Springer Geography

Courage Kamusoko

Optical and SAR Remote Sensing of Urban Areas

A Practical Guide

MOREMEDIA

Springer

Springer Geography

The Springer Geography series seeks to publish a broad portfolio of scientific books, aiming at researchers, students, and everyone interested in geographical research.

The series includes peer-reviewed monographs, edited volumes, textbooks, and conference proceedings. It covers the major topics in geography and geographical sciences including, but not limited to; Economic Geography, Landscape and Urban Planning, Urban Geography, Physical Geography and Environmental Geography.

Springer Geography—now indexed in Scopus

More information about this series at https://link.springer.com/bookseries/10180

Courage Kamusoko

Optical and SAR Remote Sensing of Urban Areas

A Practical Guide

Springer

Courage Kamusoko
Machida, Tokyo, Japan

ISSN 2194-315X ISSN 2194-3168 (electronic)
Springer Geography
ISBN 978-981-16-5148-9 ISBN 978-981-16-5149-6 (eBook)
https://doi.org/10.1007/978-981-16-5149-6

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To my mother, who taught me that consistency is the key to progress!!

Preface

Cost-effective urban land cover mapping methods are required to produce up-to-date and accurate geospatial information for the United Nations Sustainable Development Goals (SDGs) and the United Nations New Urban Agenda (NUA). The increasing volume of Earth Observing (EO) data provides opportunities to map land cover in urban areas. However, conventional image processing and analysis using only desktop computers is quite difficult given the huge volumes of EO data. Therefore, open source software applications that include both cloud and desk computing are required in order to perform cost-effective remotely sensed image processing and classification.

The workbook is designed as a practitioner's guide and a critical resource for students, researchers, and others who are interested in applying optical and SAR data in order to improve land cover mapping in urban areas. Therefore, the workbook is intended to be hands-on, where users can explore data and methods to improve land cover mapping in urban areas. Although there are many freely available EO data, I focus on land cover mapping using Sentinel-1 C-band SAR and Sentinel-2 data from the Copernicus program under the European Space Agency (ESA). All remotely sensed image processing and classification procedures are based on open source software applications such QGIS and R as well as cloud-based platforms such as Google Earth Engine (GEE). Previous experience with open source software applications and GEE is recommended but not necessary. However, experience with GIS and remote sensing as well as basic statistical analysis is required to complete the laboratory exercises.

How is this Workbook Organized?

This workbook is organized into six chapters. Chapter 1 introduces geospatial machine learning, which is subdivided into five sections. Section 1.1 presents a brief introduction on geospatial machine learning, while Sect. 1.2 describes the study area and data. Section 1.3 provides preliminary image processing in Google Earth Engine (GEE). Finally, Sects. 1.4 and 1.5 provide the summary and additional exercises.

Chapter 2 covers exploratory image analysis and transformation. Section 2.1 provides a brief background on exploratory image analysis and image transformation. Next, Sect. 2.2 provides laboratory exercises in QGIS and R (i.e., preparing training data, creating spectral plots, computing spectral, and texture indices). Finally, Sects. 2.3 and 2.4 provide the summary and additional exercises.

Chapter 3 focuses on mapping urban land cover using multi-seasonal Sentinel-2 imagery. Section 3.1 provides a brief background. Next, Sect. 3.2 provides laboratory exercises 1 and 2, which are done in R. The final Sects. 3.3 and 3.4 provide the summary and additional exercises.

Chapter 4 focuses on mapping urban land cover using multi-seasonal Sentinel-1 imagery. Section 4.1 provides a brief background, while Sect. 4.2 provides laboratory exercises 1 and 2. Finally, Sects. 4.3 and 4.4 provide the summary and additional exercises.

Chapter 5 focuses on mapping urban land cover using multi-seasonal Sentinel-1 and Sentinel-2 imagery as well as other derived data such as spectral and texture indices. Section 5.1 provides a brief background, while Sect. 5.2 provides laboratory exercises 1 and 2. Finally, Sects. 5.3 and 5.4 provide the summary and additional exercises.

The final chapter 6 focuses on land cover classification accuracy assessment. Section 6.1 presents an overview on land cover classification accuracy assessment. Next, Sect. 6.2 provides accuracy assessment exercises. Finally, Sects. 6.3 and 6.4 provide the summary and additional exercises. An attempt has been made to organize the workbook in a general sequence of topics. Therefore, I encourage you to read the workbook in sequence from Chapter 1.

Conventions Used in this Workbook

R and Java commands or scripts are written in *Ubuntu Mono* font size 10 in italics, while the output is written in Ubuntu Mono font size 10. Note that long output from R and Java code is omitted from the workbook to save space. In some cases, I use small font sizes in Ubuntu Mono to show how the output or results would appear. This is just for illustration purposes. Readers will of course see the whole output when they execute the commands. The hash sign (# or //) at the start of a line of code indicates that it is a comment. Finally, all explanations are written in Liberation Serif font size 12.

Data and Online Resources

Boundary and training area data (shapefiles) used in this workbook are available. Furthermore, I provide additional online resources on R, QGIS, and GEE or remote sensing in the appendix.

Machida, Japan — Courage Kamusoko

Contents

1 Geospatial Machine Learning in Urban Environments: Challenges and Prospects

Abstract

Accurate and current land cover information is required to develop strategies for sustainable development and to improve quality of life in urban areas. The past decades has seen an increased availability of earth observation satellite (EOS) sensors (e.g., Sentinel-1 and Sentinel-2) as well as machine learning (ML) techniques (support vector machines, random forests) for land cover mapping. While significant progress has made to improve land cover mapping in urban areas, challenges still remain. The purpose of this chapter is to discuss briefly about geospatial machine learning in urban environments as well as some of its major challenges and prospects. The chapter will cover an introduction to geospatial ML (remote sensing image pre-processing and ML techniques), study area and data sets, hands-on exercises, summary, and additional exercises.

Keywords

Earth observation satellite sensors • Sentinel-1 • Sentinel-2 • Machine learning • Land cover • Urban areas

1.1 Introduction

1.1.1 Background

According to the United Nations, global urban population increased from 751 million in 1950 to 4.2 billion in 2018 (United Nations 2018). In the next decade, the global urban population is expected to increase to 60% (United Nations 2020; World Health Organization 2020). About 90% of urban growth will occur in less developed regions, such as East Asia, South Asia, and sub-Saharan Africa (UNHabitat 2020). To date, rapid urbanization has resulted in the increase of informal settlements and unplanned urban sprawl, especially in sub-Saharan Africa (Seto et al. 2011; United Nations 2018; UNHabitat 2020). As a result, most local government authorities fail to provide adequate infrastructure (transport and health facilities) and basic services such as clean water and sanitation (UNHabitat 2008). Furthermore, citizens living in densely populated and informal settlements are vulnerable to the outbreak of epidemics and global pandemics such as COVID-19 (Zerbo et al. 2020). Given the unplanned nature of informal settlements, overcrowding makes it difficult to maintain high hygiene standards as well as follow recommended measures such as social distancing and self-isolation (World Health Organization 2020; UNHabitat 2020). Therefore, practical policies and actions are needed to make urban environments inclusive, safe, and resilient as well as improve the quality of life urban dwellers. This requires accurate and timely geospatial information, which is critical for

Electronic supplementary material
The online version of this chapter (https://doi.org/10.1007/978-981-16-5149-6_1) contains supplementary material, which is available to authorized users.

C. Kamusoko, *Optical and SAR Remote Sensing of Urban Areas*, Springer Geography,
https://doi.org/10.1007/978-981-16-5149-6_1

planning and implementing sustainable urban development in light of the 2030 Agenda for Sustainable Development (United Nations 2019).

Geospatial or mapping agencies in less developed countries recognize the importance of geospatial information for sustainable urban planning and development. However, efforts to produce new or update geospatial information (e.g., large-scale topographic maps and land cover maps) have been constrained by poor funding as well as the high cost of acquiring data using conventional land surveys and aerial photography (Conitz 2000). It is noteworthy that most government policy makers do not prioritize investing in geospatial technology and information despite its contribution to spatial urban planning in particular, and sustainable development in general (UNHabitat 2008). Consequently, geospatial or mapping agencies in the less developed countries fail to produce timely, reliable, and accurate geospatial information. Furthermore, official development assistance (ODA) funding for major mapping projects has declined over the past decades (Kamusoko et al. 2021).

Recently, urban land cover mapping at a regional scale has increased in some less developed regions (Seto et al. 2011). This is because medium-resolution satellite remotely sensed data such as the Landsat series and Sentinel-1 and Sentinel-2 have relatively good global coverage and are available free of charge. Furthermore, advancement in machine learning methods such as random forests (Rodriguez-Galiano et al. 2012), support vector machines (Nemmour and Chibani 2006; Pal and Mather 2003) and deep learning (Yu et al. 2017; Ma et al. 2019) has also increased urban land cover mapping applications. However, most of the urban land cover mapping studies in the less developed regions have been done in capital cities (Mundia and Aniya 2005; Gamanya et al. 2009; Griffiths et al. 2010; Forkuor and Cofie 2011). This means that most cities or urban centers are still poorly quantified because mapping urban land cover at a national scale still remains difficult despite the availability of free satellite imagery and advanced ML techniques (Goldblatt et al. 2018). For example, high cost of collecting reliable reference data sets (training or validation) from field surveys and high spatial resolution imagery inhibits land cover mapping at a national scale. In addition, spectral confusion and mixed pixel problems still persist given the heterogeneous nature of the urban landscapes and the fragmented spatial configuration of small cities (Stefanov et al. 2001; Xian and Crane 2005). In most African urban cities, spectral confusion is a major problem because gravel (dirt) roads in informal settlements have similar spectral responses to those of bare vacant plots and croplands (Kamusoko et al. 2013). Nonetheless, recent technological advancement such as the availability of very high-resolution satellite imagery on Google Earth, cloud computing (e.g., Google Earth Engine) as well as the availability of microwave satellite imagery (Sentinel-1) are encouraging. This can be used to improve land cover mapping in urban areas.

1.1.2 Cloud and Desktop Open Source Software Applications

Earth observing satellites are increasing due to investments in large satellite missions and development of small satellite missions such as CubeSat. As a result, there is an increase in the volume of Earth Observing (EO) data that can be used to develop environmental applications (e.g., mapping and monitoring urban areas). However, conventional image processing and analysis using desktop computers is quite difficult given the huge volume of EO data. Therefore, open source software applications that include both cloud and desk computing are required in order to perform cost-effective remotely sensed image processing and classification. In this workbook, remotely sensed image processing and classification procedures are based on open source software applications such as QGIS and R as well as on Google Earth Engine (GEE) platform. QGIS is an open source desktop GIS application that provides data viewing, editing, and analysis capabilities. R is a free software programming language and software environment for statistical computing and graphics (R Development Core Team 2005). GEE is a cloud-based platform for planetary-scale geospatial analysis with massive computational capabilities (Gorelick et al. 2017).

In this workbook, we are going to focus on image processing and geospatial machine techniques. In this context, geospatial machine learning refers to the use of remotely sensed imagery and machine learning classifiers to map land cover in urban areas. This will also include a detailed accuracy assessment based on recommendations by Olofsson et al. (2014). The idea is to provide a hands-on approach in order to improve urban land cover mapping based on multi-seasonal Sentinel-1 and Sentinel-2 imagery, derived variables (spectral and texture indices), and machine learning classifiers such as random forests. An attempt is made to focus on common problems related to land cover classification such as spectral confusion, machine learning model performance, and accuracy assessment.