of the case study; **b**) `optical` scheme probabilities; **c**) `fusion_5×5` scheme probabilities; **d**) `fusion_11×11` scheme probabilities. Satellite imagery courtesy of Google.

There is some cases where data fusion is less beneficial. Figure 3.5 shows the probabilistic output of the optical, SAR, and fusion schemes in Bukavu, D.R. Congo. This case study, located in a mountainous area, presents two major obstacles for SAR data: dense vegetation in the north-west and steep slopes in the south-east. Mapping the probabilistic output of the SAR-based classifier reveals the confusion occurring in these areas. As a result, the probabilistic output of the fusion-based classifier appears nearly as a copy of the optical-based one.

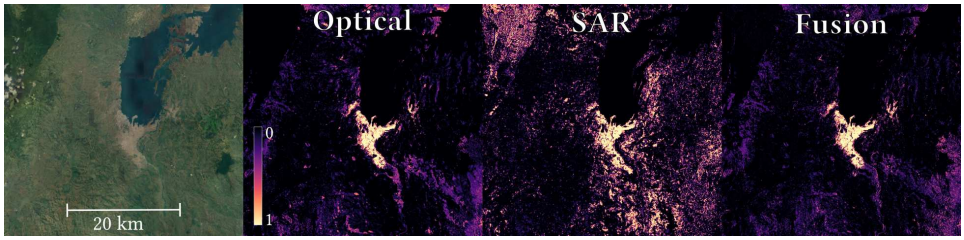

Figure 3.5: Random Forest class probabilities in Bukavu, D. R. Congo. **a**) Aerial view of the case study; **b**) `optical` scheme probabilities; **c**) `sar_vv_vh_5×5_11×11` scheme probabilities; **d**) `fusion_11×11` scheme probabilities. Satellite imagery courtesy of Google.

3.4 Conclusion

With the increasing availability of free imagery from multiple sensors such as Sentinel-1 and Landsat 8, data fusion is one of the main challenges in remote sensing. The objective of this paper was to assess the combined use of both Landsat 8 and Sentinel-1 imagery with a fusion scheme that relies on a simple pixel-based classifier. The main expectation was a better discrimination between built-up and bare soil areas in the context of urban mapping in Sub-Saharan Africa. The presented results suggest that the complementarity between medium resolution optical and SAR sensors can be exploited in the context of a supervised pixel-based classification. However, to make the pixel-based approach effective, textural information must be extracted from SAR backscattering in order to reduce the speckle noise and to provide contextual information at the pixel level.

Classification schemes including both optical and SAR features reached the highest scores in 11 case studies out of 12. Single-sensor classifiers making use of GLCM textures derived from the VV polarization outperformed the ones based on the VH polarization. Nevertheless, combining both polarizations consistently increased the classification performance. Likewise, large GLCM window sizes (9×9 or 11×11) provided a slight improvement of the classification performance both in the context of single-sensor classification and in fusion schemes. However, contrary to an hypothesis that we formulated, combining textures derived from multiple GLCM window sizes—in order to include spatial information from multiple scales at the pixel level, did not lead to a better classification performance. The visual interpretation of the results obtained suggests that small GLCM window sizes favor the detection of isolated settlements and buildings, whereas larger window sizes lead to a better differentiation between built-up areas and bare lands. In the context of this study, the RF classifier was not able to take advantage of both.

The assessment of the classifiers performances in specific land covers confirmed the high level of complementarity between the two sensors. Single-sensor SAR-based classifications presented high accuracies in bare soil areas, but suffered from a confusion between dense vegetation and buildings. On the contrary, optical-based

classifiers showed a high ability to discriminate between vegetation and built-up, but a low differentiation between bare soil and built-up—especially in the most arid landscapes such as in Gao or Katsina. This complementarity was correctly modeled by the RF classifier and, as a result, the fusion schemes presented high accuracies in both bare lands and vegetated areas.

Nevertheless, the fusion was less beneficial in case studies characterized by the presence of dense vegetation and steep slopes—for instance in a mountainous and subtropical urban area such as Bukavu. However, in this case, the RF classifier was able to learn from the training dataset that SAR data were not reliable. This suggests that including features derived from a DEM—for instance slope and aspect, could improve the model ability to quantify the reliability of SAR data at the pixel level. Such a strategy could also take place at the decision level. Further work is also required to assess the reliability of the presented approach in the context of similar sensors such as Sentinel-2, ERS-1 and ERS-2.

References

Bai, X. *et al.* (2012) 'Health and wellbeing in the changing urban environment: Complex challenges, scientific responses, and the way forward', *Current Opinion in Environmental Sustainability*, 4(4), pp. 465–472. doi: 10.1016/j.cosust.2012.09.009.

Belgiu, M. and Drăguț, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, pp. 24–31. doi: 10.1016/j.isprsjprs.2016.01.011.

Benediktsson, J. A. and Kanellopoulos, I. (1999) 'Classification of multisource and hyperspectral data based on decision fusion', *Geoscience and Remote Sensing, IEEE Transactions on*, 37(3), pp. 1367–1377. doi: 10.1109/36.763301.

Braun, A. and Hochschild, V. (2015) 'Combined use of SAR and optical data for environmental assessments around refugee camps in semiarid landscapes', *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3, pp. 777–782. doi: 10.5194/isprsarchives-XL-7-W3-777-2015.

Brockmann, D. and Helbing, D. (2013) 'The Hidden Geometry of Complex, Network-Driven Contagion Phenomena', *Science*, 342(6164), pp. 1337–1342. doi: 10.1126/science.1245200.

Chen, M. *et al.* (2008) 'Assessing objective techniques for gauge-based analyses of global daily precipitation', *Journal of Geophysical Research*, 113(D4). doi: 10.1029/2007JD009132.

Clauss, K. *et al.* (2018) 'Sentinelsat'. Zenodo. doi: 10.5281/zenodo.595961.

Clerici, N., Valbuena Calderón, C. A. and Posada, J. M. (2017) 'Fusion of Sentinel-1A and Sentinel-2A data for land cover mapping: A case study in the lower Magdalena region, Colombia', *Journal of Maps*, 13(2), pp. 718–726. doi: 10.1080/17445647.2017.1372316.

Didan, K. (2015) 'MOD13C2 MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V006', *NASA LP DAAC*. doi: 10.5067/MODIS/MOD13C2.006.

Dye, C. (2008) 'Health and Urban Living', *Science*, 319(5864), pp. 766–769. doi: 10.1126/science.1150198.

ESA (2018) 'Sentinel Application Platform (SNAP)'.

Fauvel, M., Chanussot, J. and Benediktsson, J. A. (2006) 'Decision Fusion for the Classification of Urban Remote Sensing Images', *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), pp. 2828–2838. doi: 10.1109/TGRS.2006.876708.

Forget, Y. (2018) 'Landsatxplore'. Zenodo. doi: 10.5281/zenodo.1291423.

Forster, B. C. (1993) 'Coefficient of variation as a measure of urban spatial attributes, using SPOT HRV and Landsat TM data', *International Journal of Remote Sensing*, 14(12), pp. 2403–2409. doi: 10.1080/01431169308954044.

Gamba, P. (2014) 'Image and data fusion in remote sensing of urban areas: Status issues and research trends', *International Journal of Image and Data Fusion*, 5(1), pp. 2–12. doi: 10.1080/19479832.2013.848477.

Gamba, P. and Herold, M. (2009) *Global mapping of human settlement: Experiences, datasets, and prospects*. Boca Raton: CRC Press (Taylor & Francis series in remote sensing applications).

Gillies, S. (2007) 'Shapely: Manipulation and analysis of geometric objects'. toblerity.org.

Gillies, S. (2013) 'Rasterio: Geospatial raster I/O for Python programmers'. Mapbox.

Gislason, P. O., Benediktsson, J. A. and Sveinsson, J. R. (2006) 'Random Forests for land cover classification', *Pattern Recognition Letters*, 27(4), pp. 294–300. doi: 10.1016/j.patrec.2005.08.011.

Gotlieb, C. C. and Kreyszig, H. E. (1990) 'Texture descriptors based on co-occurrence matrices', *Computer Vision, Graphics, and Image Processing*, 51(2), pp. 70–86. doi: 10.1016/S0734-189X(05)80063-5.

Griffiths, P. *et al.* (2010) 'Mapping megacity growth with multi-sensor data', *Remote Sensing of Environment*, 114(2), pp. 426–439. doi: 10.1016/j.rse.2009.09.012.

Grimm, N. B. *et al.* (2008) 'Global Change and the Ecology of Cities', *Science*, 319(5864), pp. 756–760. doi: 10.1126/science.1150195.

Grizonnet, M. *et al.* (2017) 'Orfeo ToolBox: Open source processing of remote sensing images', *Open Geospatial Data, Software and Standards*, 2(1). doi: 10.1186/s40965-017-0031-6.

Haralick, R. M., Shanmugam, K. and Dinstein, I. (1973) 'Textural Features for Image Classification', *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), pp. 610–621. doi: 10.1109/TSMC.1973.4309314.

Joshi, N. *et al.* (2016) 'A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring', *Remote Sensing*, 8(1), p. 70. doi: 10.3390/rs8010070.

JPL, N. (2013) 'NASA Shuttle Radar Topography Mission Global 1 arc second', *NASA EOSDIS Land Processes DAAC*. doi: 10.5067/MEaSUREs/SRTM/SRTMGL1.003.

Kottek, M. *et al.* (2006) 'World Map of the Köppen-Geiger climate classification updated', *Meteorologische Zeitschrift*, 15(3), pp. 259–263. doi: 10.1127/0941-2948/2006/0130.

Kupidura, P. (2016) 'Comparison of Filters Dedicated to Speckle Suppression in SAR Images', in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Prague, Czech Republic, pp. 269–276. doi: 10.5194/isprsarchives-XLI-B7-269-2016.

Lee, J. S. *et al.* (1994) 'Speckle filtering of synthetic aperture radar images: A review', *Remote Sensing Reviews*, 8(4), pp. 313–340. doi: 10.1080/02757259409532206.

Lemaître, G., Nogueira, F. and Aridas, C. K. (2017) 'Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning', *Journal of Machine Learning Research*, 18(17), pp. 1–5.

Li, H. *et al.* (2017) 'Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index', *Remote Sensing*, 9(12), p. 249. doi: 10.3390/rs9030249.

Linard, C. *et al.* (2012) 'Population Distribution, Settlement Patterns and Accessibility across Africa in 2010', *PLoS ONE*, 7(2). doi: 10.1371/journal.pone.0031743.

Linard, C., Tatem, A. J. and Gilbert, M. (2013) 'Modelling spatial patterns of urban growth in Africa', *Applied Geography*, 44, pp. 23–32. doi: 10.1016/j.apgeog.2013.07.009.

McGranahan, G. *et al.* (2009) *Africa's urban transition and the role of regional collaboration*. 5. London: International Institute for Environment; Development (IIED).

Mertes, C. M. *et al.* (2015) 'Detecting change in urban areas at continental scales with MODIS data', *Remote Sensing of Environment*, 158, pp. 331–347. doi: 10.1016/j.rse.2014.09.023.

Moore, M., Gould, P. and Keary, B. S. (2003) 'Global urbanization and impact on health', *International Journal of Hygiene and Environmental Health*, 206(4-5), pp. 269–278. doi: 10.1078/1438-4639-00223.

Oliphant, T. E. (2007) 'Python for Scientific Computing', *Computing in Science & Engineering*, 9(3), pp. 10–20. doi: 10.1109/MCSE.2007.58.

Pal, M. (2005) 'Random forest classifier for remote sensing classification', *International Journal of Remote Sensing*, 26(1), pp. 217–222. doi: 10.1080/01431160412331269698.

Pedregosa, F. *et al.* (2011) 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research*, 12(Oct), pp. 2825–2830.

Potere, D. and Schneider, A. (2007) 'A critical look at representations of urban areas in global maps', *GeoJournal*, 69(1), pp. 55–80. doi: 10.1007/s10708-007-9102-z.

Potere, D. *et al.* (2009) 'Mapping urban areas on a global scale: Which of the eight maps now available is more accurate?', *International Journal of Remote Sensing*, 30(24), pp. 6531–6558. doi: 10.1080/01431160903121134.

Puissant, A., Hirsch, J. and Weber, C. (2005) 'The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery', *International Journal of Remote Sensing*, 26(4), pp. 733–745. doi: 10.1080/01431160512331316838.

Qian, J., Zhou, Q. and Hou, Q. (2007) 'Comparison of pixel-based and object-oriented classification methods for extracting built-up areas in arid zone', in *ISPRS Workshop on Updating Geo-Spatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs*, pp. 163–171.

Rasul, A. *et al.* (2018) 'Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates', *Land*, 7(3), p. 81. doi: 10.3390/land7030081.

Rubel, F. *et al.* (2017) 'The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 18002100', *Meteorologische Zeitschrift*, 26(2), pp. 115–125. doi: 10.1127/metz/2016/0816.

Shao, Z. *et al.* (2016) 'Mapping Urban Impervious Surface by Fusing Optical and SAR Data at the Decision Level', *Remote Sensing*, 8(11), p. 945. doi: 10.3390/rs8110945.

Small, C. (2001) 'Multiresolution analysis of urban reflectance', in *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*. Rome, Italy: IEEE, pp. 15–19. doi: 10.1109/DFUA.2001.985717.

Small, C. (2005) 'A global analysis of urban reflectance', *International Journal of Remote Sensing*, 26(4), pp. 661–681. doi: 10.1080/01431160310001654950.

Small, D. (2011) 'Flattening Gamma: Radiometric Terrain Correction for SAR Imagery', *IEEE Transactions on Geoscience and Remote Sensing*, 49(8), pp. 3081–3093. doi: 10.1109/TGRS.2011.2120616.

Small, D. and Shubert, A. (2008) 'Guide to ASAR Geocoding', *RSL-ASAR-GC-AD*, (1.01).

Soergel, U. (2010) *Radar remote sensing of urban areas*. Dordrecht ; New York: Springer (Remote sensing and digital image processing, v. 15).

Tupin, F. (2010) 'Fusion of Optical and SAR Images', in Soergel, U. (ed.) *Radar remote sensing of urban areas*. Dordrecht ; New York: Springer (Remote sensing and digital image processing, 15), pp. 133–159.

UN-Habitat (2015) *Sustainable Urban Development in Africa*. Nairobi, Kenya: UN-Habitat.

Vermote, E. *et al.* (2016) 'Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product', *Remote Sensing of Environment*, 185, pp. 46–56. doi: 10.1016/j.rse.2016.04.008.

Waske, B. and van der Linden, S. (2008) 'Classifying Multilevel Imagery From SAR and Optical Sensors by Decision Fusion', *IEEE Transactions on Geoscience and Remote Sensing*, 46(5), pp. 1457–1466. doi: 10.1109/TGRS.2008.916089.

Weng, Q. (2014) *Global urban monitoring and assessment through earth observation*. Crc Press.

Wentz, E. *et al.* (2014) 'Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing', *Remote Sensing*, 6(5), pp. 3879–3905. doi: 10.3390/rs6053879.

Worldpop (2016) 'Africa 1km Population', *University of Southampton*. doi: 10.5258/SOTON/WP00004.

Zhang, C., Chen, Y. and Lu, D. (2015) 'Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery', *International Journal of Remote Sensing*, 36(17), pp. 4483–4500. doi: 10.1080/01431161.2015.1084552.

Zhang, H., Zhang, Y. and Lin, H. (2012) 'A comparison study of impervious surfaces estimation using optical and SAR remote sensing images', *International Journal of Applied Earth Observation and Geoinformation*, 18, pp. 148–156. doi: 10.1016/j.jag.2011.12.015.

Zhang, Y., Zhang, H. and Lin, H. (2014) 'Improving the impervious surface estimation with combined use of optical and SAR remote sensing images', *Remote Sensing of Environment*, 141, pp. 155–167. doi: 10.1016/j.rse.2013.10.028.

Zhu, Z., Wang, S. and Woodcock, C. E. (2015) 'Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 47, 8, and Sentinel 2 images', *Remote Sensing of Environment*, 159, pp. 269–277. doi: 10.1016/j.rse.2014.12.014.

Zhu, Z. and Woodcock, C. E. (2012) 'Object-based cloud and cloud shadow detection in Landsat imagery', *Remote Sensing of Environment*, 118, pp. 83–94. doi: 10.1016/j.rse.2011.10.028.

Zhu, Z. *et al.* (2012) 'Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data', *Remote Sensing of Environment*, 117, pp. 72–82. doi: 10.1016/j.rse.2011.07.020.

Chapter 4

A dataset of urban expansion in 46 urban areas of SSA

This chapter is based on the following publication:

Forget, Y., Shimoni, M., Gilbert, M., & Linard, C. (2019). "Mapping 20 years of urban expansion in 46 urban areas of Sub-Saharan Africa". *Forthcoming*.

Abstract. By 2050, half of the net increase in the world's population is expected to take place in Sub-Saharan Africa (SSA), driving high urbanization rates and drastic land cover changes. But the data-scarce environment of SSA limits our understanding of the urban dynamics in the region. In this context, Earth Observation (EO) is an opportunity to gather accurate and up-to-date spatial information on urban extents. During the last decade, the adoption of open-access policies by major EO programs (CBERS, Landsat, Sentinel) allowed the production of several global high resolution (10–30 m) maps of human settlements. However, mapping accuracies in SSA are usually lower, limited by the lack of reference datasets to support training and validation of the classification models. Here we propose a mapping approach based on multi-sensor satellite imagery (Landsat, Sentinel-1, Envisat, ERS) and volunteered geographic information (OpenStreetMap) to solve the challenges of urban remote sensing in SSA. The proposed mapping approach is assessed in 17

case studies for an average F1-score of 0.93, and applied in 45 urban areas of SSA to produce a dataset of urban expansion from 1995 to 2015. Across the case studies, built-up areas averaged a compound annual growth rate of 5.5% between 1995 and 2015. The comparison with local population dynamics reveals the heterogeneity of urban dynamics in SSA. Overall, population densities in built-up areas are decreasing. But the impact of population growth on urban expansion differs depending on the size of the urban area and its income class.

4.1 Introduction

According to the latest predictions of the United Nations, the world population will increase from 7.4 billions in 2015 to 9.8 billions in 2050. More than half of this global increase is expected to take place in Africa (United Nations, 2017), driving high urbanization rates. Over the same period, the urban population of Sub-Saharan Africa (SSA) is expected to increase by 235% (United Nations, 2015). Such a rapid urbanization has already been experienced in the past by other countries during the first half of the 20th century (Cohen, 2004). However, the scale of change in SSA is unprecedented: by 2050, the urban areas will have to absorb nearly 900 millions of new dwellers (United Nations, 2018a), transforming the continent's surface. Another major difference from the experience of Europe or the United States is that the urbanization of SSA is occurring at low levels of per capita income, and in countries which are vulnerable to the global economy (Cohen, 2004). Since the beginning of the 1980s, trade liberalization, high unemployment rates and structural adjustment programs led to the deterioration of the living conditions in the urban areas of SSA, a lack of public infrastructures and services and the rise of the informal sector (Bocquier, 2003; McGranahan *et al.*, 2009; Andersson Djurfeldt, 2015). Although urban environments have been associated with a lower disease burden at the global or regional scales (Hay *et al.*, 2005; Tatem *et al.*, 2008; Dye, 2008), rapid and unplanned urban growth in a context of urban poverty is the source of many health hazards: indoor and outdoor air pollution, unsafe water, lack of sanitation structures, vector-borne diseases or physical hazards (traffic, accidental fires, floods) (Moore, Gould and Keary, 2003; McGranahan *et al.*, 2009; Bai *et*

al., 2012; Brockmann and Helbing, 2013). Additionally, urbanization is still one of the primary driver of habitat and biodiversity loss (Seto *et al.*, 2011). Environment degradation can occur far beyond the local scale through sewage contamination, rivers pollution, unregulated burning emissions, or industrial pollution (McGranahan *et al.*, 2009). Up-to-date and accurate geographic information on the urbanization dynamics is a prerequisite for understanding the phenomenon, modeling its causes and anticipating its consequences.

Traditionally, urbanization studies have been based on census data published by each country (United Nations, 2017). Beyond data availability issues, the definition of what constitutes and urban area is not consistent over both spaces and time. In Angola, a locality with at least 2,000 inhabitants is classified as urban, whereas the threshold would be 10,000 in Benin (Cohen, 2004). When, in the 1980s, China lowered the threshold to qualify a locality as urban, a massive increase in the urbanization has been observed—although only caused by administrative changes (Lin, 2002; Cohen, 2004). In this context, satellite remote sensing enables the study of the urbanization process in its physical dimension through the detection of built-up areas at different periods (Wentz *et al.*, 2014). The increasing availability of open satellite imagery datasets (Wulder *et al.*, 2012)—together with the reduction of the computing costs, allowed the production of several global built-up maps: GlobCover (Arino *et al.*, 2007) or MOD-500 (Schneider, Friedl and Potere, 2009) in the 2000s, and, more recently, the Global Human Settlement Layer (M. Pesaresi *et al.*, 2016), the Global Land Cover (Chen *et al.*, 2015), or the Human Built-up and Settlements Extent (Wang *et al.*, 2017). Those datasets have supported studies in a wide range of fields such as population mapping, urban planning, disease burden estimation, resource allocation, disaster management or environmental impact assessment (Hay *et al.*, 2005; Tatem *et al.*, 2007; Linard, Tatem and Gilbert, 2013; Pesaresi, Ehrlich and Freire, 2014). In parallel, numerous studies made use of satellite imagery to analyze urbanization dynamics at the local scale (Schneider, 2012; Li, Gong and Liang, 2015; Patel *et al.*, 2015; Rahman, 2016). For instance, (Angel *et al.*, 2016) mapped and analyzed the evolution of the urban extent of 200 cities between 1990 and 2014, and Schneider and Mertes (2014) identified multi-temporal

urban land extents in 142 Chinese cities between 1978 and 2010.

Still, despite decades of scientific progress, the detection of built-up areas remains a challenge because of the intraurban and interurban heterogeneity that characterizes the urban environment (Herold *et al.*, 2004; Small, 2005). Classifications based on optical sensors are characterized by lower accuracies in urban areas located in arid or semi-arid climates because of the spectral confusion occurring between bare soil and built-up elements (Zhang, Chen and Lu, 2015; Li *et al.*, 2017; Rasul *et al.*, 2018). Low-income countries also suffer from a lack of reference datasets in both quantity and quality, inducing higher rates of misclassifications (Potere *et al.*, 2009) or omissions of rural and suburban settlements. Today, the growing availability of geographic data brings new opportunities to tackle the aforementioned issues. Combining optical imagery with Synthetic Aperture Radar (SAR) backscattering can lead to a better separation between bare soil and built-up (Tupin, 2010; Zhang, Zhang and Lin, 2014; Gamba, 2014; Braun and Hochschild, 2015; Forget *et al.*, 2018). Likewise, the lack of reference datasets from government or commercial agencies can be compensated by the use of crowd-sourced geographic databases (Goodchild, 2007), such as its most prominent project OpenStreetMap (OSM) (Haklay and Weber, 2008; Mooney and Minghini, 2017), in order to support the training of the classification models (Schultz *et al.*, 2017; Yang *et al.*, 2017; Forget, Linard and Gilbert, 2018).

The purpose of this study is twofold. The first objective is to produce a reliable multi-temporal dataset of built-up maps for a sample of 45 urban areas in Sub-Saharan Africa at 5 different dates: circa 1995, 2000, 2005, 2010 and 2015. By leveraging both multi-sensor data fusion to improve built-up detection and the integration of OSM data to support the training of the classification models, we propose an automated and low-cost approach which may be appropriate at larger scales. In this paper, we present and assess the proposed methodology and discuss its limitations. The final aim is to provide a preliminary interpretation of the results obtained, through the analysis of built-up expansion and its relationship with population growth.

4.2 Material and methods

4.2.1 Case studies

The most consistent spectral characteristic of all urban areas in the world is their heterogeneity (Small, 2005). Because of environmental, historical, or cultural variations across Sub-Saharan Africa, a method developed for a given urban area is not guaranteed to be effective in another. As a matter of fact, a built-up area in the periphery of Kampala (Uganda, tropical rainforest) is a completely different spectral object than a settlement in the city center of Gao (Mali, hot desert). Likewise, urbanization rates and history differ according to the demographic, economic, and political dynamics of a given urban area.

Table 4.1: Climate, topography and population estimates of each case study. Values are aggregated for the area of interest of each case study, *i.e.* the 20 km rectangular buffer around the city centers. Climate data are derived from the Koppen-Geiger classification (Kottek *et al.*, 2006). Mean slope and elevation are computed from the Shuttle Radar Topographic Mission (SRTM) 30 m (JPL, 2013). Population is estimated using the AfriPop/WorldPop dataset (Linard *et al.*, 2012; Worldpop, 2016).

City (Country)	Population	Climate	Elevation	Slope
Antananarivo (MDG)	2,454,009	Subtropical highland	1319.6	14.8
Bouake (CIV)	836,441	Tropical savanna	290.7	6.1
Brazzaville (COG)	7,858,583	Tropical savanna	327.3	9.8
Bukavu (COD)	1,068,012	Tropical savanna	1756.1	22.8
Chimoio (MOZ)	457,422	Humid subtropical	612.6	8.3
Dakar (SEN)	3,308,199	Hot semi-arid	12.5	2.3
Dodoma (TZA)	481,263	Hot semi-arid	1139.9	6.6
Freetown (SLE)	1,196,714	Tropical monsoon	121.0	5.4
Gao (MLI)	161,019	Hot desert	272.1	4.5
Ikirun (NGA)	1,323,133	Tropical savanna	394.7	8.1
Iringa (TZA)	252,164	Humid subtropical	1576.9	10.3
Johannesburg (ZAF)	4,816,594	Subtropical highland	1611.0	7.5

City (Country)	Population	Climate	Elevation	Slope
Kabwe (ZMB)	255,667	Humid subtropical	1168.7	3.7
Kampala (UGA)	3,477,053	Tropical rainforest	1171.0	7.5
Kaolack (SEN)	447,639	Hot semi-arid	14.2	3.9
Katsina (NGA)	1,019,434	Hot semi-arid	495.5	4.2
Kayamandi (ZAF)	1291104	Warm-summer med.	281.2	16.8
Kinshasa (COD)	8265198	Tropical savanna	319.6	9.1
Kisumu (KEN)	1183345	Tropical rainforest	1292.6	6.9
Libreville (GAB)	744131	Tropical monsoon	18.2	4.8
Lusaka (ZMB)	2557066	Humid subtropical	1216.4	4.4
Mbeya (TZA)	665390	Subtropical highland	1791.6	20.0
Mekele (ETH)	452457	Hot semi-arid	2143.1	15.5
Monrovia (LBR)	1381459	Tropical monsoon	16.8	2.9
Nairobi (KEN)	5175740	Temperate oceanic	1738.6	7.9
Ndola (ZMB)	637717	Humid subtropical	1289.0	5.1
Nelspruit (ZAF)	164982	Humid subtropical	853.9	16.4
Nzerekore (GIN)	339140	Tropical savanna	468.3	11.2
Obuasi (GHA)	375931	Tropical savanna	196.4	12.0
Okene (NGA)	983744	Tropical savanna	298.9	9.5
Onitsha (NGA)	2593562	Tropical savanna	74.6	5.9
Ouagadougou (BFA)	2239604	Hot semi-arid	306.5	3.8
Owo (NGA)	427986	Tropical savanna	271.7	7.7
Pietermaritzburg (ZAF)	617133	Temperate oceanic	867.6	15.2
Pietersburg (ZAF)	205025	Hot semi-arid	1303.5	5.1
Saint-Louis (SEN)	297477	Hot desert	5.9	2.0
San Pedro (CIV)	113641	Tropical savanna	31.1	7.3
Shaki (NGA)	395163	Tropical savanna	393.3	5.6
Tamale (GHA)	498597	Tropical savanna	148.4	5.3
Toamasina (MDG)	333439	Tropical rainforest	45.4	7.6
Tulear (MDG)	305710	Hot desert	82.7	9.1
Umuahia (NGA)	1450588	Tropical monsoon	104.2	7.9

City (Country)	Population	Climate	Elevation	Slope
Windhoek (NAM)	383456	Hot desert	1819.4	16.5
Yamoussoukro (CIV)	358063	Tropical savanna	196.7	6.1
Ziguinchor (SEN)	293083	Tropical savanna	14.9	3.9

The objective of the selection step was to ensure that various types of urban areas were represented for both the validation of the built-up classification method and the multi-temporal analysis of the urbanization dynamics. The sample of 45 case studies, shown in Table 4.1, has been selected to maximize the diversity in terms of climate, population size, topography, and economy.

4.2.2 Data acquisition and preprocessing

In order to cover the entire 1995–2015 period with both optical and SAR images, data have been acquired from 7 different sensors: Landsat 5 TM (125 products), Landsat 7 ETM+ (66), Landsat 8 (60), ERS-1&2 (127), Envisat (54) and Sentinel-1 (49), that is, 481 scenes in total. However, as shown in Figure 4.1, the coverage was far from being complete. Both optical and SAR were only available in 71% of cases. Overall, the coverage is lower as we go back in time. Thanks to the systematic global acquisition approach of today's satellites such as Landsat 8 or Sentinel-1, coverage was complete in 2015. In the past, global acquisitions were not systematic. For instance, only 6% of the scenes in the Landsat 5 TM catalog are over Africa, against 12% for Landsat 7 ETM+ (Roy *et al.*, 2010). This left many of African locations without any Landsat acquisition before 1998, and a similar issue is observed with the catalogues of ERS-1&2 or Envisat. Cloud cover above tropical areas was also a major issue. As a matter of fact, the average cloud cover of the Landsat acquisitions over Kinshasa was 76%. Across the entire catalog, only 5 scenes with a cloud cover less than 10% were available over Kinshasa, and none before 2000. The issue was partially mitigated by the use of SAR acquisitions, which are not sensitive to cloud cover or weather variations.

The partial availability of historical satellite imagery over Africa requires a thorough

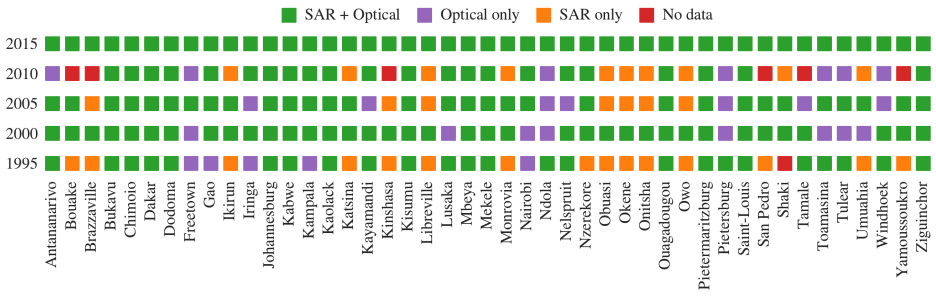

Figure 4.1: Availability of optical and SAR imagery for each case study.

imagery selection process. Three main criteria were taken into account: (1) the image quality, especially in terms of cloud cover, (2) the temporal distance between the optical and the SAR acquisitions, and (3) the seasonal distance between the acquisitions of a given case study for multi-temporal comparability. The automation of the imagery selection and acquisition were allowed by open-source software such as `sentinelsat` to access the Copernicus Open Access Hub (Clauss *et al.*, 2018), `pylandsat` to access the Google Cloud Landsat Public Dataset (Forget, 2019), and `asarapi` to access the ESA Online Catalog (Forget, 2018a).

Landsat scenes were acquired as Level-1 data products, thus radiometrically calibrated and orthorectified. SAR products were acquired as Ground Range Detected (GRD) Level-1 products, therefore multi-looked and projected to ground range using an Earth ellipsoid model. SAR imagery preprocessing was performed using the Sentinel Application Platform (SNAP) (ESA, 2018), including the following steps: (1) orbit files updating, (2) calibration to σ nought, (3) Range-Doppler terrain correction (Small and Shubert, 2008), (4) spatial subset over the area of interest, and (5) 2% histogram cutting.

Geographic extracts of the OSM database were acquired from the Geofabrik website[1] and imported into a local PostGIS database. Four categories of OSM objects were then extracted for each case study: (1) `highway` polylines (the road network), (2) `landuse`, `leisure`, and `natural` polygons (potential non-built-up training samples),

[1] http://www.geofabrik.de

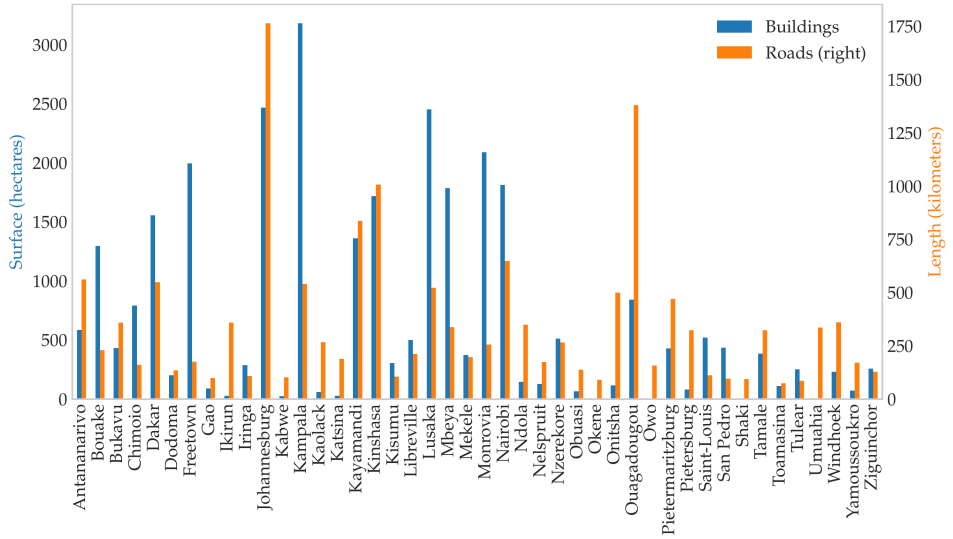

Figure 4.2: Availability of OSM building footprints (total surface in hectares) and roads (total length in kilometers) for each case study.

(3) `building` polygons (built-up training samples), and (4) `natural=water` polygons to allow the creation of a land mask. Once again, as shown in Figure 4.2, some urban areas of Sub-Saharan Africa suffer from low data availability. No buildings footprints were available in Okene, Shaki, or Owo. Likewise, only a few building footprints were available for several urban areas with more than 1 million inhabitants such as Umuahia, Katsina, Ikirun or Onitsha. On the contrary, road segments were available for all the case studies. For instance, at least 70 km of roads were extracted for small urban areas such as Toamasina, Tulear, San Pedro or Gao.

4.2.3 Definitions

Potere and Schneider (2007) stated the need for an uniform definition of what constitutes a built-up area. However, since then, no consensus has been reached in the field of urban mapping. This leads to high variations in the global measure of built-up areas. For instance, the global extent of built-up areas is estimated to be 3,524,109 km^2 according to the Global Rural Urban Urban Mapping Project (GRUMP), 308,007 km^2 in the Global Landcover 2000 (GLC00), and 774,000

km^2 in the Global Human Settlement Layer (GHSL) (Potere and Schneider, 2007; Pesaresi *et al.*, 2017).

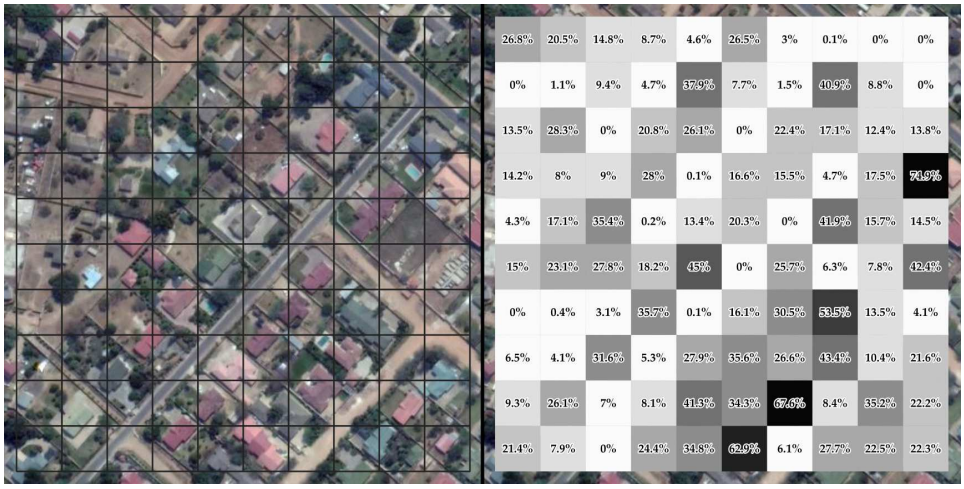

Figure 4.3: Percentage of Built-Up Areas per 30 m pixel in a suburban area of Lusaka, Zambia.

Schneider, Friedl and Potere (2010) and Mertes *et al.* (2015) defined a built-up area as a location dominated by constructed surfaces—in other words, surfaces that are covered by at least 50% of constructions, including asphalted roads. However, such a definition may lead to the omission of a large amount of built-up areas. shows the percentage of buildings in a grid of 30×30 m^2 cells, which is the spatial resolution of a Landsat image. According to the aforementioned definition, only 4 pixels out of 100 would be classified as built-up, that is 4% of the area of interest. Such a high threshold can only imply the omission of most of the suburban and rural settlements. To overcome the issue, the GHSL defined built-up as a continuous measure corresponding to the proportion of building footprint area within the total size of a cell (Pesaresi *et al.*, 2017). That would correspond to a built-up value equal to 18% over the area of interest. However, because of the complexity of the urban environment, quantifying the proportion of buildings in a given pixel is not a solved problem in urban remote sensing.

In the end, in the context of a supervised classification based on satellite imagery,

the definition of a built-up area is limited by (1) the abilities of the sensor, and (2) the training and validation samples that can be collected. In this study, the use of binary training samples from OSM (built-up or non-built-up) required a binary definition. However, to allow the detection of suburban and rural settlements, the threshold was set to 25% instead of 50%. Furthermore, because of the combined use of optical and SAR data, built-up elements only included elevated constructions—thereby excluding roads and other paved surfaces.

4.2.4 Classification of built-up and non-built-up areas

The classification of built-up areas for each case study and each data consists of : (1) the collection of both built-up and non-built-up training samples from OSM, (2) the filtering of the samples extracted from OSM for historical periods, (3) the extraction of features from both optical and SAR imagery, and (4) a pixel-based supervised classification based on the Random Forest (RF) algorithm. The implementation of the processing chain was based on Python and various scientific libraries such as NumPy (Oliphant, 2015), SciPy (Oliphant, 2007), Rasterio (Gillies, 2013), Shapely (Gillies, 2007), Scikit-Learn (Pedregosa *et al.*, 2011), or Pandas (McKinney, 2010).

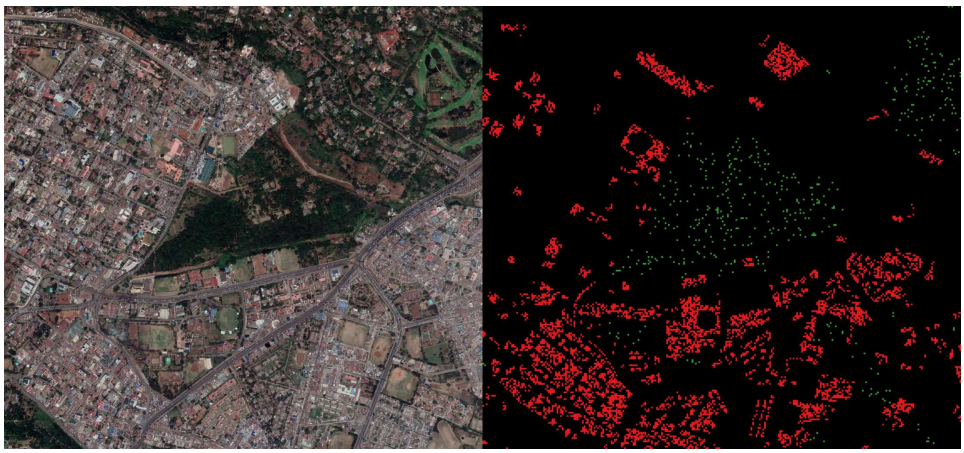

Figure 4.4: Training samples collected from OSM in Nairobi, Kenya. **a)** Aerial view of the area of interest, courtesy of Goole Earth; **b)** built-up (red) and non-built-up (green) training samples over the same area of interest.

Both built-up and non-built-up training samples were extracted from the OSM database, as proposed and detailed in a previous publication (Forget, Linard and Gilbert, 2018). Built-up training samples consisted of building footprints and urban blocks derived from the road network. Non-built-up training samples comprised a variety of `natural`, `landuse` or `leisure` features satisfying one of the following value: sand, farmland, wetland, wood, park, forest, nature reserve, golf course, greenfield, quarry, pitch, scree, meadow, orchard, grass, grassland, garden, heath, bare rock or beach. Since the availability of these features were not consistent across the case studies, additional non-built-up samples were randomly selected in areas without any building footprint or road segment in a 250 m buffer. Overall, the wide availability of roads in the OSM database allowed the collection of both built-up and non-built-up training samples in the entire set of case studies. Figure 4.4 shows a sample of the training dataset used for the classification of Nairobi, Kenya.

GLCM textures were computed with a 7×7 window size, an interpixel distance of 1 and 32 levels of quantization using the Orfeo Toolbox (Grizonnet *et al.*, 2017), according to the recommendations of previous studies (Zhang, Li and Wang, 2014; Braun and Hochschild, 2015) and results from Chapter 3. GLCMs were constructed in four direction angles (0°, 45°, 90°, and 135°) and averaged. In the case of Sentinel-1 imagery, for which two polarizations were available (VV and VH), textures were computed independently for each polarization. For each polarization, 18 textures were extracted: energy, entropy, correlation, inertia, cluster shade, cluster prominence, Harralick correlation, mean, variance, dissimilarity, sum average, sum variance, sum entropy, difference of entropies, difference of variances, information measures of correlation 1&2. That is, 36 features in the case of Sentinel-1, or 18 features in the case of ERS-1&2 or Envisat for which only one polarization was available. In order to reduce the high dimensionality of the SAR feature space, a Principal Component Analysis (PCA) was performed for each polarization. Since only the first six PCA components were retained, this reduced the feature space from 36 to 12 dimensions in the case of Sentinel-1, or from 18 to 6 dimensions in the case of ERS1&2 and Envisat.

The feature space of the classification depended on the year and the availability of

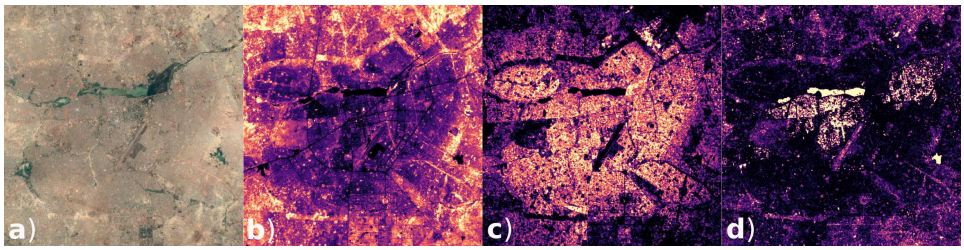

Figure 4.5: Subset of three features over the same area of interest in Ouagadougou, Burkina Faso: **a)** aerial view of the area of interest, courtesy of Google Earth, **b)** Near Infrared Landsat band, **c)** Sentinel-1 VH 7×7 GLCM Mean, **d)** Sentinel-1 VV 7×7 GLCM Dissimilarity.

satellite imagery. In 2015, the availability of Sentinel-1 allows the use of GLCM textures computed for two different polarizations, which led to a feature space of 20 dimensions (12 PCA components and 8 Landsat bands). For earlier periods, the number of dimensions was lower: 14 (6 PCA components, 8 Landsat bands), 6 in cases where only SAR imagery was available, or 8 if only Landsat was available. A subset of those features is shown in Figure 4.5 for Ouagadougou, Burkina Faso.

The classification task was performed using the RF classifier which has been shown to be effective in the classification of multisource and multimodal data (Pal, 2005; Gislason, Benediktsson and Sveinsson, 2006; Belgiu and Drăguț, 2016). The RF ensemble was constructed with 100 trees and a maximum number of features per tree equal to the square root of the total number of features—as suggested in previous studies (Gislason, Benediktsson and Sveinsson, 2006). Additionally, imbalances in the training dataset were mitigated by performing a random over-sampling of the minority class (Lemaître, Nogueira and Aridas, 2017).

In earlier periods (circa 2010, 2005, 2000, and 1995), built-up training samples extracted from the OSM database require further analysis. The OSM database does not include any information on the construction date of a building or a road segment. In a context of high urban growth, a significant amount of the extracted built-up samples are not guaranteed to be valid for earlier periods. Therefore, a filtering step was applied, by making use of a classification model fitted on the satellite imagery of 2015. This intermediary model was fitted with a simplified

feature space composed of features available in both cases. In order to reduce the influence of atmospheric and illuminations variations (Angiuli and Trianni, 2014), four spectral indices were used in place of the raw Landsat bands: the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Bareness Index (NDBal) (Zhao and Chen, 2005), the Normalized Difference Built-Up Index (NDBI) (Zha, Gao and Ni, 2003), and the Modified Normalized Difference Water Index (MNDWI) (Xu, 2006). Additionally, a set of four GLCM textures with high significance but low correlation were selected: energy, mean, dissimilarity, and cluster shade. OSM training samples and the 8 aforementioned features from 2015 were used to fit the classification model. In earlier periods, built-up training samples were filtered based on this intermediate prediction.

4.2.5 Post-processing

Three successive post-processing routines were applied on the probabilistic output of the RF classifier: (1) spatial filtering based on a mean filter, (2) scaling of the RF class probabilities to ensure multi-temporal comparability, and (3) temporal filtering.

At medium and higher spatial resolutions, pixel-based image classifications tend to produce noise—also known as the "salt and pepper" effect. This can be overcome by integrating spatial features such as GLCM textures at the classification step, or by a refinement of the classified image. Filtering-based approaches are the most widely used classification post-processing methods (Huang *et al.*, 2014). They are based on a moving window where the value of the central pixel is determined by considering the values of all pixels within it. For this study, RF class probabilities were post-processed using a simple mean filter with a 3×3 window size. This allowed a partial removal of noise, illumination artifacts, and roads.

To obtain binary built-up maps, RF class probabilities must also be thresholded. In most of the implementations, the RF classifier defaults to a value of 0.5. However, it is also dependent on the precision vs. recall trade-off required by the analysis. In the context of this study, the initial threshold was set to 0.75 in order to maximize the precision of the binary classification. Still, depending on the input sen-

sor, training data or atmospheric conditions, the optimal threshold value can vary among single-date classifications of a given case study. To overcome the issue, the optimal threshold for a given year y was defined as the threshold that maximizes the agreement with the binary classification of the year $y + 1$. In order to take into account class imbalance, the agreement was measured using the F1-Score. Then, to allow for comparability, RF class probabilities were transformed accordingly using a linear transformation.

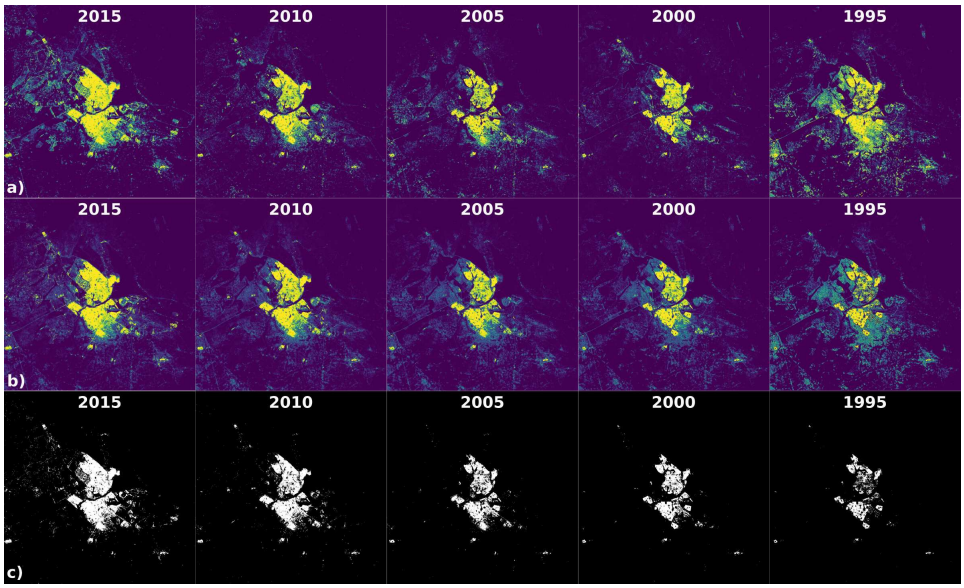

Figure 4.6: Post-processing of Ndola, Zambia. **a)** Raw RF built-up probabilities; **b)** Built-up probabilities after post-processing; **c)** Binary map (thresholded probabilities).

Moreover, single-date classifications may be characterized by high uncertainties due to sensor-specific issues, acquisition conditions, or spectral confusions. This can lead to unreasonable land cover changes that can be easily identified using a temporal consistency check (Li, Gong and Liang, 2015). Therefore, under the assumption that the transition from built-up to non-built-up is not likely in a context of urban expansion (Schneider and Mertes, 2014; Li, Gong and Liang, 2015), a temporal filtering was also applied on the RF class probabilities. In our case, we focused on two unreasonable trajectories: (1) pixels classified as built-up in a given year and as

non-built-up in the future, and (2) pixels classified as non-built-up and as built-up in the past. For the pixels concerned by one of the aforementioned state, the original value was replaced by the average probability between (1) the year of interest and the last year, or (2) the first year and the year of interest. Equation 4.1 and Equation 4.2 summarize the procedure applied in both cases:

$$Prob_t = \sum_{i=t}^{t_{max}} Prob_i \frac{1}{(t_{max} - t)} \tag{4.1}$$

$$Prob_t = \sum_{i=t_{min}}^{t} Prob_i \frac{1}{(t - t_{min})} \tag{4.2}$$

where $Prob_t$ is the modified probability, t the time step of interest, t_{min} the earliest time step, and t_{max} the latest. In practice, this allowed for a conservative filtering of the most obvious inconsistencies without over-estimating the built-up expansion dynamics.

4.2.6 Validation

The performance of the classification models was assessed using two different approaches: (1) an assessment based on independent validation samples collected from Google Earth, and (2) a K-fold cross-validation (CV) based on the training dataset extracted from OSM. Manual digitizing of samples from very high resolution imagery is work-consuming, therefore the first approach was carried out for a representative subset of the 44 case studies: Antananarivo, Bukavu, Chimoio, Dakar, Dodoma, Gao, Johannesburg, Kampala, Katsina, Kinshasa, Nairobi, Okene, Onitsha, Ouagadougou, Saint-Louis, Umuahia and Windhoek. Furthermore, samples were only collected for the periods available in the Google Earth's historical imagery catalog. To assess the accuracy of the binary built-up maps produced by the method, three metrics were computed: F1-score, precision and recall.

In parallel, K-fold CV was used to estimate the performance of all the classification models. The training dataset extracted from OSM was divided into $k = 10$ folds

of approximately equal size, with each fold being used as a validation set against the remaining $k-1$ folds. In a spatial context, spatial autocorrelation must be accounted for in order to not over-estimate the performance of the model (Brenning, 2012). Therefore, folds were not randomly produced but originated from the spatial clustering of the training samples using the K-Means algorithm. Two metrics were computed in each case: the average F1-score over the ten iterations of the CV, and its standard deviation.

4.2.7 Measuring and characterizing urban expansion

To measure the growth of built-up areas, compound annual growth rates (CAGR) were calculated according to the following equation:

$$CAGR(t_0, t_1) = (\frac{Builtup_{t_1}}{Builtup_{t_0}})^{\frac{1}{t_1 - t_0}} - 1 \tag{4.3}$$

with t_0 and t_1 being the initial and final years, and $Builtup_y$ the total surface covered by built-up areas in a given year y.

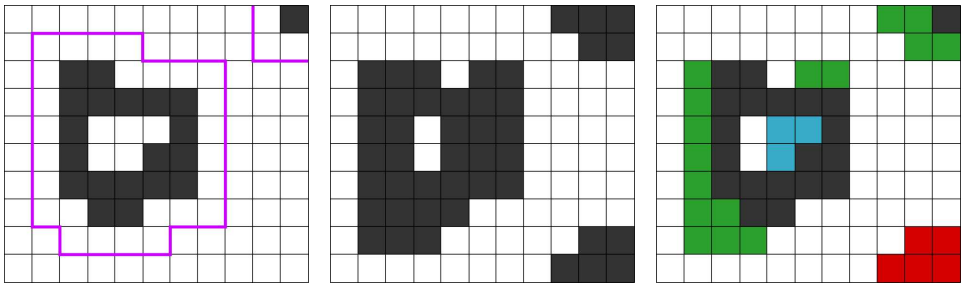

Figure 4.7: Schematic example of characterizing newly built-up areas in 200 m grid cells: **a)** built-up areas (in black) and initial urban clusters boundaries (in purple) at $t_{initial}$, **b)** built-up areas at t_{final}, and **c)** characterized newly built-up areas (existing in black, infill in blue, extension in green, leapfrog in red).

The expansion of built-up areas is a complex phenomenon that cannot be understood without analyzing it in its spatial dimension. In other words, depending on where the expansion occurs, the environmental, social, or economic consequences differ. Based on the approach proposed in the Atlas of Urban Expansion (Angel *et*

al., 2016), newly built-up areas for each period were divided into three categories: (1) *infill*, *i.e.* areas already included in an existing urban cluster, (2) *extension*, *i.e.* areas extending an existing urban cluster in a contiguous way, and (3) *leapfrog*, *i.e.* areas unattached to any existing urban cluster—an urban cluster being defined as the contiguous space that contains built-up areas less than 200 m apart. Figure 4.7 provides a schematic example of the procedure. Compound annual sprawl rates (CASR) were calculated by ignoring *infill* change, as shown in Equation 4.4.

$$CASR(t_0, t_1) = \left(\frac{Leapfrog_{(t_0,t_1)} + Extension_{(t_0,t_1)}}{Builtup_{t_0}}\right)^{\frac{1}{t_1 - t_0}} - 1 \tag{4.4}$$

Another fundamental aspect of built-up expansion is its relationship with population growth. Per-pixel population estimates for circa 2015, c. 2010, c. 2005 and c. 2000 were acquired from the Worldpop project (Worldpop, 2016). The built-up areas density for a given area of interest was defined as the ratio between the total surface of the built-up areas and the population estimates over the same area of interest. Additionally, sprawl per new dweller (SD, Equation 4.5) was computed to measure the urban extent's extension for each new dweller.

$$SD(t_0, t_1) = \frac{Leapfrog_{(t_0,t_1)} + Extension_{(t_0,t_1)}}{Pop_{t_1} - Pop_{t_0}} \tag{4.5}$$

4.3 Results and discussion

4.3.1 Assessment of the classification models

Table 4.2 presents the F1-scores obtained in the 17 case studies for which an independent validation dataset was collected. Across the 32 classifications, the average F1-score reaches 0.93 and ranges from 0.81 (Kinshasa) to 0.98 (Saint-Louis). From the entire set, Bukavu and Kinshasa reached the lowest scores. In Bukavu, the classification model appears to have been affected by the landscape—highly mountainous and densely vegetated. Most of the misclassifications occurred in high slope areas that SAR sensors confused with built-up areas. Overall, the average F1-score decreases as we go back in time: 0.94 in 2015, 0.93 in 2010, 0.92 in 2005 and 2010.

This suggests that the classification method is less efficient for past periods, due to missing satellite imagery, lower quality sensors, or less training samples. This observation is also attested by the results of the CV, with a mean score of 0.95 in 2015, 0.92 in 2005, and 0.90 in 1995. Likewise, CV standard deviation was, in average, higher in 1995 (4.8 points) than in 2015 (3.3 points).

Table 4.2: F1-scores obtained by a sample of case studies based on an independent validation dataset.

Case study	2000	2005	2010	2015
Antananarivo	.	0.88	.	0.93
Bukavu	.	0.87	0.87	0.87
Chimoio	.	0.91	.	0.95
Dakar	.	0.91	.	0.96
Dodoma	.	.	.	0.95
Gao	0.90	.	.	0.93
Johannesburg	.	0.95	.	0.95
Kampala	.	0.92	.	0.94
Katsina	0.92	.	.	0.97
Kinshasa	.	0.90	.	0.81
Nairobi	.	.	0.97	0.95
Okene	.	.	.	0.97
Onitsha	.	.	.	0.96
Ouagadougou	0.94	.	0.94	0.95
Saint-Louis	.	0.97	.	0.98
Umuahia	.	.	.	0.94
Windhoek	.	0.95	.	0.91

The comparison of the classification results with the GHSL reveals that both datasets reach a similar estimate of the total surface occupied by built-up areas across all the case studies: 6,295 km^2 against 6,662 km^2 according to the GHSL. The two datasets reach a mean agreement of 0.95. However, large differences are

observed in some case studies. Figure 4.8 shows the built-up areas maps from both datasets in Chimoio and Obuasi, where the highest variations occur. In Chimoio, the GHSL predicts 13.3 km^2 of built-up areas against 101.5 km^2. Most of the disagreement occurs in the periurban area, where the GHSL classifies low density residential areas as non-built-up. On the contrary, the GHSL estimates the built-up areas of Obuasi to reach 113.6 km^2, compared to 62.9 km^2 in MAUPP. Here, most of the disagreement occur in bare lands, which the optical sensor used by the GHSL confuses with built-up areas. Thanks to the combined use of both optical and SAR sensors, the proposed methodology appears as less sensitive to the issue.

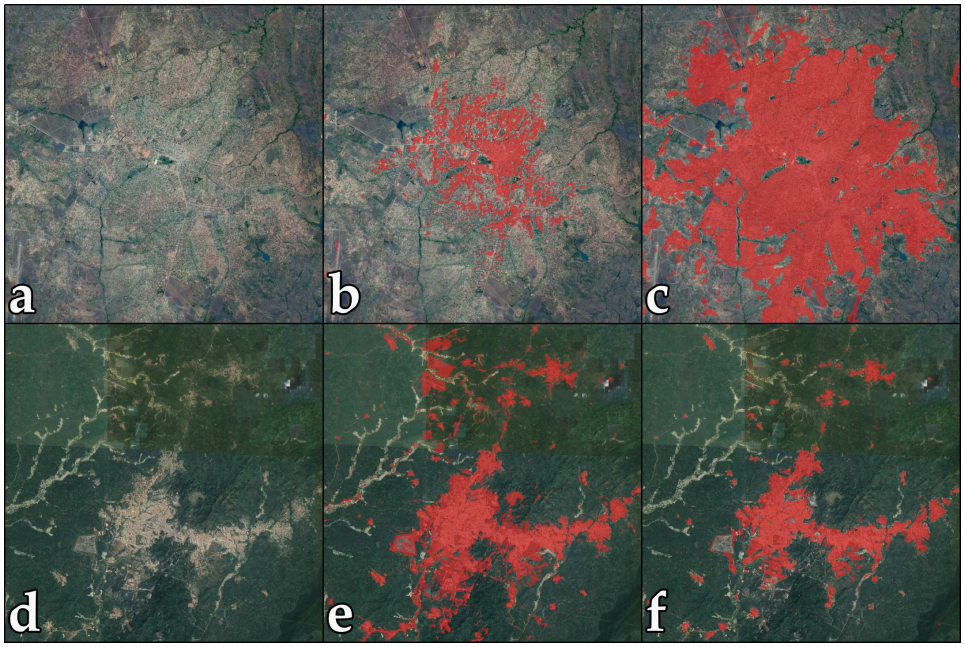

Figure 4.8: Comparison with the GHSL in Chimoio, Mozambique and Obuasi, Ghana: **a)** Aerial view of Chimoio, courtesy of Google Earth; **b)** GHSL built-up areas (in red); **c)** MAUPP built-up areas (in red); **d)** Aerial view of Obuasi, courtesy of Google Earth; **e)** GHSL built-up areas (in red); **f)** MAUPP built-up areas (in red).

4.3.2 Growth rates of built-up areas

Over all the case studies, the average compound annual growth rate (CAGR) of built-up areas between 1995 and 2015 was 5.5%. This is a significant increase compared to the estimate of 2.3% reported by the GHSL over the sames areas of interest. As previously stated, most of the disagreement between the two datasets occurs in periurban and rural areas, leading to high variations in terms of growth rates. However, the result is consistent with the one obtained in the Atlas of Urban Expansion which estimated an average CAGR of 5.1% for urban areas in Sub-Saharan Africa. Considering the population growth rates of Sub-Saharan Africa, the number appears as lower than expected. For instance, according to the Atlas of Urban Expansion, the urban areas of the United States averaged a similar CAGR (5%) between 1990 and 2000.

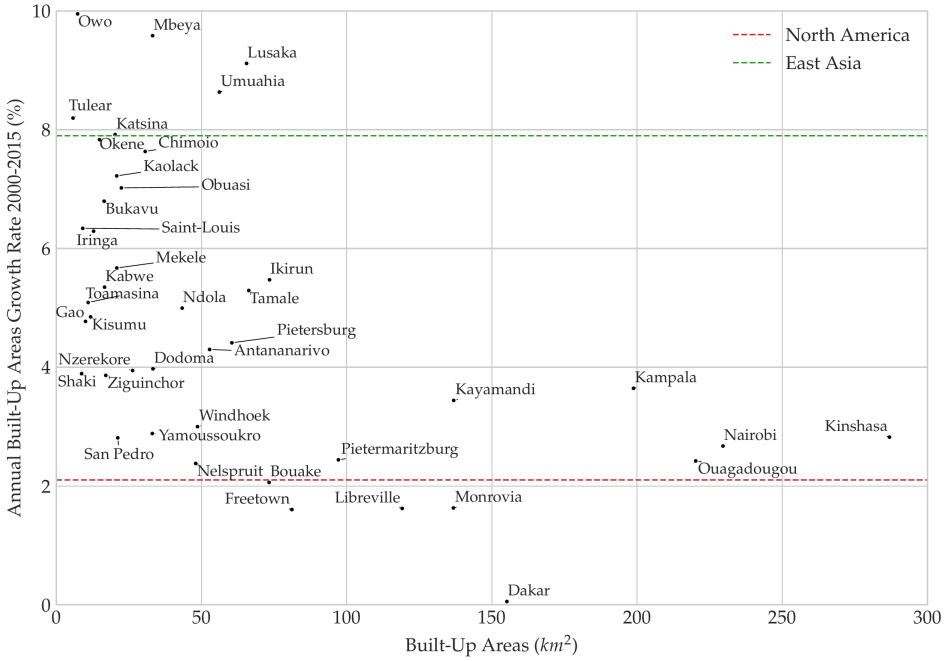

Figure 4.9: Total surface of built-up areas in 2000 and annual built-up areas growth rates between 2000 and 2015 for each case study. Johannesburg (725 km^2 of built-up areas, annual growth rate of 1.1%) and Onitsha (missing data) are excluded from the graph.

As shown in Figure 4.9, there is a relationship between the size of an urban area and its growth rates. In terms of built-up areas, the ten largest case studies were: Johannesburg, Kinshasa, Nairobi, Ouagadougou, Kampala, Dakar, Monrovia, Kayamandi, Libreville, and Pietermaritzburg—all of them have a CAGR lower than 4%. On the contrary, smaller urban areas appear to grow faster. The categorization of the case studies based on the size of their population leads to a similar conclusion: the average CAGR of built-up areas is 3.2% in large urban areas (more than 1,000,000 inh. in 2000), 4.6% in medium urban areas (between 500,000 and 1,000,000 inh.) and 5.4% in the small ones (less than 500,000 inh.). In large urban areas, the CAGRs are affected by the lack of available space, displacing the growth towards nearby peri-urban or rural areas uncovered by our areas of interest. Likewise, lower CAGRs are observed in urban areas constrained by the natural environment (Dakar, Libreville, Monrovia, Freetown).

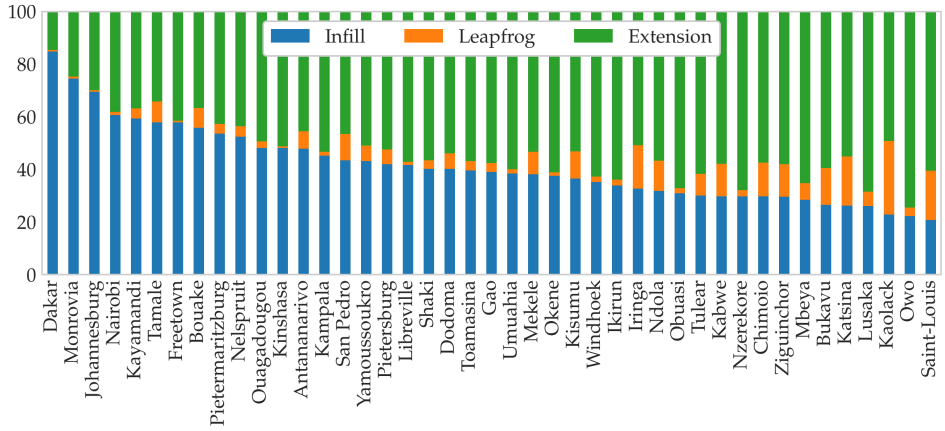

Figure 4.10: Share of each type of newly built-up area between 2000 and 2015 for each case study.

Analyzing the newly built-up areas with respect to their spatial context delivers another dimension of urban expansion. Figure 4.10 shows, for each case study, the share of each expansion type: infill, extension or leapfrog. Overall, infill expansion reaches an average share of 41%. However, large variations are observed across the case studies. In some large urban areas such as Dakar, Monrovia, Johannesburg or Nairobi, more than 60% of the growth actually occurs inside the existing urban

extent. On the contrary, in small urban areas such as Saint-Louis, Owo, or Kaolack, infill expansion makes up for less than 20% of the total growth. Data aggregation based on the size of the urban area confirms the trend: on average, the share of infill expansion reaches 54% in large urban areas, 44% in medium-sized urban areas, and 35% in the small ones. In other words, the sprawl's share (infill and extension) is higher in smaller urban areas. Defining sprawl as the combination of extension and leapfrog—or, in other words, as the newly built-up areas that increase the urban extent—allows the calculation of annual sprawl rates. In small and medium-sized urban areas, the average annual sprawl rate reaches respectively 4.1% and 3.4%, whereas it is only 2.1% in large urban areas.

4.3.3 Population densities of built-up areas

Over the case studies, the population density reaches an average of 11,031 people per km^2 of built-up area in 2015. This is consistent with the results from the Atlas of Urban Expansion, which estimated the average density of Sub-Saharan African urban areas to be 12,000 people per km^2. As expected, the population density of built-up areas is decreasing: from 16,113 people per km^2 in 2000 to 11,030 in 2015, for a CAGR of -2.5%. However, as shown in Figure 4.11, the dynamics of population density differ depending on the size of the urban area. The density of small and medium-sized urban areas is decreasing at a higher rate than in large urban areas, where density is more stable. Furthermore, the observed densities are still considerably higher than in Europe (5,000 people per km^2) or in North America (2,200 people per km^2).

At the world's scale, the built-up areas' density in a country is dependent on its income class (Angel *et al.*, 2016). Figure 4.12 shows that a similar relationship is observed across our case studies. The density of urban areas located in higher income countries—such as Johannesburg, Pietersburg, Nelspruit, Libreville, or Windhoek, do not exceed 6,000 people per km^2. On the contrary, higher densities are observed in lower income countries, especially in urban areas constrained by their natural environment (Dakar, Antananarivo, Bukavu, Kinshasa). On average, the density in low and lower-middle income urban areas is 12,295 people per km^2, whereas it is

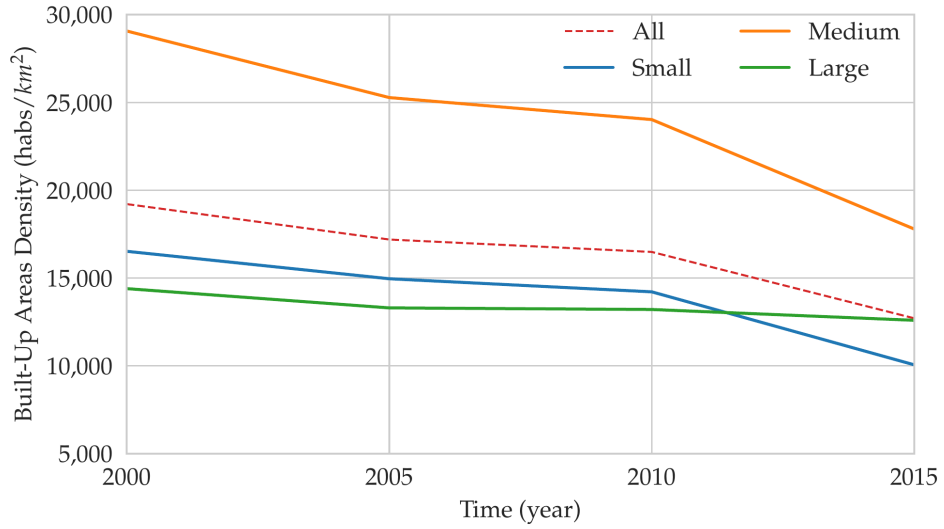

Figure 4.11: Evolution of built-up areas density (people per km^2) between 2000 and 2015 for small (less than 500,000 people in 2000), medium (between 500,000 and 1,000,000 people) and large (more than 1,000,000 people) urban areas.

only 4,348 people per km^2 in upper-middle income countries. Likewise, population size seems to affect the observed densities: large and medium-sized urban areas are, on average, 64% more dense than small urban areas.

Table 4.3 summarizes the previous observations by aggregating the sprawl per new dweller—that is, how much the urban extent is increased for each new inhabitant, with respect to population size and income class. Those numbers reveal the high heterogeneity that characterizes the dynamics of urban expansion in Sub-Saharan Africa. According to our results, the amount of sprawl for one new dweller in small urban areas of upper-middle income countries (Nelspruit, Pietersburg) is 14 times higher than in low-income large urban areas (Kinshasa, Ouagadougou, Kampala, Dakar, Antananarivo).

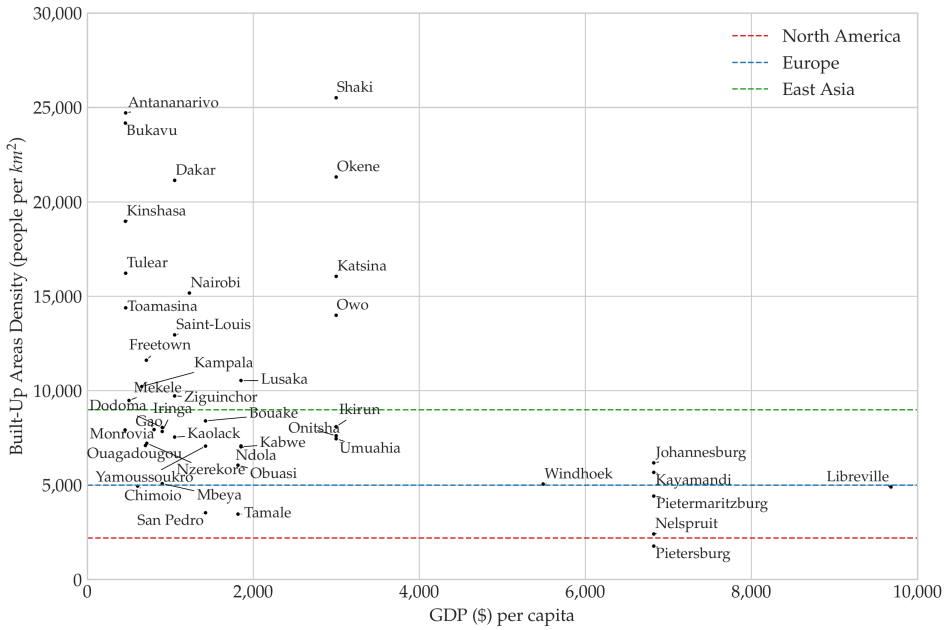

Figure 4.12: Relationship between built-up areas density (people per km^2) and GDP per capita ($ per people) of the country in 2015.

Table 4.3: Sprawl per new dweller (m^2 per people) between 2000 and 2015 depending on the size of the urban area and the country's income class (in brackets the number of case studies). Income classes are from the WorldBank's classification. Population size classes: small (less than 500,000 inh.), medium (between 500,000 and 1,000,000 inh.) and large (more than 1,000,000 inh).

Income class				
	Small-sized	Medium-sized	Large-sized	Mean
Low	141.62 (12)	33.62 (3)	37.04 (5)	99.28 (20)
Lower-Middle	157.70 (8)	92.80 (6)	89.04 (2)	124.78 (16)
Upper-Middle	522.90 (3)	153.03 (2)	107.19 (2)	298.45 (7)
Mean	196.95 (23)	87.61 (11)	64.19 (9)	

4.4 Discussion

Studying urbanization dynamics in their spatial dimension requires reliable, accurate and consistent multi-temporal maps of built-up areas. To provide insights for large-scale mapping of urban expansion, we leveraged the increasing availability of (1) open-access satellite imagery datasets from both optical and SAR sensors, and (2) crowd-sourced geographic information databases. Those decisions imply their own set of strengths and weaknesses.

Multi-sensor data fusion allowed better classification performances in arid and semi-arid regions, where moderate resolution optical sensors suffer from the spectral confusion between bare soil and built-up areas. In case studies located in an arid climate (Gao, Saint-Louis, Windhoek, and Tulear), the total surface of the detected built-up areas was 50% higher than in the GHSL. Additionally, SAR imagery allowed better data availability in tropical areas where a very low amount of optical products is available due to cloud cover. However, the use of data produced by seven different sensors (Landsat OLI, Landsat ETM+, Landsat TM, ERS-1&2, Envisat and Sentinel-1) also means varying spatial resolutions, spectral ranges, polarizations and incidence angles. Those variations are translated into inconsistencies regarding the physical objects that are detected as built-up areas.

For instance, roads were not included in our definition of a built-up area. However, the largest asphalted roads were occasionally classified as such when SAR data was not available to discriminate them based on ground texture. Moreover, previous studies have shown that the delineation of built-up areas with SAR largely depends on its spatial resolution, polarizations and incidence angles (Corbane *et al.*, 2009). To tackle the issue, the proposed methodology relies on single-date supervised classifications—in other words, one classification model for each combination of date and location. Post-processing of the resulting classifications also participates in harmonizing the multi-temporal stack of maps. Still, those inconsistencies propagates to the final results. For instance, higher annual growth rates have been observed between 2010 and 2015 than between 2000 and 2010, which suggests that built-up areas could have been either over-estimated or better detected because of the use of Sentinel-1.

Relying on OSM to collect both built-up and non-built training samples enabled the complete automation of the methodology. Thanks to the efforts of many contributors across the world, more and more building footprints are available. This allowed higher rates of detection in peri-urban areas. However, for historical periods, training samples extracted from OSM must be filtered in order to drop areas that were not constructed at the time. In the process, learning information is lost compared to circa 2015. Furthermore, ambiguous samples corresponding to rural or peri-urban settlements may be erroneously dropped in the process. Consequently, classification performances are lower as we go back in time, and low density built-up areas may suffer from higher rates of omission. Apart from the aforementioned issues, data available in OSM is not exhaustive, thereby training samples are not guaranteed to be representative of the complex urban mosaic in a given city. For instance, most of the built-up area samples in Katsina, Lusaka and Johannesburg were located in the city center and in the densest neighborhoods. This led to an under-representation of low density and vegetated suburban areas, and therefore higher rates of omissions in these areas. In the end, those issues may have induced an over-estimation of urban growth.

Despite the reported uncertainties, the results of the present study are consistent

with the findings of the Atlas of Urban Expansion (Angel *et al.*, 2016). Urbanization is not uniform across Sub-Saharan Africa, and urbanization rates reported at the regional or national scale can have little meaning at the local scale. Across the case studies, built-up areas averaged a CAGR of 4.8% between 2000 and 2015, with high variations depending on the size of the urban area: from an average of 3.2% in large urban areas to 5.4% in the smaller ones. Despite the common assumption that built-up areas grow at higher rates than the urban population (Angel *et al.*, 2016), the average CAGRs of built-up areas were not that far from the urban population annual growth rate of 4.1% reported by the UN for the region (United Nations, 2017). This suggests that the common assumption that cities are expanding their territories faster than their populations (Angel *et al.*, 2016) is not ubiquitous in Sub-Saharan Africa. In fact, in the context of this study, population densities in built-up areas did not decrease between 2000 and 2015 in urban areas such as Dakar, Freetown, Monrovia, Johannesburg, Libreville, Ouagadougou, Kinshasa, Nairobi or Kampala. Nevertheless, this is without taking into account that the ecological footprint of an urban area can be hundreds of times larger that the extent of its built-up areas (Grimm *et al.*, 2008). Likewise, the 1,600 km^2 areas of interest used in this study did not cover the whole urban areas of large cities such as Dakar, Johannesburg or Nairobi. Since most of the growth may occur in peri-urban and surrounding localities due to a lack of space, the actual growth of built-up areas may have been under-estimated in that case.

The measure of the amount of sprawl for each new dweller depending on population size and income classes revealed interesting trends. In average, the surface of sprawl per new dweller was three times higher in small urban areas (197 m^2 per new dweller) than in large ones (64 m^2). Similarly, it was three times higher in urban areas located in upper-middle income countries (299 m^2) than in low income countries (99 m^2). Nevertheless, drawing any general conclusion at the regional scale would require a larger sample size.

References

Andersson Djurfeldt, A. (2015) 'Urbanization and linkages to smallholder farming in sub-Saharan Africa: Implications for food security', *Global Food Security*, 4, pp. 1–7. doi: 10.1016/j.gfs.2014.08.002.

Angel, S. *et al.* (2016) *Atlas of Urban Expansion: The 2016 Edition, Volume 1: Areas and Densities*. Lincoln Institute of Land Policy.

Angiuli, E. and Trianni, G. (2014) 'Urban Mapping in Landsat Images Based on Normalized Difference Spectral Vector', *IEEE Geoscience and Remote Sensing Letters*, 11(3), pp. 661–665. doi: 10.1109/LGRS.2013.2274327.

Arino, O. *et al.* (2007) 'Globcover-a global land cover service with MERIS', in *Proceedings of the ENVISAT Symposium*, pp. 23–27.

Bai, X. *et al.* (2012) 'Health and wellbeing in the changing urban environment: Complex challenges, scientific responses, and the way forward', *Current Opinion in Environmental Sustainability*, 4(4), pp. 465–472. doi: 10.1016/j.cosust.2012.09.009.

Belgiu, M. and Drăguț, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, pp. 24–31. doi: 10.1016/j.isprsjprs.2016.01.011.

Bocquier, P. (2003) 'Analyzing Urbanization in Sub-Saharan Africa', in *New Forms of Urbanization*. London: Routledge, pp. 133–150.

Braun, A. and Hochschild, V. (2015) 'Combined use of SAR and optical data for environmental assessments around refugee camps in semiarid landscapes', *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7/W3, pp. 777–782. doi: 10.5194/isprsarchives-XL-7-W3-777-2015.

Brenning, A. (2012) 'Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest', in. IEEE, pp. 5372–5375. doi: 10.1109/IGARSS.2012.6352393.

Brockmann, D. and Helbing, D. (2013) 'The Hidden Geometry of Complex, Network-Driven Contagion Phenomena', *Science*, 342(6164), pp. 1337–1342. doi: 10.1126/science.1245200.

Chen, J. *et al.* (2015) 'Global land cover mapping at 30m resolution: A POK-based operational approach', *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp. 7–27. doi: 10.1016/j.isprsjprs.2014.09.002.

Clauss, K. *et al.* (2018) 'Sentinelsat'. Zenodo. doi: 10.5281/zenodo.595961.

Cohen, B. (2004) 'Urban Growth in Developing Countries: A Review of Current Trends and a Caution Regarding Existing Forecasts', *World Development*, 32(1), pp. 23–51. doi: 10.1016/j.worlddev.2003.04.008.

Corbane, C. *et al.* (2009) 'Comparative Study on the Performance of Multiparameter SAR Data for Operational Urban Areas Extraction Using Textural Features', *IEEE Geoscience and Remote Sensing Letters*, 6(4), pp. 728–732. doi: 10.1109/LGRS.2009.2024225.

Dye, C. (2008) 'Health and Urban Living', *Science*, 319(5864), pp. 766–769. doi: 10.1126/science.1150198.

ESA (2018) 'Sentinel Application Platform (SNAP)'.

Forget, Y. (2018) 'ASARapi'. Zenodo. doi: 10.5281/zenodo.1291423.

Forget, Y. (2019) 'Pylandsat'. Zenodo. doi: 10.5281/zenodo.2537105.

Forget, Y., Linard, C. and Gilbert, M. (2018) 'Supervised Classification of Built-Up Areas in Sub-Saharan African Cities Using Landsat Imagery and OpenStreetMap', *Remote Sensing*, 10(7), p. 1145. doi: 10.3390/rs10071145.

Forget, Y. *et al.* (2018) 'Complementarity Between Sentinel-1 and Landsat 8 Imagery for Built-Up Mapping in Sub-Saharan Africa', *Preprints*. doi: 10.20944/preprints201810.0695.v1.

Gamba, P. (2014) 'Image and data fusion in remote sensing of urban areas: Status issues and research trends', *International Journal of Image and Data Fusion*, 5(1), pp. 2–12. doi: 10.1080/19479832.2013.848477.

Gillies, S. (2007) 'Shapely: Manipulation and analysis of geometric objects'. toblerity.org.

Gillies, S. (2013) 'Rasterio: Geospatial raster I/O for Python programmers'. Mapbox.

Gislason, P. O., Benediktsson, J. A. and Sveinsson, J. R. (2006) 'Random Forests for land cover classification', *Pattern Recognition Letters*, 27(4), pp. 294–300. doi: 10.1016/j.patrec.2005.08.011.

Goodchild, M. F. (2007) 'Citizens as sensors: The world of volunteered geography', *GeoJournal*, 69(4), pp. 211–221. doi: 10.1007/s10708-007-9111-y.

Grimm, N. B. *et al.* (2008) 'Global Change and the Ecology of Cities', *Science*, 319(5864), pp. 756–760. doi: 10.1126/science.1150195.

Grizonnet, M. *et al.* (2017) 'Orfeo ToolBox: Open source processing of remote sensing images', *Open Geospatial Data, Software and Standards*, 2(1). doi: 10.1186/s40965-017-0031-6.

Haklay, M. and Weber, P. (2008) 'OpenStreetMap: User-Generated Street Maps', *IEEE Pervasive Computing*, 7(4), pp. 12–18. doi: 10.1109/MPRV.2008.80.

Hay, S. I. *et al.* (2005) 'Urbanization, malaria transmission and disease burden in Africa: Opinion tropical infectious diseases', *Nature Reviews Microbiology*, 3(1), pp. 81–90. doi: 10.1038/nrmicro1069.

Herold, M. *et al.* (2004) 'Spectrometry for urban area remote sensingDevelopment and analysis of a spectral library from 350 to 2400 nm', *Remote Sensing of Environment*, 91(3-4), pp. 304–319. doi: 10.1016/j.rse.2004.02.013.

Huang, X. *et al.* (2014) 'New Postprocessing Methods for Remote Sensing Image Classification: A Systematic Study', *IEEE Transactions on Geoscience and Remote Sensing*, 52(11), pp. 7140–7159. doi: 10.1109/TGRS.2014.2308192.

JPL, N. (2013) 'NASA Shuttle Radar Topography Mission Global 1 arc second', *NASA EOSDIS Land Processes DAAC*. doi: 10.5067/MEaSUREs/SRTM/SRTMGL1.003.

Kottek, M. *et al.* (2006) 'World Map of the Köppen-Geiger climate classification updated', *Meteorologische Zeitschrift*, 15(3), pp. 259–263. doi: 10.1127/0941-2948/2006/0130.

Lemaître, G., Nogueira, F. and Aridas, C. K. (2017) 'Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning', *Journal of Machine Learning Research*, 18(17), pp. 1–5.

Li, H. *et al.* (2017) 'Mapping Urban Bare Land Automatically from Landsat Imagery with a Simple Index', *Remote Sensing*, 9(12), p. 249. doi: 10.3390/rs9030249.

Li, X., Gong, P. and Liang, L. (2015) 'A 30-year (19842013) record of annual urban dynamics of Beijing City derived from Landsat data', *Remote Sensing of Environment*, 166, pp. 78–90. doi: 10.1016/j.rse.2015.06.007.

Lin, G. C. S. (2002) 'Urbanization and the Changing System of Cities in Socialist China: A Historical and Geographic Assessment', *Journal of Population Studies*, 24, pp. 89–139.

Linard, C., Tatem, A. J. and Gilbert, M. (2013) 'Modelling spatial patterns of urban growth in Africa', *Applied Geography*, 44, pp. 23–32. doi: 10.1016/j.apgeog.2013.07.009.

McGranahan, G. *et al.* (2009) *Africa's urban transition and the role of regional collaboration*. 5. London: International Institute for Environment; Development (IIED).

McKinney, W. (2010) 'Data Structures for Statistical Computing in Python', in *Proceedings of the 9th Python in Science Conference*. Austin, United States, pp. 51–56.

Mertes, C. M. *et al.* (2015) 'Detecting change in urban areas at continental scales with MODIS data', *Remote Sensing of Environment*, 158, pp. 331–347. doi: 10.1016/j.rse.2014.09.023.

Mooney, P. and Minghini, M. (2017) 'A Review of OpenStreetMap Data', in *Mapping and the Citizen Sensor*. London: Ubiquity Press, pp. 37–59. doi: 10.5334/bbf.c.

Moore, M., Gould, P. and Keary, B. S. (2003) 'Global urbanization and impact on health', *International Journal of Hygiene and Environmental Health*, 206(4-5), pp. 269–278. doi: 10.1078/1438-4639-00223.

Oliphant, T. E. (2007) 'Python for Scientific Computing', *Computing in Science & Engineering*, 9(3), pp. 10–20. doi: 10.1109/MCSE.2007.58.

Oliphant, T. E. (2015) *Guide to NumPy*. Austin, Tex.: Continuum Press.

Pal, M. (2005) 'Random forest classifier for remote sensing classification', *International Journal of Remote Sensing*, 26(1), pp. 217–222. doi: 10.1080/01431160412331269698.

Patel, N. N. *et al.* (2015) 'Multitemporal settlement and population mapping from Landsat using Google Earth Engine', *International Journal of Applied Earth Observation and Geoinformation*, 35, Part B, pp. 199–208. doi: 10.1016/j.jag.2014.09.005.

Pedregosa, F. *et al.* (2011) 'Scikit-learn: Machine learning in Python', *Journal of Machine Learning Research*, 12(Oct), pp. 2825–2830.

Pesaresi, M., Ehrlich, D. and Freire, S. (2014) 'The Global Human Settlement Layer (GHSL)New Tools and Geodatasets for Improving Disaster Risk Assessment and Crisis Management', in *Proceedings of the 11th International ISCRAM Conference*. University Park, Pennsylvania, USA.

Pesaresi, M. *et al.* (2016) 'The global human settlement layer from landsat imagery', in. IEEE, pp. 7276–7279. doi: 10.1109/IGARSS.2016.7730897.

Pesaresi, M. *et al.* (2017) *Atlas of the Human Planet 2016: Mapping Human Presence on Earth with the Global Human Settlement Layer*. Joint Research Centre, European Commission.

Potere, D. and Schneider, A. (2007) 'A critical look at representations of urban areas in global maps', *GeoJournal*, 69(1), pp. 55–80. doi: 10.1007/s10708-007-9102-z.

Potere, D. *et al.* (2009) 'Mapping urban areas on a global scale: Which of the eight maps now available is more accurate?', *International Journal of Remote Sensing*, 30(24), pp. 6531–6558. doi: 10.1080/01431160903121134.

Rahman, M. (2016) 'Detection of Land Use/Land Cover Changes and Urban Sprawl in Al-Khobar, Saudi Arabia: An Analysis of Multi-Temporal Remote Sensing Data', *ISPRS International Journal of Geo-Information*, 5(2), p. 15. doi: 10.3390/ijgi5020015.

Rasul, A. *et al.* (2018) 'Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates', *Land*, 7(3), p. 81. doi: 10.3390/land7030081.

Roy, D. P. *et al.* (2010) 'Accessing free Landsat data via the Internet: Africa's challenge', *Remote Sensing Letters*, 1(2), pp. 111–117. doi: 10.1080/01431160903486693.

Schneider, A. (2012) 'Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach', *Remote Sensing of Environment*, 124, pp. 689–704. doi: 10.1016/j.rse.2012.06.006.

Schneider, A., Friedl, M. A. and Potere, D. (2009) 'A new map of global urban extent from MODIS satellite data', *Environmental Research Letters*, 4(4), p. 044003. doi: 10.1088/1748-9326/4/4/044003.

Schneider, A., Friedl, M. A. and Potere, D. (2010) 'Mapping global urban areas using MODIS 500-m data: New methods and datasets based on "urban ecoregions"', *Remote Sensing of Environment*, 114(8), pp. 1733–1746. doi: 10.1016/j.rse.2010.03.003.

Schneider, A. and Mertes, C. M. (2014) 'Expansion and growth in Chinese cities, 19782010', *Environmental Research Letters*, 9(2), p. 024008. doi: 10.1088/1748-9326/9/2/024008.

Schultz, M. *et al.* (2017) 'Open land cover from OpenStreetMap and remote sensing', *International Journal of Applied Earth Observation and Geoinformation*, 63, pp. 206–213. doi: 10.1016/j.jag.2017.07.014.

Seto, K. C. *et al.* (2011) 'A Meta-Analysis of Global Urban Land Expansion', *PLoS ONE*. Edited by J. A. Añel, 6(8), p. e23777. doi: 10.1371/journal.pone.0023777.

Small, C. (2005) 'A global analysis of urban reflectance', *International Journal of Remote Sensing*, 26(4), pp. 661–681. doi: 10.1080/01431160310001654950.

Small, D. and Shubert, A. (2008) 'Guide to ASAR Geocoding', *RSL-ASAR-GC-AD*, (1.01).

Tatem, A. J. *et al.* (2008) 'Human population, urban settlement patterns and their impact on Plasmodium falciparum malaria endemicity', *Malaria Journal*, 7(1), p. 218. doi: 10.1186/1475-2875-7-218.

Tatem, A. J. *et al.* (2007) 'High resolution population maps for low income nations: Combining land cover and census in East Africa', *PLoS One*, 2(12), p. e1298. doi: 10.1371/journal.pone.0001298.

Tupin, F. (2010) 'Fusion of Optical and SAR Images', in Soergel, U. (ed.) *Radar remote sensing of urban areas*. Dordrecht ; New York: Springer (Remote sensing and digital image processing, 15), pp. 133–159.

United Nations (2015) 'World Urbanization Prospects: The 2014 Revision'.

United Nations (2017) *World Population Prospects: The 2017 Revision*. New York.

United Nations (2018) *World Urbanization Prospects 2018*. New York.

Wang, P. *et al.* (2017) 'Human Built-up And Settlement Extent (HBASE) Dataset From Landsat'. NASA Socioeconomic Data; Applications Center (SEDAC).

Wentz, E. *et al.* (2014) 'Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing', *Remote Sensing*, 6(5), pp. 3879–3905. doi: 10.3390/rs6053879.

Worldpop (2016) 'Africa 1km Population', *University of Southampton*. doi: 10.5258/SOTON/WP00004.

Xu, H. (2006) 'Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery', *International Journal of Remote Sensing*, 27(14), pp. 3025–3033. doi: 10.1080/01431160600589179.

Yang, D. *et al.* (2017) 'Open land-use map: A regional land-use mapping strategy for incorporating OpenStreetMap with earth observations', *Geo-spatial Information Science*, 20(3), pp. 269–281. doi: 10.1080/10095020.2017.1371385.

Zha, Y., Gao, J. and Ni, S. (2003) 'Use of normalized difference built-up index in automatically mapping urban areas from TM imagery', *International Journal of Remote Sensing*, 24(3), pp. 583–594. doi: 10.1080/01431160304987.

Zhang, C., Chen, Y. and Lu, D. (2015) 'Mapping the land-cover distribution in arid and semiarid urban landscapes with Landsat Thematic Mapper imagery', *International Journal of Remote Sensing*, 36(17), pp. 4483–4500. doi: 10.1080/01431161.2015.1084552.

Zhang, J., Li, P. and Wang, J. (2014) 'Urban Built-Up Area Extraction from Landsat TM/ETM+ Images Using Spectral Information and Multivariate Texture', *Remote Sensing*, 6(8), pp. 7339–7359. doi: 10.3390/rs6087339.

Zhang, Y., Zhang, H. and Lin, H. (2014) 'Improving the impervious surface estimation with combined use of optical and SAR remote sensing images', *Remote Sensing of Environment*, 141, pp. 155–167. doi: 10.1016/j.rse.2013.10.028.

Zhao, H. and Chen, X. (2005) 'Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+', in. IEEE, pp. 1666–1668. doi: 10.1109/IGARSS.2005.1526319.

Chapter 5

Conclusion

In the future, urbanization will primarily take place in data-scarce regions. In other words, accurate and up-to-date data are the least available where they matter the most, thereby limiting our understanding of the urbanization process and impairing decision making. In this context, EO offers the unique opportunity of observing urban growth from space at a relatively low cost. Yet, EO-based urban maps also suffer from lower accuracies in low- and middle-income countries. In SSA, the lower quality of EO-based products may be explained by multiple factors, such as: the lack of reference datasets to support model training and calibration, the lower availability of historical satellite imagery, the high cloud cover over tropical regions, or the spectral confusion between built-up and bare soil areas in arid regions.

On the other hand, the growth of open data is starting to mitigate the issue of data scarcity. In the past, EO catalogs were biased toward the coverage of high-income countries. Now, they proceed to a systematic and global acquisition strategy (Roy *et al.*, 2010). Likewise, VGI was scarce in low- and middle-income countries, because contributors focused on where they lived (Coleman, Georgiadou and Labonte, 2009). Today, the digital divide is reducing, and humanitarian mapathons help closing the gap (Mooney and Minghini, 2017). As a result, spatial data are increasingly accessible and accurate. In this thesis, we leveraged this opportunity to improve urban mapping in SSA. To that end, we developed, implemented and assessed a map-

ping approach that integrates multi-source open-access satellite imagery (Landsat, Sentinel-1, Envisat, ERS) and VGI (OpenStreetMap). The proposed approach has been applied in 46 urban areas of SSA to produce a dataset of urban expansion covering the 1995–2015 period. In the following sections, we summarize the main findings of the thesis, discuss their limitations and draw perspectives for future work.

5.1 Main findings

5.1.1 Mining training labels from OSM

In supervised learning, the quality and representativeness of the training samples are crucial to reach high classification performances. Local studies generally rely on in situ measurements to collect ground truth data, or on the manual digitization of very high resolution imagery—for instance from Google Earth. But manual digitization is time consuming and becomes unsustainable when producing global or regional maps. In order to automate the process, global mapping methods usually extract training samples from external databases provided by institutional or commercial agencies. In SSA, the lack of reference datasets prevents this type of approach. Today, thanks to a growing collaborative effort, VGI databases are increasingly reliable—including in low- and middle-income countries. In this thesis, we proposed to take advantage of the OSM database to collect training data, as formulated by the 1st hypothesis:

> *In the data-scarce context of SSA, VGI can enable large-scale mapping of built-up areas by supporting the training of the classification models without reducing their performances.*

Our results suggest that such approach is reliable. In Chapter 2, we compared two strategies for training data collection: by digitizing samples from Google Earth, and by extracting information from OSM. Built-up (positive) training samples were derived from the building footprints available in the OSM database. Likewise, non-built-up (negative) training samples were obtained by rasterization of various land use, leisure, and natural objects—such as parks, gardens, forests, or farmland. How-

ever, the objects were not available in all case studies. If the OSM database is increasingly complete, small urban areas suffered from a lower data availability. To complete the training dataset, we made use of the road network, which is the most exhaustive feature in the OSM database. Additional positive samples were collected by extracting urban blocks, *i.e.* the polygons formed by the intersections of the road network. Likewise, supplementary negative samples were randomly picked from the most remote areas. Across the 10 case studies, the OSM-based classification models averaged a F1-score of 0.93, whereas the manual approach reached 0.92. These results suggest that OSM can be a reliable data source for training data collection, especially in the data-scarce context of SSA where few alternatives are available. More importantly, the quality of OSM-based training data will improve as the database grows.

5.1.2 Optical-SAR synergies

Relying on optical imagery to detect built-up areas may suffer from several flaws. First, optical sensors are highly sensitive to atmospheric conditions; this can be problematic in regions characterized by high cloud cover such as Middle Africa. Second, distinct objects may share a similar spectral signature at medium resolution (10–40 m). For instance, spectral confusion may occur between a densely vegetated urban block and a forest, or between built-up and bare soil areas. In contrast, SAR sensors are independent from illumination and atmospheric variations. Moreover, they are sensitive to surface roughness, and can easily distinguish between plane land covers (bare soil) and elevated constructions (built-up)—regardless of the spectral characteristics of the construction materials. In this thesis, we implemented and assessed an approach that takes advantage of both optical and SAR imagery in order to better discriminate built-up areas from their surrounding environment, as proposed by the 2nd hypothesis:

> *Using SAR sensors (ERS, Envisat, Sentinel-1) in combination with multispectral data (Landsat) can lead to a more consistent detection of the built environment, especially in arid regions.*

In Chapter 3, we assessed the complementarity of two specific sensors: Landsat 8

and Sentinel-1. The results obtained confirmed the high level of complementarity between optical and SAR imagery. Still, the "speckle noise" inherent to SAR images was an obstacle for pixel-level classification. The extraction of textural features from SAR backscattering mitigated the issue and preserved most of the information related to the roughness of the terrain. Several classification schemes with varying input features (optical only, SAR only, different polarizations) were assessed in 12 urban areas of SSA. In 11 case studies out of 12, multi-source classifiers outperformed the single-source ones. Overall, single-source optical-based classifiers underperformed in distinguishing built-up from bare soil areas. On the contrary, single-source SAR-based classifications suffered from a confusion between dense vegetation and elevated constructions. But combining both data sources led to high accuracies in both land covers. These results suggest that complementing optical data with SAR increases the performance of the classification models when it comes to the detection of built-up areas. Interestingly, this complementarity can even be leveraged with a simple pixel-level classification approach—provided that GLCM textures are computed beforehand.

5.1.3 Scaling up and going back in time

The two aforementioned methodological contributions were integrated to produce a multi-temporal dataset of urban expansion. In Chapter 2 and 3, implementations were limited in space (~10 case studies) and in time (circa 2015). In Chapter 4, we scaled up the analysis to 46 urban areas and applied it for five distinct years: 1995, 2000, 2005, 2010 and 2015. In doing so, the objective was twofold: (1) assessing the scalability of the proposed mapping approach, and (2) analyzing urban dynamics in SSA based on a representative sample of cities.

Scaling up raised issues related to data availability. First, historical periods (before 2015) lacked coverage by EO catalogs. Full coverage (availability of both SAR and optical data) was reached in only 64% of the cases. This caused fallbacks to single-source classification schemes which have been shown to be less effective. Second, OSM-based training labels are not guaranteed to be valid in the past—especially in the rapidly growing cities of SSA. To mitigate the issue, labels were filtered based

on satellite data, leading to reductions of the training dataset and, thereby, lower classification performances.

Independent validation samples were collected for 17 distinct case studies. Overall, the average F1-score reached 0.93 with a standard deviation of 0.04. But variations where observed depending on the year. In the earliest years, the lower availability of both satellite and training data decreased the classification performances. Furthermore, in post-classification change detection, the accuracy of the observation may be as low as the product of each map accuracy (Singh, 1989). That is, considering an average F1-score of 0.93, $0.93^2 \simeq 0.87$ with two data points, $0.93^3 \simeq 0.80$ with three, or $0.93^4 \simeq 0.75$ with four. Consequently, despite the availability of five data points in most case studies (*i.e.* 1995, 2000, 2005, 2010 and 2015), urban dynamics were mainly analyzed using two data points in order to minimize uncertainty.

In total, we relied on satellite imagery acquired from seven different sensors: Landsat 5, Landsat 7, Landsat 8, ERS-1, ERS-2, Envisat and Sentinel-1, each of them with various spatial resolutions, spectral characteristics, polarizations, and incidence angles. In other words, depending on data availability, built-up areas were characterized by different spectral properties. In the proposed approach, the issue was mitigated by fitting a distinct classification model for each year and case study. Still, inconsistencies were observed in the temporal dimension. For instance, the availability of Sentinel-1 data led to a better detection of small isolated settlements, along with an increase in the misclassification rate. In the end, sensor differences propagate and influence the statistics derived from the multi-temporal maps. In the context of this study, this translated to an over-estimation of urban expansion between 2010 (without Sentinel-1) and 2015 (with Sentinel-1).

5.1.4 The heterogeneous dynamics of urban expansion in SSA

Across the 46 case studies, the CAGR of built-up areas was 5.5%. This is consistent with the estimates of the Atlas of Urban Expansion (5.1%). However, this is significantly higher than what the GHSL estimated: 2.3% over the same areas of interest. Most of the disagreement between both datasets occurred in peri-urban areas, where the GHSL was less tolerant on what constitutes a built-up area. Re-

gardless of accuracy differences, this suggests that definitions can have a significant impact on the interpretation of urban growth. On the other hand, such growth rates are consistent with the estimations provided by the Atlas of Urban Expansion for the urban areas of SSA (5.1%). Nevertheless, aggregated averages makes little sense in the context of SSA, where urban areas are highly heterogeneous. High variations were observed across the case studies depending on population sizes. For instance, the CAGR of built-up areas between 2000 and 2015 was significantly higher in small urban areas (less than 500,000 people; 5.4% on average) than in the large ones (more than 1,000,000 people; 3.2%).

Even more heterogeneity was observed when interpreting the growth of built-up areas with respect to population dynamics. Across the case studies, population densities decreased between 2000 and 2015: from ~16,000 to ~11,000 people per km^2 of built-up area. But the variations were high depending on population sizes and income classes. On average, large and medium urban areas were 64% more dense than small urban areas. Likewise, a lower income was associated with higher population densities. In 2015, the population densities in upper-middle income countries averaged ~4,300 people per km^2 of built-up area, against ~12,300 in low and lower-middle income countries. These observations confirmed the findings of previous studies, according to which the population densities in built-up areas are highly influenced by the income class (Angel *et al.*, 2016).

Measuring the sprawl per new dweller—that is, the expansion of the urban extent for one new inhabitant—summarizes the previous findings. Indeed, over the 2000–2015 period, sprawl per new dweller largely depends on population sizes and income classes. On average, the extents of small urban areas increased by 197 m^2 per new dweller, against 64 m^2 in the large ones. Likewise, sprawl per new dweller was three times higher in urban areas located in upper-middle income countries (298 m^2) than in low income countries (99 m^2). In other words, population increase is not a sufficient predictor of how much a city may grow, as the relationship between population growth and urban expansion differ according to demographic and economic variables.

5.2 Perspectives

Altogether, the proposed approach can be generalized into an EO-based, open and automated mapping framework, as illustrated in Figure 5.1. This framework makes use of two types of input data: (1) open-access satellite imagery, multi-sensor or not, and (2) crowd-sourced geographic information—for instance OSM. The objective is to extrapolate discrete human knowledge (2) to a continuous space based on sensor-based measurements (1).

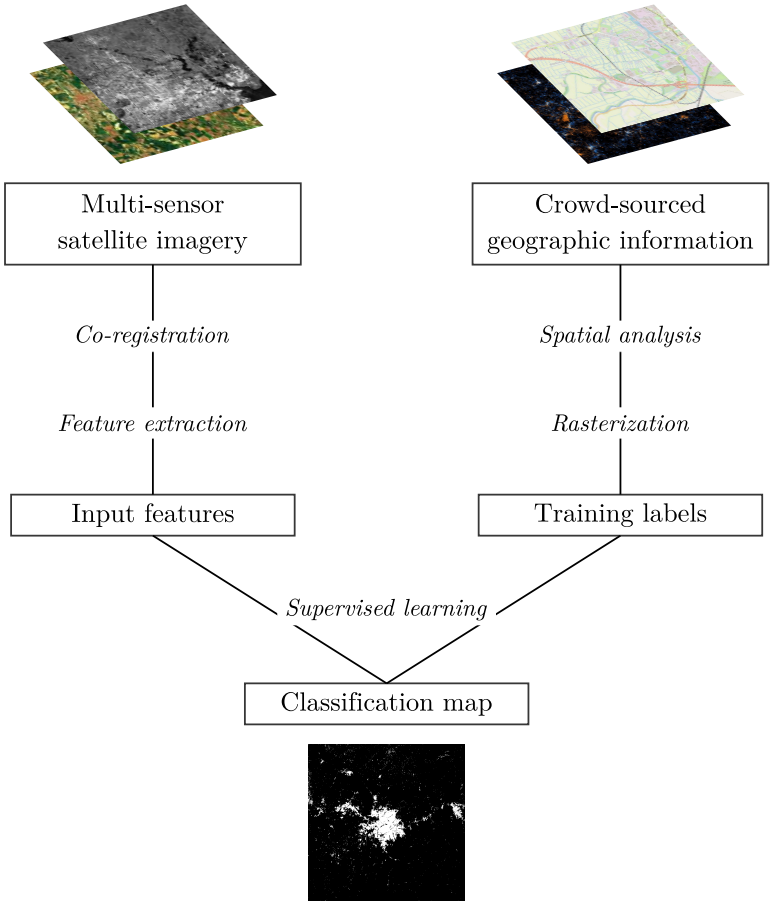

Figure 5.1: A framework for the classification of built-up areas from EO data and crowd-sourced geographic information.

Satellite imagery provides *unlabeled*, *continuous*, *exhaustive*, and *sensor-based* mea-

surements of the Earth's surface, radiance, reflectance, or backscattered intensity. The process of transforming raw input data into features that maximize the performance of machine learning algorithms is called *feature engineering*. In the proposed framework, it may refer to preprocessing routines (orthorectification, atmospheric correction, cloud detection, speckle noise removal), feature extraction (spectral indices, GLCM textures), and coregistration (spatial alignment of multi-temporal or multi-sensor images).

Contrary to satellite imagery, crowd-sourced geographic databases offer *labeled*, *discrete*, *incomplete*, and *human-based* land cover information, roads, buildings, parks, forests, etc. Most of the times, volunteered geographic information is distributed as vector data; the process of transforming vector objects into a raster grid is called *rasterization*. Depending on the mapping objectives, input data may require additional *spatial analyses*, for instance to extract urban blocks from the road network or to filter objects based on their properties (*e.g.* their OSM tags, their sizes).

Finally, the statistical relationships between sensor-based measurements (satellite imagery) and the target geographic objects (built-up and non-built-up areas) are determined through *supervised learning* in order to produce a classification map. Various machine learning algorithms may be used to that end; but the choice of algorithm must take into account the low signal-to-noise ratio caused by the use of crowd-sourced geographic databases to fetch the training labels.

Improvements to the proposed framework can refer to (1) extracting features from multi-source EO data, (2) mining training labels from crowd-sourced geographic databases, or (3) classifying input data through supervised learning. In the following subsections, related perspectives are explored and discussed.

5.2.1 Feature engineering and data fusion

> *At the end of the day, some machine learning projects succeed and some fail. What makes the difference? Easily the most important factor is the features used.* (Domingos, 2012)

The growing availability of satellite imagery from various sensors makes data fusion

increasingly crucial in urban remote sensing. But mixing optical and SAR data at the pixel level generates a large range of issues. First, coregistration errors (*i.e.* misaligned pixels) may reduce the performance of the classification models. Novel methods for automatic coregistration of multi-sensor images are still being developed (Costantini *et al.*, 2018), and should increase the accuracies of multi-sensor land cover classifications. Second, SAR backscattering is less suited than optical reflectance to pixel-level classification, namely because of speckle noise. The issue can be alleviated by computing texture features, which integrate statistical information of the neighboring pixels. In this thesis, GLCM textures were used to that end. Alternative methods are still being developed and may be more suited to the detection of human settlements (Cheng and Li, 2005; Pesaresi, Gerhardinger and Kayitakire, 2008; Zhang, Li and Wang, 2014; Safia and He, 2015). Finally, switching from a *pixel-level* to a *feature-level* fusion scheme could mitigate most of the aforementioned issues (Gamba, 2014), provided that progresses are made regarding automatic image segmentation and object-based image analysis at moderate spatial resolution.

EO studies are data-driven; regardless of the methodological advances, they primarily depend on the availability of satellite imagery. In other words, progresses in the field are driven by the improvements of satellite sensors—mainly in terms of spatial, spectral and temporal resolution. Compared to Landsat, the Sentinel program provides higher resolution optical imagery: 10 m for the visible and near-infrared bands, against 30 m with Landsat. This is a nine-fold increase in the amount of information. In urban areas, where objects of interest may be relatively small (buildings, road segments, gardens, pools), the increase in spatial resolutions leads to a lower amount of mixed pixels. Recent studies already showed the benefits of Sentinel-2 imagery to map built-up areas, such as a lower omission rate for isolated settlements and a better separation from artificial open spaces, gravel or sand (Martino Pesaresi and Christina Corbane *et al.*, 2016).

Still, no satellite will be able to go back in time. To map the historical dynamics of land cover changes, we will continue to rely on past EO programs. In this case, we cannot expect spatial and spectral resolutions to increase. In other words, the only

way to improve the mapping of historical changes is to develop better classification methods. Today, the main advantages at our disposal are the lower computation and acquisition costs, along with a set of algorithms taking advantage of it—machine learning. However, applications are limited by the uneven availability of historical satellite data across the globe. For instance, novel methods based on dense time-series (Schneider, 2012) cannot always be applied in low-income regions because of lower data availability. Likewise, in multi-sensor classification, data scarcity forces a higher temporal gap between image pairs, which can confuse classification models in rapidly changing environments. In this regard, initiatives such as the Landsat Global Archive Consolidation (Wulder *et al.*, 2016), which recovered more than 4 million historical images, are crucial.

5.2.2 The growth of crowd-sourced geographic information

> *More data beats clever algorithms.* (Halevy, Norvig and Pereira, 2009)
>
> *[…] but better data beats more data.* (Rogati, 2012)

The proposed mapping framework heavily relies on the availability of crowd-sourced geographic data, both in terms of quantity and quality. Numerous studies showed the benefits of a larger training dataset with machine learning algorithms, especially in low signal-to-noise contexts (Mellor *et al.*, 2015). In this thesis, OSM was used to collect both built-up and non-built-up training samples. A large range of geographic objects were extracted to that end (building footprints, road network, land use, land cover); and their availability largely influenced the performance of the classification models. Today, the database is growing fast. In 2017, the road network in OSM was estimated to be more than 80% complete at the global scale (Barrington-Leigh and Millard-Ball, 2017), which is more than any authoritative or commercial dataset. Between 2014—the beginning of this thesis—and 2019, the size of the database for Africa increased by a factor of six. The increase in contributions along with mapping campaigns focusing on low-income countries make OSM increasingly reliable for the collection of training samples, and the performance of the proposed approach should improve with the growth of the database.

What is even more interesting in the growth of OSM is the increasing semantic completeness. In other words, the range of geographic objects which are mapped is increasingly large. Their attributes—or *tags*—are also more accurate, thanks to initiatives that harmonize the classification system (Ali *et al.*, 2016). Spatial information related to natural land covers (forests, grasslands, soils) and farmlands is more and more available, including in low-income regions. In the context of the proposed framework, this could enable a switch from a binary (built-up vs. non-built-up) to a multi-class land cover classification. Likewise, spatial analysis of social and mobile data can provide additional information on land use and population dynamics (Deville *et al.*, 2014), which could be leveraged to collect training or testing samples.

5.2.3 Novel methods in machine learning

In machine learning, it is commonly assumed that "*more and better data beat cleverer algorithms*". Provided that they are properly parametrized and that the training dataset is representative enough, most supervised algorithms are able to produce accurate predictions (Li *et al.*, 2014; Maxwell, Warner and Fang, 2018). In this thesis, the Random Forest algorithm was selected because of its computing efficiency and its resistance to training data noise. Still, as most of machine learning algorithms, it is sensitive to class imbalance in the training dataset. To alleviate the issue, random under-sampling and over-sampling strategies were adopted. But more sophisticated methods are being developed to limit the loss of information or the introduction of biases (Lemaître, Nogueira and Aridas, 2017).

Progress could also be made regarding the classification products. More and more applications could benefit from continuous outputs, *i.e.* built-up proportion or density at the pixel scale. For instance, population mapping could make use of building height and density metrics. In the GHSL, a built-up area is expressed as "*a continuous value representing the proportion of building footprint area within a cell*" (Pesaresi *et al.*, 2017), but output products only provide categorical values. The GUF alleviates the issue thanks to a higher spatial resolution—from 30 to 12 m. The high resolution binary product is then derived into secondary, coarser products of density and

settlement patterns (Esch *et al.*, 2018). However, the lower spatial resolution of historical satellite imagery limits the application of such approach. Physical modeling approaches, such as spectral mixture analysis (Adams and Gillespie, 2006), provide insights to quantify the proportion of land covers at the subpixel scale. But their integration into multi-sensor classification schemes has not been investigated as of now. Once again, another limit is the availability of training and validation data; binary labels are scarce, continuous labels are even more so.

Finally, the potential contribution of deep learning to large-scaled land cover mapping is still to be determined. In the deep learning terminology, the process of assigning a label to each pixel of an image is called *semantic segmentation*. In the recent years, deep learning have been widely applied to remotely-sensed data, mainly on the basis of commercial high resolution imagery (Zhang, Zhang and Du, 2016; Zhu *et al.*, 2017; Ma *et al.*, 2019). However, the fine structural information found in high resolution images is lacking in open-access medium resolution imagery (Sentinel, Landsat), thereby limiting semantic segmentation approaches. At the classification level, country-scale applications did not show any gain over more traditional machine learning algorithms such as Random Forest (Inglada, Vincent and Thierion, 2019), while computing costs are dramatically increased. Moreover, deep learning models are less interpretable than more traditional methods such as tree-based machine learning or generalized linear models. As of now, in regards to land cover classification from medium-resolution imagery, deep learning appears to mostly benefit image preprocessing and feature engineering. For instance, promising results were obtained in cloud detection (Li *et al.*, 2019), image fusion (Liu *et al.*, 2018), image registration (Wang *et al.*, 2018), or resolution enhancement (Lanaras *et al.*, 2018).

5.2.4 Last words: towards more inclusive urban maps

By nature, satistical learning tends to highlight the norm and dismiss the outliers. In urban remote sensing, densely populated city centers and traditionnal building blocks have always been characterized by high accuracies. The availability of quality reference datasets and the lower occurence of complex mixed pixels make them

easier to detect with supervised learning classification models. However, today, the main challenges towards urban sustainability are urban sprawl and urban poverty (United Nations, 2018b)–two phenomena taking place in the marginal spaces of our cities: peri-urban areas and informal settlements. They are also the areas suffering from the lowest accuracies in built-up maps. Sparsed and isolated settlements are even harder to detect from space—especially when going back in time because of the lower spatial resolution of older sensors. But they are potential nodes of future urban networks. In other words, despite the high accuracies in existing global products or in the maps produced in the present thesis, reliable data are still lacking for a better understanding of long-term temporal dynamics of urbanization.

The main objective of this thesis was to provide insights and propose novel methods towards more inclusive urban maps, in order to better integrate these marginal spaces in the analyses of the recent urbanization dynamics. The use of OpenStreetMap is central to this approach: thanks to initiatives such as HOTOSM[1] which focus on undermapped communities, informal and isolated settlements are included as training samples and cease to appear as outliers in classification models.

The lower computing costs and the opening of historical satellite imagery catalogs are an opportunity for the remote sensing community: we can go back in time and map the past again, with better reference data, higher computing power and modern tools. This thesis does not solve the aforementioned issues, but I hope the proposed methods will support further research in this direction.

References

Adams, J. B. and Gillespie, A. R. (2006) *Remote sensing of landscapes with spectral images: A physical modeling approach*. Cambridge ; New York: Cambridge University Press.

Ali, A. *et al.* (2016) 'Guided Classification System for Conceptual Overlapping Classes in OpenStreetMap', *ISPRS International Journal of Geo-Information*, 5(6), p. 87. doi: 10.3390/ijgi5060087.

Angel, S. *et al.* (2016) *Atlas of Urban Expansion: The 2016 Edition, Volume 1: Areas and Densities*.

[1] https://www.hotosm.org

Lincoln Institute of Land Policy.

Barrington-Leigh, C. and Millard-Ball, A. (2017) 'The world's user-generated road map is more than 80% complete', *PLOS ONE*. Edited by M. Ali, 12(8), p. e0180698. doi: 10.1371/journal.pone.0180698.

Cheng, T. and Li, P. (2005) 'Multivariate variogram-based multichannel image texture for image classification', in *Geoscience and Remote Sensing Symposium, 2005. IGARSS'05. Proceedings. 2005 IEEE International*. IEEE, pp. 3830–3832. doi: 10.1109/IGARSS.2005.1525744.

Coleman, D. J., Georgiadou, Y. and Labonte, J. (2009) 'Volunteered Geographic Information: The nature and motivation of produsers', *International Journal of Spatial Data Infrastructures Research*, 4, p. 27. doi: 10.2902/1725-0463.2009.04.art16.

Costantini, M. *et al.* (2018) 'Automatic Coregistration of SAR and Optical Images Exploiting Complementary Geometry and Mutual Information', in *2018 IEEE International Geoscience and Remote Sensing Symposium*. Valencia: IEEE, pp. 8877–8880. doi: 10.1109/IGARSS.2018.8519242.

Deville, P. *et al.* (2014) 'Dynamic population mapping using mobile phone data', *Proceedings of the National Academy of Sciences*, 111(45), pp. 15888–15893. doi: 10.1073/pnas.1408439111.

Domingos, P. (2012) 'A few useful things to know about machine learning', *Communications of the ACM*, 55(10), p. 78. doi: 10.1145/2347736.2347755.

Esch, T. *et al.* (2018) 'Where We LiveA Summary of the Achievements and Planned Evolution of the Global Urban Footprint', *Remote Sensing*, 10(6), p. 895. doi: 10.3390/rs10060895.

Gamba, P. (2014) 'Image and data fusion in remote sensing of urban areas: Status issues and research trends', *International Journal of Image and Data Fusion*, 5(1), pp. 2–12. doi: 10.1080/19479832.2013.848477.

Halevy, A., Norvig, P. and Pereira, F. (2009) 'The Unreasonable Effectiveness of Data', *IEEE Intelligent Systems*, 24(2), pp. 8–12. doi: 10.1109/MIS.2009.36.

Inglada, J., Vincent, A. and Thierion, V. (2019) 'Mapping France's land-cover at 10 m every year. Lessons learned and future improvements.', in *International Geoscience and Remote Sensing Symposium (IGARSS)*. Yokohama, Japan.

Lanaras, C. *et al.* (2018) 'Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network', *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, pp. 305–319. doi: 10.1016/j.isprsjprs.2018.09.018.

Lemaître, G., Nogueira, F. and Aridas, C. K. (2017) 'Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning', *Journal of Machine Learning Research*, 18(17), pp. 1–5.

Li, C. *et al.* (2014) 'Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery', *Remote Sensing*, 6(2), pp. 964–983. doi: 10.3390/rs6020964.

Li, Z. *et al.* (2019) 'Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors', *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, pp. 197–212. doi: 10.1016/j.isprsjprs.2019.02.017.

Liu, Y. *et al.* (2018) 'Deep learning for pixel-level image fusion: Recent advances and future prospects', *Information Fusion*, 42, pp. 158–173. doi: 10.1016/j.inffus.2017.10.007.

Ma, L. *et al.* (2019) 'Deep learning in remote sensing applications: A meta-analysis and review', *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, pp. 166–177. doi: 10.1016/j.isprsjprs.2019.04.015.

Maxwell, A. E., Warner, T. A. and Fang, F. (2018) 'Implementation of machine-learning classification in remote sensing: An applied review', *International Journal of Remote Sensing*, 39(9), pp. 2784–2817. doi: 10.1080/01431161.2018.1433343.

Mellor, A. *et al.* (2015) 'Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin', *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, pp. 155–168. doi: 10.1016/j.isprsjprs.2015.03.014.

Mooney, P. and Minghini, M. (2017) 'A Review of OpenStreetMap Data', in *Mapping and the Citizen Sensor*. London: Ubiquity Press, pp. 37–59. doi: 10.5334/bbf.c.

Pesaresi, M. *et al.* (2016) 'Assessment of the Added-Value of Sentinel-2 for Detecting Built-up Areas', *Remote Sensing*, 8(4), p. 299. doi: 10.3390/rs8040299.

Pesaresi, M., Gerhardinger, A. and Kayitakire, F. (2008) 'A Robust Built-Up Area Presence Index by Anisotropic Rotation-Invariant Textural Measure', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1(3), pp. 180–192. doi: 10.1109/JSTARS.2008.2002869.

Pesaresi, M. *et al.* (2017) *Atlas of the Human Planet 2016: Mapping Human Presence on Earth with the Global Human Settlement Layer*. Joint Research Centre, European Commission.

Rogati, M. (2012) 'The Model and the Train Wreck: A Training Data How-to'. Santa Clara, California, United States.

Roy, D. P. *et al.* (2010) 'Accessing free Landsat data via the Internet: Africa's challenge', *Remote Sensing Letters*, 1(2), pp. 111–117. doi: 10.1080/01431160903486693.

Safia, A. and He, D.-C. (2015) 'Multiband compact texture unit descriptor for intra-band and inter-band texture analysis', *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, pp. 169–185. doi: 10.1016/j.isprsjprs.2015.04.003.

Schneider, A. (2012) 'Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach', *Remote Sensing of Environment*, 124, pp. 689–704. doi: 10.1016/j.rse.2012.06.006.

Singh, A. (1989) 'Review Article Digital change detection techniques using remotely-sensed data', *International Journal of Remote Sensing*, 10(6), pp. 989–1003. doi: 10.1080/01431168908903939.

United Nations (2018) *World Urbanization Prospects: The 2018 Revision, Methodology*. Working Paper No. ESA/P/WP.252. New York.

Wang, S. *et al.* (2018) 'A deep learning framework for remote sensing image registration', *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, pp. 148–164. doi: 10.1016/j.isprsjprs.2017.12.012.

Wulder, M. A. *et al.* (2016) 'The global Landsat archive: Status, consolidation, and direction', *Remote Sensing of Environment*, 185, pp. 271–283. doi: 10.1016/j.rse.2015.11.032.

Zhang, J., Li, P. and Wang, J. (2014) 'Urban Built-Up Area Extraction from Landsat TM/ETM+ Images Using Spectral Information and Multivariate Texture', *Remote Sensing*, 6(8), pp. 7339–7359. doi: 10.3390/rs6087339.

Zhang, L., Zhang, L. and Du, B. (2016) 'Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art', *IEEE Geoscience and Remote Sensing Magazine*, 4(2), pp. 22–40. doi: 10.1109/MGRS.2016.2540798.

Zhu, X. X. *et al.* (2017) 'Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources', *IEEE Geoscience and Remote Sensing Magazine*, 5(4), pp. 8–36. doi: 10.1109/MGRS.2017.2762307.