

Urban Sprawl in Nairobi, Kenya

Geospatial Machine Learning in Remote Sensing

Yujin Song & Lu Yee Wong

April 2025

Start

01

Introduction

Area of Interest

About Urban Sprawl

Use Case

Exploratory Analysis

02

Project Methodology

Data Collection

Labeling

Model Training

03

Initial Findings

Finish

04

Next Steps

01 Introduction



- Nairobi is the capital and largest city in Kenya
- Rapid urban growth
 - Population increasing from 137,000 in 1950 to 5.7 million in 2025 (Marcotrends)

Map of Kenya (CDC)



Image of Nairobi's Downtown (Brookings)

- Nairobi is the capital and largest city in Kenya
- Rapid urban growth
 - Population increasing from 137,000 in 1950 to 5.7 million in 2025 (Marcotrends)



Yujin Song, Lu Yui Wong

What is Urban Sprawl?

Rapid expansion driven by population growth and desire for space but linked to increased energy use, pollution, traffic congestion and loss of vegetation



The Royal Nairobi Golf Course, next to Kibera slum, one of the most famous slums in Africa

Use Case

Motivation:

Traditional monitoring methods, such as census and survey data, are often outdated and lack spatial detail

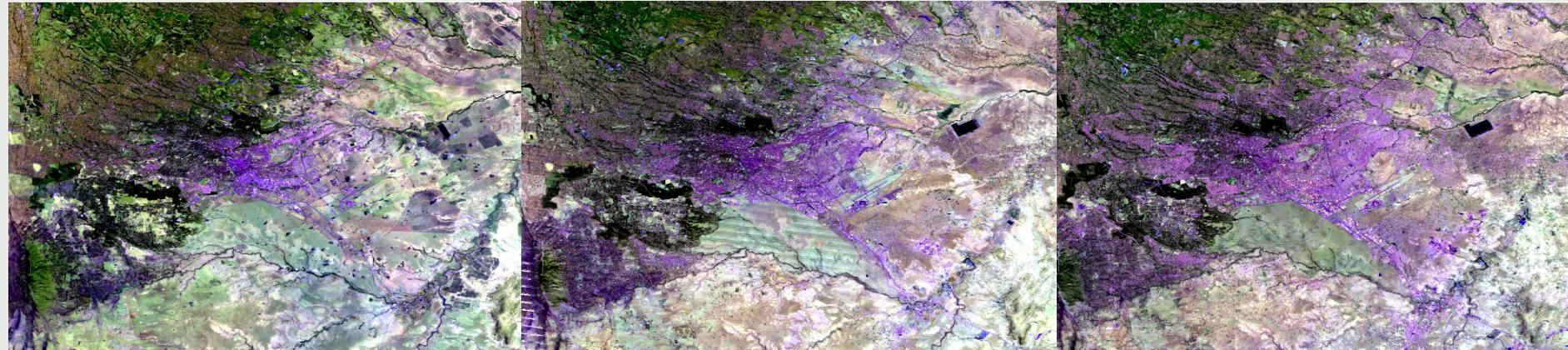
Target Users:

Nairobi stakeholders and departments responsible for urban development

Use Case:

Improved resources to inform approaches to address urban issues such as public health and water management

Exploratory Analysis



1984

2009

2019

02 Project Methodology

Pulling Data

- Acquired Landsat images at 5-year intervals (1984–2023)
- Selected three key years for training—each matched to a Landsat sensor generation

Landsat Sensor	Training Year	Applied to Years
Landsat 5	1984	1994
Landsat 7	2009	1999, 2004
Landsat 8	2019	2014, 2023

Year	Landsat	Dataset ID
1984	Landsat 5	LANDSAT/LT05/C02/T1_L2
1994*	Landsat 5	LANDSAT/LT05/C02/T1_L2
1999*	Landsat 7	LANDSAT/LT05/C02/T1_L2
2004	Landsat 7	LANDSAT/LE07/C02/T1_L2
2009	Landsat 7	LANDSAT/LE07/C02/T1_L2
2014	Landsat 8 OLI	LANDSAT/LC08/C02/T1_L2
2019	Landsat 8 OLI	LANDSAT/LC08/C02/T1_L2
2023	Landsat 8 OLI	LANDSAT/LC08/C02/T1_L2

*cloud cover up to 30% applied

Labeling Data

01

Labeling for Landsat Versions

- Landsat 5: 1984
- Landsat 7: 2009
- Landsat 8 OLI: 2019

02

Labeling for Performance

- Labeled a total of 2,700 points:
- 900 per image
 - 300 per class per image

03

Labeling for Relevant Classes

- Three Classes:
- Vegetation
 - Informal
 - Formal

Labeling Data

04

Validating with Ground Truths

- Knowledge-based Validation
- Remote Validation

Informal Settlement	Approximate Establishment	Key Growth Period
Kibera	1904s	1960s-1990s
Mathare	1920s-1930s	1950s-1970s
Korogocho	1970s	1980s-1990s
Mukuru Kwa Njenga	1980s	1990s-2000s

Labeling Data

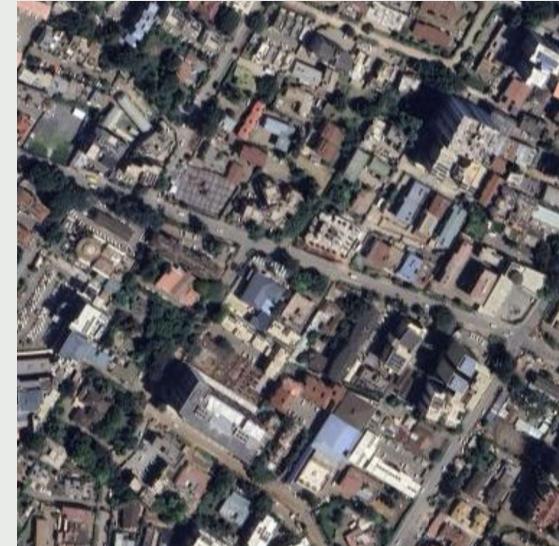
04

Validating with Ground Truths

- Knowledge-based Validation
- Remote Validation



Informal Settlement -
Kibera



Formal Settlement -
Westlands

Labeling Data

04

Validating with Ground Truths

- Knowledge-based Validation
- Remote Validation



Informal Settlement -
Kibera



Formal Settlement -
Westlands

Training Model

Image Classification

Classifier: Random Forest Model

Training Data: Selected Land Cover Classes

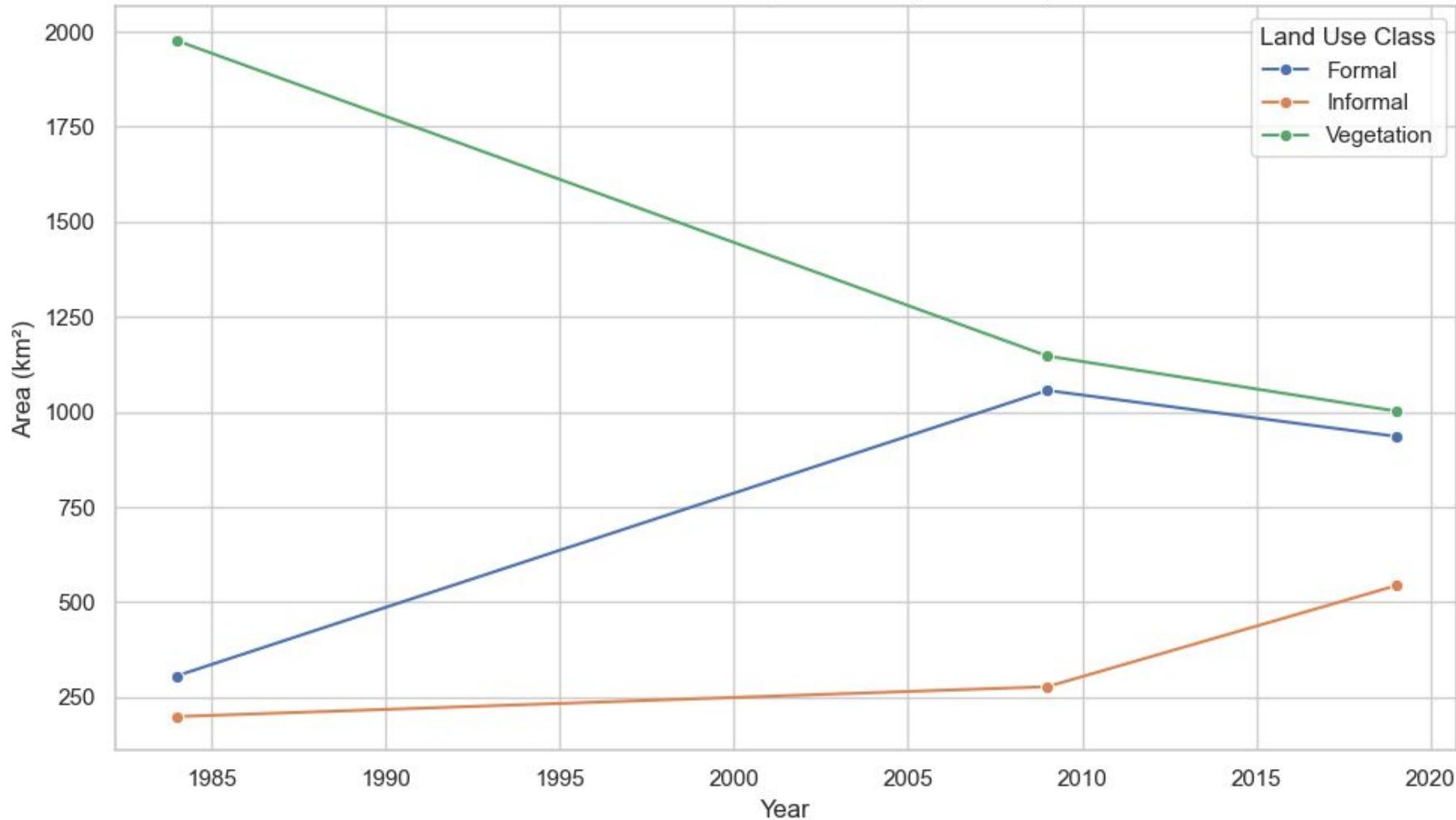
Process:

- Feature extraction using spectral bands
- Model Training using labeled years

Land Cover Change: Quantified class transition across years

03 Initial Findings

Urban Land Use Change in Nairobi (1984–2019)



04 Next Steps

Validation & Model Comparison

1. Validate the Model

- Evaluate our selected -Random Forest Model using accuracy, precision, recall, and F1 score
- Use confusion matrix and compare predictions to ground truth (1984, 2009, 2019)

2. Explore Alternative Models

- Test SVM and k-NN on same datasets for comparison
- Compare models using same metrics