UNIVERSITY OF CAPE TOWN
GEOMATICS DIVISION
SCHOOL OF ARCHITECTURE, PLANNING AND GEOMATICS
ENGINEERING AND THE BUILT ENVIRONMENT

Geospatial Analysis of Informal Settlement Development in Cape Town

By Toufeeq Fisher
FSHTOU001

Supervised by
Mr. Kaveer Singh

Plagiarism Declaration

I, Toufeeq Fisher, declare that this report titled, **Geospatial analysis of informal settlement development in Cape Town**, and the work represented in it is my own. I declare that:

- I know that plagiarism is wrong. Plagiarism is the use of another's work and pretending that it is one's own.
- I have used the Harvard styled referencing and citations. All contributions from others have been referenced and cited.
- This report is my own.
- I have not allowed and will not allow anyone to copy my work with the intention of passing it off as their own work.

Signature: Signed by candidate

Date: 31 May 2023

ACKNOWLEDGEMENTS

Firstly, and most importantly, I would like to acknowledge and thank Mr Kaveer Singh, who has shown unwavering dedication, extreme patience, and support through the process of conducting this project as a supervisor and mentor. The level of commitment he has given me throughout certain periods means a great deal. His invaluable insight into the topic has also helped shape my thinking into solving the problems outlined within this project. For all this and even more, I will forever be thankful.

Next, I'd like to thank the Geomatics department and the University of Cape Town for providing the necessary opportunity, funding, and any resources I needed to complete this project.

I'd also like to acknowledge my mother, Faldelah Fisher, who has constantly supported and motivated me through every challenge I have ever faced growing up and even now. My thanks are extended to the rest of my family, including my brother Abdul Naeem Fisher, and my grandmother, Moerida Adams.

Lastly, I'd like to thank my friends and colleagues without whom I wouldn't have the spirit to constantly smile throughout the process of finishing this project. Special thanks go out to my colleagues, Mr. Brent Kotzee and Mrs Suleiga Schroder.

ABSTRACT

A significant factor in the growth of cities is the development of informal settlements. Informal settlements are associated with negative socio-economic factors such as unemployment and a lack of secure land tenure. Over one billion people live in such settlements all around the world, and consequently, informal settlers are exposed to the effects of negative socio-economic factors. This research focuses on understanding how informal settlements develop within the City of Cape Town using spatial metrics. By understanding the development, informed steps can be taken to improve the quality-of-life informal dwellers are exposed to. The development of three informal settlements was monitored: Imizamo Yethu, Langa and Siqalo. Initially, machine learning techniques were used to determine the current development, complexity, and compactness of informal housing within settlements. High-resolution imagery was used to classify shacks in the targeted informal settlements. An accuracy assessment was conducted to validate any subsequent analysis that was completed from classified imagery. The overall accuracies ranged between 88-96%. Thereafter, change detection analysis would be used to understand how each informal settlement developed and would be compared to each other. Using the combination of change detection, linear regression, and ordinary least squares analysis across the selected informal settlements, results from this study showed that the major development characteristic was the densification of shacks. This densification followed along major formal and external transport routes, as well as informal and internal transport networks. Densification was also heavily driven by open space. There were also individual and unique internal development dynamics in each informal settlement. These were driven by slopes, employment opportunities, and neighbouring income areas. The most statistically significant factor that influenced development across all the informal settlements was open space. This was determined through ordinary least squares.

Contents

Table of Figures

List of Tables

List of Formulae

ACRONYMS

GIS	Geographic Information System
MR	Multi-resolution
OOIA	Object-oriented Image Analysis
OOIC	Object-oriented Image Classification
OLS	Ordinary Least Squares
RGB	Red, Green, and Blue

1. Introduction

1.1. Background and Motivation

In modern-day South Africa, socio-economic issues provide a host of challenges. One of these challenges is the proliferation and development of informal settlements (Wekesa et al, 2010). As a form of housing, it has been well documented that informal settlements are inadequate in their capacity to provide basic services to the people residing in them (Onyekachi, 2014). To minimize the negative effects of this form of housing, it is vital to understand why informal settlements establish and developed (Hofmann et al, 2017; Kohli et al, 2016).

The Department of Human Settlements has a mandate that aims to alleviate the conditions informal settlements are subjected to (Huchzermeyer, 2003). This mandate is realized through section 26 of the Constitution of the Republic of South Africa, 1996:

> "26. Housing. – (1). Everyone has the access to adequate housing. (2) The state must take reasonable legislative and other measures, within its available resources, to achieve the progressive realization of this right."

The Department of Human Settlements continues to state that it is the state's responsibility to provide sustainable and affordable housing to poorer communities and individuals. People residing in informal settlements are classed as having a low economic status (Manal et al, 1998).

It has been established that informal settlements and their residents are facing tough socio-economic related issues. It has also been established that there is a responsibility to alleviate the issues associated to informal settlements (Hofmann et al, 2017; UN-Habitat, 2020; WCISSF, 2016). The aforementioned serve as a motivation to determine how and why informal settlements establish and develop. The results of such a project may prove useful in mitigating the negative effects associated with informal settlements.

1.2. Research Problem

Informal settlements in the City of Cape Town are affected by a host of socio-economic problems that lead to uncontrolled informal development. Currently, there is a management plan in place to incrementally formalize informal settlements through upgrading. This would result in informal residents gaining secure land tenure and housing. For the current management plan to have an efficient formalization process, it must be validated holistically on how the informal settlement in question has developed taking the socio-economic and spatial factors over time into account. This may reduce the negative consequences seen after formalization is complete. This project will show that a geospatial analysis can provide the necessary information to allow for the efficient formalization of informal settlement proliferation and development in the City of Cape Town.

1.3. Research Aims, Objectives and Questions

This research aims to describe how informal settlements have developed within the City of Cape Town over time with respect to major physical features in those settlements. The following research questions has been developed with the intention to achieve the aim of this project:

1. Which spatial metrics best quantify the development of informal settlements?
2. What are the most influential socio-economic and physical factors that impact the development of informal settlements in Cape Town?
3. Can the development of informal settlements in the City of Cape Town be described using a statistical model based on the change of major physical features within these settlements?

The research questions are answered using geospatial analysis through the following objectives:

a) To carry out a change detection analysis of physical features using aerial imagery and machine learning techniques.
b) Once the physical features have been detected, the development of these features can be determined using change detection techniques.
c) Change detection in a spatial capacity can be evaluated using spatial metrics of each given scene. How these metrics change over time may describe the development of informal settlements.
d) The factors that are most significant to the development of physical factors would then be determined.

Using the outcome of the objectives, the research questions can be answered.

1.4. Project Organization

- Chapter 1: Introduction – This chapter provides a brief context of the project. It discusses the aims and objectives that are to be reviewed, as well as how the research may achieve the aims and objectives set.
- Chapter 2: Literature Review – The literature review discusses all the relevant pieces of work that relate to achieving the aims of this project. Through an analysis of the literature review, a method was developed to achieve the objectives of this project.
- Chapter 3: Methodology – This chapter outlines the processes by which informal settlement development may be described, under advisement of the literature review. For this project, the methodology will provide the determination of establishment of informal settlements, the steps in determining an accurate supervised object-oriented image classification, the application of change detection techniques and the determination of spatial metrics describing the features in informal settlements in aerial imagery.
- Chapter 4: Results – The results are a product of the methodology chapter and provide the information needed to make an analysis. In this project, the factors describing informal settlement establishment, the classified aerial imagery of informal settlements and the spatial metrics describing the context of features in aerial imagery of informal settlements would be determined.
- Chapter 5: Discussions – In the discussions section, change detection methods and the determination of the most relevant spatial metrics describing informal settlement development are analyzed.
- Chapter 6: Conclusions – This chapter provides a conclusion based upon the discussions in the previous chapter. Any recommendations to provide for further research would then be made based on the research process followed in this project.

2. Literature Review

2.1. Introduction

This chapter is a review of past research which aims to provide an understanding of what informal settlements are, how they are established, their development, changes in shack characteristics within it, and to provide a local land tenure perspective from Cape Town. This chapter also explores the informal settlements that are to be analyzed. The method of object-oriented image classification will also be discussed as the major technique used to describe physical changes in informal settlements over time. Further analysis will be conducted into how the physical changes may be quantified through spatial metrics. The metrics will be used to describe informal settlements through geographic information systems and statistical modelling. The statistical model – linear regression – will be used to examine metrics that are most significant in describing the development of informal settlements. This section will answer the following questions:

- Why are these settlements forming?
- How are these settlements forming?
- Why is it important to monitor the proliferation and growth of these settlements?
- What are the future implications that settlements could cause within cities?

2.2. Informal Settlements

According to the United Nations (UN-Habitat, 2020), informal settlements are defined by four parameters. These parameters are viewed as areas where:

1) its inhabitants have no legal security of land or the dwellings they inhabit,
2) the neighbourhoods of these settlements lack access to basic infrastructure and services,
3) the housing may not conform to the current planning and building regulations, and
4) the housing within the settlements is situated in geographically and environmentally hazardous areas.

Upon reviewing other pieces of literature (Augustijn-Beckers et al, 2011; Dovey, 2015; Kohli et al, 2016; Manal et al, 1998), it has become widely accepted that informal settlements are defined according to the four parameters stated above.

More than one billion people are living in informal settlements around the world (Beardsley et al, 2008, Dovey, 2015). It is also estimated that within the urban population alone, 24% of this population are informal dwellers (UN-Habitat, 2020). This 24% is largely associated with developing regions (Sub-Saharan Africa, Southern and South-Eastern Asia as well as Central and Southern Asia). Globally, it can therefore be determined that informal settlements are influential in their position and effects within cities (Augustijn-Beckers et al, 2011).

In a local context, the Western Cape in its most recent statistics has 1 in every 6 people living in informal settlements. Between 2001 and 2011, there has been an increase from 166 153 to 210 778 informal households in informal settlements. In recent years, this increase in informal households is estimated to have increased. Therefore, it can be concluded that within the Western Cape, informal settling has had and continues to have a greater influential impact within the province. More specifically, Cape Town is the locus of informal settlement development within the province (WCISSF, 2016). This local phenomenon agrees with the global phenomenon, in which urban space is the hotbed for informal settling.

The common challenges caused by the proliferation and growth of informal settlements are as follows:

- The accelerated growth of informal settlements due to inability of local governments to provide affordable housing (Augustijn-Beckers et al, 2011; Dovey, 2012; Manal et al, 1998).
- The lack of basic infrastructure and services which has a negative impact on the health and the development of productive economic activities (Beardsley et al, 2008; Dovey, 2015; Manal et al; 1998).
- A lack of spatial information to help crisis management within informal settlements (Kohli et al, 2016) – wildfires and floods (WCISSF, 2016).
- The lack of motivation to upgrade informal settlements This is due to the threats posted by the land informal settlements occupy – possible flooding, fire outbreaks and unstable land (Dovey, 2015).

2.2.1. The socio-economic status of informal settlement dwellers

Generally, settlers of informal settlements follow similar criteria. These settlers are disadvantaged groups of people that are poor, and due to their non-existent or low-income economic status (Manal et al, 1998; Wekesa et al, 2010). A comprehensive view on informality within Cape Town can be found by observing the following table (Census, 2001; Census, 2011, Community Survey, 2016):

Table 2.1: Informal dwelling statistics from 2001 to 2016 (the percentage of informal dwellings to total dwellings are in parenthesis)

Year	2001	2011	2016
Informal dwellings	1 836 231 (16.4%)	1 962 732 (13,6%)	2 193 968 (13%)

The statistics in table 2.1 does provide a comprehensive context on informality (Community Survey, 2016). However, these statistics do include backyard dwellings as a form of informal dwellings. Backyard dwellings are structures that are erected by the homeowner or landlord within the boundaries of their property (Turok, 2016). It is also important to note that although it may seem that the percentage of informal dwellings are decreasing over time, the influence backyard dwellings attribute to the total informal dwellings creates a bias against the steady rise in informal dwellings in informal settlements (Lemanski, 2009). For the purposes of this project, only informal dwellings pertaining to informal settlements are important in determining their effect on the City of Cape Town. It is for this reason that the following statistics are outlined (Census, 2001; Census, 2011, Community Survey, 2016):

Table 2.2: Population statistics of informal settlements (disregarding backyard dwellings) from 2001 to 2016

Year	2001	2011	2016
Informal households	166 153 (14%)	210 778 (13%)	320 022 (13%)
Total households	1 173 304 (100%)	1 634 000 (100%)	1 933 616 (100%)
Total population	4 524 335	5 822 734	6 279 731

Demographically, a breakdown of the people in informal settlements within the Western Cape should be dominantly Black Africans, due to the designated nature of Apartheid into racial groups and the local history within Cape Town (Lemanski, 2009). This demographic data can be viewed in the following table according to the 2011 Census:

Table 2.3: Demographic statistics of informal settlements (2011)

	Black African		Coloured		Asian		White		Other		Total	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Informal Dwelling	134 914	93,8	7531	5,2	141	0,1	387	0,3	850	0,6	143 823	100

It has now been determined that the number of dwellers in informal settlements are increasing within the Western Cape. However, why is this the phenomenon? According to the Census 2011, migration is the reason for this. It has been well documented that informal settlement growth is a direct consequence of migration in other countries (Barry et al, 2001; El-Batran et al, 1998; Shuvo et al, 2013). The 2011 census does outline statistics regarding migration into the Western Cape between 2006 and 2011. There is an estimated net migration into the province of 95 556 during this period (Census, 2011). The overall ratio of migration in versus migration out is nearly 2:1. This further exacerbates the degree to which migration affects the Western Cape. The specific migration statistics of the Western Cape can be viewed in the following table:

Table 2.4: Estimated migration statistics between 2006 and 2011 of the Western Cape

ESTIMATED PROVINCIAL MIGRATION STREAMS OF PEOPLE IN THE WESTERN CAPE: 2006 - 2011				
	Out-migration		In-migration	
	Province in 2011	Percentage	Province in 2006	Percentage
Gauteng	40 097	36%	48 951	24%
Eastern Cape	29 899	27%	104 215	50%
KwaZulu-Natal	13 196	12%	17 416	8%
North Cape	9 559	9%	12 941	6%
Free State	5 923	5%	9 098	4%
Limpopo	4 490	4%	4 776	2%
North West	4 309	4%	3 319	2%
Mpumalanga	3 464	3%	5 777	3%
Total	110 937	100%	206 493	100%

The most in-depth income statistics within informal settlements in the Western Cape could only be collected in 2011 (Census, 2011). The statistics suggest that 85% of informal settlement households earn less R3 184 per month, which is below the R3 500 minimum salary a person may earn per month in South Africa (WCISSF, 2016). Further analysis shows that within the Western Cape's informal settlements, 46% of the average income are received through government grants, with the remaining 54% earned through employment (a combination of formal and informal employment opportunities in each household). A breakdown of the average household income in informal settlements describes that the low economic status of informal dwellers is further exacerbated by how much lower than the minimum salary per person each household earns per month. This breakdown can be viewed through the following figure (Census, 2011; WCISSF, 2016):

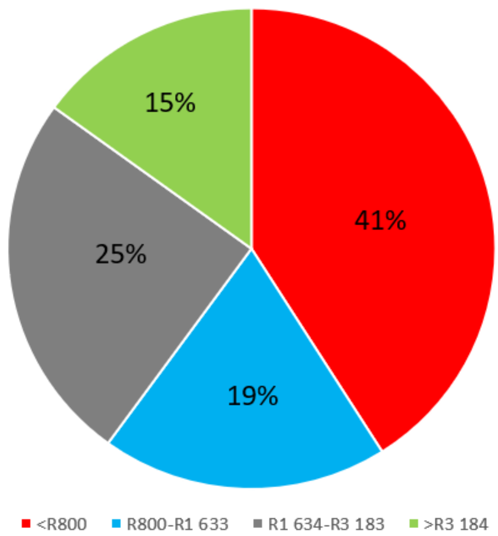


Figure 2.1: Household income statistics in Western Cape informal settlements

To understand why there is such an extreme little-to-no income earned by those that live in informal settlements in the Western Cape, the poverty statistics regarding the province should be discussed. Relative to the rest of South Africa, the Western Cape is not as challenged with regards to poverty (StatsSA, 2017). The average poverty rates across the country do exceed that of the Western Cape. However, the poverty rates within the province are still high. The context of this can be viewed in the following table:

Table 2.5: Poverty statistics between 2006 and 2015 of the Western Cape relative to South Africa

	Headcount (Percentage)			
Year	2006	2009	2011	2015
Western Cape	50,2%	41,3%	33,7%	37,1%
South Africa	**66,6%**	**62,1%**	**53,2%**	**55,5%**

The social context of informal settlers' correlates to their economic status. Since it has been established that informal settlers have little-to-no income, the way in which these people live, and the challenges faced are similar globally as well. An example of a social characteristic being enforced by economic challenges can be found in the search for employment opportunities. People generally move into informal settlements to live near employment opportunities (Dovey, 2015). Most of these people migrate from rural to urban spaces. It has been found that these opportunities require basic skills and are often informal in nature (WCISSF, 2016). Another cause for the rapid population growth in informal settlements are accredited to the rapid natural population increase within cities (Manal et al, 1998). A consequence of a low socio-economic status is found in the structures that compose the settlement – shacks. These are also called slums or favelas. These structures could only be afforded to be made with any readily available material which are often cheap and of bad quality. The material varies metals/tines, cardboard, plastics, and wood (Dovey, 2015; Hofmann et al, 2001; Kohli et al, 2016; Wekesa et al, 2010).

2.2.2. The physical characteristics of informal settlements

Informal settlements are characterized by their location and what is found within it. Across the globe, the location has similar properties as informal settlements are established on any land available and that is within proximity to employment opportunities (Dovey, 2015). A positive consequence of this allows for these areas to be established along major transport networks (Kohli, 2013). However, this is as good of a property as informal settlements may have. The spaces that informal settlements occupies does not consider the usual planning regulations and would suffer setbacks such as steep slopes and marshy areas (Augustijn-Beckers et al, 2011; Dovey, 2015; Kohli et al, 2016). There are more statistics by which informal settlements are characterized. These will be viewed at three different levels, the environment level, the settlement level and at the object level. The environment level will account for a settlement's location and its neighbourhood characteristics. The settlement level accounts for the shape and density of the informal settlement. The object level will account for the shacks and access network within the informal settlement (Dovey, 2015; Dubovyk et al, 2010; Kohli, 2013; Kohli et al, 2016; Kuffer et al, 2011; Naorem et al, 2016). These levels are further discussed in the following table:

Table 2.6: Physical characteristics of informal settlements at the environment, settlement, and object level

Level	Indicator	Phenomenon describing indicator	Statistic
Environment	Location	• Proximity to major roads • Proximity to rivers • Proximity to steep slopes • Proximity to marshy land • Proximity to power lines • Proximity to train lines • Proximity to essential services	Distance Distance Distance Distance Distance Distance Distance
	Neighbourhood Statistics	• Proximity to employment opportunities • Proximity to industrial areas • Proximity to middle-to-high socio-economic status areas	Distance Distance Distance
Settlement	Shape	Irregular and linear shape Slope	Aggregation Index Slope %
	Denisty and Compactness	• High dense structures (>70% roof coverage) • Low vegetation and open space (<20% coverage)	Patch Density NDVI
Object	Building	• Variable shack shape • Variable shack size (10-40m^2) • Irregular building orientation • Variable building colour	Shape Index Mean Area Shape Index Mean Spectral Value
	Access Network	• Irregular access shape • Variable street and paved/unpaved width	Shape Index Distance

2.3. The Geometric Stages of Informal Settlement Development

It has been established through the analysis of past literature that focuses on the establishment and development of informal settlements, that growth patterns exist (Augustijn-Beckers et al, 2011; Dovey, 2015; Sluizas, 1988). These spatial patterns follow a complex structure that is influenced by several physical, cultural, and economic factors (Sobreira et al, 2001; Zhang et al, 2020). In this section, the development of informal settlements will be discussed through its geometric stages. The geometric stages are divided into the establishment, development, the formalisation, and post-formalisation processes:

2.3.1. Establishment

The establishment of informal settlements are mainly accredited to the rapid rate of rural-urban migration and the inability of the urban poor to afford existing urban housing (Manal et al, 1998; Wekesa et al, 2010). To overcome the challenges faced socially, culturally, and economically, informal settlers look for any available land that may favour the attainment of employment opportunities (Manal et al, 1998). For this reason, informal settlements are established according to the following physical factors:

- Along major transport networks – informal settlements are established close to highways, major roads, railways, rivers, and canals (Dovey, 2015; Dubovyk et al, 2010; Kohli et al, 2016).
- Close to employment opportunities – informal settlements are established in spaces close to a city's CBD, industrial areas, or middle-to-high socio-economic status neighbourhoods (Dovey, 2015; Dubovyk et al, 2010; Kohli et al, 2016).
- Although the proximity to transport networks and employment opportunities are viewed as the most important variables to the establishment of informal settlements, a whole host of other proximity metrics exist which are viewed as less integral (Dovey, 2015; Dubovyk et al, 2010; Kohli et al, 2016; Kuffer et al, 2011; Naorem et al, 2016). These proximity metrics have been outlined at the environment level of the physical characteristics of informal settlements (section 2.2.2). The metrics include proximity to rivers, steep slopes, marshy lands, power lines and train lines (Dovey, 2015; Dubovyk et al, 2010; Kohli, 2013; Kohli et al, 2016; Kuffer et al, 2011, Naorem et al, 2016).
- On environmentally hazardous areas – informal settlements are built in spaces that are generally unattractive for planned development (Dovey, 2011; Kohli et al, 2016). These are usually on slopy or marshy land (Augustijn-Beckers et al, 2011; Kohli, 2013). Since planned development has already made use of attractive spaces, only the unattractive spaces are available to build shacks for the urban poor or urban migrators (Ishtiyaq et al, 2011).

2.3.2. Development

The development of shacks in informal settlements have been monitored extensively in projects of object-orientated image classification (OOIC) and within either agent-based (ABM) or cellular automata-based (CA) modelling approaches (Augustijn-Beckers et al, 2011, Shuvo et al, 2013). The findings of these investigates indicate that the following factors are essential in understanding how shacks are built over time in informal settlements:

- There is a densification of shacks built in the development stage from the initially established shacks. This densification is described as infilling, or the creation of new shacks in the immediate vicinity of shacks already in existence (Augustijn-Beckers et al, 2011). Density is used to quantify this change in the development of informal settlements over time (Sliuzas, 2004; Yonda, 2006).
- There is also a densification of shacks along informal transport networks within informal settlements – dirt roads and footpaths (Augustijn-Beckers et al, 2011; Yonda, 2006), and formal transport networks outside informal settlements (Kohli et al, 2016; Poelmans et al, 2009). This can be analyzed by creating vector layers of the transport networks and using proximity metrics between constructed shacks and vector layers within GIS software (Kohli et al, 2016).
- Enlargement of shacks are also observed within informal settlements over time (Augustijn-Beckers et al, 2011). This contributes as an attributing factor as to how shacks take space within informal settlements. The average size of the shacks is used to quantify this change in the development of informal settlements (Sliuzas, 2004).
- Extension of the informal settlement. Once all the suitable land is taken up within an informal settlement and a population threshold is reached, built-up areas are extended to less favourable land in the immediate vicinity (Augustijn-Beckers et al, 2011; Shuvo et al, 2013; Zhang et al, 2020). This is quantified by the total area that the informal settlement occupies over time (Verburg et al, 2004). A visual depiction of this phenomenon can be viewed by overlaying an informal settlement over time and observing new shacks over time (Shuvo et al, 2013).

2.3.3. Formalization

The effects of formalization have not been a focal point across many pieces of literature that deal with monitoring the physical development of informal settlements. Due to this reason, there is simply not enough research to understand how the make-up of informal settlements changes over time. However, there are investigations that have focused on what the effect of formalization has in a socio-economic context (Dovey, 2011; Huchzermeyer, 2003; Massey, 2013; Zhang et al, 2020).

During the formalization process, the building of formal housing requires the temporary relocation of informal dwellers (Dovey, 2011; Huchzermeyer, 2003). Governmental projects would fund this relocation, and subsequently this would allow for the construction of formal houses (Huchzermeyer, 2003).

2.3.4. Post-formalization

Post-formalization is another stage within informal settlement development that is not focused on within the physical development context of past research. However, the context in which the physical environment results may prove to be useful in determining whether-or-not the correct implementation of upgrading has been used for the informal settlement. This way, any problems that were faced by an informal settlement can now be analyzed to determine whether it has been solved. The problems have been outlined in the literature review already.

2.4. Upgrading Policies for the Western Cape Informal Settlements

2.4.1. The Western Cape Informal Settlement Strategic Framework

The Western Cape Informal Settlement Strategic Framework (WCISSF) is a strategy document released by the Western Cape government, in which a plan is determined to curb the informal development of informal settlements in the province. The document outlines various reasons for the need to formally develop informal settlements and discusses associated problems, such as:

- Informal settlements being an outcome of uncontrolled urbanization, migration, and the failure of the current housing market.
- Informal settlements are explicitly vulnerable to climate change, disease outbreaks, violent crimes, and unemployment.
- There is a lack of trust between informal settlements and municipalities in relation to governance and service provision.
- The current public sector investment is not being allocated optimally to minimize the negative effects of informal settlements.

The vision and mission of the WCISSF is to improve the quality of life for informal dwellers by enabling access to financial, tenure, land, and economic opportunities and at the same time provide incremental housing opportunities that is people-centered and partnership-based.

As this investigation analyzes mainly a spatial-temporal context, the strategies discussed in the WCISSF would be focused on the physical changes planned within the strategic framework. The major strategies outlined within the document, with a spatial reference are to (WCISSF, 2016):

a) **Upgrade informal settlements through access to basic services and infrastructure, as well as incremental housing opportunities** – use the Red Book (see 2.4.2) standards for provision of basic and emergency services, have new settlement formation based on expected growth, make optimal use of land, have initial focus on highly dense areas to accommodate possible health hazards and use Red Book for new housing sizes and regulations.
b) **Enhance the quality of life and active citizenship** - incorporate neighbourhood planning procedures to promote consistent housing relative to neighbouring settlements, incorporate local economic and social opportunities.
c) **Strengthen sector capability, governance, and resources** – Develop strong incentives for residents to maintain formal living (avoid backyard dwelling creation).

2.4.2. The Neighbourhood Planning and Design Guide – The Red Book 2019

The Red Book is a series of guideline documents which has the aim of improving the quality of settlement planning and design for Town Planners in South Africa. This document will be discussed with respect to the methods regarding upgrading informal settlements. This is relevant as the settlement and design planning is essential in eradicating problems associated with informal settlements in a spatial or physical environmental context.

Upgrading of informal settlements is considered essential by the South African government in delivering housing. As vital as housing objectives are, another key intention of upgrading is to support the provision of basic services and security of tenure. These objectives are encouraged to be followed in a staged process. This is started by allowing for in-situ upgrading as far as possible – development in the settlement to minimize the relocation of residents. In-situ upgrading also allows for strong community cohesion to stay in place, and if any important structures are within the settlement, it may be left undisrupted. In-situ will be avoided if the location poses danger to the health and lives of the settlement's residents. However, once the method of upgrading has been selected, the upgrading itself will follow incrementally to allow the residents to (Red Book, 2019):

- Settle in the area where they are and make a living within their current location.
- Build on what is needed within the settlement with regards to basic services.
- Be integrated into the surrounding region, town, or city.

In terms of the considerations to planning the settlement upgrading, the following are discussed:

- Water supply
- Public open space
- Solid waste management
- Sanitation
- Neighbourhood layout and structure
- Electrical energy
- Stormwater
- Transport and road pavements
- Housing and social facilities

It is important to assess the current context of the Western Cape's policies of upgrading. However, monitoring the development of the informal settlements within the province is necessary to implement these policies (WCISSF, 2016). Machine learning provides the opportunity to monitor such development. This involves using remotely sensed data to detect informal settlements within cities and doing subsequent analysis on the detected features (Coppin et al, 2004; Fallatah et al 2022; Huang et al, 2019; Lu et al, 2004; Valente et al, 2020; Zhu et al, 2019).

2.5. Remote Sensing Data

Airborne platforms provide image data through remotely sensed applications (Benz et al, 2004; Zhu et al, 2019). This image data found can be useful when at it's at a high quality. The quality of the imagery is defined or quantified by the resolution (Lu et al, 2004; Wulder et al, 2004).

Resolution describes the amount of detail that a remotely sensed image holds. Its relation to remote sensing can be described as the ability of a remote sensor to capture and display the objects on the ground. This plays a significant role in change detection, as the ability to accurately detect objects from remotely sensed imagery will influence the analysis in the change of those objects (Coppin et al, 2004; Huang et al, 2019; Lu et al, 2004; Zhu et al, 2019). The types of resolution that is useful for the context of this project include the spatial resolution, spectral and temporal resolution. These are discussed in the following:

2.5.1. Spatial Resolution

The spatial resolution of an image is determined by closely capturing area which contains objects on the ground and its relation to other objects. It's quantified by the size of each pixel in the image (Turner et al, 2003; Witharana et al, 2018; Valente et al, 2020). The spatial resolution also plays a major role in the spectral resolution of an image as these two phenomena are closely related (Price et al, 1997). This will become apparent through the description of the spectral resolution of an image.

2.5.2. Spectral Resolution

The spectral resolution of an image refers to the remote sensor's ability to retrieve reflectance values of objects found on the ground (Gardner et al, 2002). It then provides clarity as to why a greater spatial resolution relates to a greater spectral resolution. If an object can be seen clearly on the ground, the ability of a sensor to retrieve reflectance values from that object can be retrieved on a more accurate basis (Price et al, 1997). The concept of spectral reflectance originates initially from the Sun's electromagnetic (EM) radiation. This EM radiation reflects off the surface of all objects on the ground towards the remote sensor. The remote sensor will then transform this into a value to represent the wavelength at which the objects reflected the EM radiation. Different objects would then have different reflective values. Although the spectral properties of urban materials are limited in terms of knowledge, shacks are made from relatively similar materials (Herold et al, 2003). This makes it manageable for a program to classify objects based on the spectral values (Hofmann et al, 2017; Kohli et al, 2016; Shekhar, 2012).

2.5.3. Temporal Resolution

Temporal resolution is recognised as the interval at which a given scene can be imaged (Berne et al, 2004). This is an important phenomenon as the time at which the images over a scene are taken can be used to detect changes in the scene (Lu et al, 2004).

2.6. Change Detection

Change detection is defined as the process where differences can be detected either within objects or phenomenon by observing it over time (Singh, 1989). This definition is reiterated throughout object-based investigations in which change is investigated with respect to land-cover types (Lu et al, 2004; Mas, 1999; Walter, 2013). Change detection is further popularized specifically in the change of informal settlement development, as the growing need to understand the development of these settlements are demanded within the current urban context (Abebe et al, 2019; Hofmann, 2001; Kohli et al, 2016; Mason et al, 1997; Mudau et al, 2022).

Using the aforementioned literature, change detection is comprised using four steps:

- Pre-processing that may be required between the datasets,
- The image classification techniques to be selected,
- The classified results having been validated through an accuracy assessment, and
- Two or more classified images being analyzed to determine any changes.

Within the context of informal settlement development, object-oriented image classification is a technique used to define objects within informal settlements (Hofmann, 2001; Kohli et al, 2016; Mason et al, 1997). The way these objects are compared for change post-classification varies across different pieces of literature (Augustijn-Beckers et al, 2011; Kohli et al, 2016). Shacks may be analysed by their accumulative number over time to analyse development (Augustijn-Beckers et al, 2011; Hofmann, 2001; Kohli et al, 2016). Internal movement networks may be analysed by the space it occupies and its relative location to other objects within the settlement over time (Hofmann et al, 2017). Geographic information-based solutions are recommended for the extraction of movement networks (Kahraman et al, 2018). For the purposes of this investigation, spatial metrics of the objects will be analysed to gain an understanding of its development over time – the method by which change is detected (Grippa et al, 2018; Kuffer et al, 2011). To gain a further appreciation of which metrics are most relevant in understanding change within the objects of the informal settlements, linear regression will be used to compare the metrics to factors related to the development of informal settlements.

2.7. Machine Learning

Machine Learning is defined as a system that has the capability to learn to solve problems (Alpaydin, 2020). These problems are solved in the real world by using a database. A database would be analyzed to find a solution to a problem. Machine Learning is no different. By optimally analyzing a database, the system can automatically learn what the most optimal parameters are. These optimal parameters can then be used to find a solution to a problem. A popular example that has been used across decades comes in the form of recognizing faces. In the real world, humans can identify different faces if they have seen it before. Even if the faces that are to be identified are under different lighting, in a different pose or contain different hair styles, this is an effortless task for humans to consciously solve. However, this conscious problem solving is difficult to create in a computer program as it is difficult to explain our conscious processes. This is where Machine Learning has its advantages in problem solving. We know that a face is not just a random collection of pixels in an image. It has symmetry. There are shapes of eyes, a mouth, and a nose in certain locations on the face. There are also shades of colour in different shapes. This can be viewed as the database by which a machine may solve the problem of recognizing a face. The use of a sample face will give the machine the necessary information to understand the composition of symmetry, shapes and colour (Bishop, 2011). This composition is known as a pattern found in the sample data. A learning program can then aim to find the face given the sample data by checking for this pattern in any other given image. In Machine Learning, this is called pattern recognition. In conclusion, Machine Learning is a system or a model that uses the optimal set of parameters to solve a problem. This optimal set of parameters are trained within the program to find patterns given in sample data. There are two major advantages of Machine Learning (Alpaydin, 2020):

- It can be used in a predictive nature (to predict future events), and
- It can be used in a descriptive nature (gain knowledge of current events)

The type of learning that is used in the example above is known as supervised learning or classification. In a simple definition, supervised learning or classification is choosing sample data, whereby the samples are given the correct classification assigned to it (Sathya et al, 2013). The most pragmatic software that can be of use for this project that incorporates machine learning is e-Cognition and this will be investigated in the next section. Other machine learning algorithms investigated for this project included random forest, support vector machine and naïve Bayes (Li et al, 2019; Mahesh, 2018; Prins et al, 2020; Saravanan et al, 2018; Ward et al, 2021; Wu et al, 2023).

2.8. Object-oriented Image Classification Process

Object-oriented image classification (OOIC) or analysis (OOIA) is a technique that is widely used to define settlements or objects within settlements (Walter, 2003). This is due to its great potential for informal settlement mapping – or more specifically shack detecting – as the ability to include spatial, spectral, and contextual characteristics are available (Kohli et al, 2016; Shekhar, 2012). This ability follows the similar human cognitive image interpretation discussed in the Machine Learning section (Bishop, 2011). OOIC can therefore be used as a form of supervised classification, where the combination of the local ontology and available remote sensing data can be used as a sample to detect patterns to solve complex and time-consuming problems (Kohli et al, 2016). Instead of defining OOIC, a general process is followed to classify objects based upon sample data. The process has been found to work optimally using a program called eCognition (Baatz et al, 2004; Fallatah et al 2022). eCognition first conducts a multi-resolution segmentation to break an image down into segments that are to be analyzed according to the user's convenience (Karakis et al, 2006). Then, by using features that are optimal for the detection of a specified object in a scene, these objects are trained to be classified by the program (Fallatah et al, 2019; Kohli et al, 2016; Shekhar, 2012). A more in-depth discussion of these tools is discussed below, as their relevance to the OOIC procedure can be clearer to understand:

2.8.1. Multi-resolution Segmentation

Across various pieces of literature, multi-resolution (MR) segmentation in eCognition is the first and viewed as the main step of the OOIC process (Hofmann et al, 2017; Karakis et al, 2006; Kohli et al, 2016; Shekhar, 2012; Zhang et al, 2010). Segmentation generally refers to the searching of homogenous regions and creating fragments within the imagery (Tian et al, 2006). In eCognition, the MR tool starts by considering each pixel as a separate object. Subsequently, pairs of the image objects which are homogenous are merged to form larger objects until a user specified threshold of heterogeneity is reached (Darwish et al, 2003). This threshold is set by the scale parameter in the MR segmentation tool. The merging of homogenous regions is based on the similarity of shape and brightness values of the objects segmented in the imagery (Zhang et al, 2010). To break these two properties down further, shape refers to the compactness (the ratio of the total fragment border length to the square root of the number of pixels in a fragment) and smoothness (ratio of the total fragment border length to the shortest border length of the fragment) of the segmented fragments. The brightness refers spectral heterogeneity of the red, green, and blue (RGB) band values on the image (Baatz et al, 2004). The shape and brightness values are recommended to follow the local ontology (Kohli, 2013).

There have been multiple pieces of literature that has had the aim of identifying the optimal set of parameters. A study was conducted by Darwish and Reinhardt in 2003 which aimed at determining the optimum scale parameters in classifying land types. The research was focused on Algiers, which is the capital of Algeria. For built-up land types, it was found that higher scale parameters produced better results in terms of the accuracy of the classification. The same conclusion was produced in a separate paper which also focused on segmenting built-up land (Karakis et al, 2006). The results of both studies determined that the level of over-segmentation produced by the scale values allowed for more accurately classified segments. Over-segmentation occurs when an object that is semantic in nature is further fragmented into multiple fragmented pieces. The results of under-segmentation have also been researched. Under- segmentation occurs when more than one class is covered within one specific fragment. **F**indings have determined that under-segmentation in aerial imagery resulted in errors becoming significantly larger the more under-segmented images were (Lui et al, 2010). To achieve a preferred result, the entire process of the MR segmentation is done on a trial-and-error approach (Benz et al, 2004; Zhang et al, 2010).

2.8.2. Feature Detection

Feature detection is defined as the retrieval of meaningful image objects based upon features describing those objects (Tian et al, 2006). Once the MR segmentation process is completed in eCognition, the next step in the overall OOIC process is to describe the features by which image segments would be classified (Shekhar, 2012). Since the objective of the OOIC is to detect meaningful objects – shacks, the process of feature detection will be to describe the features that will allow for the accurate classification of shacks within the aerial imagery (Hofmann et al, 2017; Kohli et al, 2016). In classifying shacks from remotely sensed imagery, there are challenges. These structures are often geometrically and texturally complex (Kohli et al, 2013). It is suggested that 90% of shacks are approximately rectangular in shape and has flat roofs, with the remaining 10% extremely varied. The structures are also texturally complex due to the diversity in the materials used to construct the shacks (Mason et al, 1997). These key features are to be explored based on the following characteristics according to past literature, however it is recommended to use the local ontology to assist in the feature detection process (Kohli, 2013):

Geometry

In a study based on shack shape typology in the Marconi Beam informal settlement in Cape Town, South Africa, it has been found that 85% of the shacks classified are 4-sided and a further 11% were 6-sided (Mason et al, 1997). This confirms that shacks follow simple geometric characteristics with regards to rectangularity. It is vital to note that shacks that deviated from rectangularity were not uncommon. However, it has been determined that rectangularity would be a useful feature to be descriptive of shacks in the classification procedure.

Area representing the spatial dimensions of shacks is seen as an important feature in the classification of shacks relative to other image objects in an aerial imagery scene (Hofmann et al, 2001; Kohli et al, 2016).

Colours of the roofs

In a study on detecting informal settlements, the colour of roofs is a key feature in detecting informal settlements (Hofmann et al, 2001; Shekhar, 2012). In addition, Mason and Baltsavias (1997) says that although the roofing of shacks is constructed from a diverse range of materials (plastic, iron sheeting, timber), the use of red, green, and blue (RGB) colour bands may be used to detect a distinct shack. This is derived in their study of the detection of shacks in informal settlements from satellite imagery.

Density

Mason and Baltsavias (1997) mentions that the density of shacks is also a key consideration in the feature characteristics of shacks. In their study, they found a general separation of 2-3m between the shacks. This, therefore, is quite densely built shacks per given area. Kohli (2013) agrees with this statement as a study in identifying informal settlements has roof coverage greater than 70%. Density can now be associated as a feature unique to shacks compared to other objects in the scene (examples may be motor vehicles, rubbish bins and open spaces). In the study of the general detection of informal settlements by Hofmann, density is also considered as a major consideration in the classification process. Formal settlements follow a highly sparse density compared to that of informal settlements (Dovey, 2015; Hofmann et al, 2001).

Vegetation and it's spatial relation in informal settlements

Vegetation may be an influence when classifying shacks, as according to a study, there is generally a lack of vegetation in informal settlements, it may be situated in a position which might alter the differentiation of two or more shacks (Mason et al, 1997). Vegetation is a factor in Hofmann's study in the same way as Mason and Baltsavias' study. Another consideration that will be vital in the classification of shacks with be the contextual influence of vegetation. If vegetation does cover contiguous shacks, its spatial relation to shacks would have to be a feature in differentiating shacks in the classification procedure. To detect and eliminate the influence of vegetation in a scene, it has been established that using the customised arithmetic RGB ratio of G/(R+G+B) may be useful in removing it from a remotely sensed imagery within eCognition (Kohli, 2013). There is an agreement with this phenomenon in another study to different shacks and vegetation using the ratio of the RGB band values (Shekhar, 2012).

Shadows and it's spatial relation in informal settlements

Shadows within OOIC projects has been covered extensively as it influences the quality of the classification of image objects. It has been established that low values of RGB brightness and the mean difference of RGB brightness values can be used to detect and eliminate shadows on remotely sensed imagery for informal settlement structures (Kohli, 2013). Another study agrees with this point as highly built-up structures does project shadow-related features in an image (Grippa et al, 2018). Low RGB band values are further discussed as a prominent feature to different shacks with the rest of the informal settlement (Shekhar, 2012).

2.8.3. Classifier for the classification

The classifier used in eCognition to classify image objects is the nearest neighbour classifier algorithm. This algorithm allows for a classification based on comparing fragmented segments from the MR segmentation and classifying them based on input features in a feature space (Cunningham et al, 2007). To fully represent how the nearest neighbour is applied, the main idea is that fragments that are similar in feature space are close to each other compared to fragments that are represented differently. The impression that the word 'close' implies is that it may fall within a specified quantified variable distance in feature space (Muja et al, 2009). An example to describe this may be found in RGB band values, which will be represented as k in the figure below. The idea can be followed that objects of similar RGB band values are close compared to other more varied RGB band values:

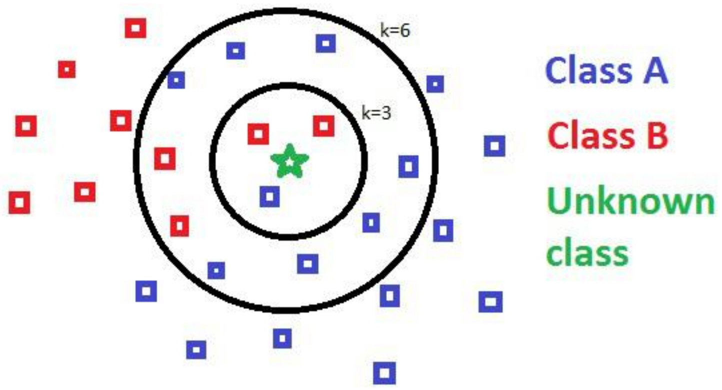


Figure 2.2: Use of the K-nearest neighbour

The OOIC process is a supervised one, where the image object to be determined is known beforehand. The nearest neighbour classifier uses user-specified training samples of the image object to be determined, and the feature space specified in the feature detection process to conduct the nearest neighbour algorithm in eCognition (Yu et al, 2006).

According to a project analysing the K-nearest neighbour classifier by Cunningham and Delany (2007), following has been determined:

- The advantages of applying the nearest neighbour classifier in eCognition includes the software conducting noise reduction techniques in the classification process. This will allow for greater accuracy in the classification process. Analysis techniques for post classification would also be aided by the nearest neighbour classifier as an explanation of objects being assigned under a certain class. For the context of this project, this is very useful as the reasoning for shacks being classified can be explained using the nearest neighbour classifier.
- However, there are disadvantages associated to using the nearest neighbour algorithm in eCognition. One involves the sensitivity of the algorithm through using irrelevant or redundant features in assigning objects to classes. This may result in erroneous results. An example of this is using RGB band values. If shacks are being classified using the band values, rubbish on the surface may be classified as shacks. It should be stated that a variety of features will be contributing to the classification and would aid in minimizing this error. If an equal weighting is associated to all the features, this will assist as well.

2.9. Geographic Information Systems

Geographic Information Systems (GIS) uses software that create, manage, analyse, and visualize geographic data (Steiniger et al, 2009). Together with remote sensing, GIS is viewed to be an ever useful and evolving tool in determining and evaluating long-term changes in land use due to urbanisation (Balogun et al, 2011). This is especially useful as urbanisation is a direct factor that contributes to the proliferation and development of informal settlements (WCISSF, 2016).

GIS has been used in various pieces of literature under the context of OOIC (Benz et al, 2004; Hofmann et al, 2017). Its use has lied in validating whether classified objects within images are accurate and through analysing the change through which these objects have undergone over time and its use in analysing change detection (Hofmann et al, 2017). This proves that GIS is a valuable tool to describe how informal settlements have developed in Cape Town. Its specific use in this investigation will lie in its ability to provide spatial information based on feature characteristics.

2.10. Validity testing for classified imagery

An accuracy assessment is conducted is to express the accuracy of images or maps as a percentage of the mapped area that has been classified correctly after it has been compared with ground truth or reference data (Congalton et al, 1986). It has also been established that in a GIS context, it is important to incorporate an accuracy assessment into work that requires an analysis of pre-processed data (Lunetta et al, 2006). The accuracy assessment's importance is further explained below (Congalton, 2001):

- To assess output by evaluating success and failure rates
- To quantitively compare methods; and
- To use the output from the analysis to make informed decision.

An accuracy assessment uses a classified image compared to that of an original image. The percentage of points that is placed in the correct classes in the comparison determines the accuracy of a classification (Patil et al, 2013). This accuracy is produced in the form of a confusion matrix or contingency table from the comparison of the classified and original image (Congalton et al, 1986). A confusion matrix or contingency table is generally used to assess the thematic accuracy of classes (Stehman, 1997). Its configuration includes the major diagonal describing the level of agreement between the two sets of data. The accuracy of the classified image is then calculated by dividing the number of correct classifications (found on the major diagonal) by the total number of samples taken. These samples can be represented as random points that can be distributed across the scene in question.

There are four different outcomes that be interpreted as a product of a confusion matrix (Faweett, 2006):

- True positive – the outcome is a positive and it is classed as a positive
- False negative – the outcome is a positive, however it is classed as a negative
- True negative – the outcome is a negative, however it is classed as a positive
- False positive – the outcome is a negative, however it is classed as a positive

These outcomes are represented in the following two-by-two confusion matrix:

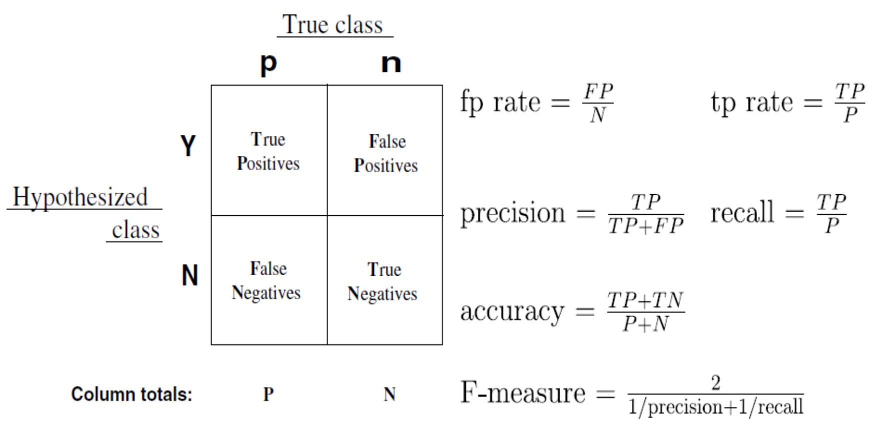


Figure 2.3: An example of a Confusion Matrix

Another measure of accuracy is contained in the confusion matrix – the kappa coefficient. The k or kappa coefficient represents the overall accuracy of the classified image. This kappa coefficient statistic is a multi-variant, which results in an unbiased outcome as it considers the entire confusion matrix and not an overall accuracy (Fletcher et al, 2011; Navulur, 2007). The kappa coefficient is calculated using the following formula (Lillesand et al, 2008):

$$k = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} X_{i+}X_{+i}}{N^2 - \sum_{i=1}^{r} X_{i+}X_{+1}} \qquad \text{(2.1 Kappa coefficient)}$$

where k is the kappa coefficient, r are the number of rows in the confusion matrix, x_{ii} are the number of observations in row i and column i in the major diagonal, X_{i+} are the total number of observations in row i, X_{+i} are the total observations in column i, and N is the total number of observations in the confusion matrix

The value in the kappa value may be rated according to the following table (Fletcher et al, 2011; Lillesand et al, 2008; Navulur, 2007):

Table 2.7: Rating of Kappa values

Kappa	Rating
< 0.00	poor
0.0 – 0.20	slight
0.21 – 0.40	fair
0.41 – 0.60	moderate
0.61 – 0.80	substantial
0.81 – 1.00	close to perfect

2.11. Spatial Metrics

Spatial metrics is defined as the statistics attained from analysing spatial data and is used in finding patterns in digital imagery (Saura et al, 2001). The metrics would be based upon the Island biogeographical model (McGarigal, 2002). In this model, the emphasis will be on a single patch type (classified shacks in informal settlements). This emphasis would describe how these patches behave within a certain boundary, its spatial character and distribution. The major advantage of this model is that it is relatively simple compared to describing the spatial characteristics of varying patch types, making this easy to represent the structure of the shacks against background data (rest of the informal settlement). The major disadvantage of this model is that it may be an oversimplification of how patches interact in the general capacity compared to other informal settlements (McGarigal, 2002). It is vital then to use the ideal metrics to describe the spatial character of shacks in different informal settlements. Spatial metric determination is also sensitive to spatial resolution (Moody et al, 1995) and the extent of the study area (Turner et al, 1989). To combat these limitations, high spatial resolution remotely sensed imagery is recommended (Kupfer, 2012). Additionally, the level of accuracy in classifying the imagery is important, as acceptable low misclassified results can still lead to magnified errors in the determination of spatial metric values (Langford et al, 2006). These limitations can be addressed through correct data manipulation and analysis (Kupfer, 2012).

FRAGSTATS has been the suggested tool to retrieve spatial metrics in many landscape ecology investigations (Kuffer et al, 2014). This software can determine statistics that quantifies objects across a landscape, both in the composition and configuration of patches (McGarigal, 2002). Since it has been established that shacks are the objects or patches of interest, shacks will be analysed in the context of its spatial patterns within informal settlements.

To clearly define patch metrics, these values represent the spatial context of each individual patch in the landscape. To gain an understanding of all patches of a single patch type, class metrics are of interest. Class metrics considers the spatial distribution and pattern of all patches within a class simultaneously (McGarigal et al, 2009). For this investigation, class metrics are of interest, as the spatial context for all shacks across the landscape are under inspection. To understand the value of class metrics, first the patch metrics will be discussed. Once the spatial context of individual patches is understood, the class metrics describing all patches can be determined. The metrics are explored in further detail in the following tables (McGarigal, 2002):

Table 2.8: Patch-based metrics to describe shacks in informal settlements

Metric	Formula	Variables	Description	Unit	Range of values
Area	$AREA = a_{ij}(\frac{1}{10\ 000})$	a_{ij} = area of patch ij	AREA (m^2) of the patch, divided by 10 000 to convert to hectares	Hectares	AREA > 0, without limit The range of the AREA is limited to the extent of the image
Perimeter	$PERIM = p_{ij}$	p_{ij} = perimeter of patch ij	PERIM is the perimeter of the patch (m), which includes the internal holes of the patch	Meters	PERIM > 0, without limit
Perimeter-Area Ratio	$PARA = \frac{p_{ij}}{a_{ij}}$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij	PARA is a measure of the ratio of the shape complexity of patch ij. A limitation of PARA is that shape complexity varies in patch size and PARA does not compensate for this	None	PARA > 0, without limit
Shape Index	$SHAPE = \frac{.25p_{ij}}{\sqrt{a_{ij}}}$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij	SHAPE measures shape complexity for patch ij in an adjusted square standard compared to that of PARA	None	SHAPE ≥ 1, without limit When SHAPE = 1, the patch is a square and SHAPE increases without limit as patch shape becomes more irregular
Fractal Dimension	$FRAC = \frac{2\ln(.25p_{ij})}{\ln a_{ij}}$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij	FRAC measures shape complexity across a range of spatial scales (patch sizes), and thus overcomes a major limitation of PARA as a measure of shape complexity	None	1 ≤ FRA ≤ 2 FRAC approaches 1 with very simple perimeters such as squares. FRAC approaches 2 for shapes that are highly convoluted, plane-filling perimeters
Euclidean Nearest-Neighbour Distance	$ENN = h_{ij}$	h_{ij} = distance (m) from patch ij to nearest neighbouring patch of the same class (shacks)	ENN is a measure of patch context and is used extensively to quantify patch isolation	Meters	ENN > 0, without limit As ENN approaches 0, the distance to the nearest neighbour decreases

Table 2.9: Class-based metrics to describe shacks in informal settlements

Metric	Formula	Variables	Description	Unit	Range of values
Total Area	$CA = \sum_{j=1}^{n} a_{ij}(\frac{1}{10\,000})$	a_{ij} = area of patch ij	CA is the sum of the patches (m^2) of the corresponding patch type (shacks), divided by 10 000 to convert to hectares	Hectares	CA > 0, without limit CA approaches 0 as the corresponding patch type (shacks) becomes increasingly rare on the landscape
Percentage of Landscape	$PLAND = P_i = \frac{\sum_{j=1}^{n} a_{ij}}{A}(100)$	P_i = proportion of landscape of patch I a_{ij} = area of patch ij A = total landscape area (m^2)	PLAND is the percentage of landscape comprised by a corresponding patch type (shacks)	Percent	0 < PLAND ≤ 100 PLAND approaches 0 when the corresponding patch type (shacks) becomes increasingly rare on the landscape and approaches 100 when the entire landscape consists of a single patch type (shacks)
Mean Patch Area	$\text{MN\_}AREA = [\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}(\frac{1}{10\,000})]/NP$	a_{ij} = area of patch ij NP = number of patches	MN_AREA (m^2) is the average area of the sum of patches, divided by 10 000 to convert to hectares	Hectares	MN_AREA > 0, without limit The range of the MN_AREA is limited to the extent of the image
Mean Perimeter-Area Ratio	$\text{MN\_PARA} = (\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{p_{ij}}{a_{ij}})/NP$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij NP = number of patches	MN_PARA is a measure of the averaged ratio of the shape complexity of the sum of patches in the landscape. A limitation of MN_PARA is that the mean shape complexity varies in mean patch size and MN_PARA does not compensate for this	None	MN_PARA > 0, without limit
Mean Shape Index	$\text{MN\_}SHAPE = (\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{.25p_{ij}}{\sqrt{a_{ij}}})/NP$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij NP = number of patches	MN_SHAPE measures average shape complexity for patches for the landscape in an adjusted square standard compared to that of MN_PARA	None	MN_SHAPE ≥ 1, without limit When MN_SHAPE = 1, the mean patch is a square and MN_SHAPE increases without limit as patch shape becomes more irregular
Mean Fractal Dimension	$\text{MN\_}FRAC = (\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{2\ln(.25p_{ij})}{\ln a_{ij}})/NP$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij NP = number of patches	MN_FRAC measures average shape complexity across a range of spatial scales (patch sizes), and thus overcomes a major limitation of MN_PARA as a measure of shape complexity	None	1 ≤ MN_FRA ≤ 2 MN_FRAC approaches 1 with very simple perimeters such as squares. MN_FRAC approaches 2 for shapes that are highly convoluted, plane-filling perimeters
Mean Euclidean Nearest-Neighbour Distance	$\text{MN\_}ENN = (\sum_{i=1}^{m} \sum_{j=1}^{n} h_{ij})/NP$	a_{ij} = area of patch ij p_{ij} = perimeter of patch ij NP = number of patches	MN_ENN is a measure of mean patch context and is used extensively to quantify patch isolation	Meters	MN_ENN > 0, without limit As MN_ENN approaches 0, the distance to the nearest neighbour decreases
Number of patches	$NP = n_i$	n_i = number of patches in the landscape of patch type i	NP represents the measure of the sum of a patch type (shacks) on the landscape. NP is also a simple measure of the fragmentation of a patch type	None	NP ≥ 1, without limit NP = 1 when the entire landscape contains only 1 of the corresponding patch-type
Patch Density	$PD = \frac{n}{A}(10\,000)(100)$	n_i = number of patches in the landscape of patch type I A = total landscape area (m^2)	PD is the number of patches per unit area	Number per 100 hectares	PD > 0, constrained by cell size

2.12. Linear Regression Analysis

Linear regression is used to determine values of parameters for a linear function, whereby the linear function is used to describe the best fit of a set of observations (Du et al, 2004; Naoum et al, 2004; Ra et al, 2012). This can be translated according to the following formula:

$$Y = X\beta \qquad \text{(2.2. Linear Regression)}$$

where Y represents the dependent variable, X represents the independent variable, and β represents the parameters that needs to be determined, such that formula (2.2) may best fit a given data set D.

Once the parameters are determined, it is possible to determine the independent variable Y, given $X_1, \ldots,$ X_{n+D}. This determination of Y is known as the prediction of the linear regression model. If there is a perfect fit of the model's function in the given data, Y in the data set D will be exactly equal to the predicted value. However, this will not always be the case as Y will differ from a predicted value of Y for values of X. This difference is known as an error of the estimate, or residual value. The major outcome of a regression analysis is to determine the values of the parameters β, such that the sum of the squared residual values is minimized which is known as the least squares regression fit (Du et al, 2004). The least squares regression fit calculates the parameters β based on the following formula:

$$\beta = (X^T X)^{-1} X^T\, Y \qquad \text{(2.3 Least Squares Fit)}$$

The least squares regression fit or line of best fit for the formula 2.2 can be quantified by a correlation coefficient (Naoum et al, 2004). The correlation coefficient is a value used to assess the strength and direction of the linear relationships between variables (Aggarwal et al, 2017, Du et al, 2004, Mukaka, 2012). The general rule of thumb in interpreting the size of a correlation coefficient is described in the following table (Aggarwal et al, 2017):

Table 2.10: Interpretation of the size of a correlation coefficient

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to –1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	negligible correlation

By using linear regression between spatial metrics of separate informal settlements, it is possible to determine whether there is a causal relation of certain metrics in the development of the relative individual informal settlements. This relation may prove to be valuable, as this topic has not been explored yet in the informal settlement context. By understanding how informal settlements develop through spatial metrics, upgrading or formalization of these settlements can use the rate of development of these metrics to guide the approach of upgrading.

2.13. Ordinary Least Squares Analysis

Ordinary Least Squares (OLS) is a statistical linear regression technique by which one or more independent variables and a dependent variable are compared to determine the relation between the both the independent variable(s) and dependent variable (Pohlmann et al, 2003). This multi-variate technique can be used to test the relation of the spatial metrics of the informal settlements (dependent variables) to the macro and micro socio-economic factors (independent variables) within the City of Cape Town. This covers an important aspect of this project since one of the research questions is to find out if any such significant relation exists.

OLS has key determinants to test the relation of the independent and dependent variables. These determinants are given and defined below (McAuliffe, 2015; Thiese et al, 2016):

- t-statistic: the t-test or a resultant t-statistic is a value that is determined when comparing the mean values between two groups of data – the independent and dependent variables (De Winter, 2013; Kim, 2015). This value will aid in determining whether there is a relationship between the two groups of data. The t-test formula is as follows:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{n_1}+\frac{1}{n_2})}} \qquad \text{(2.4. t-statistic)}$$

  where t is the t-value, $\bar{x}_1$ and $\bar{x}_2$ are the respective means of the two groups that are being compared, s^2 is the standard error (statistical accuracy of measurements that is quantified by the standard deviation of a dataset) of the two groups and the n_1 and n_2 values are the number of observations for the respective groups. Higher t-values indicate that a large difference exists between the two groups of data, and the inverse is also true. Together with the p-value, these results are crucial in understanding whether a significant relationship exists between two groups of data.
- p-value: a p-value is a way to measure the probability that a random chance can occur from an observed difference. Therefore, the smaller the value of p, the greater the statistical significance between the different groups of data. This value of p is calculated using integral calculus. However, based on the t-statistic, software can be used to generate values of p and run a probability test to determine the significance between different datasets.

2.14. Summary of the Literature Review

In this chapter, information regarding the establishment factors of informal settlements is outlined. This makes it possible to determine a methodology whereby the detection of these factors is tested for the informal settlements of this investigation. Classification techniques to detect the development of informal settlements are also discussed. OOIC will be used to detect shack image objects. GIS methods will be used to detect movement networks in the informal settlements. The determination of spatial metrics regarding shacks and movement networks will assist in the understanding of the context of these objects in each scene. Change detection techniques outlined in past papers will then be used to deduce how informal settlements have developed in the City of Cape Town. The way the development can be explained will be tested using linear regression analysis. Any meaningful relationship discovered in the statistics from the establishment and development phases will be determined through linear regression analysis and ordinary least squares.

3. Methodology

3.1. Overview of the research methodology

The methodology outlines a framework in which the objectives of this research can be achieved. This framework is broken down into 7 phases, with the aim of collectively determining the change of internal settlement dynamics and development (objectives A and B), what factors are the most statistically significant in describing the development in each (local level) and across all (global level) targeted informal settlements (objective C), and the investigation on how certain macro and micro socio-economic factors have been more significant than others in influencing informal settlement dynamics (objective D).

Phase 1 involves the collection of available aerial imagery which contains the informal settlements. This data includes imagery where the same epochs were investigated across all informal settlements. Phase 2 involves pre-processing of the available aerial imagery. This process involves outlining the extent of the informal settlements, to maximize the efficiency of data processing in the following phases. In phase 3, a supervised object-oriented image classification was conducted. The purpose of this phase is to accurately produce image objects of shacks in each informal settlement, and across the available epochs. Once a sufficient level of accuracy is achieved in the OOIC, phase 4 was conducted. Phase 4 involves two processes. Firstly, the determination of spatial metrics of the shacks in each informal settlement across all epochs is conducted. Using the spatial metrics determined, the internal development dynamics for each informal settlement was investigated. And secondly, linear regression analysis of the spatial metrics was conducted. The process of linear regression analysis related metrics to provide a deeper insight to spatial trends, such as the rate of development, the density, and complexity of shacks across all epochs. At phase 5, the process of change detection by using GIS analysis was conducted. This phase was used to add additional insight in determining spatial trends within the internal boundaries of the informal settlements investigated. Phase 5 partially achieves objective C. Phase 6 involves two separate processes. Firstly, statistical trend analysis of macro and micro socio-economic data was determined. Secondly, the statistical trends determined were compared to the spatial trends determined in phase 5. This comparison was conducted using an Ordinary Least Squares (OLS) approach. The results and analysis of the OLS of individual settlements were then compared to each other in phase 7. Phase 7 determines whether there are macro and micro socio-economic factors that are more significant than others when related to the development of informal settlement. This final phase achieves objectives C and D. A practical overview into achieving all three objectives is illustrated in the Figure 3.1:

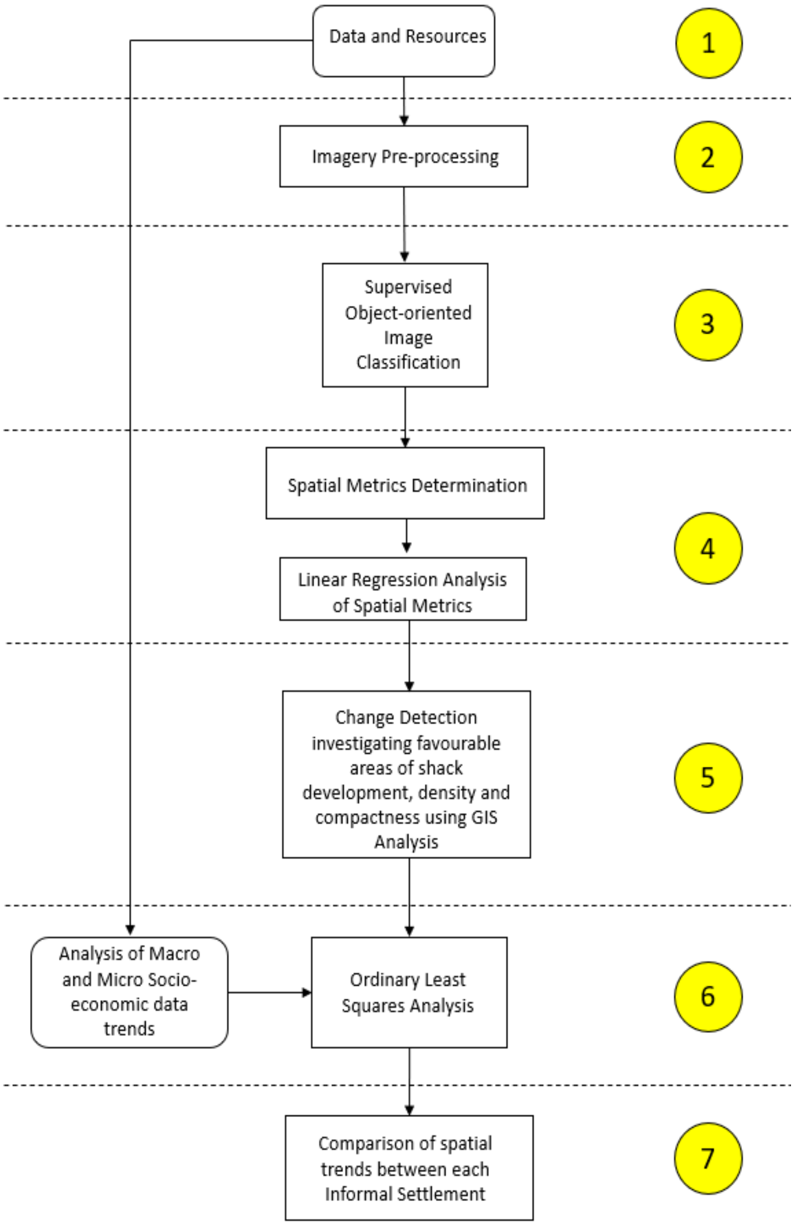


Figure 3.1: Overview of Methodology

3.2. Phase 1: Data Acquisition and Resources

The data required to fulfill the research methodology objectives includes that of high-resolution aerial imagery, attributive spatial data, and software to conduct the necessary processes to produce results and analyze those results. These resources are explored further in the following:

3.2.1. Aerial Imagery and Study Area

Aerial imagery has been made available by the University of Cape Town (Geomatics Division in the Engineering and Built Environment faculty) for the years 2011, 2014, 2017, and 2019. Each image has a spatial resolution of 0.08m and is geo-referenced. Each image also contains and is limited to the three RGB bands. The aerial imagery was initially acquired from the City of Cape Town. The informal settlements that are analyzed are contained within the blocks corresponding to the national grid (W36C11, W36C12, W45D17, W45D18, W56C7, W56C12).

The sites investigated are Imizamo Yethu, Joe Slovo (Langa south-east) and Siqalo informal settlements. Their location within the City of Cape Town is mapped with respect to the CBD in the following figure:

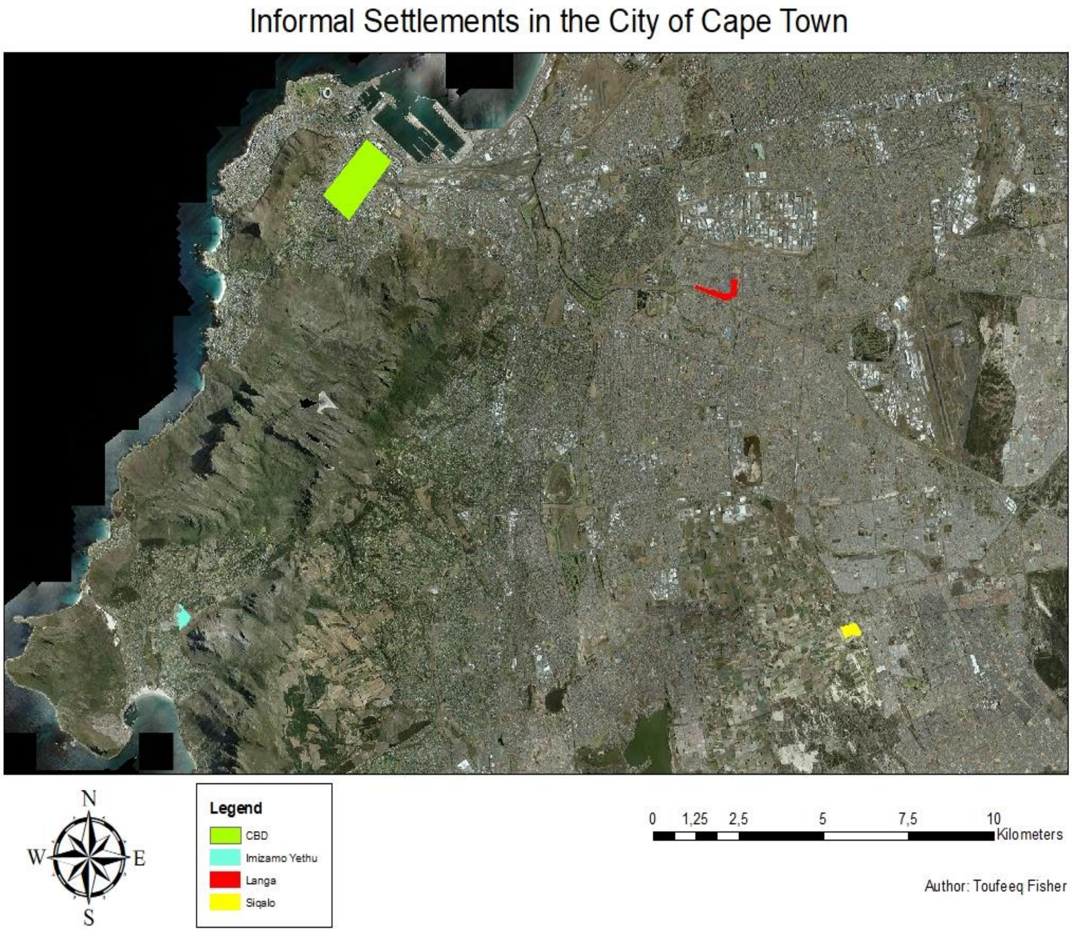


Figure 3.2: Selected study area sites of informal settlements relative to CBD in the City of Cape Town

a) Imizamo Yethu (east) Informal Settlement

Imizamo Yethu (east) is situated in Hout Bay, on the western side of Table Mountain in the City of Cape Town. It is approximately 22 km away from the CBD. The settlement was established in 1991. It has a rich and tragic history in the city, as it is infamously known for being constantly prone to fire breakouts (Smith et al, 2005). This is due to the combination of the general susceptibility of fires in informal settlements and the famous south-easterly wind that is trapped along the hills of the settlement (Harte et al, 2009). Another distinct quality of this Imizamo Yethu is that the settlement is notoriously known for the lack of basic services and infrastructure available to those who reside in it.

Figure 3.3: Imizamo Yethu Informal Settlement

b) Joe Slovo (Langa south-east) Informal Settlement

Langa is, geographically, the closest informal settlement to the CBD of the City of Cape Town - approximately 12 km. The advantage of this location lies in providing an attractive destination for people with a low socio-economic status in search for jobs. Langa is also the oldest black township in the City of Cape Town – having been established in 1990 (Jacobs et al, 2015). The informal settlement located parallel to a national road (National Road 2), a canal and within proximity to an industrial area (containing employment opportunities) and other services. According to the available aerial imagery, upgrading has partially taken place within Langa south-east, therefore it has the distinct quality of being useful in understanding how formalization has changed the physical dynamics of the informal settlement.

Langa Informal Settlement

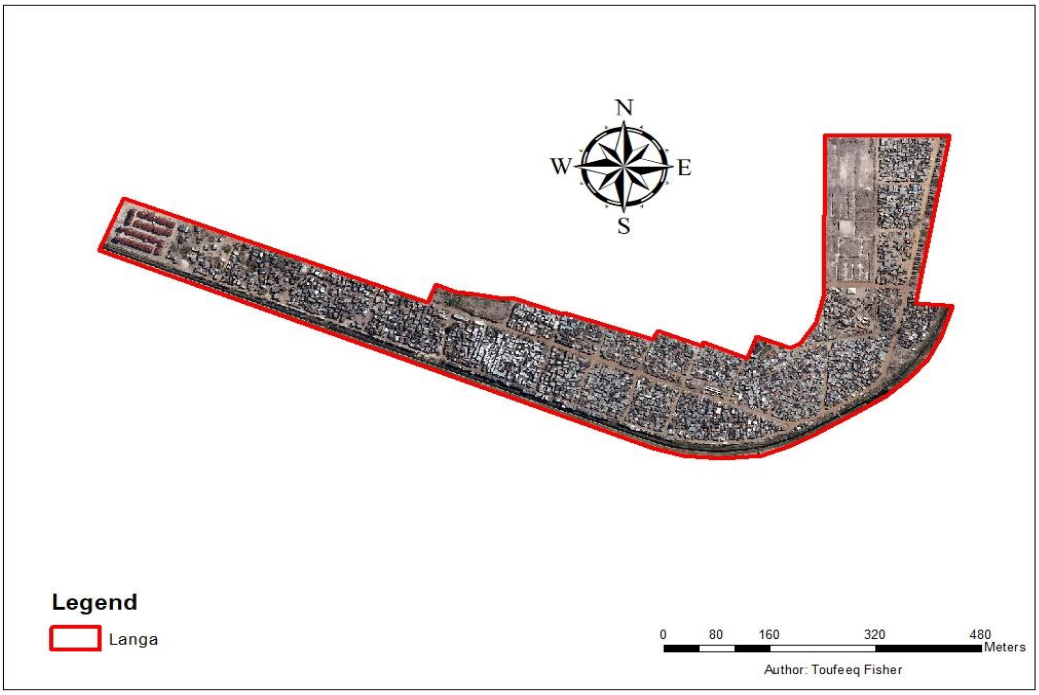


Figure 3.4: Langa Informal Settlement

c) Siqalo

Siqalo offers a different perspective to observing informal settlement development within the City of Cape Town. Geographically, Siqalo is considered quite far from the CBD (approximately 20 km). However, the informal settlement does have the added challenge of being located along routes containing heavy traffic to and from the CBD (unlike Imizamo Yethu which is roughly the same distance away from the CBD). The settlement is located along a major road on its east side - Jakes Gerwal Drive and agricultural land on the remaining sides. There are also clear internal movement networks within the settlement.

Siqalo Informal Settlement

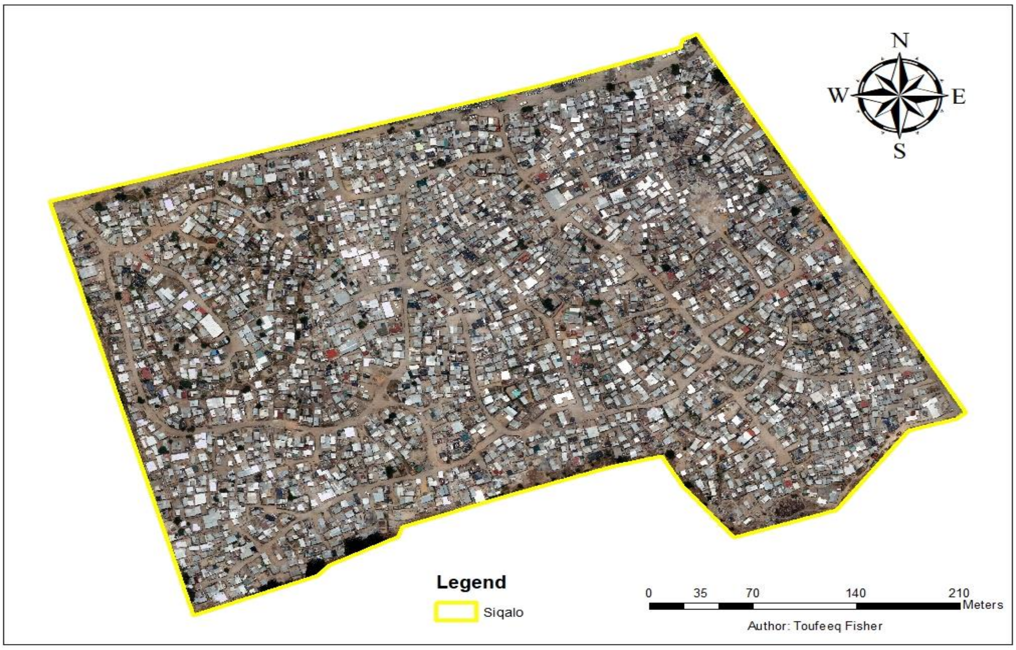


Figure 3.5: Siqalo Informal Settlement

3.2.2. Attributive Spatial Data

The use of attributive spatial data has its use in this investigation in the determination of establishment factors for informal settlements. The following outlines the data that was of use:

- Google Maps Service data – The data included from this source includes shapefiles in which service locations are mapped.
- Open data portal – This source provides DEMs which describe height data for the City of Cape Town.
- Census 2011 – The 2011 Census data is the most comprehensive set of in-depth statistics that is available for the past decade with respect to income, population, and demographic statistics.

3.2.3. Software

The software choices that were used in the process of this investigation is based upon previous research:

- ArcGIS 10.5.1 was used to conduct contextual analysis to determine the spatial factors relating to the establishment of the informal settlements in the City of Cape Town.
- SuperCROSS 4.5.5 was used to query the Census 2011 data.
- eCognition 9.0 was used to conduct the OOIC to detect and classify shacks in aerial imagery. ArcGIS 10.5.1 was also be used to conduct the pre-processing, local shack feature detection, and accuracy assessment as part of the OOIA.
- Fragstats 4.2 was used for spatial metric determination of shacks for each informal settlement.

3.3. Phase 2: Image Pre-processing

Image pre-processing is required to extract each informal settlement from the aerial imagery. This imagery has been georeferenced. The coordinate system in which the imagery is geo-rectified follows the official system for South Africa – The Gauss Conform coordinate system. This coordinate system is based on the World Geodetic 1984 ellipsoid (commonly known as WGS84), hence any processing conducted with two or more images will be valid. To extract each informal settlement in each year of the available imagery, ArcGIS 10.5.1 is used. The following process is used to extract each informal settlement from the aerial imagery:

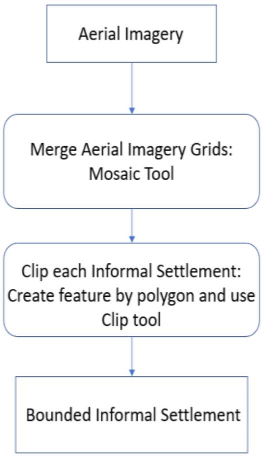


Figure 3.6: Aerial Image Pre-processing

Since each informal settlement is not bounded within one grid of the sourced aerial imagery, grids were combined to describe a bounded informal settlement. This is possible in ArcGIS 10.5.1 using the mosaic function. In this investigation, two grids were merged for each informal settlement: Imizamo Yethu (W56C11 and W56C12), Langa (W45D17 and W45D18), and Siqalo (W36C11 and W36C12).

To extract each informal settlement from the merged aerial imagery, polygons were manually created which capture the perimeter of each settlement. Using the mask tool in ArcGIS 10.5.1, the clipped area of each informal settlement for the years 2011, 2014, 2017, and 2019 were produced. The results can be viewed in the figures depicting each settlement in the study area section (3.3) above.

3.4. Phase 3: Supervised Object-oriented Image Classification of Informal Settlements

The classification model used to conduct the classification is a supervised classification as the class classified is known beforehand – with regards to shacks (Sathya el al, 2013). The classification that was conducted using the eCognition 9.0 software as it has been established that it optimally classifies objects in aerial imagery from available software packages (Baatz et al, 2004). The steps in conducting an OOIC in eCognition involve processing a MR segmentation on the aerial imagery, feature detection for the class classified, as well as the classification processing to determine the shacks in the scene (Hofmann et al, 2017; Karakis et al, 2006; Kohli et al, 2016).

3.4.1. Multi-resolution Segmentation

Since each scene for each informal settlement has the same spatial resolution of 0,08m, the same parameters for a scene can be used across each epoch. The parameters that are user-specific are the scale parameter, shape, and compactness. Scale parameter contributes to the overall size of the expected fragments. The lower the scale parameter, the larger the resultant fragmented segments. The converse also applies. The shape works in correlation to the brightness values in the imagery. The lower the shape value, the higher the influence of RGB band values when segmenting an image. The shape definition in eCognition refers to the ratio of the total fragment border length to the square root of the number of pixels in a fragmented segment. Compactness refers to the overall compactness of the resultant fragments. The lower the compactness value, the less compact the resulting fragment will be (Baatz et al, 2004; Imani et al, 2020). To achieve a level of over-segmentation, through a trail-and-error process, the following parameters were used for each informal settlement (Zhang et al, 2010):

Table 3.1: MR Segmentation Parameters used for each Informal Settlement

Informal Settlement	Scale Parameter	Shape	Compactness
Imizamo Yethu	56	0.1	0.7
Langa	60	0.1	0.5
Siqalo	52	0.1	0.6

Since there are 11 aerial images were processed (Siqalo 2011 contains no shacks), 11 pre-processed images were segmented. An example of the resulting segmented imagery can be viewed for each informal settlement below:

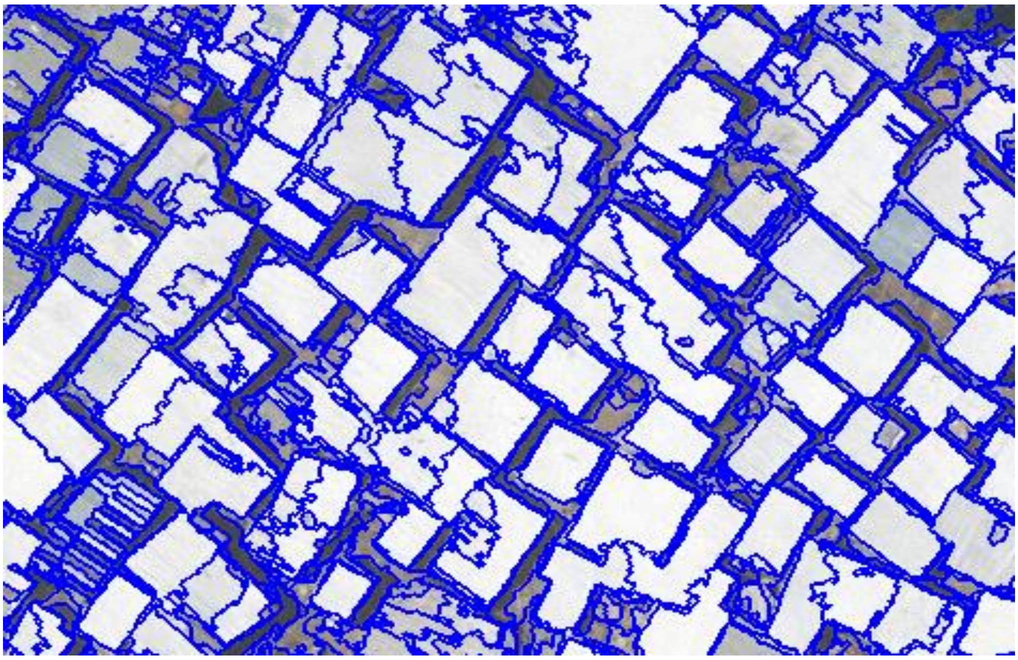


Figure 3.7: Segmented Image of Imizamo Yethu (focused onto the center of the image)

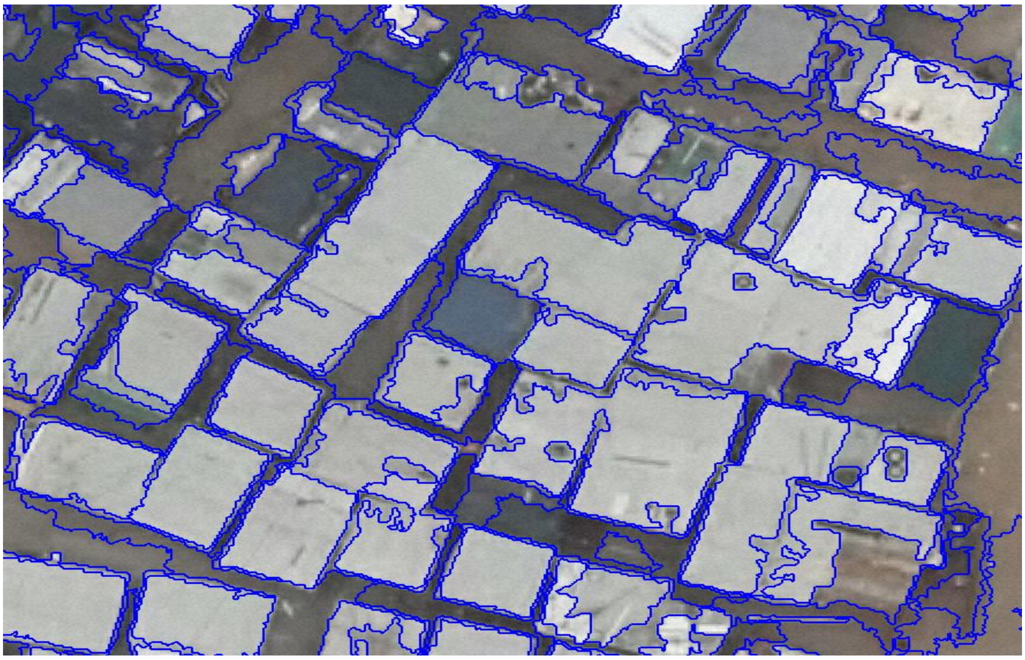


Figure 3.8: Segmented Image of Langa (focused onto the center of the image)

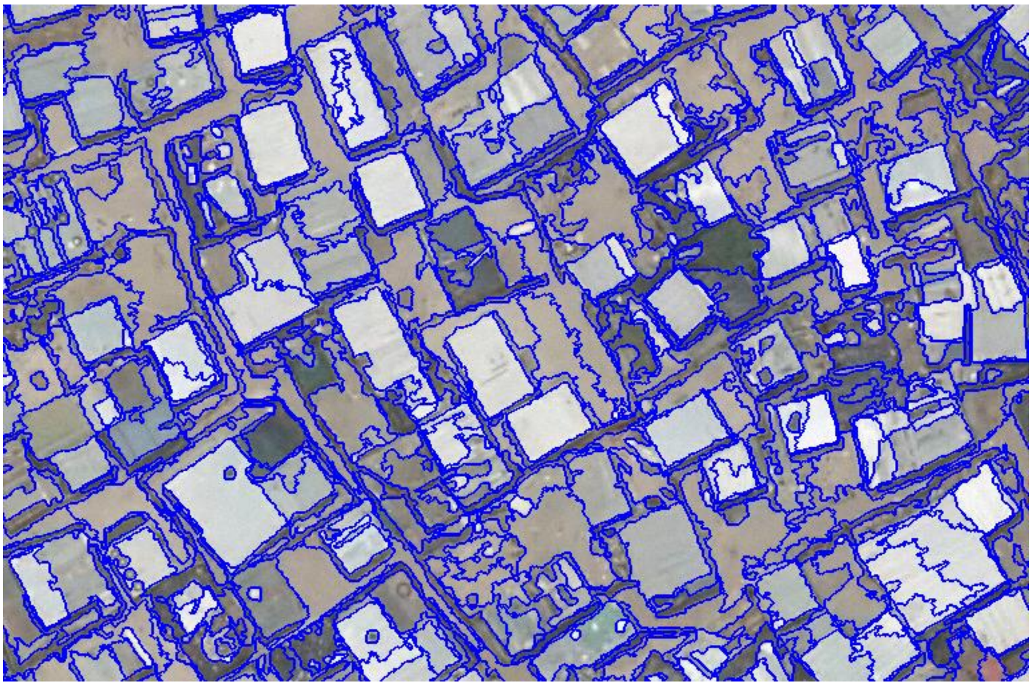


Figure 3.9: Segmented Image of Siqalo (focused onto the center of the image)

3.4.2. Feature Detection

In the OOIC process, once the MR segmentation has produced the necessary acceptable results, the next step is to conduct a feature detection process in eCognition. Feature detection is the description of features that contributes to the retrieval of meaningful objects in an image (Tian et al, 2007). As has been established in past pieces of literature, features that accurately depicts shacks as image objects in a scene include the geometry of shacks, the RGB band values of the roofs of shacks in aerial imagery, the dense nature of shacks in a scene and the differentiation of vegetation and shadows relative to shacks (Grippa et al, 2018; Hofmann et al, 2001; Kohli et al 2016; Kohli, 2013; Mason et al, 1997; Shekhar, 2012). However, in those pieces of literature, the determination of features for the feature detection process is motivated by the local ontology. Using a combination of features used in past literature and in the local ontology will, therefore, prove more advantageous in classifying shacks in the informal settlements in this project. After visually analyzing each informal settlement, it has been determined that the scenes are similar in their composition of shacks. Therefore, the local ontology will be observed using only one image – that of Siqalo in 2017. This image contains all the features that have been described in past literature. To conduct an analysis of the local ontology, the software ArcGIS 10.5 will be used. To extract the necessary features of shacks, the digitization of the objects will be conducted and analyzed with respect to their RGB band value ranges as most features described in past literature contains the use of this (Kohli et al, 2016). The shack area information and total shack count will also be included as part of the analysis. The following digitized image was produced:

Figure 3.10: Digitized shacks of Siqalo 2017

The RGB band value data is then masked by the digitized shacks to produce the following:

Figure 3.11: RGB band value extraction using digitized shacks

The following information was extracted from the red, green, and blue band values of the shacks in the scene:

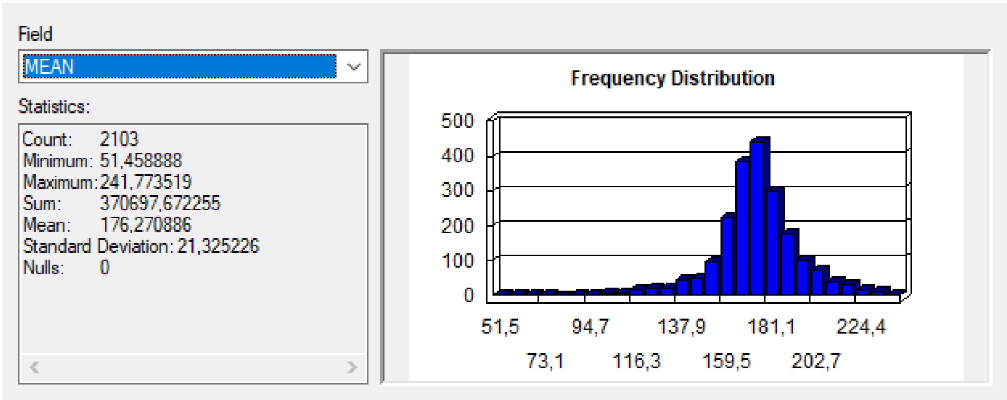


Figure 3.12: Statistical distribution of the Mean Red Band Values of Shacks in Siqalo 2017

The range of the red band value within the digitized shacks is roughly between 51 and 242, with the average being 176 (standard deviation 21). Although the range is quite varied, there is a dense range between 140 and 220 for the red band value.

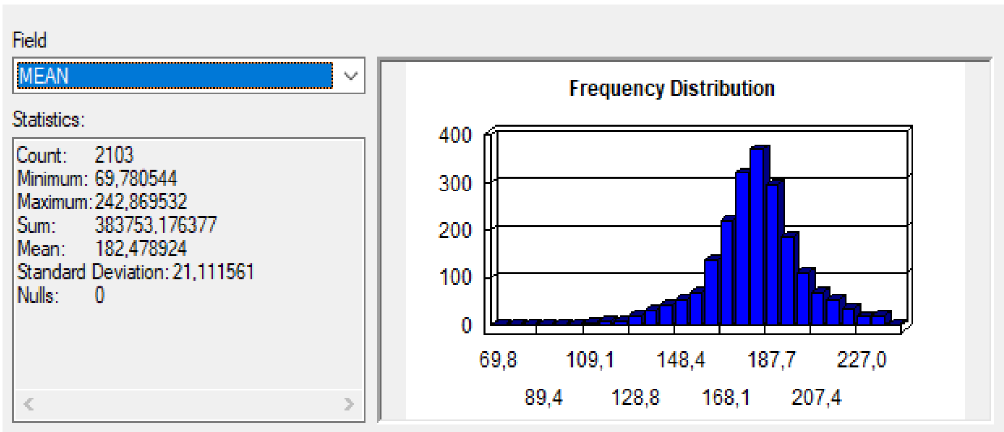


Figure 3.13: Statistical distribution of the Mean Green Band Values of Shacks in Siqalo 2017

The range of the green band value within the digitized shacks is roughly between 70 and 243, with the average being 182 (standard deviation 21). Although the range is quite varied, there is a dense range between 120 and 220 for the blue band value.

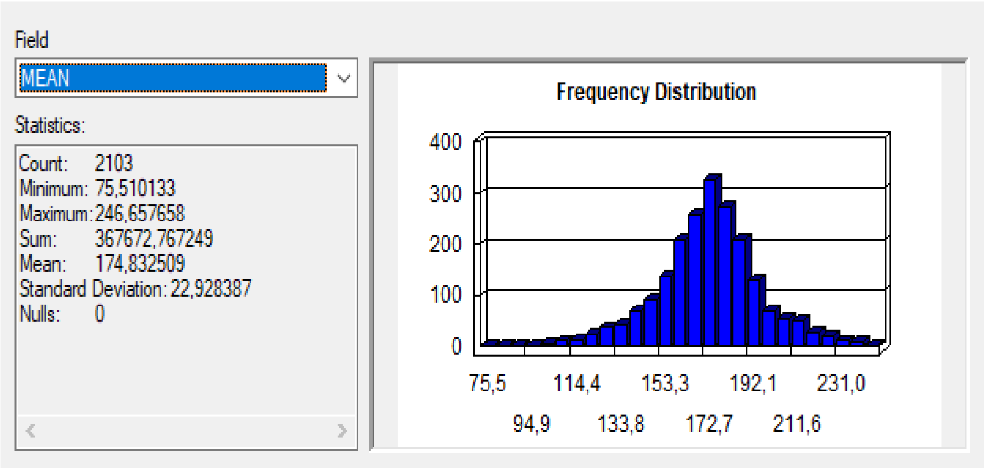


Figure 3.14: Statistical distribution of the Mean Blue Band Values of Shacks in Siqalo 2017

The range of the blue band value within the digitized shacks is roughly between 75 and 247, with the average being 174 (standard deviation 23). Although the range is quite varied, there is a dense range between 115 and 220 for the blue band value.

The following shack area information was extracted from the shacks in the scene (the total shack count is also represented in the figure – 2103 shacks):

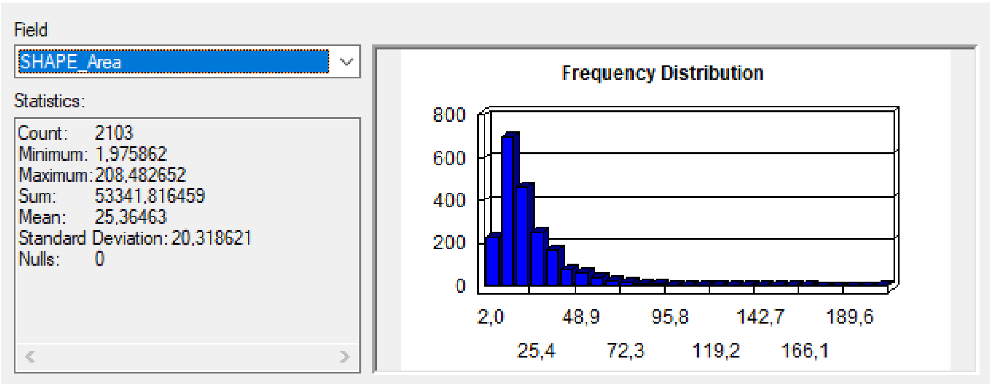


Figure 3.15: Statistical distribution regarding the area of shacks in Siqalo 2017

The average shack area in Siqalo during 2017 is 25m^2, with a standard deviation 20. This knowledge may be used as a feature in the feature detection process of shacks in informal settlements.

Qualitatively, it has been observed through the digitization process that the shacks are rigid polygons or rectangular in nature. This observation will allow to conclude that the geometry of shacks will greatly influence the accuracy of the classification. This is supported in previous research (Mason et al, 1997). Another observation has been made relating to the density of the shacks. Through the digitization of shacks, the density that has been observed is high and proves to be a useful feature in accurately classifying shacks. Previous research supports this point as well (Kohli, 2013).

By using the combination of the features addressed in past literature and the local ontology, image segments were classified based on this resulting feature space:

- Red band values with the range 140 – 220 (local ontology analysis)
- Green band values with the range 120 – 220 (local ontology analysis)
- Blue band values with the range 115 – 220 (local ontology analysis)
- Mean area value of $25m^2$, with a standard deviation of 20 (local ontology analysis)
- Rigid geometric shape (local ontology analysis)
- Shadow RGB band values greater than 115 (minimum of either red, green, or blue range) (Grippa et al, 2018; Kohli, 2013; Shekhar, 2012)
- Minimize vegetation values using band ratio – Green/(Red+Green+Blue) (Kohli, 2013; Mason et al, 1997; Shekhar, 2012)

3.4.3. Classification Processing

The classification procedure in eCognition made use of the nearest neighbour classifier to classify image objects. This classifier uses a combination of user-specified training samples of the object to be classified and specified features in the feature detection process (Cunningham et al, 2007). This process is optimized using the feature space optimization tool in the classification process. The feature space optimization function provides a method to combine the best of the features given in the feature detection processing to accurately classify image object (Baatz et al, 2004). Since this is a supervised OOIC, the object classified was known before the OOIC – shacks. This is the only image object classified in the classification. The background represented the only other class in the imagery.

3.4.4. Accuracy Assessment

To express the quality and the validity of the classification produced in eCognition, an accuracy assessment was conducted in ArcGIS. The accuracy was defined as a function of the percentage of the points that is correctly classified when comparing the classified imagery to that of the original imagery (Congalton et al, 1986; Patil et al, 2013). The extent to which the accuracy was determined is found in the production of a confusion matrix or contingency table (Faweet, 2006; Stehman, 1997). In the confusion, a kappa coefficient was determined, which represents the overall accuracy of a classified image (Fletcher et al, 2011; Navulur, 2007). The calculation of this kappa coefficient (k) and the definition of the range of kappa values was covered in the literature review.

3.5. Phase 4: Spatial Metrics and Linear Regression Determination

The spatial metrics determined in this investigation will describe the spatial context of shacks in each informal settlement for the years 2011, 2014, 2017, and 2019. This process was conducted in Fragstats 4.2 by assessing the classified rasters displaying shacks and the background data. The specific metrics targeted for the analysis in change of shacks over time were outlined in the literature review.

The linear regression analysis for this project investigated the correlation of metrics relating to the development, complexity, and compactness of shacks in informal settlements over time. The process of linear regression analysis was conducted in Microsoft Excel by comparing explanatory metrics to response metrics. Explanatory metrics are intended to provide reasoning to the response metrics. The following will be relationships were investigated for each informal settlement site:

a) Development:
   - NP vs CA – the number of shacks classified in each epoch to the total area covered by shacks in each epoch.
   - MN AREA vs CA – the mean area of shacks in each epoch to the total area covered by shacks in each epoch.
   - MN ENN vs CA – the average distance between shacks in each epoch to the total area covered by shacks in each epoch.
   - MN SHAPE vs CA – the average shape complexity of shacks in each epoch to the total area covered by shacks in each epoch.

b) Complexity of shacks:
    - MN ENN vs MN SHAPE – the average distance between shacks in each epoch to the mean shape of shacks in each epoch.
    - NP vs MN SHAPE – the number of shacks classified in each epoch to the mean shape of shacks in each epoch.
    - MN AREA vs MN SHAPE – the mean area of shacks in each epoch to the mean shape of shacks in each epoch.

c) Compactness of shacks:
    - MN AREA vs MN ENN – the mean area of shacks in each epoch to the average distance between shacks in each epoch.
    - NP vs MN ENN – the number of shacks classified in each epoch to the average distance between shacks in each epoch.

3.6. Phase 5: Change Detection Analysis

Since spatial metrics describe the context of shacks in each epoch, the change of these metrics across consecutive epochs describes the development of shacks in the selected informal settlements. Metrics that were most useful in describing change in informal settlements include the percentage of landscape covered by shacks, the number of shacks, mean shack area, shack density, mean shack shape index, and mean shack nearest neighbour were analyzed. Any other metric has been deemed synonymous in definition when compared to one of the outlined selected metrics to describe shack development.

Although these metrics define the development of shacks in informal settlements to the most descriptive degree, there is a major consideration to make when describing the development of an informal settlement. This consideration relates to the extent of the informal settlement. When assessing the internal geographic factors relating to the extent of informal settlements, the available land to which shacks develop may be limited or extended over the period of development. The only metric that this affects is the percentage of landscape covered by shacks. However, this metric was the most important when determining the rate of change in informal settlements. It is then recommended to observe if the extent of the informal settlements has changed and compute this metric to relate to the changed informal settlement extent.

To gain an understanding as to why and how informal settlements are established in the locations it occupies, a neighbourhood suitability analysis may be performed. As has been determined in past pieces of literature, push and pull factors exist which may explain the establishment of informal settlements (Dovey, 2015; Dubovyk et al, 2010; Kohli, 2013; Kohli et al, 2016; Kuffer et al, 2011; Naorem et al, 2016). The results of a neighbourhood suitability analysis in terms of informal settlement establishment may prove useful when determining why informal settlements develop in the manner it does. A study found that the development of shacks in informal settlements have grown greater with respect to major transport networks relative to other factors (Kohli, 2013). Another study found that the development of informal settlements skewed towards employment opportunities (Wekesa et al, 2010). The aim of the neighbourhood suitability analysis is to provide the most relevant factors which may explain why and how the informal settlements develop in the context of this investigation.

By assessing the neighbourhood factors of informal settlements, the context relating to why informal settlements establish in the locations they occupy may be investigated. Subsequently, these major factors may also play a role in determining why informal settlements develop at the rate that they are. For this investigation, these factors will be analyzed in both the external and internal capacity of the informal settlements:

1) External socio-economic factors relating to the establishment and development of informal settlements.
    - It is well documented that the major reason for the establishment and development of informal settlements is rural-to-urban migration for people of low socio-economic status (Manal et al, 1998; UN-Habitat, 2020). The motivation for this migration is for the poor to find residency within a reasonable capacity to employment opportunities (Augustijn-Beckers et al, 2011). Since informal settlements host most rural-to-urban migrators, a neighbourhood factor to investigate is the distance to employment opportunities. The City of Cape Town recognizes the business districts or industrial areas as the main source of employment opportunities. 26 such districts are currently in existence around the city. ArcGIS 10.5.1 will be used to assess the distances from each informal settlement to these employment opportunities.
    - Another major reason for the establishment and subsequent development of informal settlements is the inability for the urban poor to afford housing near areas of employment opportunities (Augustijn-Beckers et al, 2011; Dovey, 2015). The 2011 Census and the 2016 Community Survey contains data at the ward level of Cape Town, which describes the salary of formal settlement residents of the areas each informal settlement investigated exists. Since it has been established what people in informal settlements earn in the literature review, a comparison of their income statistics to the formal neighbourhood may be analyzed to determine whether it is possible for rural-to-urban migrators to afford housing near employment opportunities. To conduct this comparison, ArcGIS 10.5.1 will be used to investigate the 2011 Census and 2016 Community Survey data.
    - Although the first point in investigating the external socio-economic is viewed as the most important proximity factor, other proximity factors exist which may provide clarity as to why informal settlements are established (Dovey, 2015; Kohli, 2016; Kuffer et al, 2011). These proximity factors include distances to major transport networks (availability for public transport use), distances to educational facilities (availability for schooling the urban poor), and distances to basic amenities (availability of amenities to provide for basic needs). ArcGIS 10.5.1 will be used to assess these distances from informal settlements to transport networks, educational facilities, and basic amenities.

    Since most of the external socio-economic factors relating to the establishment and development of informal settlement are proximity metrics. The distances to each external socio-economic factor will be analyzed with respect to the selected informal settlements.

2) Internal geographic factors relating to the development extent of informal settlements.

    The importance of the land informal settlements occupy is crucial in determining the rate at which informal settlements develop (Hofmann et al, 2017). As shacks represent the objects the urban poor choose to create as a housing option in aerial imagery, its change with respect to the available land will describe the rate informal settlements develop. This available land is subject to limitations in a geographic and built environment context (Kohli et al, 2016). It then becomes necessary to investigate what these limitation factors are to truly determine the rate at which informal settlements develop. Past research suggests that the following factors limit the development extent of informal settlements (Dubovyk, 2010; Kohli, 2013, Kuffer et al 2011):
    - Water bodies such as rivers, canals, lakes, and damns,
    - Extreme steep slopes,
    - Built formal movement networks, and
    - Built formal housing.

ArcGIS 10.5.1 can be used to assess the extent of informal settlement boundaries with respect to the actual internal land informal settlements occupy. By analyzing each aerial image, the factors that limit the development of informal settlements can be observed. DEM data (from Open Data Portal) providing heights of the internal and external informal settlement surface can be used to compare the suitability of land to build shacks (Hofmann et al, 2011; Kohli et al, 2016; Kuffer et al, 2011). Any other physical feature that can be detected using aerial imagery will be created manually through digitization, as it is a pragmatic solution compared to detecting too many objects by means of machine objects.

Once the DEM data and physical features are mapped, by analyzing what is in the informal settlement and what's neighbouring it, the true extent of the informal settlement can be established. As has been recommended when discussing spatial metrics, the percentage of shacks occupying land will then be represented to its actual extent. Any subsequent analysis of the rate at which informal settlements develop can therefore be trusted.

3.7. Ordinary Least Squares (OLS) Analysis

The OLS determines if there were any statistically significant relationships between socio-economic factors and development in informal settlements. The data selected to represent the socio-economic factors include:

- The extent of open space available for the development of each informal settlement. This is the area within each informal settlement that does not contain shacks (percentage open space).
- The rate of unemployment in the City of Cape Town for each classified year (percentage unemployment).
- A poverty coefficient for the City of Cape Town.
- GDP for the City of Cape Town.

The socio-economic factors mentioned above were selected due to the availability of data during the investigative period. Since the internal dynamics of development has been established, the OLS related the development to that of the selected socio-economic factors.

This process of OLS will produce p-values, in which the selected socio-economic statistics can be tested against the informal settlement development statistics. The lower the p-values in the OLS, the more statistically significant a socio-economic statistic will be in explaining the development statistic in each informal settlement.

3.8. Phase 7: Comparison of Spatial Trends between each investigated Informal Settlement

At this point, the development of each investigated informal settlement has been determined through change detection. The relationship between the development and the outlined socio-economic factors was also determined in the OLS. Comparisons were then made between each informal settlement. This was an overview of the significant factors related to the development across all informal settlements. A comparison table was used to highlight corresponding spatial trends from all prior analysis on significant socio-economic factors that drove the development of informal settlements in Cape Town.

3.9. Summary of the Methodology

The methodology outlined the process that had to be followed to accomplish the objectives of this project. Given remotely sensed aerial image data from the City of Cape Town and using machine learning techniques, shacks within targeted informal settlements can be classified. Through using change detection techniques, the development of shacks can be monitored and quantified by determining the spatial metrics of shacks across each informal settlement at the available epochs over the investigative period (2011-2019). The spatial metrics can then be analyzed to find any significant patterns in the development of the individual informal settlements. After determining if any significant patterns exist, they can be compared to that of selected socio-economic data to determine whether there is a significant statistic that be used to explain what drives the development of informal settlements.

4. Results, Analysis and Discussions

4.1. Overview of Results, Analysis and Discussions

The results for this research include classified imagery in which shacks are detected in each informal settlement in the years 2011, 2014, 2017, and 2019, the accuracy of each classification quantified by the kappa co-efficient, the spatial metrics determined to describe shacks in each informal settlement at the given epochs, and the socio-economic data used in the OLS analysis relative to developmental statistics.

The analysis of the results and any subsequent discussions that can be made with respect to the findings of this investigation are also found in this chapter. Firstly, the change for the selected informal settlements were analyzed with respect to shack development at the class level. One of the objectives of this investigation is to understand how informal settlements develop. By detecting and analyzing the change of shacks, the objective of determining how informal settlements develop can be met. Shacks were detected to a level of accuracy that is sufficient for any subsequent analysis using OOIA methods across the range of epochs 2011, 2014, 2017, and 2019. By computing spatial metrics describing the spatial patterns the shacks follow at the class level, the development of informal settlements was analyzed. The spatial metrics that were analyzed are the Percentage of Landscape covered by shacks, the Number of Shacks, the Mean Shack Area, the Shack Density, the Mean Shack Shape Index and the Mean Shack Nearest Neighbour. Secondly, once each of these metrics have been analyzed to explain how shacks have developed in informal settlements, the development of shacks can be related to the neighbourhood factors determined in the results. The neighbourhood factors include the average height at which shacks occupy in each epoch, the distance to employment opportunities, the distance to health care services, the distance to educational facilities, and the distance to amenities. By analyzing the growth within informal settlements to the neighbourhood factors, the reason why informal settlements were established where they are may become apparent – which is another objective of this investigation. And thirdly, any relationships that have been deduced from the first two objectives were discuss the relationship between the selected informal settlements during the period of 2011 and 2019.

4.2. Results of Object-oriented Image Classification and Accuracy of Classifications

In this section, the classification of each informal settlement within the investigative period were outlined. The accuracies of each classified image were also investigated to test for their validity in future analysis.

4.2.1. Object-oriented Image Classification of Informal Settlements

The classification of each informal settlement for the years 2011, 2014, 2017, and 2019 were mapped in the following:

a) Imizamo Yethu

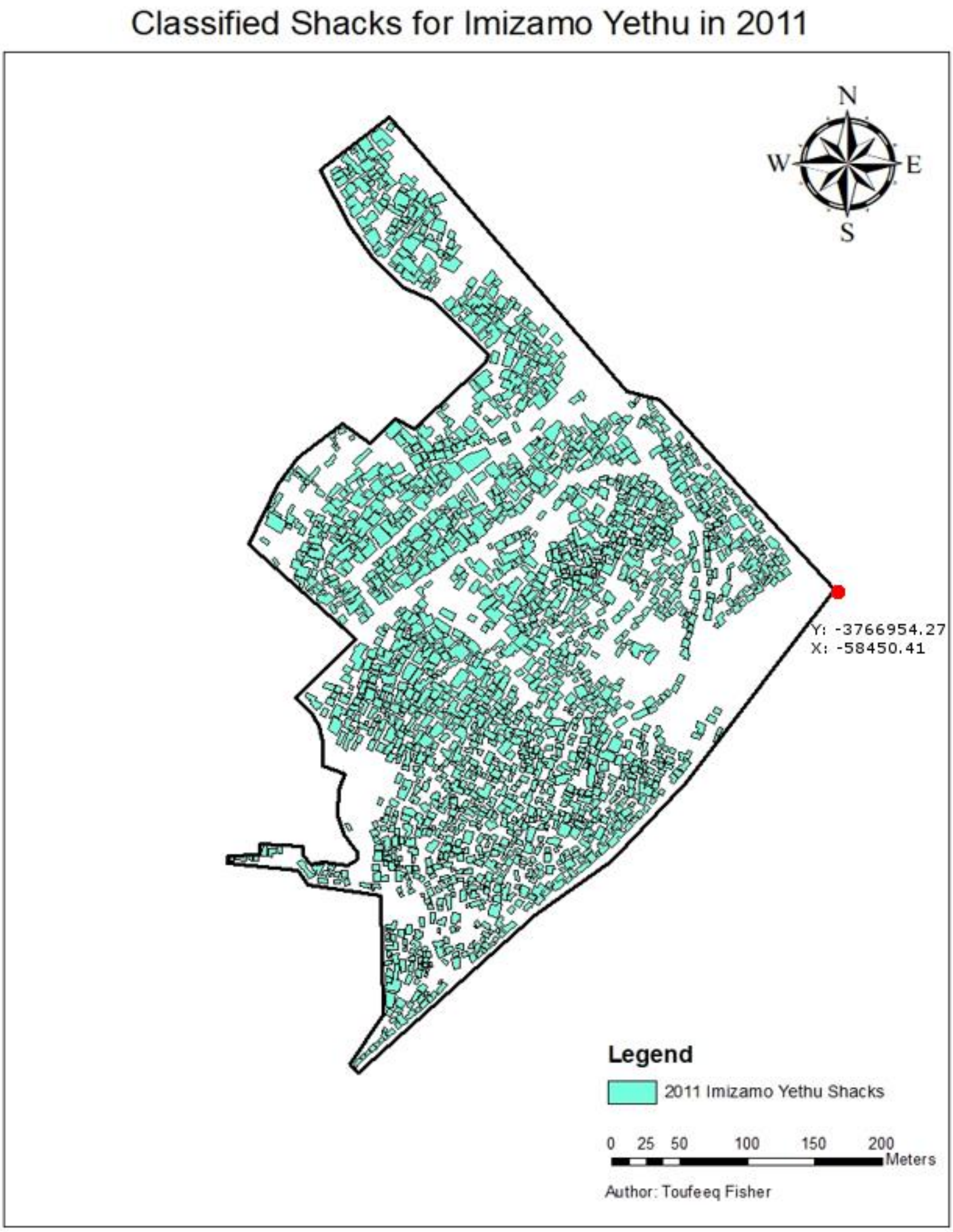


Figure 4.1: Classified Shacks for Imizamo Yethu 2011

Classified Shacks for Imizamo Yethu in 2014

Figure 4.2: Classified Shacks for Imizamo Yethu 2014

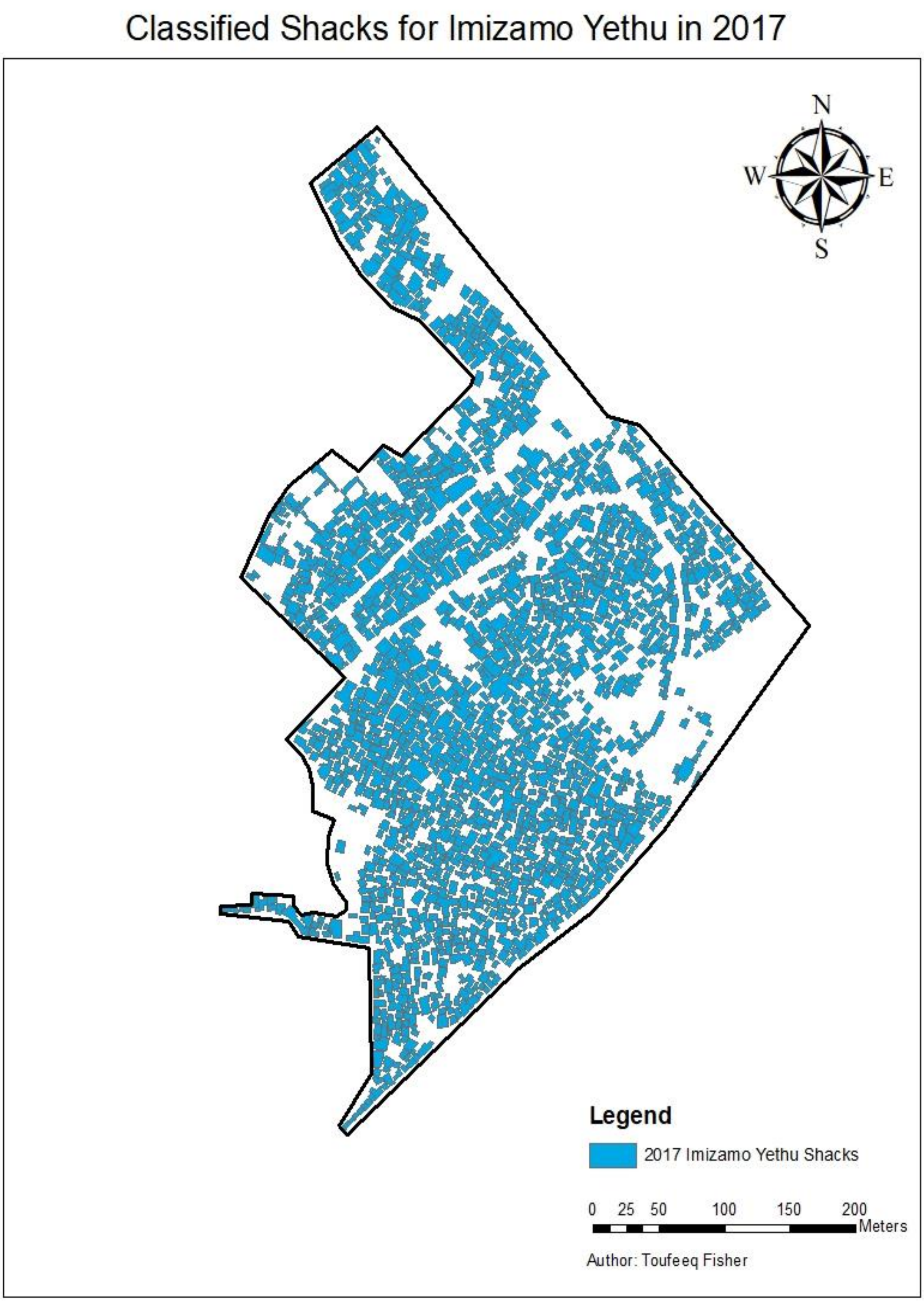


Figure 4.3: Classified Shacks for Imizamo Yethu 2017

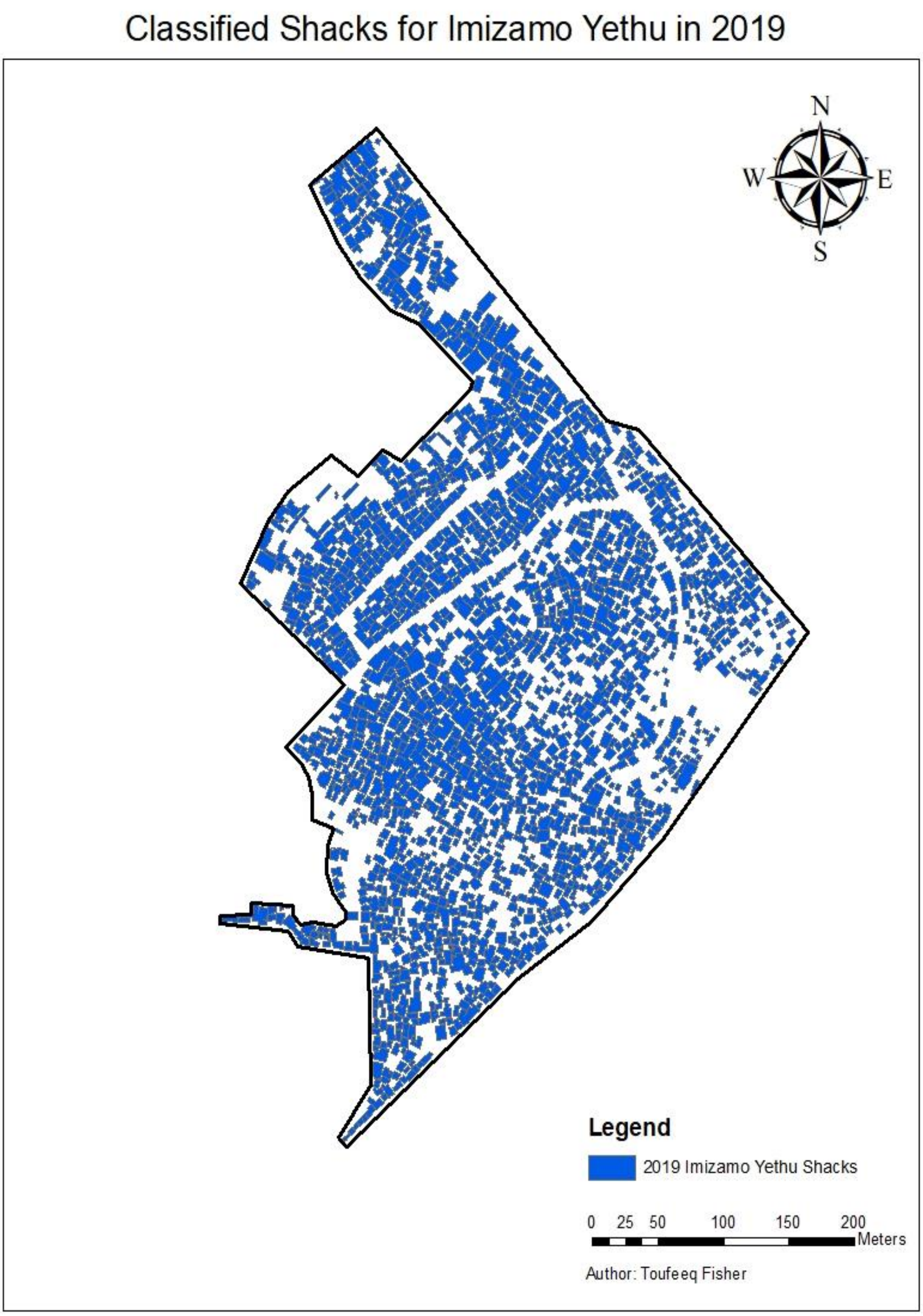


Figure 4.4: Classified Shacks for Imizamo Yethu 2019

b) Langa

Figure 4.5: Classified Shacks for Langa 2011

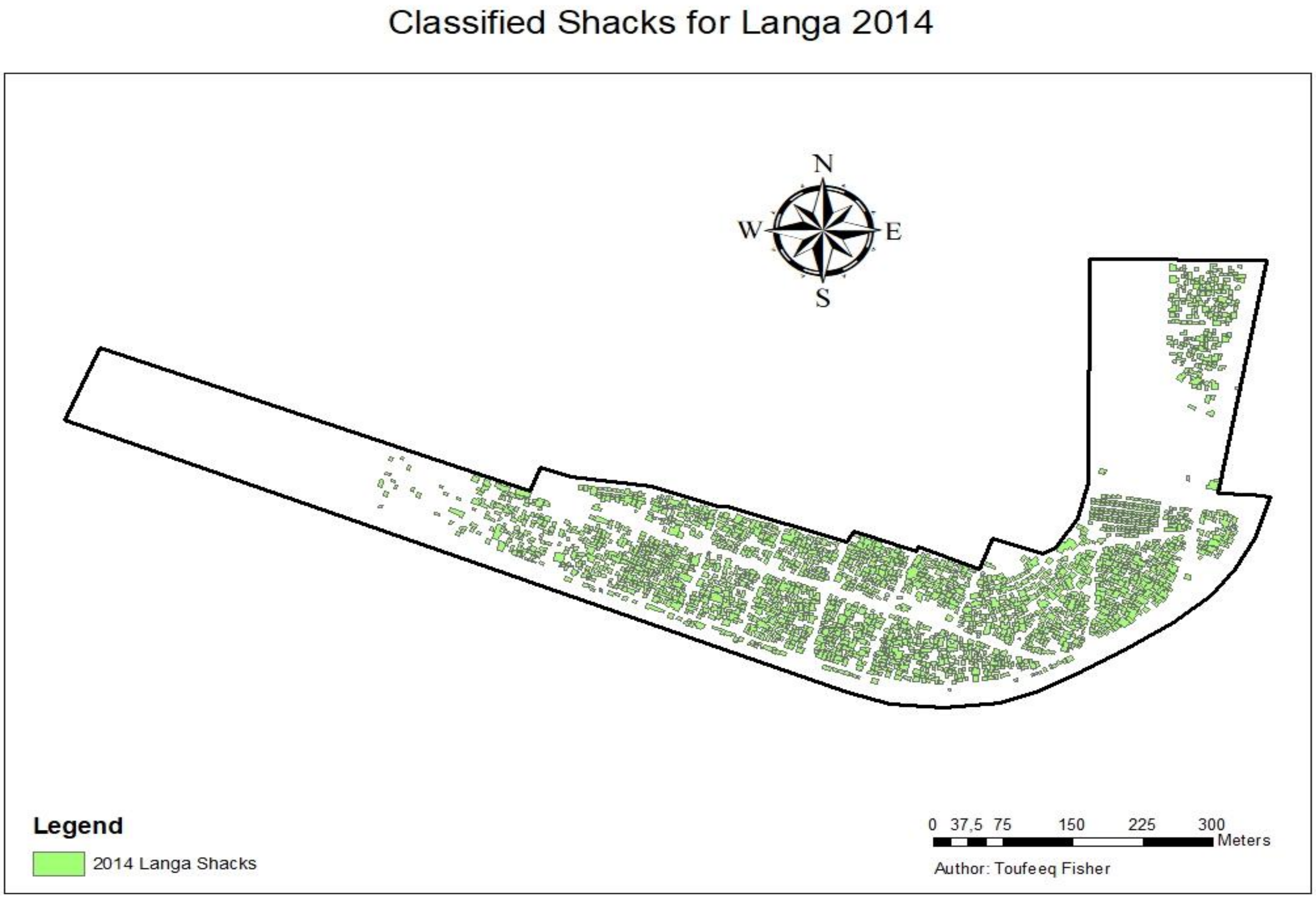


Figure 4.6: Classified Shacks for Langa 2014

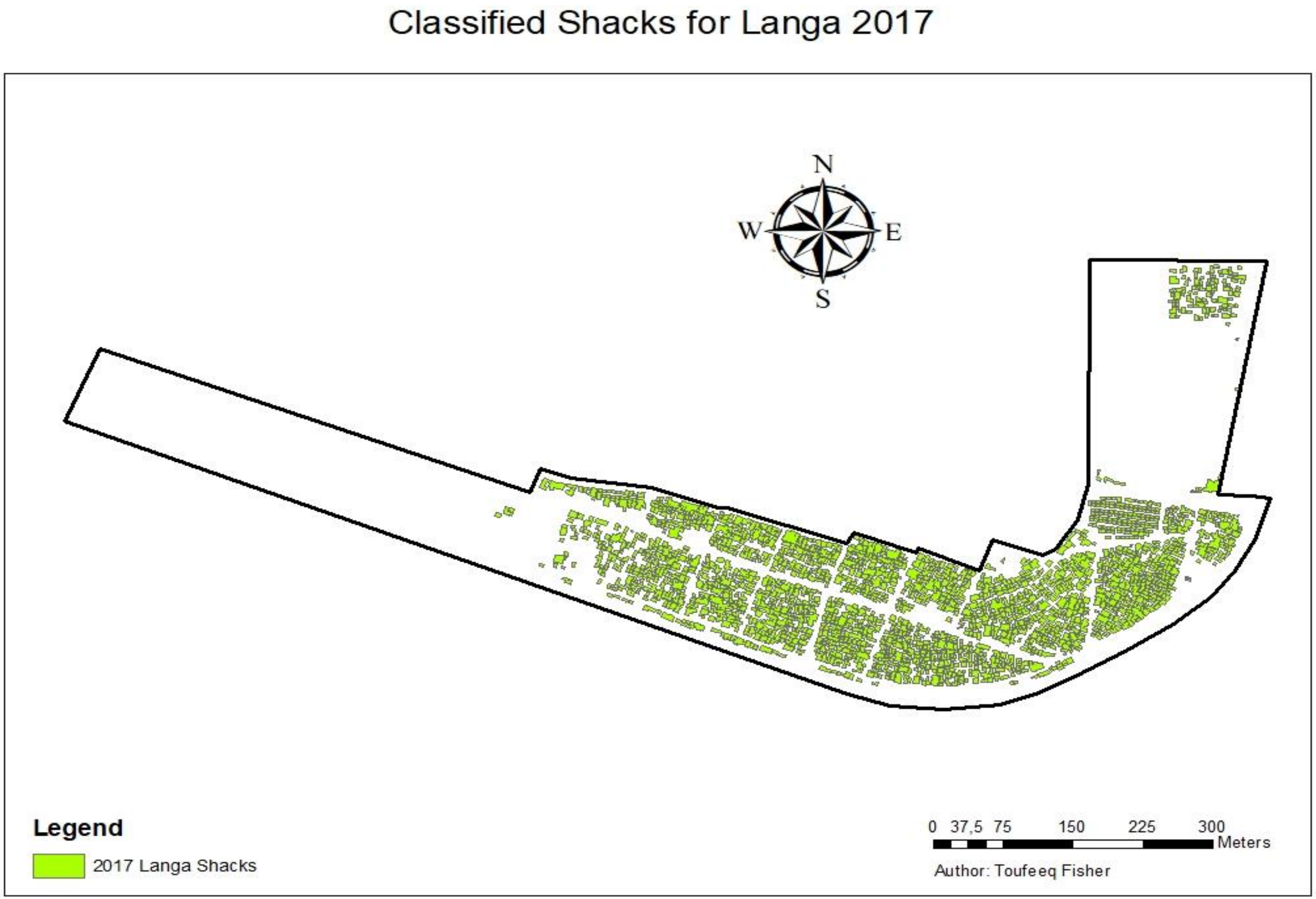


Figure 4.7: Classified Shacks for Langa 2017

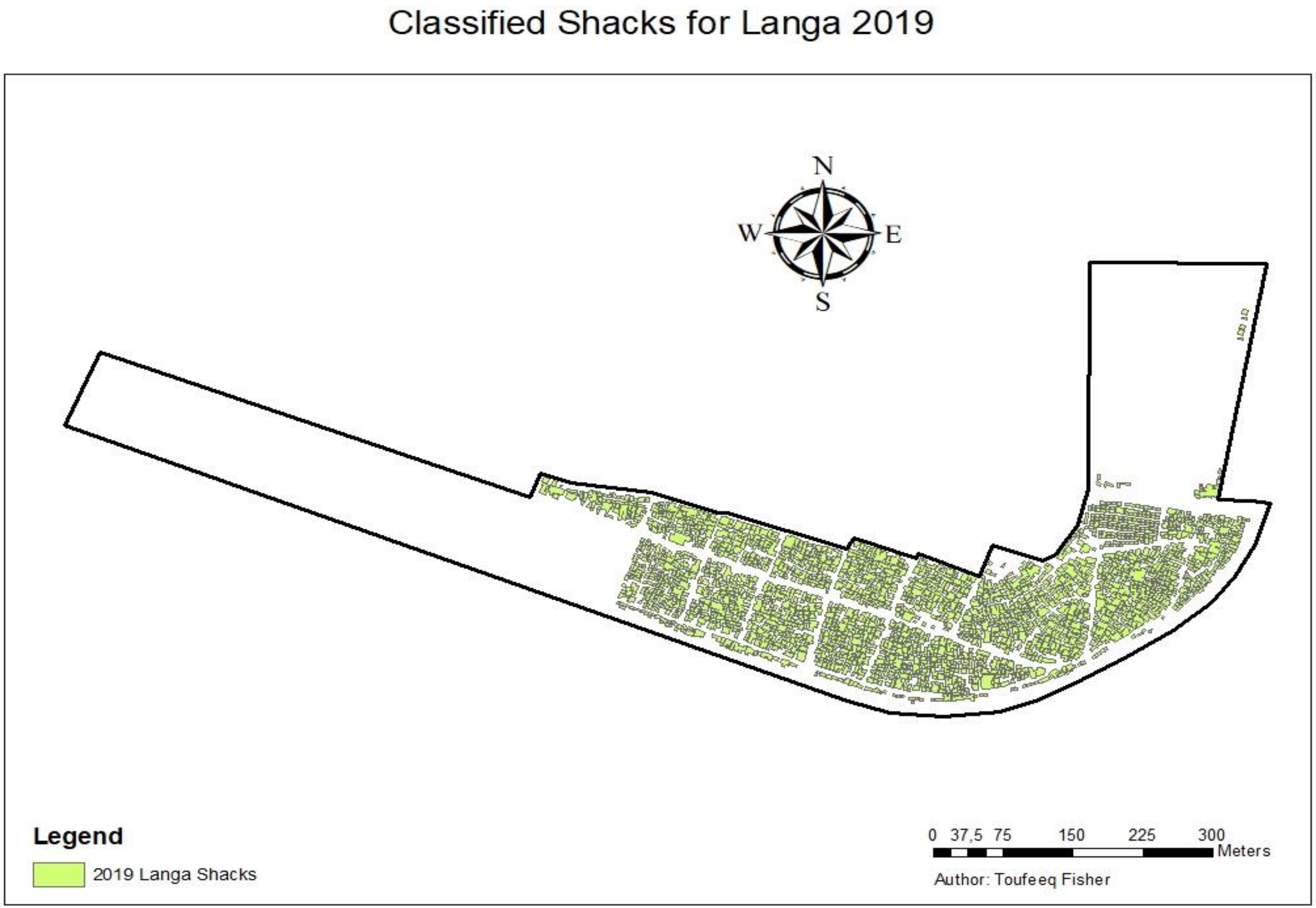


Figure 4.8: Classified Shacks for Langa 2019

c) Siqalo

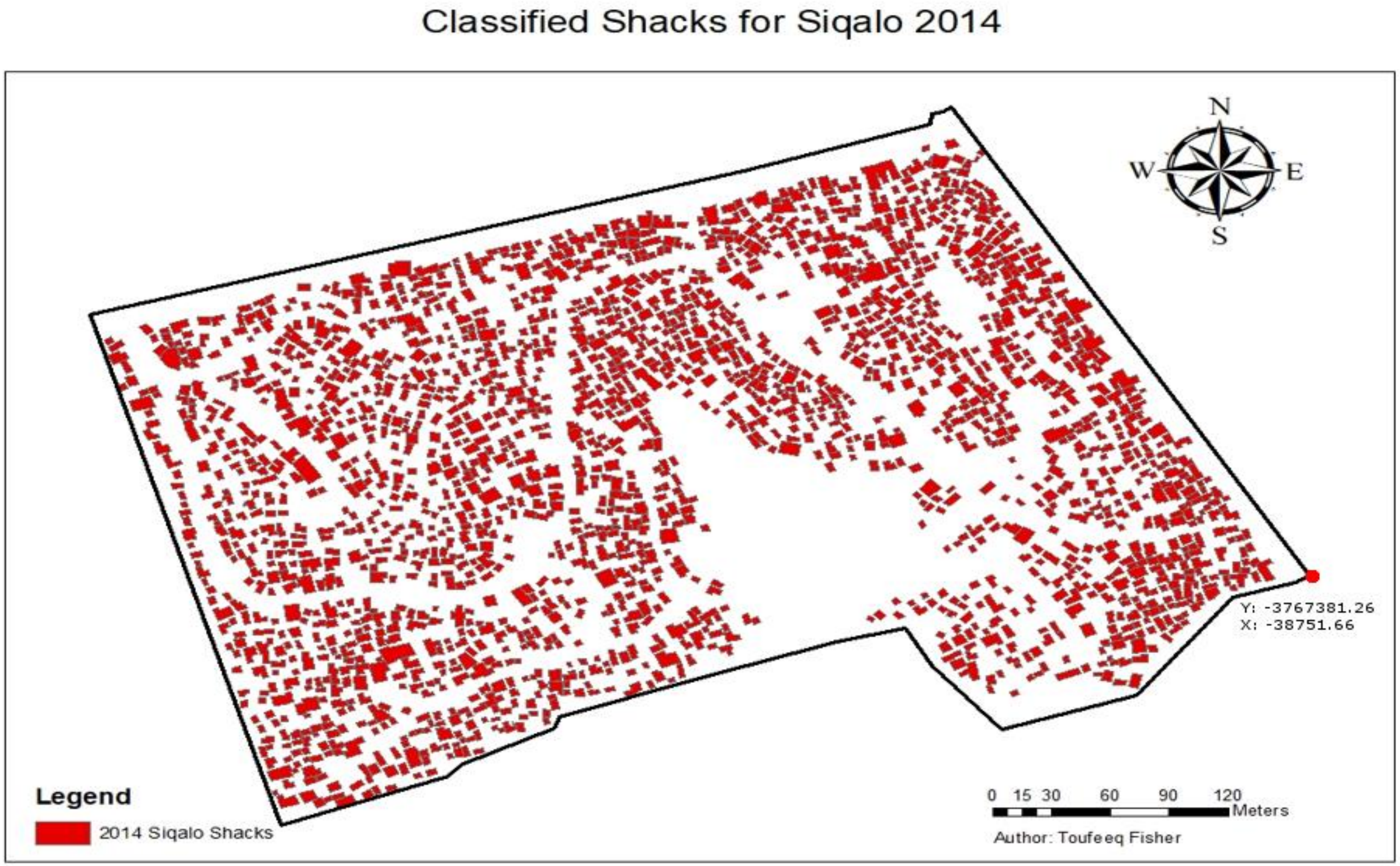


Figure 4.9: Classified Shacks for Siqalo 2014

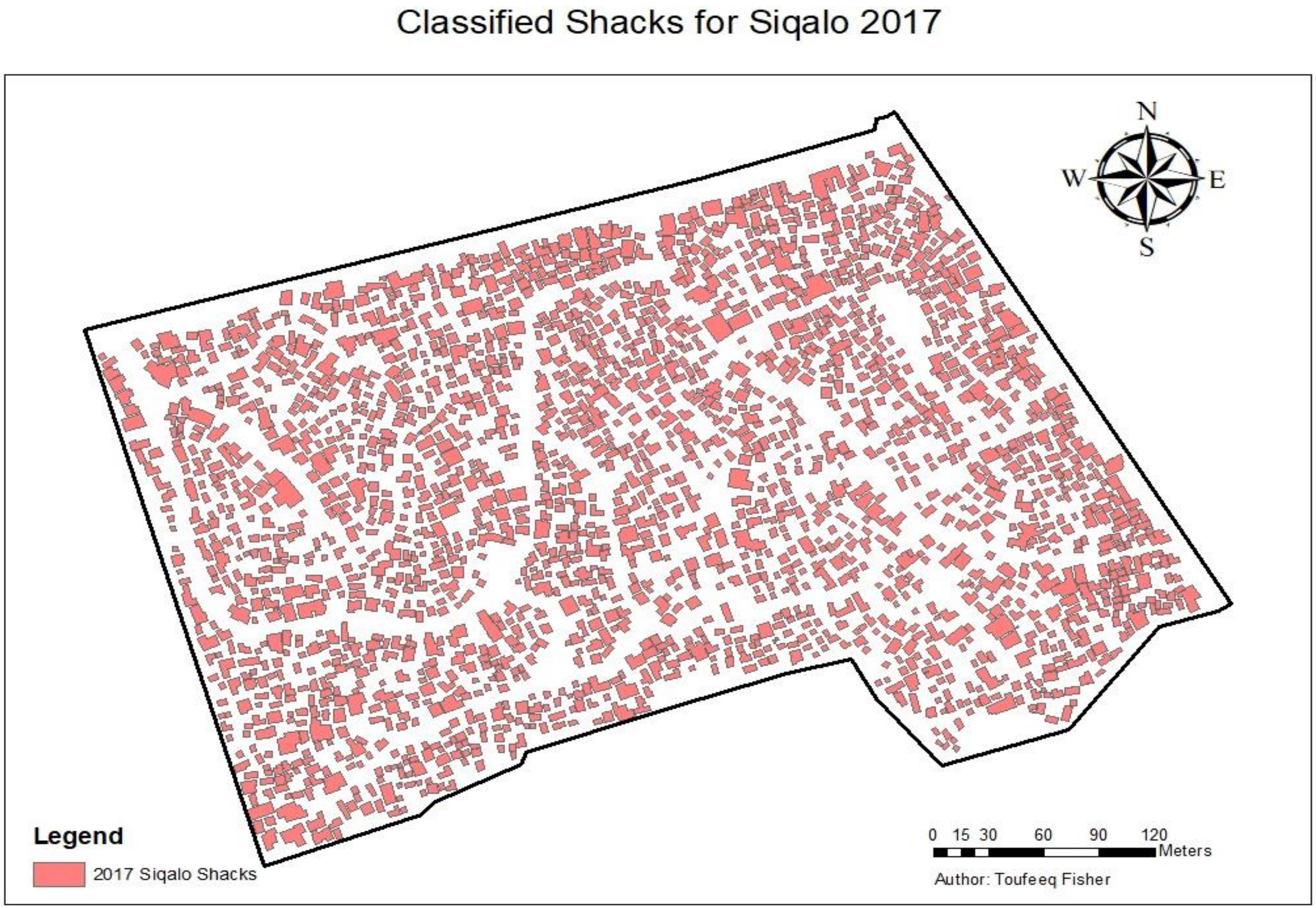


Figure 4.10: Classified Shacks for Siqalo 2017

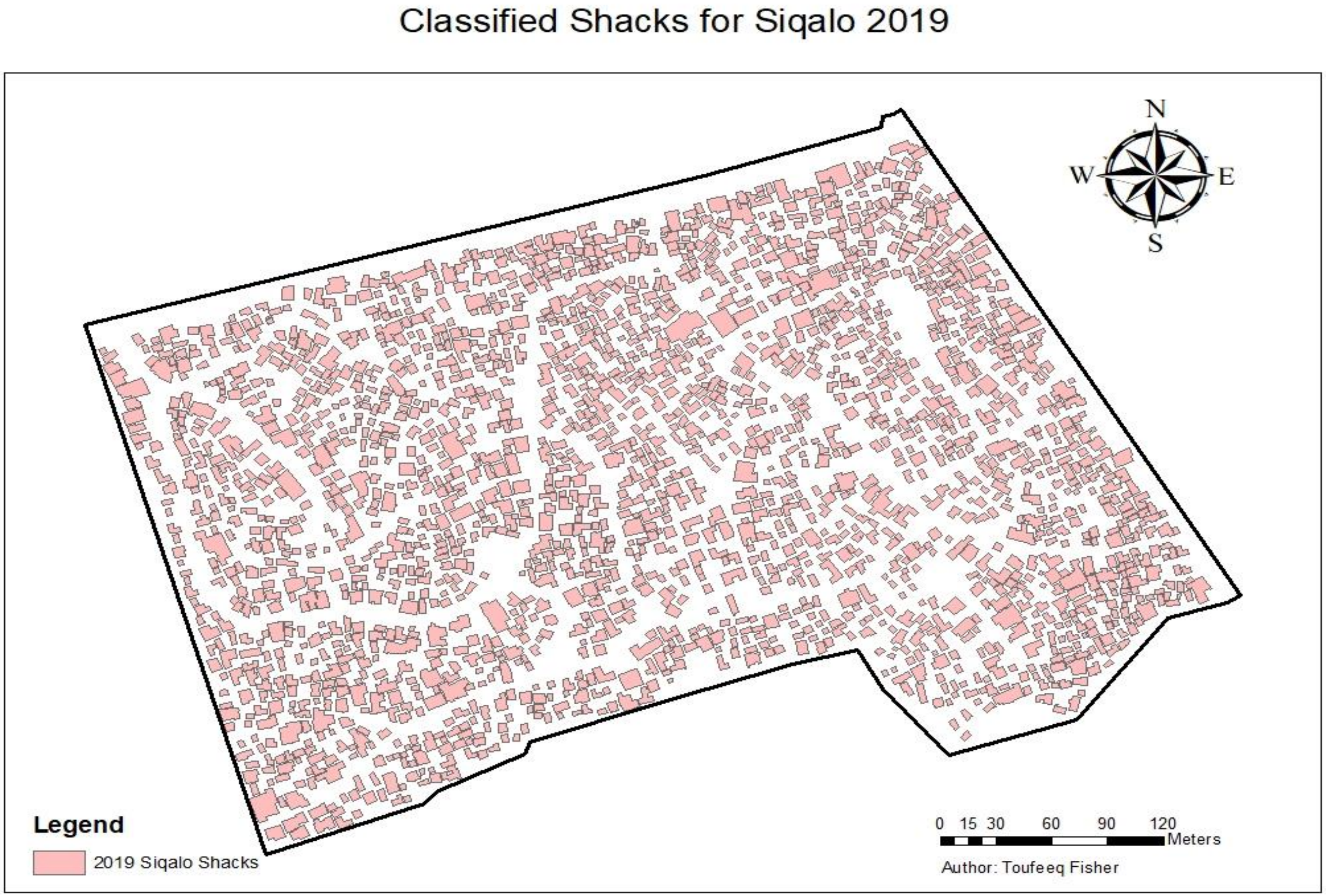


Figure 4.11: Classified Shacks for Siqalo 2019

4.2.2. Accuracy Assessment Results

The accuracies determined for all the classified imagery is as followed:

Table 4.1: Accuracy (k) determined for the classified imagery

Classified Informal Settlement (Year)	Kappa value
Imizamo Yethu (2011)	0,88
Imizamo Yethu (2014)	0,96
Imizamo Yethu (2017)	0,88
Imizamo Yethu (2019)	0,92
Langa (2011)	0,91
Langa (2014)	0,89
Langa (2017)	0,95
Langa (2019)	0,90
Siqalo (2014)	0,94
Siqalo (2017)	0,91
Siqalo (2019)	0,96

The full confusion matrices can be found in **Appendix A**. All the kappa coefficient falls in the 'close to perfect' range (0.81 – 1.00) (Fletcher et al, 2011; Lillesand et al, 2008; Navulur, 2007). Due to the satisfactory result, great confidence can be given to conduct further analysis on the classified imagery.

4.3. Spatial Metrics Results

The selected spatial metrics for this investigation have been outlined in the methodology. These metrics for each informal settlement for the years 2011, 2014, 2017, and 2019 are given in the following table:

Table 4.2: Spatial Metrics of each informal settlement

	Informal Settlement										
	Imizamo Yethu				Langa				Siqalo		
	2011	2014	2017	2019	2011	2014	2017	2019	2014	2017	2019
CA (m^2)	54751	60967	66759	68288	67471	51239	51535	54311	44389	53972	56615
PLAND (%)	38,95	43,38	47,50	48,59	30,16	22,90	23,03	24,28	27,17	33,03	34,65
MN AREA (m^2)	36	40	55	48	31	28	35	36	22	27	28
MN PARA	11582,5	11913,9	10491,4	10904,5	11799,5	11878,8	11755,6	12053,7	12727,5	12434,8	13226,8
MN SHAPE	1,4649	1,511	1,6001	1,5354	1,3669	1,3615	1,4346	1,4568	1,3368	1,3674	1,3668
MN FRAC	1,222	1,2306	1,2379	1,2309	1,1907	1,1862	1,2011	1,2121	1,1964	1,2039	1,1995
MN ENN (m)	0,65	0,54	0,50	0,56	0,59	0,65	0,55	0,39	0,95	0,96	0,87
NP	1518	1513	1221	1432	2207	1799	1467	1526	1988	2035	2022
PD ($NP/1km^2$)	10800	10765	8687	10189	9864	8041	6557	6821	12167	12454	12375

4.4. Macro and Micro Socio-economic Data for the Ordinary Least Squares (OLS) Analysis

After determining the available socio-economic data at the micro and macro level, it has been determined that open space in the settlement (percentage open space), the unemployment rate (percentage in the Western Cape for the year), the rate of poverty in the province (poverty coefficient), and the Gross Domestic Product (GDP) values will be used as the factors that will be tested against the development statistics of each informal settlement. The source of the socio-economic data was collected and interpolated from StatsSA for the investigative period. The following independent and dependent variables were related in the Ordinary Least Squares (OLS) for this project for each individual informal settlement:

a) Imizamo Yethu

Table 4.3: OLS data for Imizamo Yethu

Imizamo Yethi OLS data							
Year	CA	MN_SHAPE	MN_ENN	Open_Space	Unemployment	Poverty	GDP
2011	54751	1,4649	0,65	61,05	25	0,52	7454,7
2012	56823	1,4809	0,61	59,57	25,2	0,52	7500
2013	58895	1,4969	0,57	58,09	25,3	0,54	7564
2014	60967	1,511	0,54	56,62	23,7	0,57	7582,7
2015	62898	1,5407	0,53	55,25	23,5	0,58	7556,8
2016	64829	1,5704	0,51	53,88	21,1	0,58	7476,6
2017	66759	1,6001	0,5	52,5	23	0,61	7476,4
2018	67759	1,5681	0,53	51,41	21,2	0,62	7433,6
2019	68759	1,5354	0,56	51,41	21,2	0,63	7346

b) Langa

Table 4.4: OLS data for Langa

Langa OLS Data							
Year	CA	MN_SHAPE	MN_ENN	Open_Space	Unemployment	Poverty	GDP
2011	67471	1,3669	0,59	69,84	25	0,52	7454,7
2012	62060	1,3651	0,61	72,26	25,2	0,52	7500
2013	56649	1,3633	0,63	74,68	25,3	0,54	7564
2014	51239	1,3615	0,65	77,10	23,7	0,57	7582,7
2015	51277	1,3859	0,61	77,06	23,5	0,58	7556,8
2016	51315	1,4103	0,58	77,02	21,1	0,58	7476,6
2017	51535	1,4346	0,55	76,97	23	0,61	7476,4
2018	52832	1,4457	0,47	76,35	21,2	0,62	7433,6
2019	54311	1,4568	0,39	75,72	21,2	0,63	7346

c) Siqalo

Table 4.5: OLS data for Siqalo

Siqalo OLS Data							
Year	CA	MN_SHAPE	MN_ENN	Open_Space	Unemployment	Poverty	GDP
2014	44389	1,3368	0,95	72,83	23,7	0,57	7582,7
2015	47583	1,3470	0,95	70,88	23,5	0,58	7556,8
2016	50778	1,3572	0,96	68,93	21,1	0,58	7476,6
2017	53972	1,3674	0,96	66,97	23	0,61	7476,4
2018	55294	1,3671	0,92	66,16	21,2	0,62	7433,6
2019	56615	1,3668	0,87	65,35	21,2	0,63	7346

4.5. Change Detection Results

Spatial metrics have determined statistics describing the spatial context of shacks in the selected informal settlements. Since class metrics describes this context with respect to all shacks in each informal settlement, these metrics would be analyzed:

4.5.1. Total Area covered by shacks in informal settlements

Two of the determined class metrics describes the total land that shacks cover in informal settlements – Total Area (CA) and Percentage of Landscape (PLAND). However, these metrics are synonymous in their description and for the comparative analysis in change of informal settlements, only CA will be used. The comparison of CA for the years 2011, 2014, 2017, and 2019 is illustrated in the following figure:

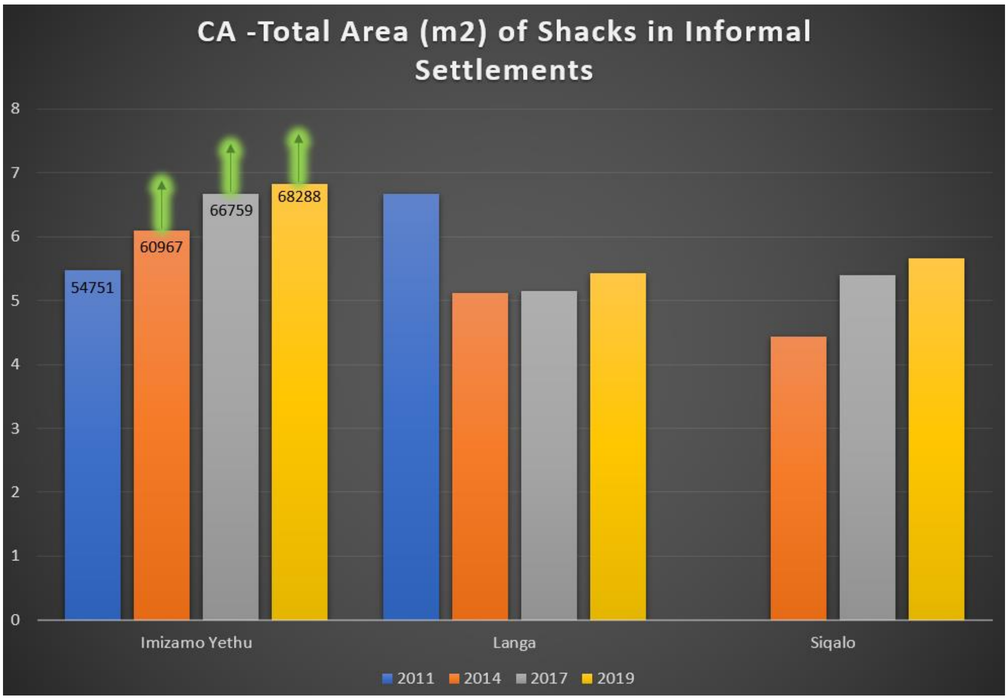


Figure 4.12: Total Area (CA) covered by shacks in informal settlements

Comparison maps were created to illustrate the location of the change in shacks in each informal settlement. The comparison is made between consecutive epochs, displaying where shacks were in the first year of the comparison versus the change in shacks in the next year of the comparison. Each informal settlement's comparison analysis can be viewed in the following figures:

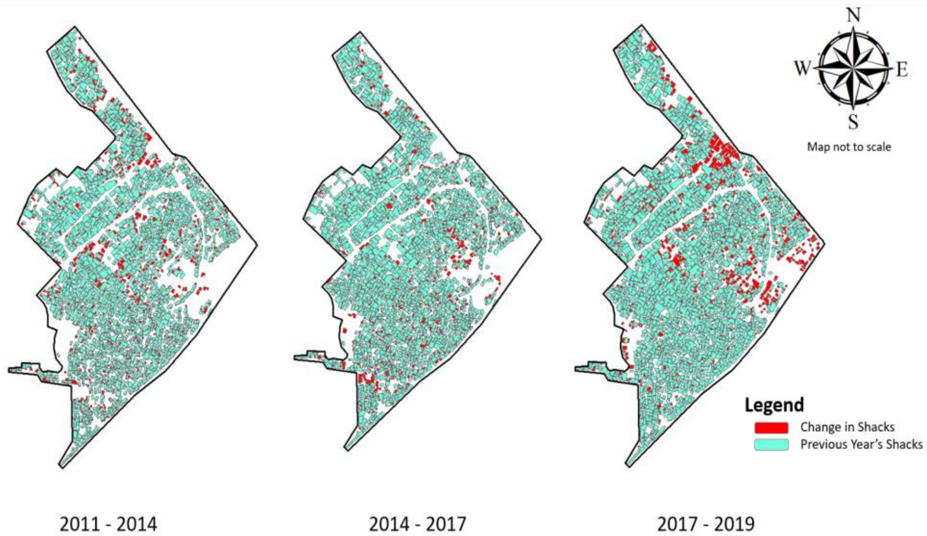


Figure 4.13: Comparative Analysis in change of shacks in Imizamo Yethu (2011-2019)

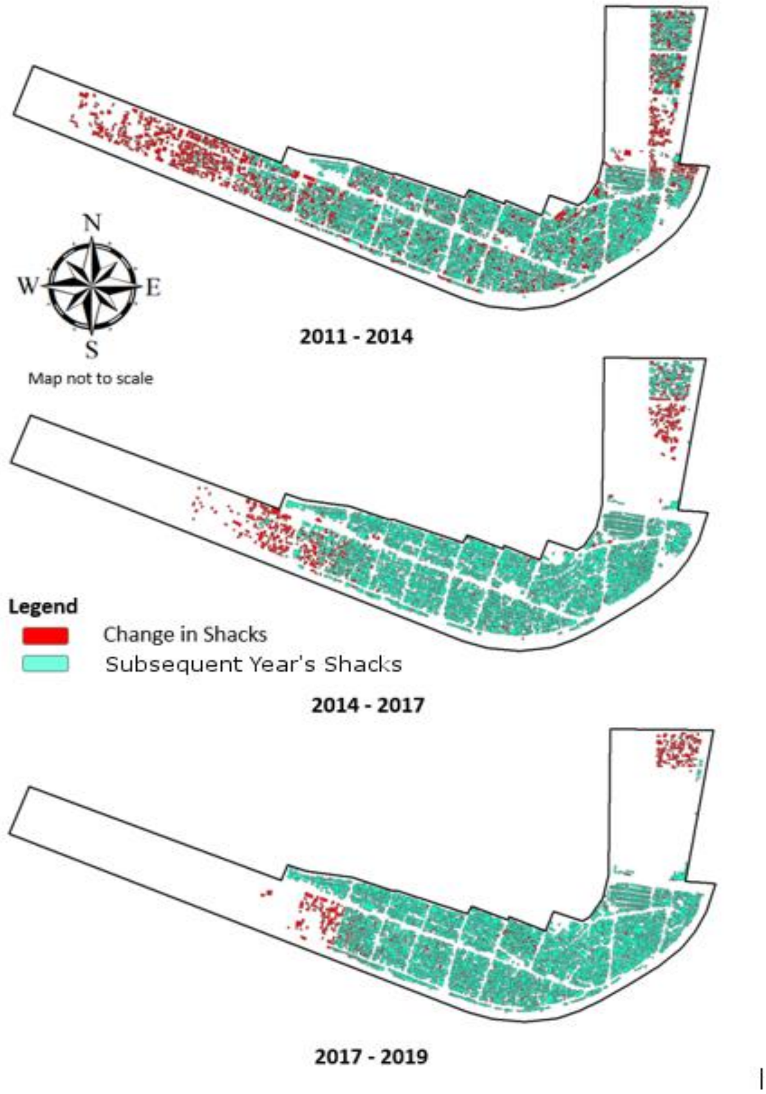


Figure 4.14: Comparative Analysis in change of shacks in Langa (2011-2019)

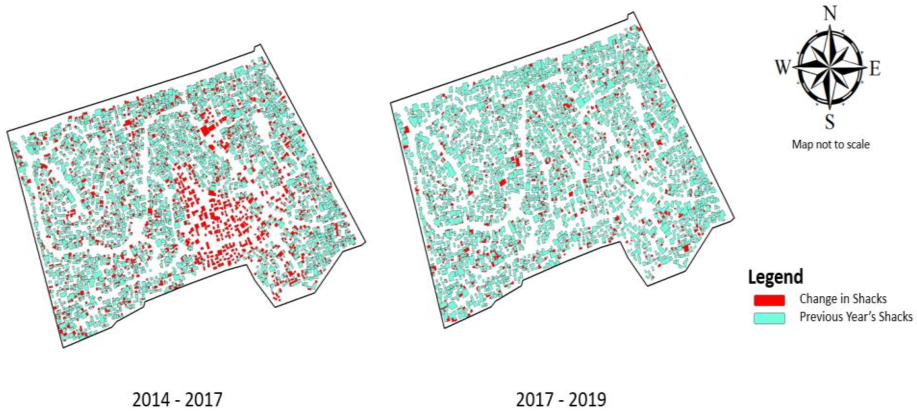


Figure 4.15: Comparative Analysis in change of shacks in Siqalo (2014-2019)

4.5.2. Number of Shacks in informal settlements

The Number of Shacks (NP) is considered a simple measure of the fragmentation within a range of shacks in a scene (McGarigal, 2002). This spatial context will be relevant in future metric analysis (such as change in shack density), hence its change in informal settlements is of interest. These changes are illustrated in the following figure:

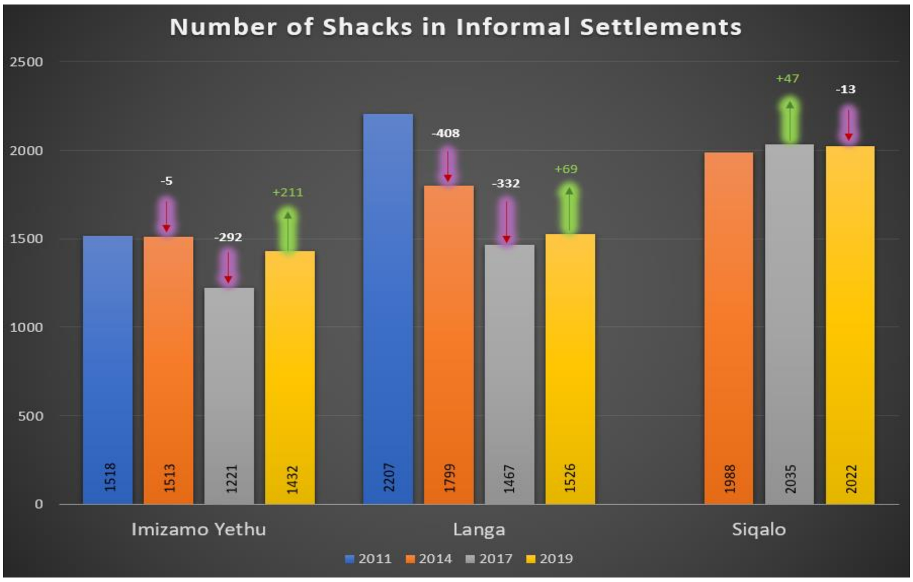


Figure 4.16: Change in the Number of shacks (NP) in informal settlements

4.5.3. Mean Shack Area (hectares) in informal settlements

The Mean Shack Area (MN AREA) describes the average area of all patches in a scene. This metric, such as NP, is descriptive in its own respect. However, it too will be relevant in future metric analysis (such as change in shack density). The change in mean shack area is illustrated in the following figure:

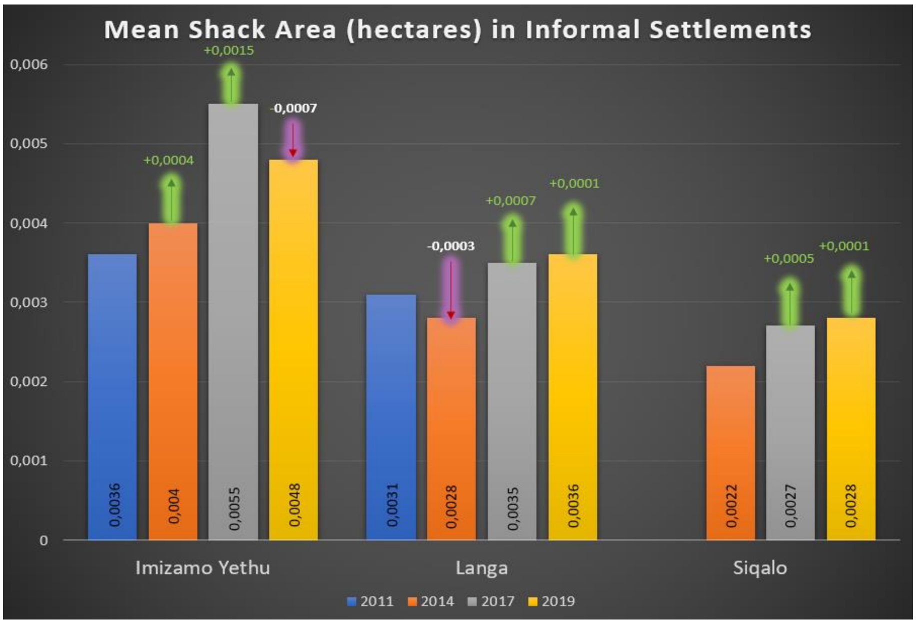


Figure 4.17: Change in the Mean Shack Area in hectares (MN AREA) in informal settlements

4.5.4. Shack Density

Shack density (PD) is described as the number shacks per 100 hectares (McGarigal, 2002). This metric considers both NP and MN Area in its determination. Hence, this metric is a simple descriptor of the overall shack composition and configuration within informal settlements. The change in PD for the selected informal settlements is illustrated in the following figure:

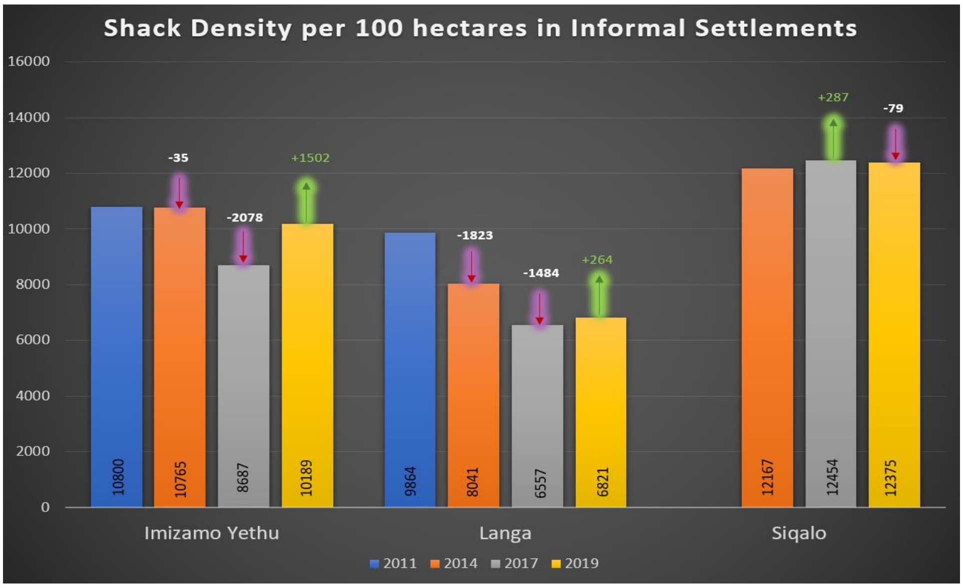


Figure 4.18: Change in the Shack Density in informal settlements

4.5.5. Mean Shack Shape Index

Three of the determined metrics measures the complexity of the shape of all shacks in informal settlements. These metrics are the Mean Perimeter-Area Ratio (MN PARA), Mean Fractal Dimension (MN FRAC) and Mean Shape Index (MN SHAPE). However, MN PARA has the limitation with the shape complexity varying on the mean shack area, and the MN PARA does not compensate for the shape complexity varying. MN FRAC and MN SHAPE does compensate for the limitation of MN PARA by measuring the average shape complexity across a range of spatial scales. However, MN FRAC and MN Shape are synonymous in their description. Hence, only MN SHAPE will be used. The change in MN SHAPE for the selected informal settlements are illustrated in the following figure:

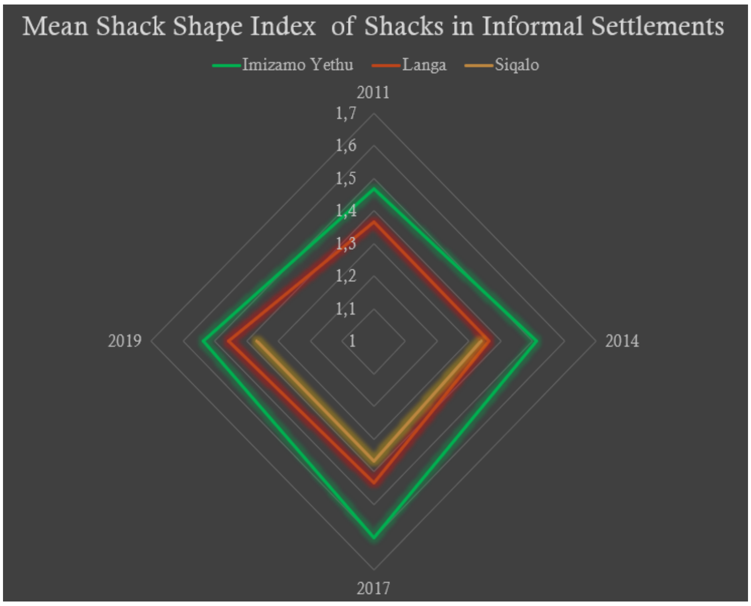


Figure 4.19: Change in the Mean Shape Index of Shacks in informal settlements

4.5.6. Mean Shack Nearest Neighbour

The Mean Shack Nearest Neighbour (MN ENN) is used to extensively quantify the isolation of shacks in informal settlements. The change in this metric across in informal settlement can be illustrated in the following figure:

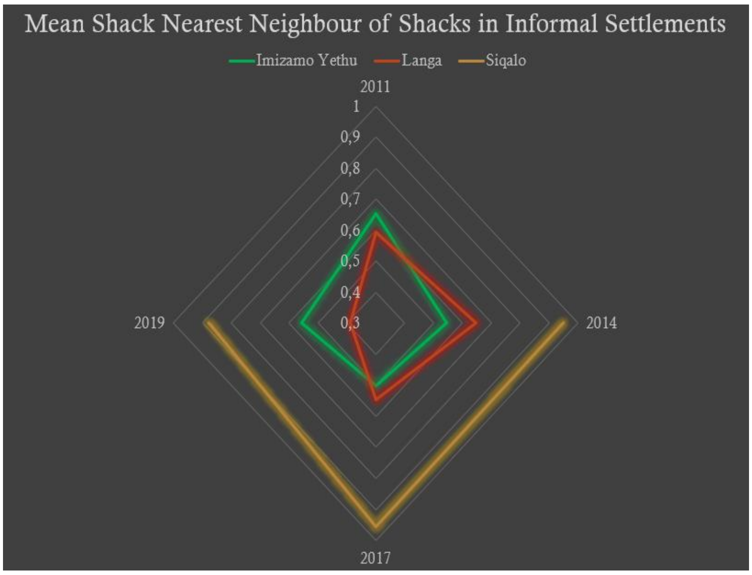


Figure 4.20: Change in the Mean Nearest Neighbour of Shacks in informal settlements

4.5.7. Change in the Spatial Context in selected informal settlements

The change in each spatial metric in informal settlements is a good approximation of how the shacks have developed. To gain a greater understanding of overall development, all the metrics used in the analysis will be compared to each other for each informal settlement. This comparison is given in the following:

1) Imizamo Yethu:
   - In 2011, the shacks are relatively sparsely distributed throughout the settlement, covering 38.96% of the landscape. This total area is distributed across 1518 shacks, resulting in a PD of 10800 shacks per 100 hectares. The MN SHAPE indicates that the shack shape complexity is relatively low in the settlement, favouring more square shaped roofs compared to more convoluted shapes. Considering the sparsely distributed shacks and the shape of shacks in Imizamo Yethu, the average distance between the nearest neighbours of shacks is relatively large. This is reflected by the relatively high MN ENN value of 0,65m.
   - Between 2011 and 2017, the total area covered by shacks increased by 8.5%, however the number of shacks decreased by 297. This combined net change is accompanied by a net decrease of 2113 shacks per 100 hectares in shack density. To compensate for this increase in total area but decrease in the number of shacks, the shape of the shacks has become more convoluted in shape to fit in the available space in the settlement. This is reflected by an increase in the MN SHAPE between 2011 and 2017. To further support this statement, the distance between shacks has decreased from 0.65m to 0.5m. This proves that the total increase in shacks is being accommodated in the same space as in 2011. This is also apparent when analyzing Figure 4.13. The change in shacks can be viewed as mostly occurring centrally within the settlement, and between shacks that were in existence.

- In 2019, the total area of shacks in the informal settlement has continued to grow, but at a decelerated rate at 1,8% from 2017 (0.9% growth per year) compared to the 8,5% between 2011 and 2017 (1.4% average growth per year). In this growth of the total area, an additional 211 shacks have been erected. This has resulted in an additional 1502 shacks in the PD per 100 hectares. These additional shacks have had an influence in the change of MN SHAPE (decrease) and MN ENN (increase). A decrease in MN SHAPE suggests that the additional shacks are square than convoluted in its roof shape. An increase in MN ENN suggests that there is more space on average between the shacks in the informal settlement. By referring to Figure 5.2, the reason for the change of all the metrics between 2017 and 2019 may be accounted for. As can be viewed in figure 4.13, the additional shacks have either been erected in available space towards the centre, or more evidently, on the periphery of the settlement. In those pockets of space, shacks having square roofs and greater distances between them would be logical relative to shacks being erected in smaller spaces, between shacks that have been erected.

2) Langa:
   - In 2011, shacks are relatively and compactly distributed throughout the settlement. This is evident by the relatively low MN ENN of 0,59m in 2011. The average shape of the roofing of shacks is relatively low, as many of the shacks conform to the square roofing provided through upgrading policies.
   - Between 2011 and 2017, there is a drastic reduction in the total area covered by shacks in the landscape by 7.1%. This is due to formalization in the settlement shown in the aerial imagery and upon the physical visitation of the site. In total, there was a net loss of 740 shacks and 3307 shacks per 100 hectares in shack density. The shacks that have remained and any additional shacks that may have been erected have contributed to an increase in MN SHAPE. This is natural as the any available land in which shacks were erected may have been small and the shacks may have had to adapted in those spaces. This is evident in Figure 5.3. The greatest loss of shacks can be seen from the west and northern areas of the settlement. The newly erected shacks can be viewed between the existing shacks between 2011 and 2017. The fact that new shacks have been erected mostly between existing shacks is supported by the decrease in MN ENN form 0,59m to 0,55m.
   - In 2019, there has been a steady increase in the total area covered by shacks in the informal settlement by 1,24% from 2017. Although upgrading is clearly forcing for the removal of shacks, this metric proves that a larger area of shacks is detected in the informal settlement. Using the spatial metrics, MN SHAPE has increased, and MN ENN has largely decreased. This means that net growth in the landscape is contributed to shacks the extension of shacks into more convoluted shapes opposed to square, as well as the fact that shacks are more densely packed per 100 hectares. This is evident in the PD decreasing from 9864 in 2011 to only 6821 shacks per 100 hectares. This is evident to a further extent by analyzing Figure 4.14, as there is very little space between shacks of the settlement.

3) Siqalo:
   - In 2014, shacks are relatively sparsely distributed throughout the informal settlement. This is evident in the relatively large value of the MN ENN of 0,95m. In total, shacks cover 27,2% of the landscape and this is distributed across 1988 shacks. This results in a PD of 12167 shacks per 100 hectares. Considering the relatively low total area covered by shacks, the MN SHAPE is expectedly and relatively low, suggesting that the shacks of Siqalo are squarer than convoluted in shape.

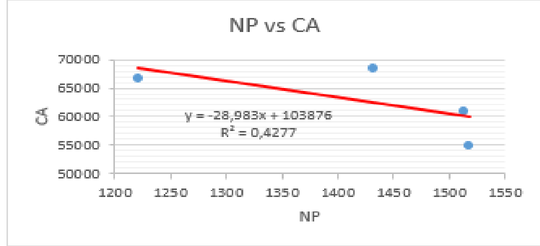
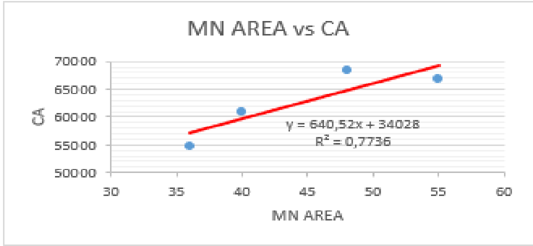
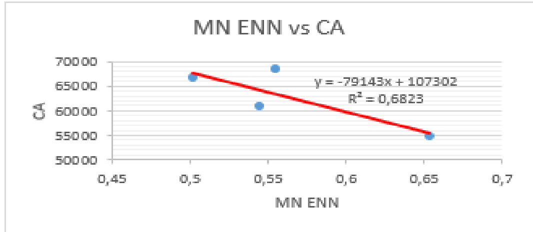
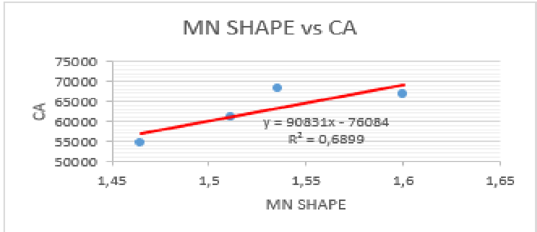
- Between 2014 and 2019, there is no great deviation in the determined spatial metrics, except in the total area covered by shacks. PLAND has increased by 7.5% in the five-year period (an average of 1,5% per year). Consequently, the MN ENN decreased from 0.95m to 0,87m. Although this is substantial, there is no great change in the deviation of the MN AREA, MN SHAPE, NP, and PD. In Figure 4.15, most of the additional shacks were erected in great amounts of open space within the settlement between 2014 and 2017. This phenomenon continues between 2017 and 2019, however to a lesser degree. Many shacks were not constructed between other shacks in the immediate vicinity, hence more square and less convoluted shape roofing (small increase in MN SHAPE). NP, and consequently PD have not changed to a great degree and the small change can be accredited to a small increase in MN AREA.

4.6. Linear Regression Determination and Analysis

4.6.1. Imizamo Yethu:

a) Development

Table 4.6: Linear Regression of Development Statistics in Imizamo Yethu

<table>
<tr><th>Metrics Analysed</th><th>Linear Fit</th><th>Correlation coefficient (r)</th><th>Interpretation of Metrics Analyzed</th></tr>
<tr><td>NP vs CA</td><td>
</td><td>-0,654</td><td>There is a moderate negative correlation between the number of shacks classified and the total area covered by shacks between 2011 and 2019. This moderate negative correlation follows an indirectly proportional relationship. The lower the number of shacks identified, the larger the total area covered by shacks in Imizamo Yethu between 2011 and 2019. This indicates densification in the settlement.</td></tr>
<tr><td>MN AREA vs CA</td><td>
</td><td>+0,880</td><td>There is a high positive correlation between the mean area of shacks and the total area covered by shacks between 2011 and 2019. This high positive correlation follows a directly proportional relationship. The larger the mean area of shacks, the larger the total area covered by shacks in Imizamo Yethu between 2011 and 2019. This indicates densification in the settlement.</td></tr>
<tr><td>MN ENN vs CA</td><td>
</td><td>-0,826</td><td>There is a high negative correlation between the average distance between shacks and the total area covered by shacks between 2011 and 2019. This high negative correlation follows an indirectly proportional relationship. The shorter the average distance between shacks, the larger the total area covered by shacks in Imizamo Yethu between 2011 and 2019. This indicates densification in the settlement.</td></tr>
<tr><td>MN SHAPE vs CA</td><td>
</td><td>+0,831</td><td>There is a high positive correlation between the average shape complexity of shacks and the total area covered by shacks between 2011 and 2019. This high positive correlation follows a directly proportional relationship. The higher the average shape complexity of shacks, the larger the total area covered by shacks in Imizamo Yethu between 2011 and 2019.</td></tr>
</table>

The development of shacks in Imizamo Yethu is further analyzed by the mean center of shacks within the settlement. This can be followed in the following figure:

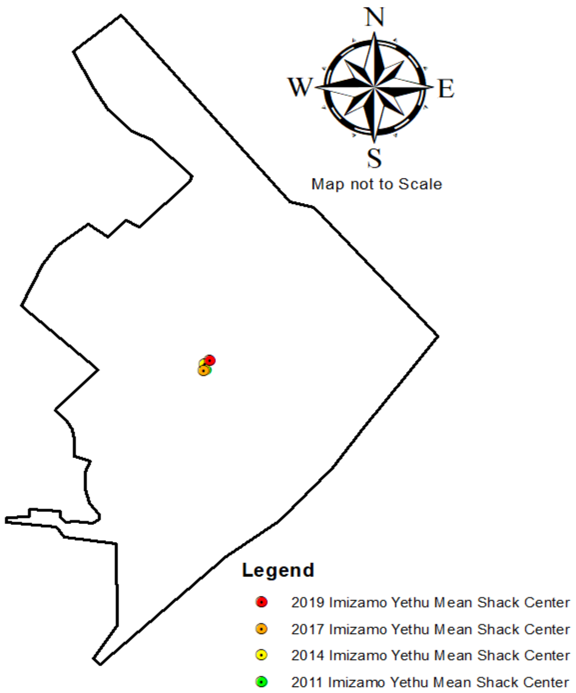
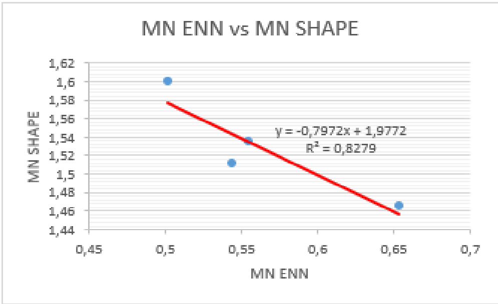
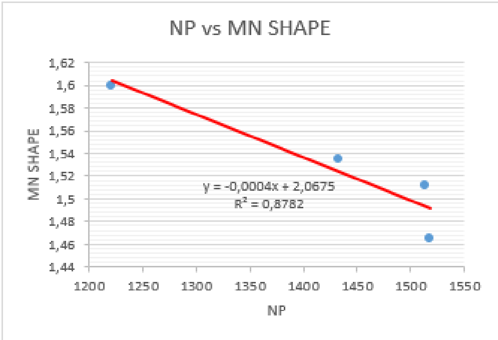
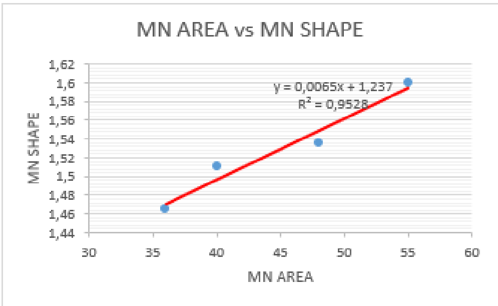


Figure 4.21: Mean Center of Shacks in Imizamo Yethu between 2011 and 2019

Between 2011 and 2019, the mean center of shacks not changed significantly within the settlement. This, in turn, suggests that there is not any significant direction of development within the settlement between 2011 and 2019. Other factors to describe the development of shacks includes regions where densification or shack removal is present, and how the complexity of shacks shape changes based upon the open space available for development. These factors will be analyzed in the compactness and complexity sections, respectively.

b) Complexity

Table 4.7: Linear Regression of Complexity Statistics in Imizamo Yethu

Metrics Analysed	Linear Fit	Correlation coefficient (r)	Interpretation of Metrics Analyzed
MN ENN vs MN SHAPE		-0,910	There is a very high negative correlation between the average distance between shacks and the average shape complexity of shacks between 2011 and 2019. This very high negative correlation follows an indirectly proportional relationship. The shorter the average distance between shacks, the higher the average shape complexity of shacks in Imizamo Yethu between 2011 and 2019. This indicates that when there is less space between shacks, that shacks are accommodated or adapted to fit less space.
NP vs MN SHAPE		-0,937	There is a very high negative correlation between the number of shacks classified and the average shape complexity of shacks between 2011 and 2019. This very high negative correlation follows an indirectly proportional relationship. The lower number of shacks classified, the higher the average shape of shacks in Imizamo Yethu between 2011 and 2019.
MN AREA vs MN SHAPE		+0,976	There is a very high positive correlation between the mean area of shacks and the average shape complexity of shacks between 2011 and 2019. This very high positive correlation follows a directly proportional relationship. The larger the mean area of shacks, the higher the average shape complexity of shacks in Imizao Yethu between 2011 and 2019. This indicates that when there is less space between shacks, that shacks are accommodated or adapted to fit less space.

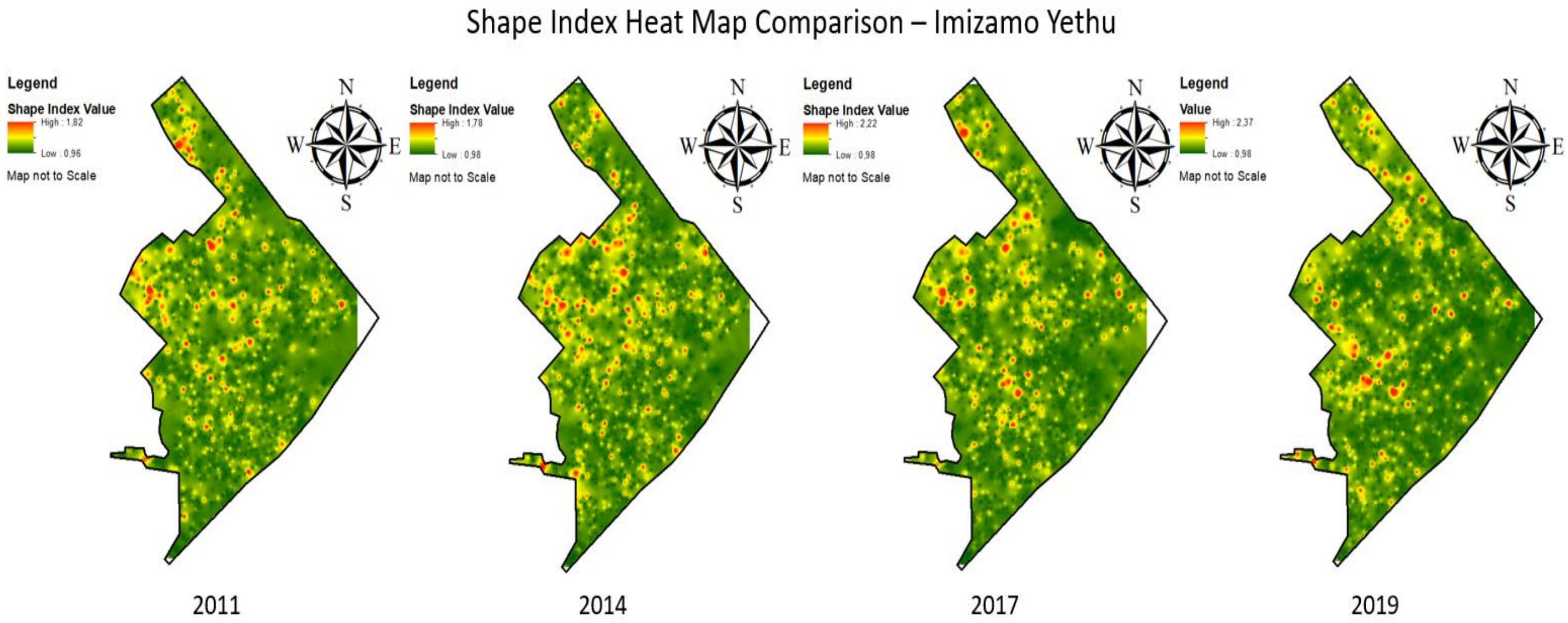
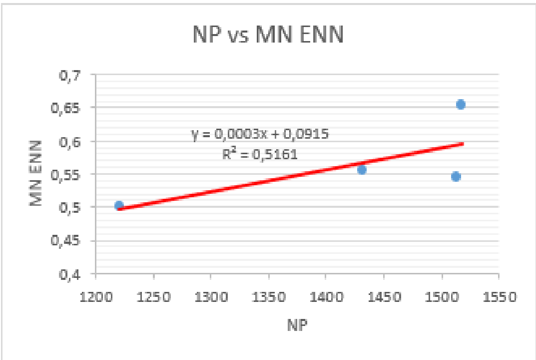
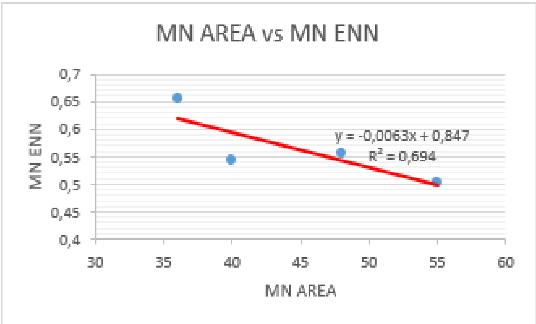


Figure 4.22: Imizamo Yethu Shape Index Heat Maps

The range of the shape complexity of shacks has increased between 2011 and 2019 (from 0,96 – 1,82 in 2011 to 0,98 – 2,37), as illustrated in the figure above. As the settlement develops over time, clusters of relatively complex shack shapes show a decreasing trend across regions of the settlement. This decreasing trend of relative shack shape complexity can be seen particularly in the western portion of Imizamo Yethu between 2011 and 2017. In 2019, larger regions indicating relative shape complexity of shacks can be seen towards the center and areas on the western boundary of the settlement. This indicates that there may have been a greater demand for the development of shacks in the western portion of the settlement. The development in these areas may have required shacks to adapt to any available space, resulting in an increased complexity of shack shapes. Coincidentally, the western portion of the settlement contains lower heights above sea level and relatively gradual slopes compared to that of the eastern half of the settlement. Socio-economic factors such major transport networks, basic amenities and educational facilities are also mainly situated west of the settlement. This indicates that both height and the proximity of socio-economic factors may influence the development of shacks in Imizamo Yethu.

c) Compactness

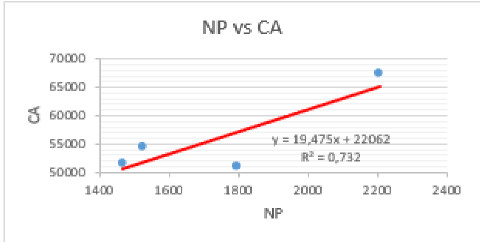
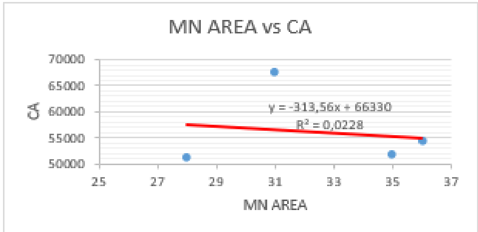
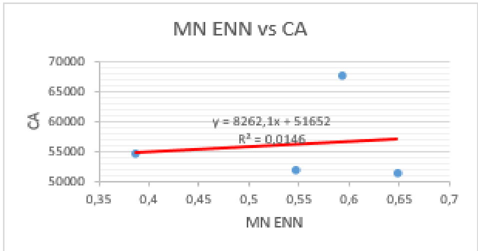
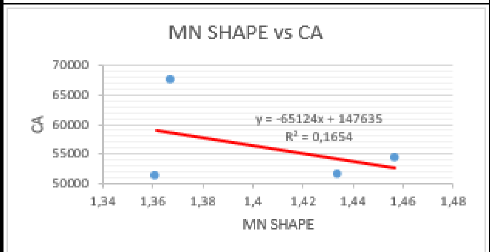
Table 4.8: Linear Regression of Compactness Statistics in Imizamo Yethu

NP vs MN ENN		+0,718	There is a high positive correlation between the number of shacks classified and the average distance between shacks between 2011 and 2019. This high positive correlation follows a directly proportional relationship. The higher the number of shacks classified, the larger the average distance between shacks in Imizamo Yethu between 2011 and 2019.
MN AREA vs MN ENN		-0,833	There is a high negative correlation between the mean area of shacks and the average distance between shacks between 2011 and 2019. This high negative correlation follows an indirectly proportional relationship. The larger the mean area of shacks, the shorter the distance between shacks in Imizamo Yethu between 2011 and 2019. This indicates that the settlement is prone to densification.

4.6.2. Langa:

a) Development

Table 4.9: Linear Regression of Development Statistics in Langa

Metrics Analysed	Linear Fit	Correlation coefficient (r)	Interpretation of Metrics Analyzed
NP vs CA		+0,856	There is a high positive correlation between the number of shacks classified and the total area covered by shacks between 2011 and 2019. This high positive correlation follows a directly proportional relationship. The higher the number of shacks classified, the larger the total area covered by shacks in Langa between 2011 and 2019.
MN AREA vs CA		-0,151	There is little if any correlation between the mean area of shacks and the total area covered by shacks in Langa between 2011 and 2019.
MN ENN vs CA		+0,121	There is little if any correlation between the average distance between shacks and the total area covered by shacks in Langa between 2011 and 2019.
MN SHAPE vs CA		-0,407	There is a low negative correlation between the average shape complexity and the total area covered by shacks in Langa between 2011 and 2019.

The development of shacks in Langa is further analyzed by the mean center of shacks within the settlement. This can be followed in the following figure:

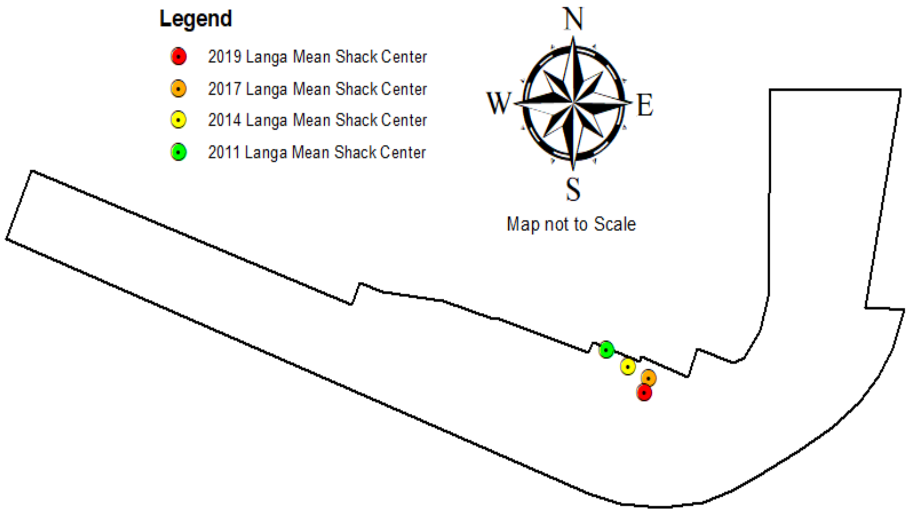
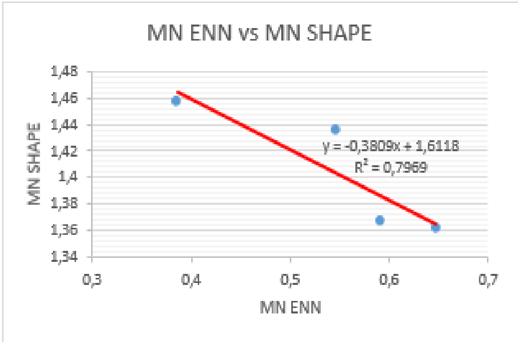
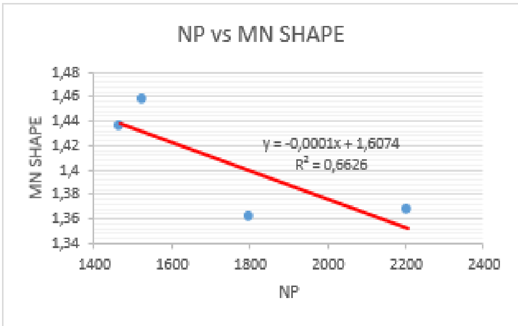
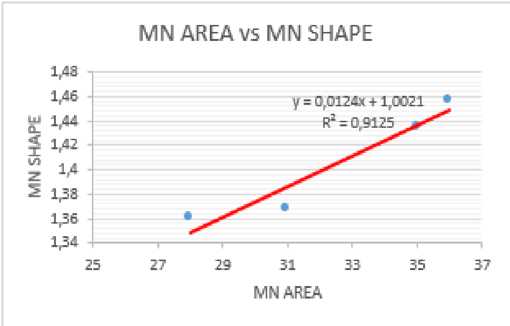


Figure 4.23: Mean Center of Shacks in Langa between 2011 and 2019

Between 2011 and 2019, the mean center of shacks has changed significantly within the settlement. The direction of development follows from the northwest into the southeast. An obvious reason for this direction of development is formalization, as can be seen in the sourced imagery in Appendix A. The direction of development will be analyzed further with respect to the compactness and complexity sections of shacks in Langa.

b) Complexity

Table 4.10: Linear Regression of Complexity Statistics in Langa

Metrics Analysed	Linear Fit	Correlation coefficient	Interpretation of Metrics Analyzed
MN ENN vs MN SHAPE		-0,893	There is a high negative correlation between the average distance between shacks and the average shape complexity of shacks between 2011 and 2019. This high negative correlation follows an indirectly proportional relationship. The shorter the average distance between shacks, the higher the average complexity shapes of shacks in Langa between 2011 and 2019. This indicates that when there is less space between shacks, that shacks are accommodated to or adapted to fit less space.
NP vs MN SHAPE		-0,814	There is a high negative correlation between the number of shacks classified and the average shape complexity of shacks between 2011 and 2019. This high negative correlation follows an indirectly proportional relationship. The lower the number of shacks classified, the higher the average shape complexity of shacks in Langa between 2011 and 2019.
MN AREA vs MN SHAPE		+0,955	There is a very high positive correlation between the mean area of shacks and the average shape complexity of shacks between 2011 and 2019. This very high positive relationship follows a directly proportional relationship. The larger the mean area of shacks, the higher the average shape complexity of shacks in Langa between 2011 and 2019. This indicates that when there is less space between shacks, that shacks are accommodated to or adapted to fit less space.

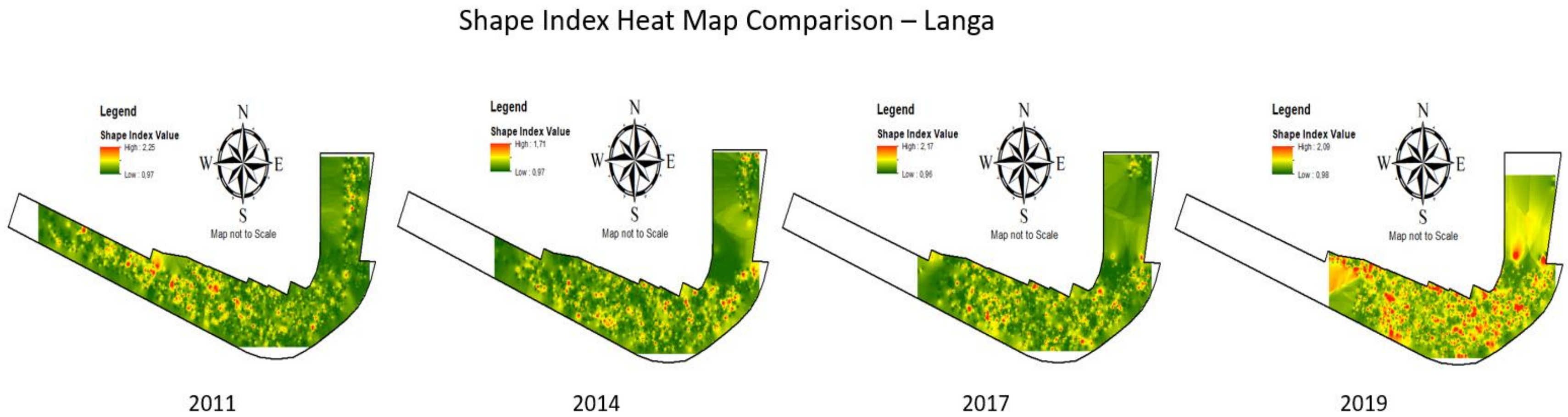
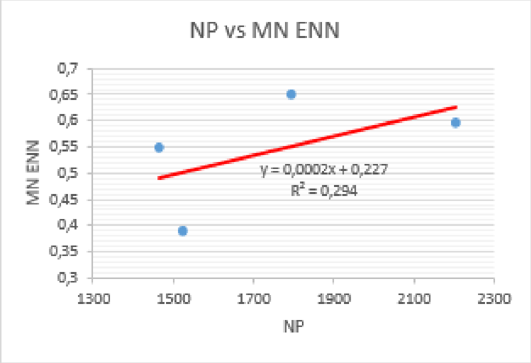
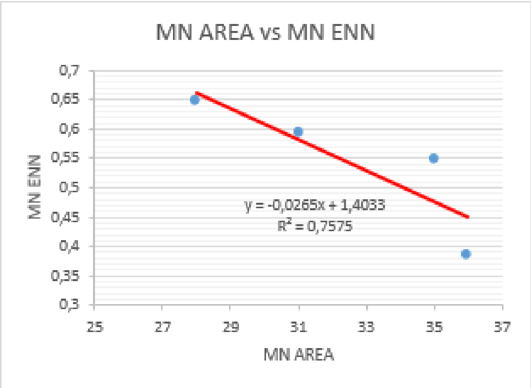


Figure 4.24: Langa Shape Index Heat Maps

The range of shape complexity of shacks has decreased between 2011 and 2019 (from 0,97 – 2,25 to 0,98 – 2,09), as illustrated in the figure above. As the settlement develops, clusters of relatively complex shack shapes vary, with the regions of relatively high shack complexity being observed in 2019 compared to that from 2011 and 2017. As indicated in the development section of Langa, formalization of Langa influences how shacks develop in the settlement. Since the available space for shack development is decreasing and the area of shacks is increasing between 2017 and 2019, it can be safely concluded that relatively high shack complexity is accommodated to compromise for the lessened available space.

c) Compactness

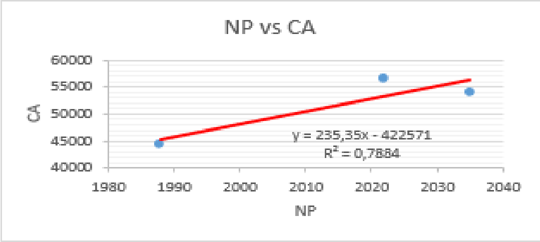
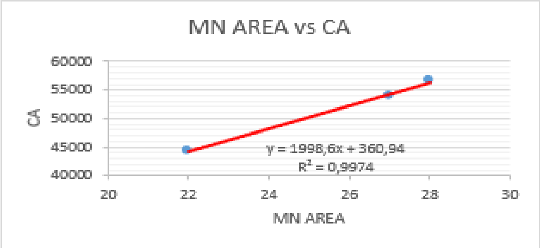
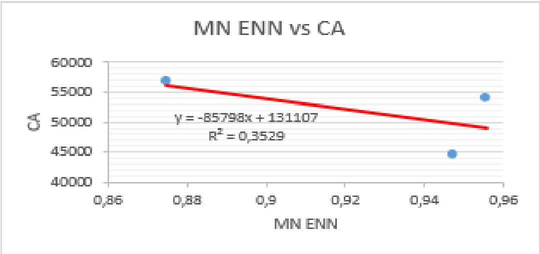
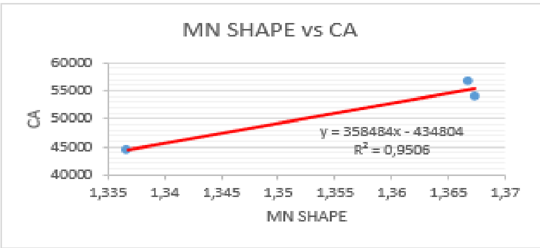
Table 4.11: Linear Regression of Compactness Statistics in Langa

Metrics Analysed	Linear Fit	Correlation coefficient (r)	Interpretation of Metrics Analyzed
NP vs MN ENN		+0,542	There is a moderate positive correlation between the number of shacks classified and the average distance between shacks between 2011 and 2019. This moderate positive correlation follows a directly proportional relationship. The higher the number of shacks classified, the larger the average distance between shacks in Langa between 2011 and 2019.
MN AREA vs MN ENN		-0,870	There is a high negative correlation between the mean area covered by shacks and the average distance between shacks between 2011 and 2019. This high negative correlation follows an indirectly proportional relationship. The larger the mean area of shacks, the shorter the average distance between shacks in Langa between 2011 and 2019. This indicates densification within the settlement.

4.6.3. Siqalo:

a) Development

Table 4.12: Linear Regression of Development Statistics in Siqalo

<table>
<tr><th>Metrics Analysed</th><th>Linear Fit</th><th>Correlation coefficient (r)</th><th>Interpretation of Metrics Analyzed</th></tr>
<tr><td>NP vs CA</td><td>
</td><td>+0,888</td><td>There is a high positive correlation between the number of shacks classified and the area covered by shacks between 2014 and 2019. This high positive correlation follows a directly proportional relationship. The higher the number of shacks, the larger the area covered by shacks in Siqalo between 2014 and 2019.</td></tr>
<tr><td>MN AREA vs CA</td><td>
</td><td>+0,999</td><td>There is a very high positive correlation between the mean area of shacks and the area covered by shacks between 2014 and 2019. This very high positive correlation follows an indirectly proportional relationship. The larger the mean area of shacks, the larger the area covered by shacks in Siqalo between 2014 and 2019. This indicates densification in the settlement.</td></tr>
<tr><td>MN ENN vs CA</td><td>
</td><td>-0,594</td><td>There is a moderate negative correlation between the average distance between shacks and the area covered by shacks between 2014 and 2019. This moderate negative correlation follows an indirectly proportional relationship. The shorter the distance between shacks, the larger the area covered by shacks in Siqalo between 2014 and 2019. This indicates densification in the settlement.</td></tr>
<tr><td>MN SHAPE vs CA</td><td>
</td><td>+0,975</td><td>There is a very high positive correlation between the average shape complexity and the area covered by shacks between 2014 and 2019. This very high positive correlation follows a directly proportional relationship. The higher the average shape complexity of shacks, the larger the average area covered by shacks in Siqalo between 2014 and 2019.</td></tr>
</table>

The development of shacks in Siqalo is further analyzed by the mean center of shacks within the settlement. This can be followed in the following figure:

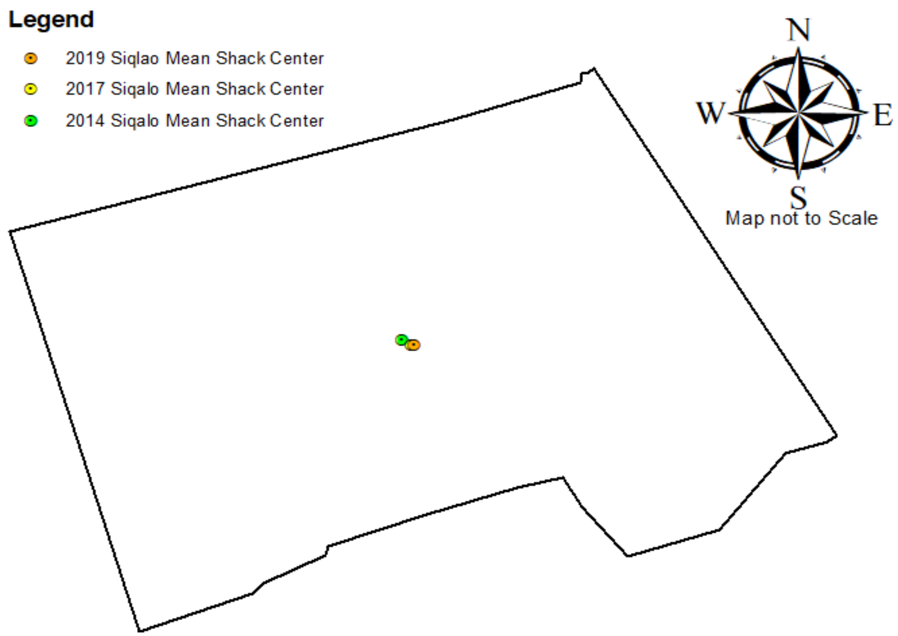
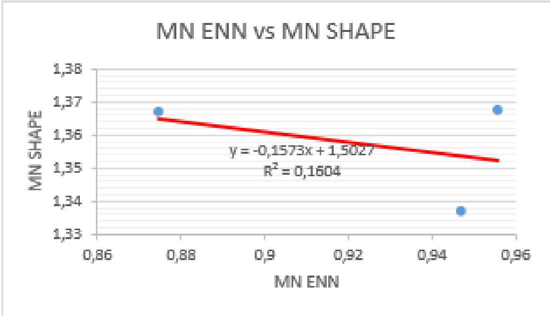
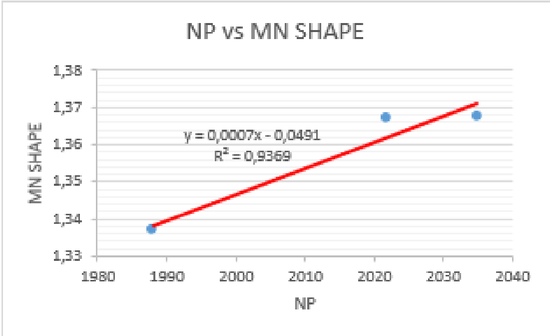
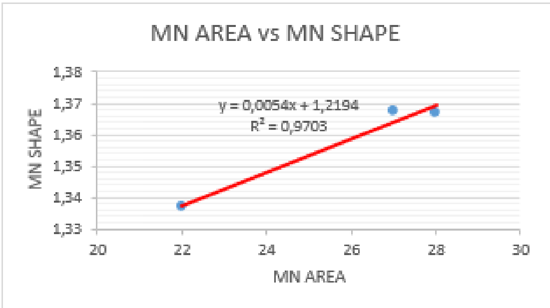


Figure 4.25: Mean Center of Shacks in Siqalo between 2014 and 2019

Between 2014 and 2019, the mean center of shacks not changed significantly within the settlement. This, in turn, suggests that there is not any significant direction of development within the settlement between 2014 and 2019. Other factors to describe the development of shacks includes regions where densification or shack removal is present, and how the complexity of shacks shape changes based upon the open space available for development. These factors will be analyzed in the compactness and complexity sections, respectively.

b) Complexity

Table 4.13: Linear Regression of Complexity Statistics in Siqalo

Metrics Analysed	Linear Fit	Correlation coefficient (r)	Interpretation of Metrics Analyzed
MN ENN vs MN SHAPE		-0,400	There is a low negative correlation between the average distance between shacks and the average shape complexity of shacks in Siqalo between 2014 and 2019.
NP vs MN SHAPE		+0,968	There is a very high positive correlation between the number of shacks classified and the average shape complexity of shacks between 2014 and 2019. This very high positive correlation follows a directly proportional relationship. The higher the number of shacks classified, the higher the average shape complexity of shacks in Siqalo between 2014 and 2019.
MN AREA vs MN SHAPE		+0,985	There is a very high positive correlation between the mean area of shacks and the average shape complexity of shacks between 2014 and 2019. This very high positive correlation follows an indirectly proportional relationship. The larger the mean area of shacks, the higher the average shape complexity of shacks in Siqalo between 2014 and 2019. This indicates that when there is less space between shacks, that shacks are accommodated or adapted to fit less space.

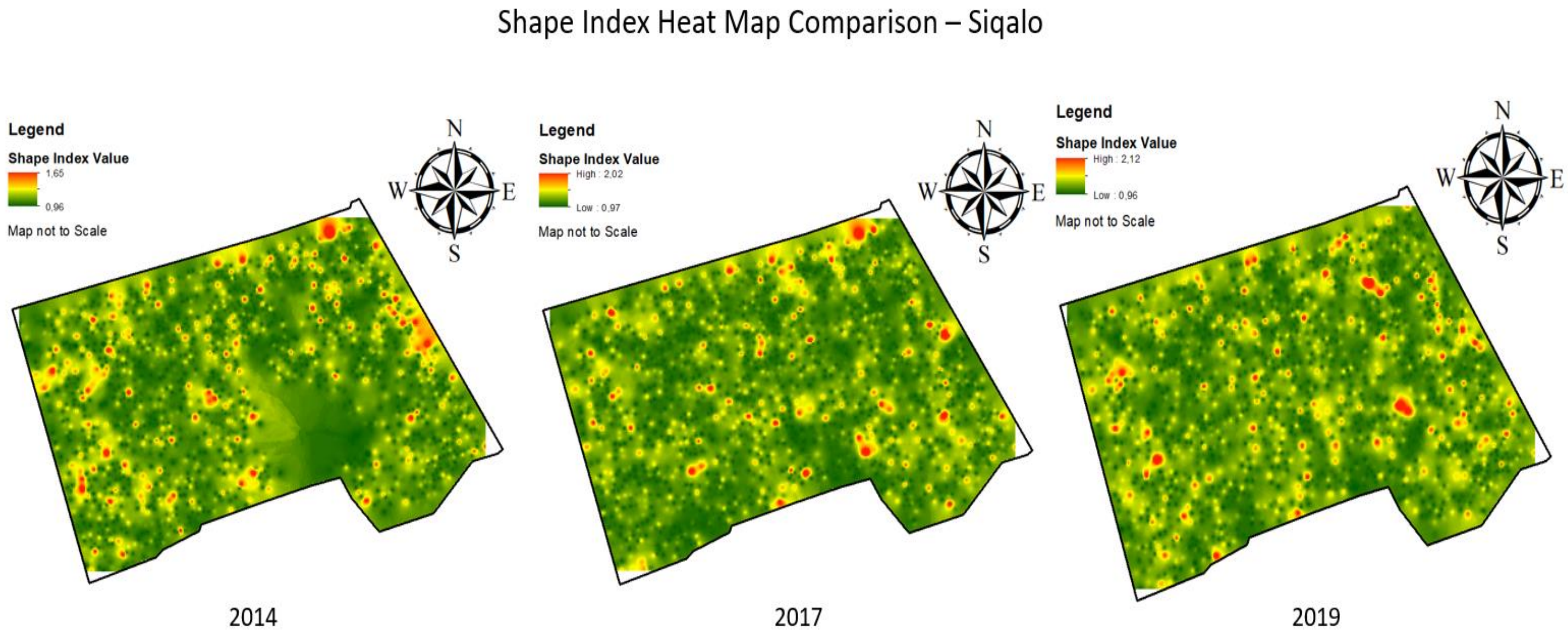
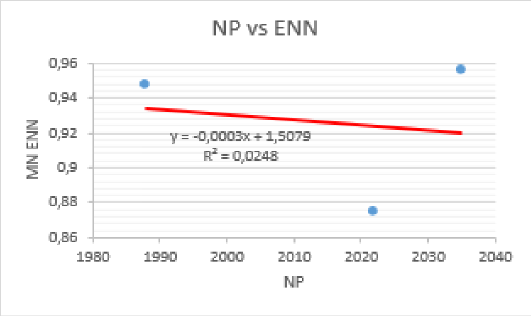
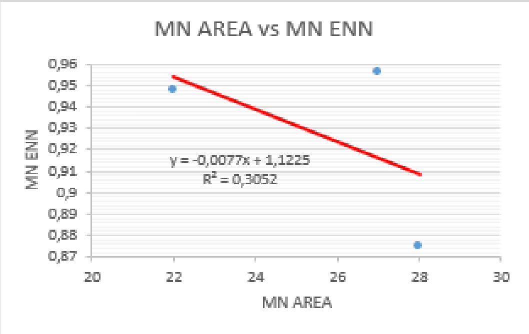


Figure 4.26: Siqalo Shape Index Heat Maps

The range of shack shape complexity has generally increased between 2014 and 2019 (from 0.96 – 1.65 in 2014 to 0.96 – 2.12 in 2019), as illustrated in the figure above. As the settlement develop, there are clusters of relatively large regions of high shack shape complexities on the eastern and northern boundary of Siqalo in 2014, and this graduate to relatively smaller regions of high shack shape complexity all throughout the settlement by 2019. This indicates that there was an initial demand for space to develop shacks along the eastern and northern boundaries, which is strongly linked to the short vicinity to major transport networks and basic amenities. By 2019, the relatively high distributed nature of shack shape complexity indicates that the development of shacks was directed all throughout the settlement, where relatively low shack shape complexity is favoured. This observation indicates that there was an easing the demand of shack development close to basic amenities and major transport networks.

c) Compactness

Table 4.14: Linear Regression of Compactness Statistics in Siqalo

Metrics Analysed	Linear Fit	Correlation coefficient (r)	Interpretation of Metrics Analyzed
NP vs MN ENN		-0,157	There is little if any correlation between the number of shacks classified and the average distance between shacks in Siqalo between 2014 and 2019.
MN AREA vs MN ENN		-0,552	There is a moderate negative correlation between the mean area of shacks and the average distance between shacks between 2014 and 2019. This moderate negative correlation follows an indirectly proportional relationship. The larger the mean area of shacks, the shorter the distance between shacks in Siqalo between 2014 and 2019.

4.7. OLS Analysis

4.7.1. OLS Statistical Results for each informal settlement

The following statistics were determined in the OLS for each informal settlement with respect to its observed independent variable (CA, MN_Shape and MN_ENN) and dependent variables (multi-variate of percentage open space, rate of unemployment, poverty coefficient and GDP):

a) Imizamo Yethu

Table 4.15: OLS Results of Imizamo Yethu's development against socio-economic factors

CA	Standard error	t-stat	p-value
Open_Space	202,765	-6,508	0,003
Unemployment	167,476	0,165	0,877
Poverty	15406,457	0,475	0,660
GDP	2,010	-0,579	0,594

According to the OLS of Imizamo Yethu's development against each socio-economic statistic, the most statistically significant factor is that of open space.

Table 4.16: OLS Results of Imizamo Yethu's shack complexity against socio-economic factors

MN_Shape	Standard error	t-stat	p-value
Open_Space	0,015	-1,631	0,178
Unemployment	0,013	0,125	0,906
Poverty	1,165	-0,896	0,421
GDP	0,000	0,776	0,481

According to the OLS of Imizamo Yethu's shack complexity against each socio-economic statistic, the most statistically significant factor is that of open space.

Table 4.17: OLS Results of Imizamo Yethu's shack compactness against socio-economic factors

MN_ENN	Standard error	t-stat	p-value
Open_Space	0,009	2,380	0,076
Unemployment	0,008	0,508	0,638
Poverty	0,704	1,093	0,336
GDP	0,000	-4,385	0,012

According to the OLS of Imizamo Yethu's shack compactness against each socio-economic statistic, the most statistically significant factor is that of GDP.

b) Langa

Table 4.18: OLS Results of Langa's development against socio-economic factors

CA	Standard error	t-stat	p-value
Open_Space	17,916	-123,863	0,000
Unemployment	16,579	3,171	0,034
Poverty	957,719	0,140	0,896
GDP	0,516	-1,309	0,261

According to the OLS of Langa development against each socio-economic statistic, the most statistically significant factor is that of open space.

Table 4.19: OLS Results of Langa's shack complexity against socio-economic factors

MN_Shape	Standard error	t-stat	p-value
Open_Space	0,007	0,348	0,746
Unemployment	0,007	0,178	0,867
Poverty	0,384	1,350	0,248
GDP	0,000	-1,391	0,237

According to the OLS of Langa's shack complexity against each socio-economic statistic, the most statistically significant factor is that of poverty.

Table 4.20: OLS Results of Langa's shack compactness against socio-economic factors

MN_ENN	Standard error	t-stat	p-value
Open_Space	0,015	0,137	0,898
Unemployment	0,014	-0,581	0,593
Poverty	0,812	-1,452	0,220
GDP	0,000	1,851	0,138

According to the OLS of Langa's shack compactness against each socio-economic statistic, the most statistically significant factor is that of GDP.

c) Siqalo

Table 4.21: OLS Results of Siqalo's development against socio-economic factors

CA	Standard error	t-stat	p-value
Open_Space	1,564	-1048,986	0,001
Unemployment	2,063	0,241	0,850
Poverty	186,277	-3,764	0,165
GDP	0,048	0,032	0,980

According to the OLS of Siqalo's development against each socio-economic statistic, the most statistically significant factor is that of open space.

Table 4.22: OLS Results of Siqalo's shack complexity against socio-economic factors

MN_Shape	Standard error	t-stat	p-value
Open_Space	0,001	-10,874	0,058
Unemployment	0,001	1,014	0,496
Poverty	0,087	-2,860	0,214
GDP	0,000	2,134	0,279

According to the OLS of Siqalo's shack complexity against each socio-economic statistic, the most statistically significant factor is that of open space.

Table 4.23: OLS Results of Langa's shack compactness against socio-economic factors

MN_ENN	Standard error	t-stat	p-value
Open_Space	0,010	-3,110	0,198
Unemployment	0,013	0,722	0,602
Poverty	1,140	-2,321	0,259
GDP	0,000	1,630	0,350

According to the OLS of Siqlao's shack compactness against each socio-economic statistic, the most statistically significant factor is that of poverty.

4.7.2. OLS Analysis between outlined socio-economic data and developmental statistics of each informal settlement

According to the OLS statistics results, the most significant statistics (lowest p-value) relating the outlined socio-economic data and that of developmental statistics can be found in the following table:

Table 4.24: Most statistically significant socio-economic factors when related to developmental statistics of each informal settlement

	Imizamo Yethu	Langa	Siqalo
lowest p value: CA	Open Space	Open Space	Open Space
lowest p value: MN_SHAPE	Open Space	GDP	Poverty
lowest p value: MN_ENN	GDP	GDP	Open Space

After evaluating each developmental statistic relative to the outlined socio-economic statistics, the following analysis was determined:

- In Imizamo Yethu, the availability of space for shack development was the most influential explanatory variable. During the investigative period, the total area covered by shacks in Imizamo Yethu has increased, indicating densification. Even though Imizamo Yethu was established 31 years ago, it has a changing physical structure of densification that follows the establishment and developmental geometric stages of informal settlement development (section 2.3.1.). The reasoning behind densification of shacks is due to the availability of employment opportunities, proximity to formal and informal transport networks, and the need to extend shack development to less favourable locations within the settlement site once a threshold of settlers was accommodated within the settlement. Imizamo Yethu follows this trend as in very close vicinity, employment opportunities and formal transport networks exist. Due to Imizamo Yethu comprising all the above characteristics, it is a very suitable site for informal dwellers. Thus, open space was of high importance relative to the other OLS factors investigated. Additionally, the shape of shacks in Imizamo Yethu has become more complex over the investigative period, according to the change detection results (section 4.5.7). In the developmental stage of Imizamo Yethu, the progression of more complex shack shapes agreed with past literature – that an increase in total shack area has made shacks more convoluted in shape when compared to the relatively square shack shapes observed in the early stages of the investigative period. Imizamo Yethu is situated in a very steep valley (section 3.2.1.). Here, informal development thrived as formal development elected to be situated on more favourable land. Initial development most likely occurred on gentler slopes, and densification occurred on steeper slopes. However, when evaluating the mean nearest neighbour of shacks in the settlement, GDP was determined as the most influential factor. This indicated that the economy within Cape Town had a higher value when shacks were closest to each other.

Settlers of Imizamo Yethu may have taken advantage of the economic opportunities presented, or more locally, within Hout Bay.

- In Langa, GDP was the most influential statistic in shack development over the investigative period. In a spatial context, Langa is the closest of all the investigated settlements to the CBD (section 3.2.1). It also has many other economic advantages. Such as the settlement's relative location and proximity to a national road (major formal transport network) and a major industrial area providing employment opportunities. These may be localised explanations as to why GDP may play as influential a role as it does in the context of shack development in Langa. Langa, relative to the other two informal settlements investigated, exhibits a unique spatial trend. According to the change detection (Figure 4.8), Langa has undergone formalization during the investigative period. This resulted in an initial drop in total area covered by shacks (Figure 4.6). However, the development of the settlement continued in a similar fashion when compared to Imizamo Yethu. Subsequent densification and an increase in the shack area occurred even though the settlement was exposed to constant formalization throughout the investigative period. The amount of space in the settlement would then have decreased, giving it less of an influence when compared to the GDP pull factors in the local area. This influence extended to the shape of shacks and the space between shacks. According to the OLS analysis, open space has had an influence in Langa. It is the most statistically significant factor related to the total area covered by shacks. This means that every available space within the settlement would have been utilized in terms of its capability to build shacks. This, in turn, provides an interesting context relative to the formalization of the site. Even though the settlement site had less space for the total area covered by shacks, the GDP pull factors may have made the open space more attractive for shacks to be built within the settlement.
- In Siqalo, the availability of space for shack development was the most influential statistic. This is the same result as Imizamo Yethu and naturally there are similarities in the explanation for why open space was the most influential statistic. However, the difference between Siqalo and Imizamo Yethu was the time of establishment. Siqalo was established within the investigative period between the years 2011 and 2014. Due to this, Siqalo exhibits more establishment factors in its geometric developmental stage. A factor includes the Siqalo being formed and developed in proximity to a major road (section 3.2.1) – major transport network. The physical location was also established near a middle-class socio-economic neighbourhood (Mitchell's Plain). The similarities between Imizamo Yethu and Siqalo include the unattractiveness of the sight. Whereas Imizamo Yethu was established on a slopy valley, Siqalo was established on marshy land (section 3.2.1). The developmental characteristics of Siqalo in a change detection context also related to Imizamo Yethu similarly (section 4.5.7). This came in the form of densification of shacks once the initial shacks were developed in the settlement. The initial shacks were developed in proximity to the formal major road and the informal or internal informal roads of the settlement. Due to the similarities between the settlements, open space was most influential in the shack development of Siqalo. The shack shape, however, was influenced by the poverty coefficient of Cape Town during the investigative period. Shack shape in Siqalo became more complex as the rate of poverty increased in Cape Town. The nearest neighbour between shacks was most influenced by open space.

By viewing the OLS analysis results, open space is clearly the dominant explanatory variable in its influence in shack development across all settlements. When reviewing past research, the most important reason for the establishment of informal settlements was that they are created as a method for the urban poor to cope with the inability to afford formal housing (Manal et al. 1998; Wekesa et al, 2010). Once these settlements are established, road networks and employment opportunities would drive the rate of development within the settlement. According to the results of the OLS, open space is determined to be the most influential factor in shack development across the investigated informal settlements. Densification was the determined method of development within the settlements, with attractive spaces being used first. Upon a threshold reached within the settlement, unattractive spaces would then be favoured for the densification of shacks.

4.8. Summary of the analysis section

After conducting a change detection of shacks across the investigated informal settlements, the following were determined:

- The most relevant statistics that allowed for the description in informal settlement development were the total area covered by shacks, the complexity in the shape of shacks, and the compactness of shacks.
- Across all the informal settlements, the area shacks covered within the extent of each informal settlement increased. This increase in the area covered by shacks influenced the mean shape and compactness of shacks. The reasoning for this can be found in section 4.5.7. for each investigated informal settlement.

Further analysis of the internal dynamics was made when conducting linear regression between spatial metrics within each informal settlement. Specific relationships between these metrics can be found in section 4.6.

Using the relevant statistics of informal settlement development and the outlined socio-economic factors collected from the results, an OLS was conducted to determine the most statistically significant socio-economic factors that related to the internal dynamics of each informal settlement. Subsequent analysis was then completed to explain the OLS results.

5. Conclusions

This section outlines the conclusions made in this research project and to determine the following; whether the aim and objectives of have been achieved, an examination of the results and any recommendations that can be made for any future project relating to research like the topics discussed in this one.

5.7. Review of Aim and Objectives

Objective A and B: What is the change of the internal settlement dynamics and development within each targeted informal settlement?

- Across each informal settlement, densification of shacks was internally the most prominent phenomenon of growth or development. This densification followed along the major formal external transport routes and internal informal movement networks. The densification was most influenced by the occurrence of open space.
- With respect to each informal settlement, unique internal dynamics included:
  a) Imizamo Yethu – initial shacks were developed near major external transport networks, employment opportunities and gentle slopes (attractive open space). Subsequent shack development and densification was made on available and steep sloped areas (less attractive open space).
  b) Langa – initially shacks were evenly distributed throughout the settlement. The process of formalization occurred during the investigative process. This resulted in an initial drop in the area covered by shacks in the informal settlement. However, subsequent densification in any available open space occurred. The settlement is surrounded by major formal transport routes and upon investigation, has not influenced shack development within the settlement.
  c) Siqalo – the establishment of Siqalo occurred during the investigative period. Initial shack development occurred near the major formal and external transport routes. The initial development also followed along the edge of the settlement and in proximity to the informal internal movement networks. Subsequent shack development occurred through densification, with initial densification occurring near the external transport routes and nearby middle-class income area (Mitchell's Plain).

Objective C: What factors are the most statistically significant in describing the development in each targeted informal settlement?

Through conducting an OLS analysis, the discussion on the most statistically significant factors describing shack development can be found in section 4.7.2. Across all the informal settlements, open space is the most significant factor. Initial establishment conditions favoured the existence of external and formal transport routes, external employment opportunities and nearby middle-to-high neighbourhoods. However, unique factors are discussed in accordance to shack development in section 4.7.2. These unique factors include the total area covered by shacks in each informal settlement investigated, the mean shape of shacks and the mean nearest neighbour of shacks in relation to socio-economic explanatory variables.

Objective D: Which macro and micro socio-economic factors are most statistically significant in describing informal settlement development within the City of Cape Town?

After producing an OLS analysis, a comparison test of the OLS analysis results was used to determine which macro and micro socio-economic factors were most statistically significant across all the investigated informal settlements. The following conclusions were determined:

- Across all the informal settlements, the availability of open space was the most statistically significant factor in explaining the development in shacks.
- Across all the informal settlements, there was no one most statistically significant factor that explained the development in shack shape over time as open space, GDP and poverty was statistically significant for Imizamo Yethu, Langa and Siqalo respectively.
- Across all the informal settlements, there was no one most statistically significant factor that explained the development in shack compactness over time. However, GDP played the most influential role in explaining shack compactness over time. This was validated by both Imizamo Yethu and Langa having GDP as the most statistically significant factor to explain shack compactness over time. In Siqalo, open space was the most statistically significant factor in explaining the development of shack compactness.

5.8. Limitation to research

Limiting factors throughout the course of this project mostly involved the availability of high temporal remotely sensed data and object-based classification software will improve a research project like this. It must be commented however that this data can be retrieved by private means, but at a cost. Similarly, more advanced software will also be costly. This may cause such a project to be less pragmatic to conduct.

The major limiting factor found throughout the course of this project was the availability of socio-economic data at the local level. The Western Province's greatest source of socio-economic data comes from mass census surveys, with the last three major surveys having been conducted in 2001, 2011 and 2016. Since this research project has an aim to determine the relationship between the current socio-economic state at the local level and the change in the internal dynamics of informal settlements, an increased frequency in the availability of detailed socio-economic data will assist in problem solving within the province.

5.9. Recommendations for further research

It is recommended that other informal settlements are investigated to understand their reasons for shack development. If this can occur in conjunction with more epochs investigated, a more detailed understanding of shack development can be determined. It is recommended that these epochs exist to the monthly level.

Further research can also be conducted within the formalization stage of the geometric development of informal settlements. By gaining a greater understanding of the influence of formalization, the understanding of informal shack growth can be enhanced in a project relating to change detection.

5.10. Concluding Remarks

According to the results and analysis of this project, shack development growth within the City of Cape Town is most influenced by the availability of open space. The open space may not always be favourable for the development of housing, however open space in conjunction with proximity to major transport networks and employment opportunities provides reasoning for shacks to develop within informal settlements.

References

Abebe, M.S., Derebew, K.T. and Gemeda, D.O. (2019). 'Exploiting temporal-spatial patterns of informal settlements using GIS and remote sensing technique: a case study of Jimma city, Southwestern Ethiopia', Environmental Systems Research, 8(6)

Alpaydin, E. (2020). 'Introduction to Machine Learning', Fourth Edition, 3-11

Aggarwal, R., Ranganathan, P., and Pramesh, C.S. (2017). 'Common pitfalls in statistical analysis: Logistic regression', Perspect Clin Res, 8(3), 148-151

Augustijn-Beckers, E., Flacke, J. and Retsios, B. (2011). 'Simulating informal settlement growth in Dar es Salaam, Tanzania: An agent-based housing model', Computers, Environment and Urban Systems, 93-103

Baatz, M., Benz, U., Dehghani, S., Heynen, M., Holtje, A., Hofmann, P. and Lingenfelder, I. (2004). eCognition professional user guide, version 4.0., Definiens Imaging GmbH, Munich, Germany

Balogun, A.L., Matori, A.N. and Lawal, D.U. (2011). 'A Gepgraphic Information System and Multi-Criteria Decision Analysis in Proposing New Recreational Park Sites in Universiti Teknologi Malaysia', Modern Applied Science. 5(3)

Barry, M. and Ruther, H. (2001). 'Data collection and management for informal settlement upgrades', International Conference on Spatial Information for Sustainable Development'

Mahesh, B. (2018). 'Machine Learning Algorithms- A review', International Journal of Science and Research, 381-386

Beardsley, J. and Werthmann, C. (2008). 'Improving informal settlements', Havard Design Magazine, 1-3

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M. (2004). 'Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information', ISPRS Journal of Photogrammetry and Remote Sensing, 239-240

Berne, A., Delrieu, G., Creutin, J. and Obled, C. (2004). 'Temporal and spatial resolution of rainfall measurements required for urban hydrology', Journal of Hydrology, 166-167

Bishop, C.M. (2011). 'Pattern Recognition and Machine Learning', Information Science and Statistics, 1-3

Balogun, I.A., Adeyewa, D.Z., Balogun, A.A. and Morakinyo, T.E. (2011). 'Analysis of urban expansion and land use changes in Akure, Nigeria, using remote sensing and geographic information system (GIS) techniques', Journal of Geography and Regional Planning, Vol. 4, 533-541

Census. (2001). ' https://www.statssa.gov.za/?page_id=3892 '.

Census. (2011). ' https://www.statssa.gov.za/?page_id=3839 '

Congalton, R.G. (2001). 'Accuracy assessment and validation of remotely sensed and other spatial information', International Journal of Wildland Fire, 10(4), 321-328

Coppin, P., Jonckheere, L. Nackaerts, K., Muys, B. and Lambin, E. (2004). 'Digital change detection methods in ecosystem monitoring: a review', International Journal of Remote Sensing, 25(9), 1565-1596

Community Survey. (2016).
'https://www.statssa.gov.za/?page_id=6283#:~:text=The%20Community%20Survey%202016(CS,Survey%20was%20conducted%20in%202007. '.

Congalton, R.G. and Mead, R.A. (1986). 'A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrices', IEEE, Vol 24, 169-174

Cunningham, P. and Delany, S. J. (2007). 'k-Nearest Neighbour Classifiers', Technical Report UCD-CSI-2007-4, 1-17

Darwish, A., Leukert, K. and Reinhardt, W. (2003). 'Image Segmentation for the Purpose of Object-based Classification', IGARSS, IEEE 2003 International GeoScience and Remote Sensing Symposium

Dovey. K. (2012). 'Informal urbanism and complex adaptive assemblage', International Development Planning Review, 34(4), 349-368

Dovey. K. (2015). 'Sustainable informal settlements', Social and Behavioral Sciences, 179, 5-13

Dovey. K. and King, R. (2011). 'Forms of Informalilty', Built Environment. 37(1), 11-29

Du, W., Han, Y.S. and Chen, S. (2004). 'Privacy-preserving multivariate statistical analysis: Linear Regression and Classification', SIAM International, 1-12

Dubovyk, O., Sliuzas, R. and Flacke, J. (2010). 'Spatio-temporal modelling of informal settlement development in Sancaktepe district, Istanbul, Turkey', ISPRS Journal of Photogrammetry and Remote Sensing 66 (2011), 235-246

El-Batran, M. and Arandel, C. (1998). 'A shelter of their own: informal settlement expansion in Greater Cairo and government responses', Environment and Urbanization, 10(1)

Fallatah, A., Jones, S., Mitchell, D. and Kohli, D. (2019). 'Mapping informal settlement indicators using object-oriented analysis in the Middle East', International Journal of Digital Earth, Vol. 12(7), 802-824

Fallatah, A., Jones, S., Wallace, L. and Mitchell, D. (2022). 'Combining object-based machine learning with long-term time-series analysis for informal settlement identification', Remote Sensing (MDPI), Vol. 14, 1226

Faweet, T. (2006). 'An introduction to roe analysis', Pattern recognition letters, 27(8), 861-874

Fletcher, R. Everitt, J.H. and Yang, C. (2011). 'Employing airborne multispectral imagery to map Brazilian pepper infestation in South Texas', Geocarto International, Vol. 26, 527-536

Gardna, M., Herold, M., Hadley, B. and Roberts, D. (2002). 'The spectral dimension in urban land cover mapping from high-resolution optical remote sensing data', 1-12

Grippa, T., Georganos, S., Zarougui, S., Bognounou, P., Diboulo, E., Forget, Y., Lennert, M., Vanhuysse, S., Mboga, N. and Wolff, E. (2018). 'Mapping Urban Land Use at Street Block Level Using OpenStreetMap, Remote Sensing Data, and Spatial Metrics', International Journal of Geo-Information, 1-21

Harte, E.W., Childs, I.R.W., and Hastings, P.A. (2009). 'Imizamo Yethu: a Case Study of Community Resilience to Fire Hazard in an Informal Settlement Cape Town, South Africa', Geologival Research, 47(2), 142-154

Herold, M., Gardner, M., Hadley, B. and Roberts, D. (2002). 'The spectral dimension in urban land cover mapping from high-resolution optical remote sensing data', 1-12

Herold, M., Gardner, M.E., Roberts, D.A. (2003). 'Spectral resolution requirements for mapping urban areas', IEEE Transactions on Geoscience and Remote Sensing, 1-3

Hofmann, P. (2001). 'Detecting Informal Settlements from IKONOS image data using methods of object oriented image analysis – an example from Cape Town (South Africa), 1-2

Hofmann, P. and Bekkarnayeva, G. (2017). 'Object-based Change Detection of Informal Settlements', Urban Remote Sensing Join Event (IEEE), 5090-5808

Huang, J., Zhang, X., Xin, Q. Sun, Y., and Zhang, P. (2019). "Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network'. ISPRS, Vol. 151, 91-105

Huchzermeyer, M. (2003). 'From "contravention of laws" to "lack of rights": redefining the problem of informal settlements in South Africa', Habitat International 28 (2004) 333-347

Jacobs, F., Jordhuis-Lier, D., and de Wet, P.T. (2015). 'The politics of knowledge Knowledge Management in Informal Settlement Upgrading in Cape Town', Urban Forum, 26, 425-441

Lemanski, C. (2009). 'Augmented informality: South Africa's backyard dwelling as a by-product of formal housing policies', Habitat International, 33(4), 472-484

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J. and Worthy, L.D. (2006). 'Land-cover change detection using multi-temporal MODIS NDVI data', Remote Sensing of Environment, 105m, 142-154

Imani, M. and Ghassemian, H. (2020). 'An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges', Information Fusion, Vol. 59, 59-83

Ishtiyaq, M. and Kumar, S. (2011). 'Typology of Informal Settlements and Distribution of Slums in the NCT, Dehli', Institute of Town Planners, India Journal, 20-33

Jacobs, F., Jordhus-Lier, D. and de Wet, P.T. (2015). 'The Politics of Knowledge: Knowledge Management in informal settlement upgrading in Cape Town', Urban Forum

Karakis, S. Marangoz, A.M. and Buyukalih, G. (2006). 'Analysis of segmentation parameters in eCognition software using high resolution quickbird MS imagery', ISPRS, 1-4

Kohli, D. (2013). 'Spatial metrics and image texture for slum detection', N-Aerus XIV, Vol 12, 1-8

Kohli, D., Sliuzas, R. and Stein, A. (2016). 'Urban slum detection using texture and spatial metrics derived from satellite imagery', Journal of Spatial Science, 404-426

Kuffer, M. and Barros, J. (2011). 'Urban Morphology of Unplanned Settlements: The Use of Spatial Metrics in VHR Remotely Sensed Images', Mapping Global Change, Procedia Environmental Sciences 152-157

Kuffer, M., Sliuzas, R.V., and Barros, J. (2014). 'The development of morphological unplanned settlement index using very-high resolution (VHR) imagery, Computers, Environment and Urban Systems, 48, 138-152

Kupfer, J.A. (2012). 'Landscape ecology and biogeography: Rethinking landscape metrics in a post-fragstats landcape', Progress in Physical Geography: Earth and Environment, 36(3)

Langford, J., Beygelzimer, A. and Kakade, S. (2006). 'Cover trees for nearest neighbor'. 'ICML 2006', 97-104

Lu, D., Mausel, P., Broandizio, F. and Moran, E. (2004). 'Change detection techniques', International Journal of Remote Sensing, 25(12), 2365-2047

Li, Y., Hu, W., Dong, D. and Zhang, X. (2019). 'Building damage detection from post-event imagery using single shot multibox detector', Applied Sciences (MDPI), Vol. 9(6), 1128

Lillesand, T. M. Kiefer, R. W. & Chipman, J. W. (2008). 'Introduction to remote sensing and image interpretation', 6th ed., United States of America, John Wiley & sons

Lui, D. and Xia, F. (2010). 'Assessing object-based classification: advantages and limitations', Remote Sensing Letters, Vol. 1, No.4, 187-194

Lunetta, R.S. and Lyon, J.G. (2005). 'Remote sensing and GIS accuracy assessment', 1-3

Manal, E. and Arandel, C. (1998). 'A shelter of their own: informal settlement expansion in Greater Cairo and government responses', Environment and Urbanization, Vol. 10, No. 1, 217-232

Mas, J.F. (1999). 'Monitoring land-cover changes: a comparison of change detection techniques', International Journal of Remote Sensing, 20(1), 139-152

Massey, R. (2013). 'Competing rationalities and informal settlement upgrading in Cape Town, South Africa: a recipe for failure', Journal of Housing and the Built Environment, 28, 605-613

Mason, S. and Batsavias, E.P. (1997). 'Image-based reconstruction of informal settements', ETHzurich-Research Collection, 1-11

McAuliffe, R.E. (2015). 'T-statistic', 8

McGarigal, K. and Cushman, S.A. (2002). 'Comparative evaluation of experimental approaches to the study of habitat fragmentation effects', Ecological Applications, 12(2), 335-345

McGarigal, K. Tagil, S. and Cushman, S.A. (2009). 'Surface metrics: an alternative to patch metrics for the quantification of landscape structure',Landscape Ecology, 24, 433-450

Moody, A. and Woodcock, C.E. (1995). 'The influence of scale and spatial characteristics of landscape on land-cover mapping using remote sensing', Landscape Ecology, 10, 363-379

Mudau, N. and Mhangara, P. (2022). 'Towards understanding informal settlement growth patterns: Contribution to SDG reporting and spatial planning', Remote Sensing Applications: Society and Environment, Vol. 27, 100801

Muja, M., and Lowe, D.G. (2009). 'Fast approximate nearest neighbors with automatic algorithm configuration', VISAPP 2009

Mukaka, M.M. (2012). 'A guide to appropriate use of correlation coefficient in medical research', Malawi Medical Journal, 24(3)

Naorem, V., Kuffer, M., Verplanke, J. and Kohli, D. (2016). 'Robustness of rule sets using VHR imagery to detect informal settlements – a case of Mumbai, India', GEOBIA 2016

Naoum, S. and Tsanis, I.K. (2004). 'A multiple linear regression GIS module using spatial variables to model orographic rainfall', IWA, Journal of Hydroinformatics, 6.1

Navulur, K. (2007). 'Multispectral image analysis using the object-oriented paradigm', Taylor & Francis CRC Press

Onyekachi, A. F. (2014), `Prospects and challenges of informal settlements and urban upgrading in Abuja', ISSR Journals, Volume 11, No. 2 (2014), 420-426

Patil, T.R. and Sherekar, S.S. (2013). 'Performance Analysis of Naïve Bayes and J48 Classification Algorithm for Data Classification', International Journal of Computer Science and Applications, Vol. 6, 256-261

Poelmans, L. and Van Rompaey, A. (2009). 'Complexity and performance of urban expansion models', Computers, Environment and Urban Systems 34 (1), 17–27

Pohlmann, J.T. and Leitner, D.W. (2003). 'A comparison of ordinary least squares and logistic regression', The Ohio Journal of Science, 103(5)

Price, J.C. (1997). 'Spectral band selection for visible-near infrared remote sensing: spectral-spatial resolution tradeoffs', IEEE Transactions on Geoscience and Remote Sensing, Vol 35, 1277-1285

Prins, A.J. and Van Niekerk, A. (2020). 'Crop type mapping using LiDAR, Sentinel-2 and aerial imagery with machine learning algorithms', Geo-spatial Information Science, Vol. 24, I.2

Ra, P.K., Nathawat, M.S. and Onagh, M. (2012). 'Application of Multiple Linear Regression Model through GIS and Remote Sensing for Malaria Mapping in Varanasi District, India', Heath Science Journal, Vol. 6, Issue 4

Saravanan, R. and Sujatha, P. (2018). 'A state of art techniques on machine learning algorithms: A perspective of supervised learning approaches in data classification', IEEE

Sathya, R. (2013). 'Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification', International Journal of Advanced Research in Artificial Intelligence, 34-35

Saura, S. and Martinez-Millan, J. (2001). 'Sensitivity of Landscape Pattern Metrics to Map Spatial Extent', Photogrammetric Engineering and Remote Sensing, 67(9), 1027-1036

Shekhar, S. (2012). 'Detecting Slums from Quick Bird Data in Pune using an Object Oriented Approach', International Archives of the Photgrammetry, Remote Sensing and Spatial Information Sciences, Volume 39-B8, 519-524

Shuvo, F.K. and Janssen, P. (2013). ‘Modelling informal settlements using a hybrid automata approach’, Open Systems: Proceedings of the 18th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2013), 591–600

Singh, A. (1989). ‘Review Article Digital change detection techniques using remotely sensed data’, International Journal of Remote Sensing, 10 (6), 989-1003

Sluizas, R. (1988). ‘Problems in monitoring the growth of a squatter settlement: The housing process in Menzese, Dar es Salaam’, ICT, Enschede

Sliuzas, R. (2004). ‘Managing informal settlements: a study using geo-information in Dar es Salaam, Tanzania’, Ph.D. Thesis. Utrecht University

Smith, H.M. (2005). ‘The relationship between settlement density and informal settlement fires: case study of Imizamo Yethu, Hout Bay and Joe Slovo, Cape Town Metropolis’, Geo-information for Disaster Management, 1333-1355

Sobreira, F. and Gomes, M. (2001). ‘The geometry of slums: Boundaries, packing and Diversity’

StatsSA. (2017). ‘ https://www.statssa.gov.za/publications/P0302/P03022017.pdf ‘

Steiniger, S. and Weibel, R. (2009). ‘GIS Software – A description in 1000 words’, Encyclopeadia of Geography, 1-2

Stehman, S.V. (1997). ‘Selecting and interpreting measures of thematic classification accuracy’, Remote sensing of Environment 62(1), 77-89

Thiese, M.S., Ronna, B. and Ott, U. (2016). ‘P value interpretations and considerations’, 8(9), 928-931

Tian, J. and Chen, D.M. (2007). ‘Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition’, International Journal of Remote Sensing, 28:20, 4625-4644

Trimble Germany GmbH. User Guide-eCognition® Developer, Trimble: Munich, Germany, 2015

Turner, M.G. (1989). ‘Landscape Ecology: The effect of Pattern on Process’, Annu. Rev. Ecol. Syst., 20, 171-197

Turner, M.G., Pearson, S.M., Bolstad, P. and Wear, D.N. (2003). ‘Effects of land-cover change on spatial pattern of forest communities in the South Appalachian Mountains (USA)’, IEEE, 1-16

Turok, I. (2016). ‘Getting urbanization to work in Africa: the role of the urban land-infrastructure-finance nexus’, Area Development and Policy, 1(1), 30-47

UN-Habitat, (2020). ‘Caribbean strategy for informal settlement upgrading’, A guide to inclusive and resilient urbanisation.

Valente, J. Sari, B., Kooistra, L. Kramer, H. and Mucher, S. (2020). ‘Automated crop plant counting from very high-resolution aerial imagery’, Precision Agriculture, Vol. 21, 1266-1384

Verburg, P., de Nijs, T., van Eck, R., Visser and J., de Jong, K. (2004). ‘A method to analyse neighborhood characteristics of land use patterns’, Computers, Environment and Urban Systems 28 (6), 667–690

Walter, V. (2003) ‘Object-based classification of remote sensing data for change detection’, ISPRS Journal of Photogrammetry and Remote Sensing, 58 (2004), 225-238

Ward, C.M., Harguess, J. and Hilton, C. (2018). ‘Ship classification from overhead imagery using synthetic data and domain adaptation’, IEEE

Wekesa, B.W., Steyn, G.S. and Otieno, F.A.O. (2010). ‘A review of physical and socio-economic characteristics and intervention approaches of informal settlements’, Habitat International 35 (2011), 238-245

Witharana, C., Zhang, W., Liljedahl, A.K. and Kanevskiy, M. (2018). ‘Deep Convolutional Neural Networks for Automated Characterization of Arctic Ice-Wedge Polygons in Very High Spatial Resolution Aerial Imagery’, Remote Sensing, 10/9, 1487

WCISSF. (2016). ‘ https://www.westerncape.gov.za/assets/departments/human-settlements/docs/issp/western_cape_issp_strategic_framework_2016.pdf ’

Wu, N. Crusiol, L.G.T., Liu, G., Wuyun, D. and Han, G. (2023). ‘Comparing mahine learning algorithms for pixel/object-based classifications of semi-arid grassland in Northern China using multisource medium resolution’, Remote Sensing, 1-22

Wulder, M.A., Hall, R.J., Coops, N.C. and Franklin, S.E. (2004). ‘Highly Spatial Resolution Remotely Sensed Data for Ecosystem Characterisation’, BioScience, Vol. 54, 511-521

Yonder, A. (2006). ‘ISs in Istanbul, Turkey—from shacks to high rises’, In: Proc. Salzburg Congress on Urban Planning and Development, SCUPAD, Salzburg, 11–14

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M. and Schlrokauer, D. (2006). ‘Object-based detailed vegetation classification with Airborne High Spatial Resolution Remote Sensing Imagery’, Photogrammetric, Engineering and Remote Sensing, Vol. 72, 799-811

Zhang, Y., Maxwell, T., Tong, H. and Dey, V. (2010). ‘Development of a supervised software tool for automated determination of optimal segmentation parameters for eCognition’, ISPRS, Vol. 38, Part 7B, 690-696

Zhang, Z., Ren, H., Das, P. Guo, W. and Kisovi, L.M. (2020). ‘Population density and Spatial Patterns of Informal Settlements in Nairobi, Kenya’, Sustainability, 12(18), 7717

Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T.A., and Taubembok, H. (2019).’Understanding an urbanizing planet: Strategic directions for remote sensing’, Remote Sensing of Environment, Vol. 228, 164-182

Appendix A:

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	28	0	28	1	0
2	C_2	2	20	22	0,909091	0
3	Total	30	20	50	0	0
4	P_Accuracy	0,933333	1	0	0,96	0
5	Kappa	0	0	0	0	0,918033

Confusion matrix of Imizamo Yethu 2019

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	28	2	30	0,933333	0
2	C_2	1	19	20	0,95	0
3	Total	29	21	50	0	0
4	P_Accuracy	0,965517	0,904762	0	0,94	0
5	Kappa	0	0	0	0	0,876033

Confusion matrix of Imizamo Yethu 2017

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	23	1	24	0,958333	0
2	C_2	0	26	26	1	0
3	Total	23	27	50	0	0
4	P_Accuracy	1	0,962963	0	0,98	0
5	Kappa	0	0	0	0	0,959872

Confusion matrix of Imizamo Yethu 2014

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	27	0	27	1	0
2	C_2	3	20	23	0,869565	0
3	Total	30	20	50	0	0
4	P_Accuracy	0,9	1	0	0,94	0
5	Kappa	0	0	0	0	0,878049

Confusion matrix of Imizamo Yethu 2011

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	14	0	14	1	0
2	C_2	2	34	36	0,944444	0
3	Total	16	34	50	0	0
4	P_Accuracy	0,875	1	0	0,96	0
5	Kappa	0	0	0	0	0,904943

Confusion matrix of Langa 2019

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	15	0	15	1	0
2	C_2	1	34	35	0,971429	0
3	Total	16	34	50	0	0
4	P_Accuracy	0,9375	1	0	0,98	0
5	Kappa	0	0	0	0	0,953271

Confusion matrix of Langa 2017

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	11	0	11	1	0
2	C_2	2	37	39	0,948718	0
3	Total	13	37	50	0	0
4	P_Accuracy	0,846154	1	0	0,96	0
5	Kappa	0	0	0	0	0,890591

Confusion matrix of Langa 2014

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	17	0	17	1	0
2	C_2	2	31	33	0,939394	0
3	Total	19	31	50	0	0
4	P_Accuracy	0,894737	1	0	0,96	0
5	Kappa	0	0	0	0	0,913345

Confusion matrix of Langa 2011

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	18	0	18	1	0
2	C_2	1	31	32	0,96875	0
3	Total	19	31	50	0	0
4	P_Accuracy	0,947368	1	0	0,98	0
5	Kappa	0	0	0	0	0,957118

Confusion matrix of Siqalo 2019

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	16	2	18	0,888889	0
2	C_2	0	32	32	1	0
3	Total	16	34	50	0	0
4	P_Accuracy	1	0,941176	0	0,96	0
5	Kappa	0	0	0	0	0,911032

Confusion matrix of Siqalo 2017

OBJECTID *	ClassValue	C_1	C_2	Total	U_Accuracy	Kappa
1	C_1	9	0	9	1	0
2	C_2	1	40	41	0,97561	0
3	Total	10	40	50	0	0
4	P_Accuracy	0,9	1	0	0,98	0
5	Kappa	0	0	0	0	0,935065

Confusion matrix of Siqalo 2014