Comparison of spatial and spectral properties of Landsat-8 and Sentinel-2 data for mapping plant chlorophyll-*a*

By

Bongwe Vhungoho

(14009707)

A dissertation submitted in partial fulfilment of the requirements for the degree of Master of Environmental Sciences in the Department of Geography and Environmental Sciences at the University of Venda, South Africa

Supervisor: Dr O.E. Malahlela

August 2023

Declaration

I, **Vhungoho Bongwe**, declare that this thesis **"Comparison of spatial and spectral properties of Landsat-8 and Sentinel-2 data for mapping plant chlorophyll-*a*"** is my own, unaided work except for references that are attributed to their sources. It is submitted to the Department of Geography and Environmental Sciences at the University of Venda for the degree of Master of Environmental Sciences. It has not been submitted before for any degree or examination at any other University. The work has been done under the direction of the supervisor listed below:

…… ……

Mr. Vhungoho Bongwe

Acknowledgement

Firstly, I would like to thank the Almighty God for giving me the knowledge, strength, and ability to successfully conduct this research.

I would also like to express my gratitude and special thanks to my supervisor, Dr O.E Malahlela for his invaluable patience and feedback. This research project would not have been complete without his endless academic support and advice at all stages of this study.

I thank my family for their endless emotional support and encouragement. Without them, this would not have been possible.

I appreciate Mr I. Zitha from English department for his editing service, and everyone who was not mentioned above that kept on making me feel better, contributed to my work, and made my time at the University of Venda over the past few years a wonderful experience.

Contents

List of figures

List of tables

Acronyms

AIC – Aikaike's Information Criterion

ANOVA – Analysis of Variance

Car – Carotenoids

CCM – Chlorophyll Content Meter

Chl-*a* – Chlorophyll-*a*

CO_2 – Carbon dioxide

COR – Correlation

DIS – Dissimilarity

DVI – Difference vegetation index

ENT – Entropy

ENVI – Environment for Visualising Images

EWT – Equivalent water thickness

GLCM – Gray Level Co-occurrence Matrix

GPS – Ground Positioning System

HOM – Homogeneity

HSD – Honestly Significant Difference

LAI – Leaf Area Index

MAD – Mean Absolute Deviation

MODIS – Moderate Imaging Spectroradiometer

MSI – Multispectral Imagery

NDVI – Normalized difference vegetation index

NPP – Net primary productivity

O_2 – Oxygen

OLI – Operational Land Imager

PCA – Principal Component Analysis

QGIS – Quantum Geographical Information Systems

QUAC – Quick Atmospheric Correction

R^2 – R-squared

RMSE – Root Mean Square Error

RVI – Ratio vegetation index

SMLR – Stepwise multiple linear regression

SRWI – Simple ratio water index

SVM – Support Vector Machine

SWIR – Shortwave Infrared

TIRS – Thermal Infrared Sensors

USGS – United States Geological Surveys

VDM – Vhembe District Municipality

VNIR – Visible Near Infrared

Abstract

Background

Chlorophyll-*a* (Chl-*a*) is a vital parameter to assess vegetation quality in plants as an indicator of photosynthetic capacity to ensure proper flow of ecosystem services. Nowadays, with a rapid increase in human population and deforestation chl-*a* in higher plants remain at risk from degradation.

Aim

The study sought to compare the spatial and spectral properties Landsat-8 and Sentinel-2 in estimating and mapping chlorophyll-*a* (chl-*a*) concentrations in the Vhembe District Municipality (VDM), South Africa.

Methods

Landsat-8 and Sentinel-2 multispectral data were used in conjunction with field data collected in August 2017, Firstly, this study assessed the correlation between chl-*a* and satellite data. Secondly, explored the optimal spatial resolution for mapping chlorophyll-*a* with stepwise multiple linear regression, and lastly, this study mapped the concentration of plant chl-*a* across a heterogeneous landscape.

Results

When assessing the correlation between chl-*a* and satellite data there was an obvious correlation between chlorophyll-*a* and Band 5 entropy with the highest R^2 of 0.39 at 30 m spatial scale of Landsat-8. However, there was no statistical significant difference amongst the various spatial resolution. The ability of Gray Level Co-occurrence Matrix (GLCM) texture features with Landsat-8 at medium resolution 30 m with $R^2 = 0.55$, $p = 0.000006$, and RMSE = 0.17 µg/m² in estimating plant chl-*a* yielded higher performance accuracy than Sentinel-2 at 10 m resolution with $R^2 = 0.24$, $p = 0$, and RMSE = 0.46 µg/m², and 20 m resolution with $R^2 = 0.52$, $p = 0.00001$, and RMSE = 6.90 µg/m². In exploring the optimal spatial resolution, Landsat-8 at 30 m spatial resolution was optimal for mapping plant chlorophyll-*a*. Lastly, plant chl-*a* were successfully mapped with Landsat-8 multispectral data at 30 m spatial resolution using multiple linear regression. The distribution of plant chlorophyll-*a* varies across the study area and is unevenly distributed due to different species and height.

Discussions

Chlorophyll-*a* as a crucial parameter in plants and requires continuous monitoring to ensure and improve ecosystem services provided by plants. This study estimated plant chlorophyll-*a* across the Vhembe District Municipality. The correlation coefficients derived by GLCM's features demonstrated the ability of GLCM's features in predicting and mapping plant chlorophyll-*a* with Landsat-8 at 30 m spatial resolution. Several studies have successfully mapped chlorophyll-*a* from a homogeneous

landscape, for instance in agricultural crops and limited studies mapped chlorophyll-*a* calibrated from a heterogeneous landscape which motivated this study. Recent advancement in optical remote sensing data opens new avenues for mapping plant chlorophyll-*a* at various spatial resolution.

Conclusion

The use of remote sensing data at 30 m spatial resolution with GLCM features effectively predicted plant chlorophyll-*a* and enable the data processing and performance accuracy assessment. Therefore, this study highlighted the importance of Landsat-8 imagery in vegetation monitoring across a heterogeneous landscape.

Keywords: Chlorophyll-*a*; Stepwise multiple linear regression; Gray level co-occurrence matrix; Spatial resolution; Landsat-8; Sentinel-2

CHAPTER ONE:

General Introduction

1.1 Background of the study

Climate change is an issue of global concern, which cannot be ignored. It poses a lot of implications on faunal and floral diversity and directly affects net primary production (NPP). The causes of climate change include natural activities such as volcanic eruptions, ocean current movement and solar variations (Essays UK, 2018). Additionally, human activities such as deforestation, burning of fossil fuels and emissions from industries across the world bring about global climate change. With the effects of changes in climate patterns, Bellard *et al*. (2012) conducted a study about the upcoming effects of climate change on biodiversity. The study revealed that animals may no longer be adapted to the set of environmental conditions and in consequence, they will migrate. However, for sedentary vegetation ecosystems, changes in climate patterns reduces their ability to survive and provide the ecosystem services. Such services include medicine, fruits (*Sclerocarya birrea*), and habitat, regulate surface

temperature, gaseous exchange ensuring that there is enough oxygen (O_2) and reduction of carbon dioxide (CO_2).

The ability of plants to provide such services depends on their biophysical and biochemical characteristics. The biophysical characteristics, on one hand, include plant structure, canopy structure, leaf area index (LAI) and composition of plants within a certain landscape among others. On the other hand, some of the most important biochemical characteristics include equivalent water thickness (EWT) and pigments that are mostly produced by plants i.e., chlorophyll (Chl) and carotenoids (Car) which largely affect the photosynthesis (Darvishzadeh *et al*., 2019). Pigments in plants are crucial for higher plant assimilatory tissue (Lichtenthaler, 1987) that absorbs light energy and is responsible for its photochemical conversion (Solovchenko *et al*., 2014). Chlorophyll can be represented by chlorophyll *a* (chl-*a*) and chlorophyll *b*, and on the other hand carotenoids by α-carotene and β-carotene, and their concentrations are important determinant of landscape structure and function. Therefore, the accurate and timely monitoring of these parameters, chl-*a* in particular, is crucial in understanding the structure and functioning of ecosystems in a heterogeneous landscape.

Previous studies suggested that the quantity of chl-*a* can be traditionally measured in plants using wet chemical techniques that involve the use of organic solvents to extract pigments (Simon and Helliwell, 1998; Gitelson *et al*., 2009; Liang *et al*., 2017). This technique involves gathering field samples of fresh leaves and analysing them in a lab (Gitelson *et al*., 2009). Nevertheless, even though this approach has been known to be effective, some trees may lead to danger when acquiring leaf samples due to their height. This method is additionally known to be destructive, difficult to meet real-time, non-destructive, and extensive monitoring needs because it is laborious and require destructive sampling (Liu *et al*., 2019). Moreover, the method is suitable for small-scale spatial coverage, but not for large-scale analysis.

The alternative to traditional method is aerial photography which has proven to be effective, since the image data contain less atmospheric interference (due to lower altitude). For example, in their 2015 study, Tang and Shao explored the use of aerial photography in remote sensing for forestry research and practices, Singhal *et al*. (2019) conducted a study on chl-*a* machine learning-based chlorophyll estimation with a multi-spectral unmanned aerial system. The suitability of multispectral sensors on-board unmanned aerial vehicle together with machine learning techniques was found as an alternative to destruction and laborious method for estimating biophysical and biochemical characteristics of plants. However, a challenge with this method is that aerial photography tends to be expensive especially when one is covering a wide area. This method also requires a specialized mapping expertise, aerial photo interpretation skills and large extensive effort for standardizing image contrast and rectification (Fensham and Fairfax, 2002).

Alternatively, the utilization of optical remote sensing techniques and hyperspectral sensing is often preferred for monitoring chl-*a* distribution. This approach surpasses aerial photography because of its narrow contiguous bands that allow for improved feature detection. This technique has excellent potential for precise vegetation parameter retrieval. Zhang *et al*. (2008) documented a study about leaf chl-*a* content retrieval from airborne hyperspectral images. However, this technique has its own limitations and some of the limiting factors in this method include the following: (i) the cost of acquiring such data, (ii) difficult to store data in which there is excess data which may not be used, and (iii) demands high computational power (Ahmad *et al*., 2012).

Currently, the optical remote sensing is an appropriate alternative to hyperspectral remote sensing for estimating chl-*a* distribution because it is cheaper and some of the datasets are freely available and convenient, for example Landsat and Sentinel data. The availability of free datasets opens new avenues for estimating chl-*a*. Optical remote sensing also allows for a wider coverage and because all information is digital, it can be easily integrated with geospatial software. Consequently, using remote sensing images, with their extensive spatial coverage and consistent temporal frequency (Xu *et al*., 2022), providing researchers and managers with a more affordable option than aerial photography and hyperspectral remote sensing. In another study Clevers *et al*. (2017) documented a study on retrieval of LAI and chlorophyll content of a potato crop with vegetation indices using Sentinel-2. The study only focused on Sentinel-2 data and the chlorophyll studied was for a homogeneous landscape. In another study, Darvishzadeh, *et al*. (2019) used an invertible forest reflectance model to map the leaf chlorophyll content of spruce stands using data from Sentinel-2 and RapidEye.

1.2 Problem statement

Although several approaches have been extensively utilized for mapping chl-*a*, however, as far as we are aware limited research scripts assessed how to retrieve plant chl-*a* using gray level co-occurrence matrix (GLCM) method with Landsat-8 and Sentinel-2 on a heterogeneous landscape. It is of great importance for the comparison of chl-*a* mapping using these two sensors for ensuring improved flow of ecosystem services.

1.3 Aim and Objectives

The primary goal in this study focuses specifically on the application when estimating plant chlorophyll-*a* using remote sensing on a heterogeneous landscape in the Vhembe District Municipality (VDM) of Limpopo, South Africa. To fulfil this main aim, the study comprises the following objectives:

i. To assess the correlation between chlorophyll-*a* and multispectral satellite data available at various spatial resolutions.

ii. To explore the optimal spatial resolution for mapping plant chlorophyll-*a* on a heterogeneous landscape using Landsat-8 and Sentinel-2 data.

iii. To map the concentration of chl-*a* across a heterogeneous landscape of the savannah biome.

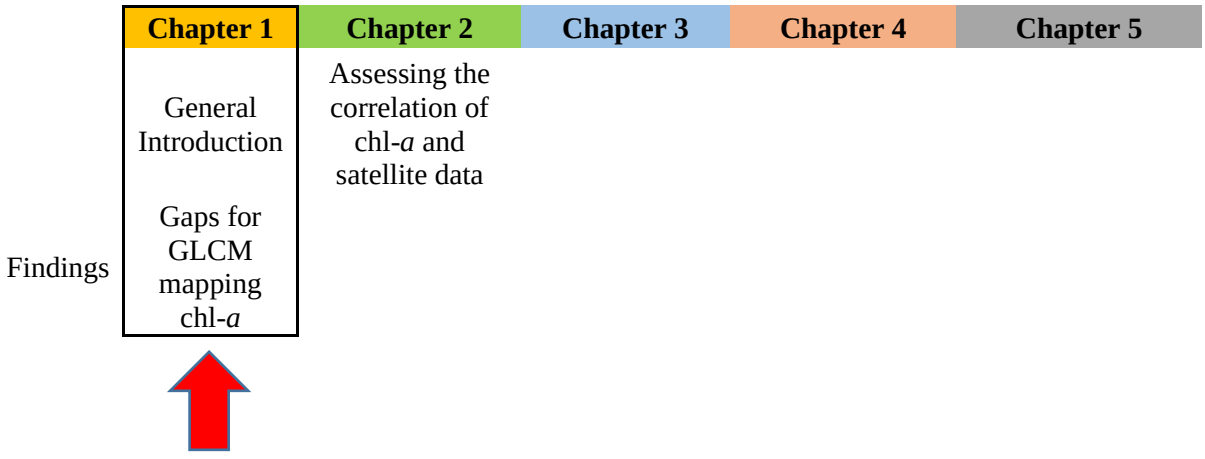
1.4 General methodology

The study used Landsat-8 and Sentinel-2 multispectral data to assess the correlation between chlorophyll-*a* multispectral data at various spatial resolutions in the Vhembe District Municipality, and the results were used to determine the ability of gray level co-occurrence matrix (GLCM) features in extracting chl-*a*. The study then used data from the field with remotely satellite data to predict chlorophyll-*a* content by employing four different GLCM features and the use of stepwise multiple linear regression to create the most effective final model that fits the task. Consequently, this study mapped chlorophyll-*a* across the study area using Landsat-8 data and GLCM features with stepwise multiple linear regression, a detailed methodology is provided in each chapter.

1.5 Research outline

This thesis consisted of five (5) chapters. With chapter 1 being background of the study, aims and objectives of the study. The research objectives i.e., (i - iii) have been addressed in each analysis Chapter. Chapter 2 established correlation between chl-*a* and multispectral satellite data at various spatial resolutions. Chapter 3 explored the optimal spatial resolution for mapping plant chlorophyll-*a* on a heterogeneous landscape using Landsat-8 and Sentinel-2 data. Chapter 4 presents the mapping of chl- *a* concentration calibrated from a heterogeneous landscape of the savannah biome. Chapter 5 presents the conclusion of the study, wherein the outcomes of each objective are discussed in depth together with the recommendations for upcoming studies.

1.6 Summary of study progress



CHAPTER TWO:

Assessing the correlation between chlorophyll-*a* (chl-*a*) and the multispectral satellite data available at various spatial resolutions

Abstract

As among the most significant pigments in higher plants, chlorophyll-*a* has a tremendous impact on the production of food through photosynthesis. Given its significance, scientists continue to exploit chlorophyll-*a* for various applications and at various scales. However, assessing chlorophyll concentrations at a heterogeneous landscape can be very challenging considering the combination of varying types of vegetation species. Additionally, the availability of various remote sensing datasets can present a challenge as to which specific spectral bands are better configured to be sensitive to chl-*a*. Therefore, this study was conducted to evaluate the correlation between chlorophyll-*a* (chl-*a)* and multispectral data of Landsat-8 and Sentinel-2 at various spatial scales at the Vhembe District Municipality in Limpopo, South Africa. In this study, we tested the use of gray level co–occurrence matrix (GLCM) and spectral bands reflectance derived from Landsat-8 (15 m to 30 m resolution) and Sentinel-2 (10 m to 20 m resolution) to estimate the distribution chlorophyll-*a* in the Vhembe District Municipality (VDM), South Africa. Amongst all the GLCM features and spectral bands used in this study, a significance difference was tested with the analysis of variance (ANOVA) tests and Tukey

honestly significant difference (HSD) test. The results revealed that there is no significant difference between all the spatial scales of Landsat-8 and Sentinel-2. The researcher tested the correlation using a correlation of determination (R^2) between chl-*a* and spectral bands, and chl-*a* with GLCM features. The results showed that on their own, spectral bands yielded very low correlations with chl-*a* while the inclusion of texture features improved correlation by 28 % (from R^2 = -0.11 with bands only to R^2 = 0.39 with GLCM features). The results also revealed a slightly higher and significant positive correlation between chlorophyll-*a* and GLCM's entropy (R^2 = 0.39, spatial resolution = 30 m) while the red and green bands yielded very low negative correlation (R^2 = -0.11, spatial resolution = 15 m). The correlation results suggest the use of GLCM features in estimating chlorophyll-*a* effectively at moderately high and medium spatial scales, this was confirmed by the correlation of chl-*a* and GLCM's features being slightly high with Landsat-8 at 30 m and Sentinel-2 at 20 m spatial resolution. Interestingly, the Tukey results suggest that it does not matter which sensor is to be used between Sentinel-2 and Landsat-8 of attempts are made to map chl-*a* at a heterogeneous landscape.

Keywords: Chlorophyll-*a*; Correlation; Gray level co-occurrence matrix; Spatial resolution

2.1 Introduction

Plants cover one-third of earth's surface with different levels of degradation (FAO, 2014), and they are an important indicator of changes to the land's ecological environment (Zhou *et al*., 2020). Plants also play a crucial role in providing ecosystem services such as food, regulating surface temperature and gaseous exchange ensuring enough oxygen to support human well-being and survival (Costanza *et al*., 1997; MA, 2005). Plants are an imperative part of the terrestrial environment not by only producing raw materials for industrial use and biomass, but also in assisting to understand climate change (Lemaire *et al*., 2005).

The relationship between the global environment and terrestrial ecosystems is of main concerns in the study of global climate change (Project, 2010). The biophysical and biochemical characteristics of plants are closely related to the ecological processes that are related to plant processes, such as photosynthesis, transpiration, and net primary productivity (NPP). Chlorophyll-*a* is a crucial index to characterize the physiological state, photosynthetic capacity, and growth conditions of plants among these parameters. Monitoring plant growth conditions and comprehending the process of material and energy exchange between plants and the environment depend greatly on its dynamic change characteristics. (Zhou *et al*., 2020).

Changes in chlorophyll-*a* directly influence biochemical procedures such as photosynthesis and primary productivity (Haboudane *et al*., 2002). Several factors impose depletion of plant chlorophyll-*a*, and these include the rapid growth in human population and climate change. Therefore, a quantitative approach to monitor plant chl-*a* has a significant role in assessing the health, nutrition, and stress levels of plants. Green plants consist of distinctive spectral features, chlorophyll, carotenoids, and water content that can be combined to compose plant spectral characteristics (Estebam *et al*., 2015). Efforts are thus made to quantify plant chlorophyll in order to understand various ecological and landscape processes associated with climate regulations, material flows and potential impacts of anthropogenic activities on ecosystems (Darvishzadeh *et al*., 2019).

One of the efforts for quantifying chlorophyll is through remote sensing approach, in which the collected samples of chlorophyll are analysed and correlated with various satellite-based data. For example, Peng *et al*. (2018) assessed leaf chlorophyll content in degraded temperate vegetation using a portable hyperspectral instrument. Rascher *et al*. (2007) also monitored the spatio-temporal dynamics of photosynthesis with a portable hyperspectral imaging system. In another study, Delegido *et al*. (2011) have predicted chlorophyll concentration and leaf area index (LAI) using Sentinel-2 red-edge bands. More recently, several studies such as Zhou *et al*. (2020), Karimi *et al*. (2022), and Xu *et al*. (2019) have retrieved leaf chlorophyll using multispectral data.

The availability of remote sensing data at various spatial resolutions presents both the opportunities and challenges for estimation of chlorophyll at landscape level. The satellite data with course resolution, such as moderate resolution imaging spectroradiometer (MODIS) with a 1 km pixel is often used for mapping vegetation biophysical and biochemical parameters at regional to global scale (Xie *et al*., 2008). Remote sensing with moderately medium and high-resolution satellite data significantly strengthen the detection of the structural properties of vegetation, making it feasible to employ the spatial combination features of ground objects (Liu *et al*., 2020). While on the other hand, the fine resolution satellite data for mapping vegetation parameters has always been the application of remote sensing's focus and challenge (Meddens *et al*., 2011).

When considering the difference between high and low spatial satellite data, it becomes evident that low spatial resolution can be effective for certain projects that don't require a high level of detail. With these differences in spatial resolution, it becomes imperative to evaluate the efficacy of Landsat-8 and Sentinel-2 in retrieving plant chlorophyll-*a* at various spatial resolutions to establish the best suitable spatial resolution for mapping vegetation chlorophyll. A study by Jaramaz *et al*. (2013) highlighted success in the derivation of vegetation biophysical and biochemical parameters from Sentinel-2 remote sensing data, thus enabling vegetation monitoring. In another study by Zhou *et al*. (2020), Landsat-8 was recommended for evaluating the amount of chlorophyll in leaves. Buma and Lee. (2020) also

highlighted the success of estimating chlorophyll-*a* concentrations using both Landsat-8 and Sentinel-2 datasets.

Various approaches have been utilised on optical remote sensing for extracting and mapping of vegetation parameters. Some such methods include amongst others image differencing (Karthik and Shivakumar, 2017), image classification (Al-doski *et al*., 2013), regression analysis (Matus-Hernandez *et al*., 2018), hybrid models (Verrelst *et al*., 2015) and texture analysis (Lottering *et al*., 2021). Perhaps one of the common image differencing methods include using the normalized difference vegetation index (NDVI) which gained popularity since 1970's (Mutanga and Skidmore, 2004). Additionally, various spectral indices were derived which utilize other spectral bands of Landsat datasets, e.g. difference vegetation index (DVI) and ratio vegetation index (RVI) (Xue and Su, 2017).

Some of image classification methods such as supervised were also adopted for mapping chl-*a* content. For example, Thiemann and Kaufmann. (2000) conducted a study determining chlorophyll content and trophic state of lakes using supervised maximum likelihood classification. The challenge with the image classification is that it requires interpretability and applicability (Small, 2021). The development of various approach with multiple linear regression as one example of a statistical and machine learning technique often has been employed in remote sensing for mapping and estimating vegetation parameters. The examples of such methods include generalized linear regression models (Lopatin *et al*., 2015), artificial neural networks (Adede *et al*., 2019), random forests (Mohammadpour *et al*., 2022) and support vector machines (Marcinkowska *et al*., 2014). There are many studies that utilized random forests and support vector machines for mapping of chl-*a* (Sha *et al*., 2019; Cheng *et al*., 2021; Nofriza *et al*., 2021; Sonobe *et al*., 2021).

The use of gray level co-occurrence matrix (GLCM) has not been fully explored for extracting chl-*a* at various spatial resolutions, particularly for a heterogeneous landscape. The spatial relationship between pixels is examined by the second order statistical texture analysis method known as GLCM, which also determines how frequently a particular combination of pixels appears in an image at a specific direction and distance (Bhargava *et al*., 2020). Recently GLCM features have shown to be more effective in detecting vegetation and measuring biological objects (Hlatshwayo *et al*., 2019) compared to other approaches used for mapping vegetation. The GLCM approach includes feature extraction wherein the initial step, the multispectral image is changed to grayscale, the second step is to acquire spatial features based on the relationship of brightness value defined by a window size (Iqbal *et al*., 2021). For example, Lottering *et al*. (2021) discriminated invasive *Solanum mauritianum* from other species using image texture and sparse PLS discriminant analysis.

In this chapter, with the availability of freely available multispectral data it becomes imperative to establish the correlation between chl-*a* and satellite data, and test for significance difference between chlorophyll-*a* correlations at various spatial resolutions in the Vhembe District Municipality, South

Africa. Therefore, this chapter attempted to assess the correlation between chl-*a* and multispectral data available at various spatial resolutions.

2.2 Materials and methods

2.2.1 Description of the study area

This research was conducted in the Vhembe District Municipality (VDM), which is situated in the northern-most region of South Africa between 23º40' S and 30º00' E (Figure 2.1). The topography of the VDM landscape is varied and includes savannah, semi-arid regions, woodland vegetation, and croplands (Rosmarin, 2013), with different floral and faunal biodiversity. According to a community survey, the district has a total land area of 21 407 square km and a population of 1 393 949 people (Stats SA, 2016). Annual average rainfall of the district is approximately 500mm, and about 87.1% of which falls between October and March, primarily because of the modified orographic rain effect of the Drakensberg mountains connecting the Soutpansberg mountains (Vhembe District Municipality, 2017/18). The distribution of plants in the study area ranges from being sparsely populated in the west to being extremely densely populated with deciduous plants in the south and east. Patches of grassland and forest biomes are found in the savannah biome of the VDM commonly known as Bushveld (VDM, 2017/2018). The plant species that are commonly found in the VDM include *Adansonia digitata (*Baobob tree*)* with simple leaves that are much longer, unlike other *Adansonia* species, *Adansonia digitata* leaflets have varying sizes and stalk lengths. (Rutherford *et al*., 2006; Baum, D.A, 1995). Forest plant species include *Xymalos monospora* (Harv.) Baill with dark green leaves and *Kiggelaria africana*

(L.) with yellow-green leaves beneath and green above (Mostert *et al*., 2008). Temperature in the Vhembe District Municipality is projected to rise by 1-3°C by the middle of the twenty-first century, and the South African National Climate Change Response found that rainfall in the Vhembe district will decline by 5-10%. (Department of Environmental Affairs and Tourism: DEAT, 2004).

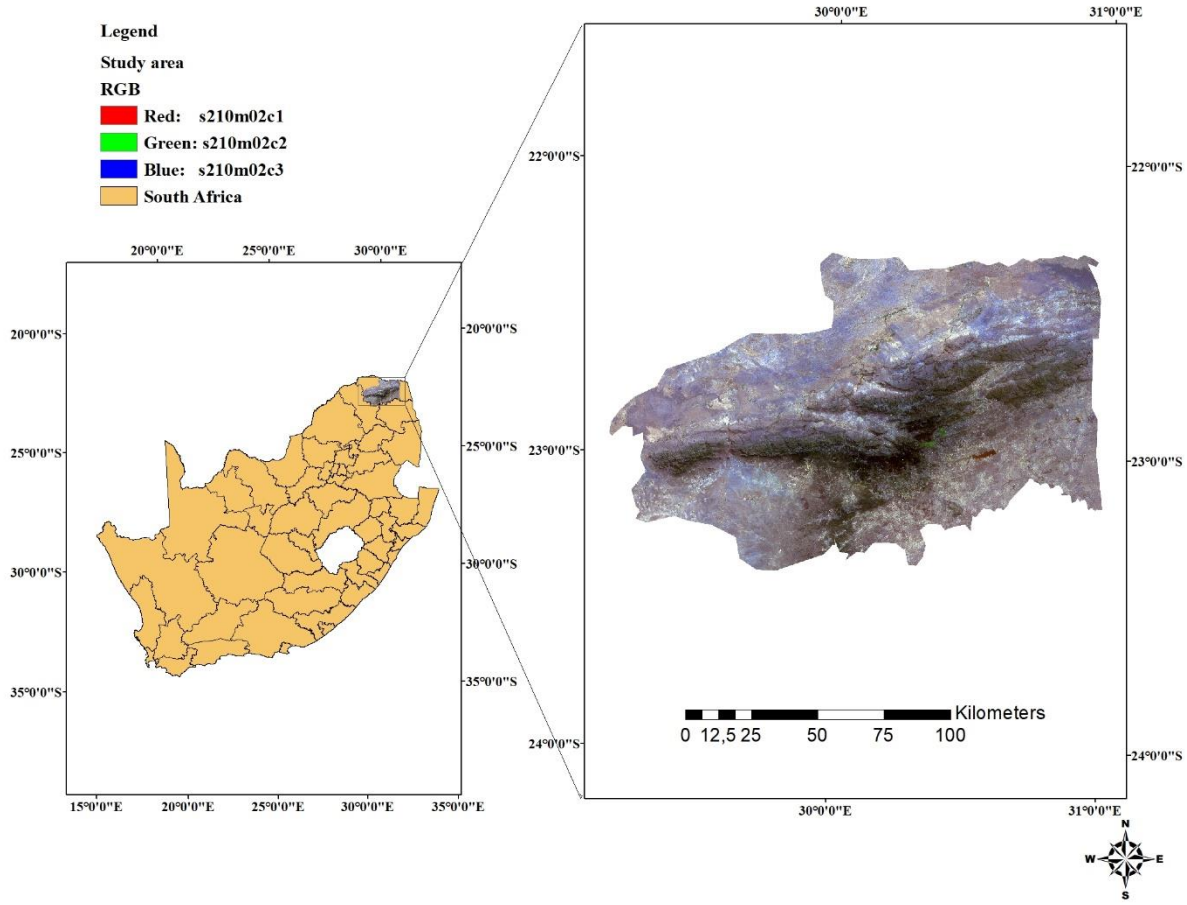


Figure 2.1 Location of the study area in the Limpopo province of South Africa.

2.2.2 Chlorophyll field data measurements

Chlorophyll-*a* measurements were collected in the Vhembe District Municipality, South Africa. The Vhembe District Municipality was considered for this study due to its varying heterogeneous landscape, which is made up of savannah, semi-arid zones, woodland vegetation, and croplands unlike studies carried out in homogeneous landscape. Field data collection was done during August 2017, wherein 112 sampling points (n = 112) were measured and recorded. Sampling points were selected using a random sampling method at the study area. The coordinates for each sampled points within the Vhembe District Municipality were captured using a hand-held standard Garmin eTrex 20™ Global Positioning System (GPS). For plant chl-*a* measurements, a CCM-300 chlorophyll content meter was used. Only the sun-lit leaves were clipped, then the chl-*a* content was recorded across five (5) leaves of individual

tree species. The five chl-*a* were then averaged to derive a single value for that location. Sampling was done within an area of 30 m × 30 m, equivalent to a pixel size of Landsat-8 multispectral data.

2.2.3 Satellite data collection and pre-processing

In this study, Landsat-8 and Sentinel-2 multispectral images were used to assess the correlation between chlorophyll-*a* and multispectral data available at various spatial resolutions. The use of Landsat and Sentinel imagery for this study was considered because of the availability of free satellite images which are convenient for estimating chlorophyll-*a*. Furthermore, Landsat-8 and Sentinel-2 were chosen for this study because the two sensors have shown to be helpful in mapping vegetation parameters (Wang *et a*l., 2022). For example, Sothe *et al*. (2017) conducted a study to map the stages of successional forests in a subtropical forest in southern Brazil, Sentinel-2 and Landsat-8 data were used.

2.2.4 Landsat-8 data collection

Landsat-8 Operational Land Imager (OLI) that covers the whole study area with less cloud cover were acquired, Landsat-8 image acquisition date was 20/08/2017 from USGS (United States Geological Surveys) https://earthexplorer.usgs.gov. Landsat consists of thermal infrared sensors (TIRS) and operational land imager (OLI). The visible near-infrared (VNIR) and shortwave infrared (SWIR) regions are made up of nine spectral bands from the OLI sensor. For the objective of this chapter only seven (7) multispectral bands were used for the estimation of chlorophyll-*a*. Table 2.1 lists the spectral bands of Landsat-8 along with the corresponding wavelengths.

Table 2.1 Multispectral bands of Landsat-8 sensor and its corresponding wavelengths.

Band	Spatial resolution (m)	Wavelength (µm)
Coastal aerosol	30	0.43 – 0.45
Blue	30	0.45 – 0.51
Green	30	0.53 – 0.59
Red	30	0.64 – 0.67
Near infrared	30	0.85 – 0.88
Shortwave infrared 1	30	1.57 – 1.65
Shortwave infrared 2	30	2.11 – 2.29

2.2.5 Sentinel-2 data collection

Multispectral imagery (MSI) of Sentinel-2 were downloaded from USGS (United States Geological Surveys) https://earthexplorer.usgs.gov where the satellite passed over the area on 30/08/2017. The Sentinel-2 data provided by the European Space Agency (ESA) comprising a total of 13 spectral bands ranging from the visible-near infrared at 10 m, red edge bands and shortwave infrared at 20 m and coastal, cirrus and water vapour at 60 m. However, for this study only eight (8) multispectral bands in the visible-near infrared, red edge region and shortwave infrared region were tested to estimate chl-*a* in the study. Table 2.2 shows the multispectral bands explored with its corresponding wavelengths.

Table 2.2 Sentinel-2 multispectral bands explored and its corresponding wavelengths.

Band	**Spatial resolution (m)**	**wavelength (µm)**
Blue	10	0.490
Green	10	0.560
Red	10	0.665
Vegetation Red Edge 1	20	0.705
Vegetation Red Edge 2	20	0.740
Vegetation Red Edge 3	20	0.783
Near infrared	10	0.842
Vegetation Red Edge	20	0.865

2.3 Image pre-processing

2.3.1 Atmospheric correction

Atmospheric correction was performed on both Landsat-8 and Sentinel-2 imagery to accurately compensate for atmospheric effects i.e., distribution of aerosols, cloud cover and amount of water vapour in the atmosphere. The Quick Atmospheric Correction (QUAC) is a technique that atmospherically corrects the multispectral and hyperspectral imagery in the visible and near-infrared through shortwave infrared bands (Exelis Visual Information Solutions, 2016). The QUAC module was used to reduce atmospheric effects from the downloaded Landsat-8 and Sentinel-2 multispectral images in ENVI version 5.3 software (Exelis Visual Information Solutions, 2016).

2.3.2 Mosaicking

Mosaicking is a method of combining images that can be acquired at various times, different viewing angles with overlapping parts to obtain a seamless, large-format, high-resolution image (Zhang *et al*., 2018). To obtain a one very large image that extends over the whole study area since the study area extends beyond the boundary of the image scene, a total of three (3) images for Sentinel-2 and two (2) images for Landsat-8 were merged in Quantum Geographical Information System (QGIS) software

(QGIS Development Team, 2016). The VDM shapefile (*.shp*) was used retrieve the final image of the research area.

2.3.3 Texture analysis

For texture analysis, gray level co-occurrence matrix (GLCM) was employed to process remote sensing data and increase the accuracy of the obtained chlorophyll map which is a widely used statistical method (Iqbal *et al*., 2021), developed by Haralick, Shanmugam and Dinstein in the 1970s. For more accurate texture orientation estimation in remote sensing imagery, a gray level co-occurrence matrix method was created. In this chapter, only four textural parameters namely Correlation (Kayitakire *et al*., 2006); Dissimilarity (Rubner *et al*., 2001); Entropy (Yuan *et al*., 1991) and Homogeneity (Tuttle *et al*., 2006) were derived from ENVI version 5.3 software. A toolbox from ENVI 5.3 with a filter of co-occurrence measures and each co-occurrence image texture parameters was controlled by a 3 × 3 moving window size parameter (Lottering *et al*., 2021). Each of the following listed textural parameters is computed using equations (1) to (4) (GLCM Equations, 2011; Iqbal *et al*., 2021) :

i. Correlation (COR): It is a measurement of the linear dependencies between the image's gray tones. and reflects the consistency of image texture and is given by equation (1).

$$\sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(i-\mu_j)}{\sqrt{(\sigma_i{}^2)(\sigma_j{}^2)}} \right] \qquad (1)$$

ii. Dissimilarity (DIS): A linear measure of local variations in an image is called dissimilarity and is given by equation (2).

$$\sum_{i,j=0}^{N-1} P_{i,j}|i-j| \qquad (2)$$

iii. Entropy (ENT) with values ranging from 0 to a log of the kernel size, statistically determines uncertainty and complexity of image texture, and is given by equation (3).

$$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right) \qquad (3)$$

iv. Homogeneity (HOM): Inverse difference moment is another name for this statistical measurement. It assesses how homogeneous an image's texture is, assuming if a high homogeneity value indicates the absence of intra-regional changes and locally homogeneous distribution in the image texture (Mirjalili and Hardeberg, 2022), and is given by equation (4).

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \qquad (4)$$

where *N* denotes the number of gray levels, μ and σ are the average and variance of the pixel elements, and ln is the natural logarithm. *P (i, j)* is the normalized value of the gray scale at positions *i* and *j* of the kernel, with a sum of 1 (Iqbal *et al*., 2021).

2.4 Correlation analysis

The correlation analysis was conducted for both spectral bands in each spatial resolution and the GLCM features. The coefficients of determination (R^2) were calculated for testing and assessing the relationship between chlorophyll-*a* and multispectral data of both Landsat-8 and Sentinel-2, the performance of the correlation was ranked from a value of -1 to +1. The higher the value of the coefficient of determination, the stronger the relationship is between the compared variables. Four (4) different gray level co-occurrence matrix features and seven (7) bands for Landsat and eight (8) bands for Sentinel were used to assess the correlation of each with chlorophyll-*a*. Analysis of variance (ANOVA) tests and Tukey Honestly Significant Difference (HSD) test were used after conducting correlation analysis to test for the significant difference using ANOVA and Tukey honestly significant difference (HSD) in R statistical software. The above-mentioned ANOVA test was performed to test for statistical differences, but ANOVA do not establish where the difference lies, wherein this is where the Tukey's HSD tests comes in to find out which specific sample's means are statistically different (Nanda *et al*., 2021).

2.5 Results

In this chapter, the correlation between chl-*a* and multispectral data of Landsat-8 and Sentinel-2 at various spatial resolution i.e., 10 m, 15 m, 20 m and 30 m is assessed. Table 2.3 shows two highest correlation results of chlorophyll-*a* and spectral bands at various spatial resolutions.

Table 2.3 Correlation co-efficient results of two highest spectral bands at each spatial resolution.

Sensor	Spatial resolution (m)	Spectral bands	Correlation coefficient (R^2)
S2	10	Red edge1	-0.16
S2	10	Red	-0.15
S2	20	Red	-0.14
S2	20	Blue	-0.13
L8	15	Red	-0.11
L8	15	Green	-0.11
L8	30	SWIR1	-0.14
L8	30	SWIR2	-0.15

The analysis of variance test was undertaken between the R^2 of the spectral bands, and table 2.4 shows the ANOVA results.

Table 2.4 The analysis of variance **(**ANOVA) table for spectral bands.

	Degrees of freedom	Sum of Squares	Mean of Squares	F value	Pr(>F)
Satellite	3	0.002	0.0006	0.918	0.470
Metrices	14	0.042	0.003	4.049	0.020 *
Residuals	9	0.007	0.0007		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

Results for Tukey HSD (Honestly Significant Difference Test) for spectral bands are presented in table 2.5. The table shows that there is no significant difference Landsat-8 and Sentinel-2 with various spatial resolutions.

Table 2.5 Tukey HSD results for spectral bands.

Satellite and resolution	Difference	Lower	Upper	P adjusted
S-2 (20 m) vs. S-2 (10 m)	0.007	-0.035	0.049	0.946
L-8 (15 m) vs. S-2 (10 m)	0.017	-0.035	0.058	0.742
L-8 (30 m) vs. S-2 (10 m)	0.022	-0.022	0.066	0.447
L-8 (15 m) vs. S-2 (20 m)	0.009	-0.042	0.062	0.937
L-8 (30 m) vs. S-2 (20 m)	0.014	-0.029	0.058	0.737
L-8 (30 m) vs. L-8 (15 m)	0.005	-0.048	0.058	0.991

Figure 2.2 shows the box plots for spectral bands at various spatial resolution.

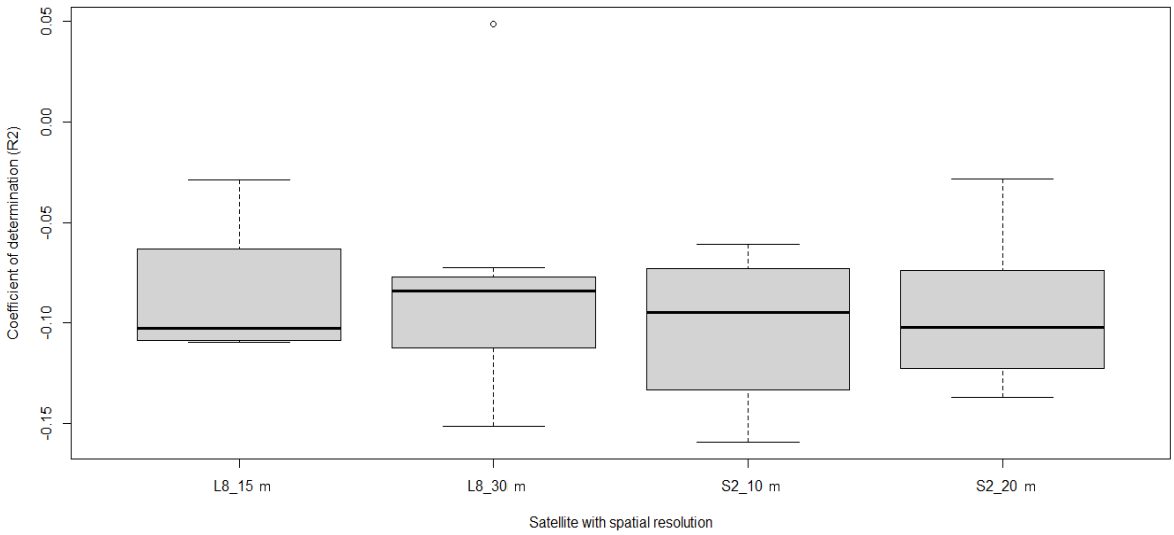


Figure 2.2 Box plots of coefficient of determination (R^2) of Spectral bands at each spatial resolution

Figure 2.3 shows the results of the highest R^2 which is Red edge band of Sentinel 2 at 10 m spatial resolution in a scatterplot.

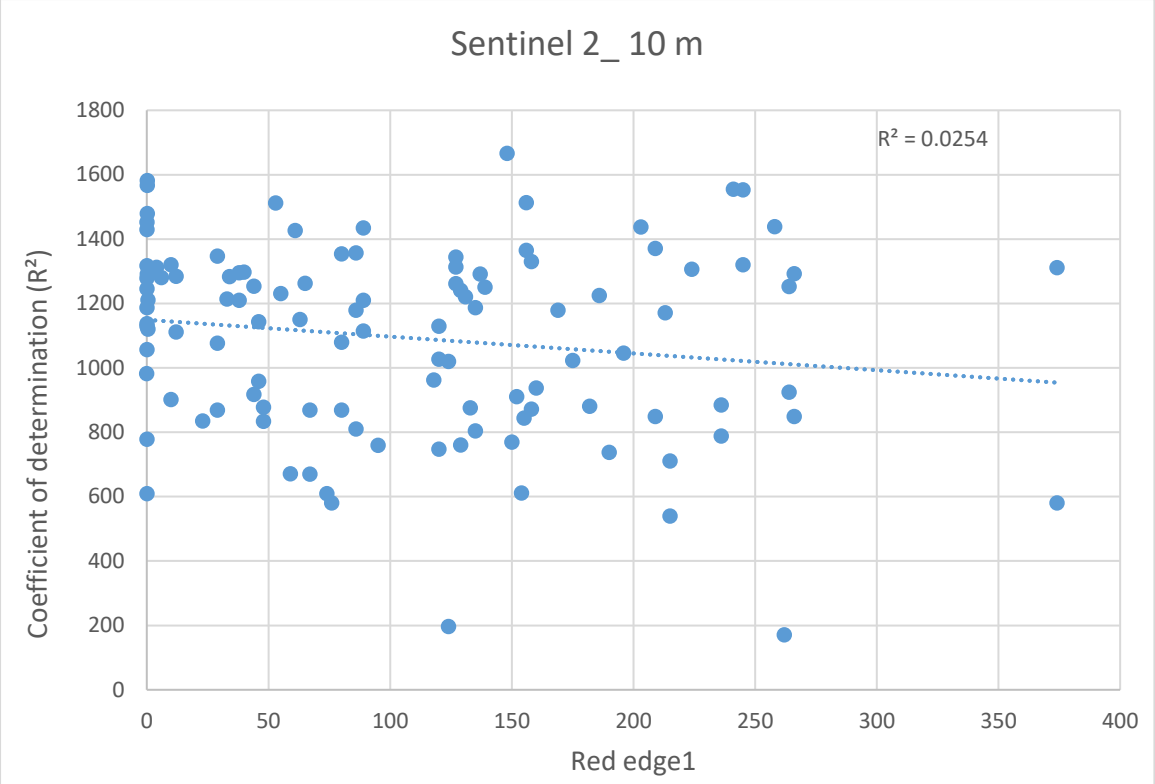


Figure 2.3 Scatterplot showing highest coefficient of determination for spectral bands of Sentinel 2 at 10 m resolution.

Table 2.6 shows two highest correlation results of chlorophyll-*a* and GLCM variables at various spatial resolution.

Table 2.6 The correlation co-efficients of two highest gray level-co-occurrence matrix (GLCM) variables at each spatial resolution.

Sensor	Spatial resolution (m)	GLCM Variables	Correlation coefficient (R^2)
S-2	10	B6 Correlation	-0.26
S-2	10	B6 Dissimilarity	-0.19
S-2	20	B6 Dissimilarity	0.38
S2	20	B6 Homogeneity	-0.37
L-8	15	B3 Correlation	0.14
L-8	15	B2 Correlation	0.11
L-8	30	B5 Entropy	0.39
L-8	30	B6 Entropy	0.35

The analysis of variance test was undertaken between the R^2 of GLCM features and table 2.7 gives the descriptive results of analysis of variance of GLCM features.

Table 2.7 ANOVA results for GLCM features

	Degrees	Sum of	Mean of	F value	Pr(>F)

	of freedom	Squares	Squares		
Satellite	3	0.048	0.016	1.817	0.159
Matrices	63	2.336	0.037	4.205	0.000 ***
Residuals	41	0.362	0.009		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The results of Tukey HSD test are shown in in Table 2.8.

Table 2.8 Tukey HSD (Honestly Significant Difference Test) results for GLCM features.

Satellite and resolution	Difference	Lower	Upper	P adjusted
L-8 (30 m) vs. L-8 (15 m)	0.016	-0.063	0.095	0.946
S-2 (10 m) vs. L-8 (15 m)	-0.039	-0.116	0.038	0.531
S-2 (20 m) vs. L-8 (15 m)	-0.012	-0.089	0.065	0.975
S-2 (10 m) vs. L-8 (30 m)	-0.055	-0.120	0.010	0.128
S-2 (20 m) vs. L-8 (30 m)	-0.028	-0.093	0.037	0.654
S-2 (20 m) vs. S-2 (10 m)	0.027	-0.036	0.098	0.658

Figure 2.4 shows the boxplots presenting the R^2 of GLCM features at each spatial resolution.

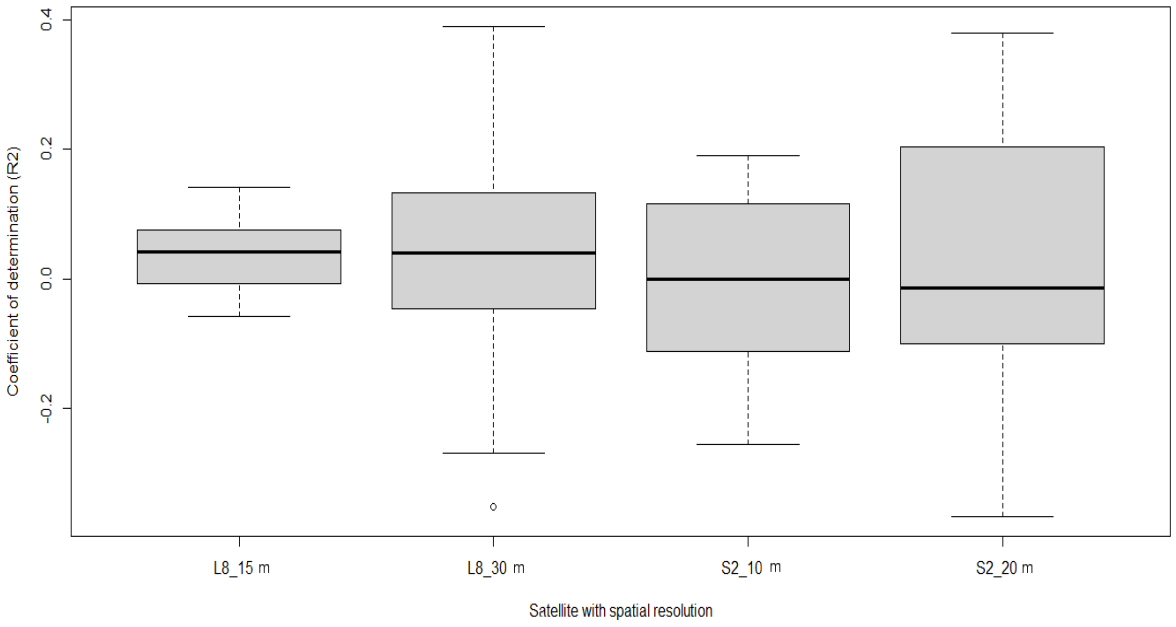


Figure 2.4 Box plots of coefficient of determination (R^2) of GLCM features at each spatial resolution from R statistical software.

Figure 2.5 shows the highest R^2 of Band 5 Entropy at 30 m resolution for Landsat-8

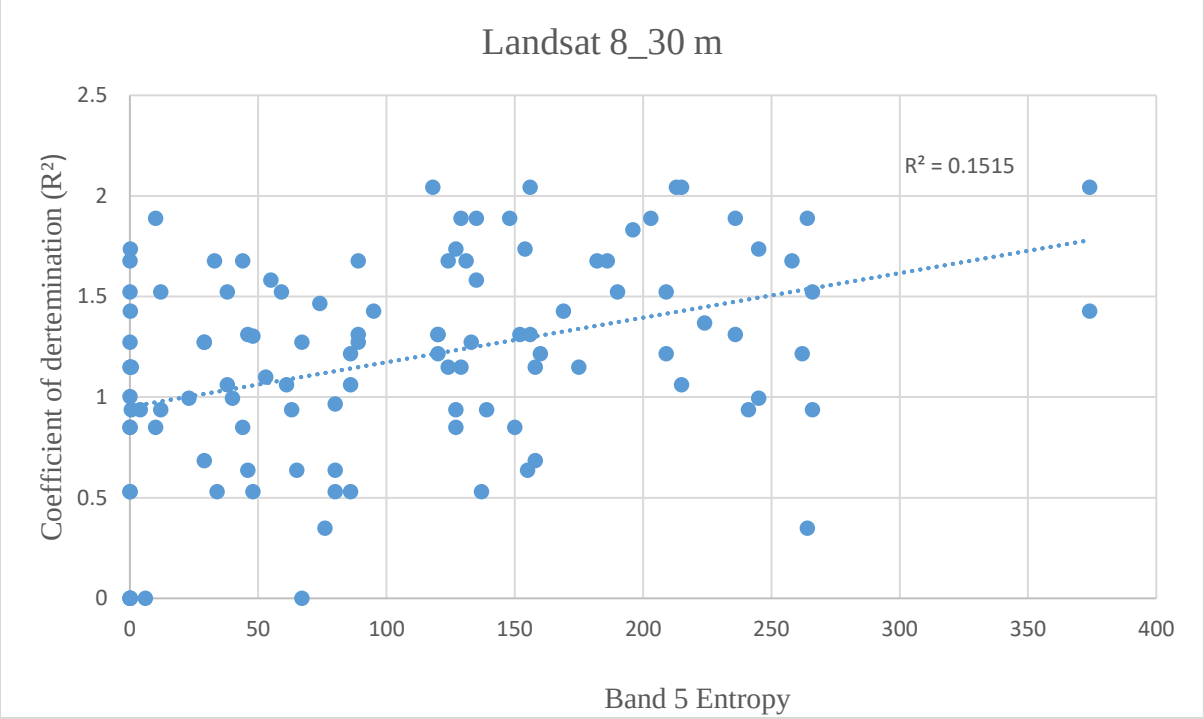


Figure 2.1 Scatterplot showing highest coefficient of determination of Band 5 Entropy for Landsat-8 at 30 m resolution.

2.6 Discussion

This study firstly assessed the correlation between chlorophyll-*a* and spectral bands, and then secondly assessed the correlation between chlorophyll-*a* and GLCM features both at various spatial resolutions of Landsat-8 and Sentinel-2 to estimate chlorophyll-*a* using spectral reflectance. Multispectral remote sensing provided an approach for estimating plant chlorophyll-*a* because of its abundance of data and different wavelengths. Many studies have proven the correlation between spectral variants and chlorophyll pigment (Damayanti *et al*., 2021). However, previous researchers have found that varying plant conditions in terms of health status, level of chlorophyll supply and nitrogen supply of different leaves are intimately connected to the environment that the plant grows in (Chen and Chen, 2008). Through the analysis of chlorophyll-*a* observation data in the study area and two multispectral images i.e., Landsat-8 and Sentinel-2 obtained for the same periods, it was proved that chlorophyll observed have slightly higher positive correlation with GLCM's features with Landsat-8 at 30 m and Sentinel-2 at 20 m and spatial scale. Nevertheless, on the other hand very low negative correlation was observed between observed chlorophyll-*a* and spectral bands at Landsat-8 at 30 m and Sentinel-2 at 20 m spatial scale.

This negative correlation in Sentinel 20 m simply implies that reflectance in the red and blue bands decreases when chlorophyll-*a* concentration increases, the more the vegetation becomes healthier the

more it absorbs the red and blue bands for photosynthesis (DeRiggi, 2017). As for Landsat-8 30 m, the negative correlation at SWIR 1 and SWIR 2 bands which are the water absorption bands it becomes clear that the amount of water within the plant leaves and chlorophyll-*a* amount are always positively correlated to each other. For instance, if a plant is much greener we expect the amount of moisture to be high, while if the plant is not green but brownish the moisture is expected to be very low (Crick, 2016). For example, in another study Keskin *et al*. (2018) successfully explored, using chromameter color values, the impact of leaf moisture content on predicting nutritional stress.

In addition, the correlation coefficients derived by GLCM's texture demonstrated the importance of GLCM's texture for improving the prediction and mapping of chlorophyll-*a*. Thus, the potential of satellite remote sensing in chlorophyll-*a* detection and mapping using readily available multispectral data and GLCM's has been proven (Mohammadpour *et al*., 2022). Numerous articles published to date have shown the estimation of chlorophyll-*a* using Landsat and Sentinel satellite images with field measurements and using GLCM's textures, but something that motivated this study was a wide range of techniques used for this objective. A study by Iqba *et al*. (2020) utilised GLCM's textures successfully in crop classification using remote sensing platforms, however this study differs from our study because it was performed on a homogeneous landscape. Numerous studies demonstrated the significance of GLMC'S texture features, Wang *et al*. (2015) successfully conducted a study about monitoring health conditions in the forest using IKONOS imagery. Zheng *et al*. (2017) suggested using GLCM textures which improves shrub land, agricultural land and barren discrimination.

ANOVA and Tukey HSD findings showed that there was no statistical significant difference amongst various spatial resolution assessed in this chapter. Even though, the results from the analysis of variance and Tukey HSD test results showed that there was no difference that was significant, but with the overall correlation it becomes apparent that Landsat-8 GLCM at 30 m is better than all other spatial resolution including standard reflectance values of spectral bands.

2.7 Conclusion

This study assessed the correlation between chlorophyll-*a* and Landsat-8 and Sentinel-2 multispectral data at various spatial resolutions using spectral bands and GLCM features approach. The results proved that there was an obvious correlation between chlorophyll-*a* and Band 5 entropy with the highest R^2 of 0.39 at 30 m spatial scale of Landsat-8. Among the four GLCM features used in this study, three features are found to have slightly higher correlation with chlorophyll-*a* including entropy, dissimilarity, and homogeneity at 30 m Landsat-8 and 20 m Sentinel-2 spatial resolution. These were the conclusions drawn:

i. Spatial resolution significantly affects the detection and mapping of vegetation parameters, with the recent developments of satellite sensors that comes with varying spatial resolutions challenges rises in the methodology of remotely sensed data analysis.

ii. Moderately high and medium resolution images are normally effective for mapping vegetation at small scale as shown with a slight significant correlation between chlorophyll-*a*, entropy, dissimilarity, and homogeneity at spatial resolution of 30 m, 20 m and 10 m it becomes much possible to extract chlorophyll-*a* values.

iii. Entropy was the most important co-occurrence texture feature that is highly correlated to chlorophyll-*a* followed by dissimilarity then homogeneity.

iv. Very low negative correlation was yielded from red and blue spectral bands of Sentinel-2 20 m spatial scale and SWIR 1, SWIR 2 Landsat-8 30 m spatial resolution.

In conclusion, this chapter has revealed the ability of GLCM texture features at moderately high spatial resolution of 10 m and medium resolution 30 m in extracting chlorophyll-*a* and also provides the baseline for exploring the optimal spatial resolution for plant chlorophyll-*a* estimation, hence, further research is suggested to comprehend the findings. In this chapter the correlations between chlorophyll-*a* and multispectral data at various spatial resolutions has been assessed using a combination of field data and remotely sensed data. These correlations were assessed to achieve the aim of the study as well as to address objective 1 of the study. The following chapter demonstrates the approach for establishing the most suitable spatial resolution between Landsat-8 and Sentinel-2 data for estimating and mapping plant chlorophyll-*a* using satellite data and field dataset from the study area. It builds from the current chapter in that it utilizes the regression modelling for exploring the possibility of mapping and estimating chl-*a* in a heterogeneous landscape by leveraging on the correlations obtained in this chapter.

2.8 Summary of study progress

	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5
Findings	General Introduction	Assessing the correlation of chl-*a* and satellite data	Exploring the optimal spatial resolution for mapping chl-*a*		
	Gaps for GLCM mapping chl-*a*	GLCM outperforms standard reflectance values of spectral bands			

CHAPTER THREE:

Exploring the optimal spatial resolution for mapping plant chlorophyll-*a* on a heterogeneous landscape using Landsat-8 OLI and Sentinel-2 MSI data

Abstract

Plant chlorophyll-*a* (chl-*a*) perform a significant role in decreasing the temperature of the earth's atmosphere, mitigating negative impacts such as depletion of natural resources, extinctions of various species and the imbalance of natural ecosystems. With recent increase in human activities such as deforestation and land degradation, chlorophyll-*a* quantity is reduced across varying landscape. Previous methods such as hyperspectral remote sensing of monitoring the quantity of chl-*a* in plants have been known to demand high computational power and costly when acquiring data. The availability of data at no cost from Landsat-8 and Sentinel-2 has long presented exciting avenues for monitoring plant chl-*a*. For this reason, this study aimed to calibrate models and compare the estimation of terrestrial chl-*a* content between Landsat-8 operational land imager (OLI) and Sentinel-2 multispectral imagery (MSI) at 10-30 m spatial resolution. The gray level co-occurrence matrix (GLCM) with a moving window of 3 × 3 were explored in this study. A stepwise multiple linear regression model was utilized for predicting plant chl-*a* content by way of texture matrix. The results have shown that Landsat-8 at 30 m spatial resolution is suitable for estimating plant chlorophyll-*a* with $R^2 = 0.55$ ($\rho = 0.000006$) while the Landsat-8 at 15 m achieved $R^2 = 0.16$ ($\rho = 0.16$). In Landsat-8 at 30 m spatial resolution GLCM features that were significant in estimating chl-*a* are Entropy and Homogeneity with $\rho < 0.05$.

Co-occurrence texture features are known to be accurate in predicting and monitoring vegetation (Lottering *et al*., 2021). Through stepwise multiple linear regression, the study concluded that Landsat-8 at 30 m and Sentinel-2 at 20 m, 10 m spatial resolution outperformed all other various spatial resolution models. Findings from this study suggest that the optimal spatial resolutions for mapping plant chlorophyll-*a* on a heterogeneous landscape is Landsat-8 at 30 m and Sentinel-2 at 10 m and 20 m.

Keywords: chlorophyll-*a*; Landsat-8; remote sensing; Sentinel-2; spatial resolution

3.1 Introduction

Plants' overall photosynthetic and physiological activities are primarily controlled by biochemical parameters (Cornelissen *et al*., 2003). Chlorophyll pigment, being one of the primary photochemicals, plays a crucial role in photosynthetic activities and is considered to be the most significant biochemical property of leaves (Mridha *et al*., 2020). According to Lichtenthaler and Buschmann, (2001), the pigments of chlorophyll are known as chlorophyll *a* and *b*, and each one has unique spectral absorption characteristics. Plants use both forms of chlorophyll pigments to convert energy. Total leaf chlorophyll-*a* content (mass per unit leaf area) is crucial for adaptation and energy absorption and is the main driver of any photosynthetic activity in the leaf (Mridha *et al*., 2020; Gitelson *et al*., 2006). Because of its function in photosynthesis, leaf chlorophyll-*a* is an essential indicator of nutritional stress and development status and is frequently used to evaluate pressure, diseases, and heavy metal contamination in plants. (Cui and Zhou, 2017).

Chlorophyll-*a* controls carbon exchanges in the forest, which is very important in recognizing how ecosystems are responding to climate change when used as an indicator of leaf photosynthesis (Croft *et al*., 2019). Leaf chlorophyll-*a* content is normally used by ecologists to study ecosystem-biotic community interactions, for instance, the health of a plant in relation to a disease or stressor in the environment (Darvishzadeh *et al*., 2019). Additionally, previous research emphasized the significance of chlorophyll-*a* pigments in determining the quality of a habitat and identifying tree species (Delegido *et al*., 2014), precision farming and estimation of crop net primary productivity (Navarro-Cerrillo *et al*., 2014). Foliar chlorophyll-*a* is one of the crucial biodiversity indicators that remote sensing can monitor (Skidmore *et al*., 2015).

A number of in-situ laboratory and remote sensing methods were used in the development of estimating plant chlorophyll-*a*. In the past, plant chlorophyll-*a* content was determined by wet chemical analysis in laboratories (Buschmann and Lichtenthaler, 2001). This method involves collecting fresh leaf samples from the field and analysing them in the lab. Despite being the most accurate way to gauge how much chlorophyll-*a* is present in plants, this method is destructive and expensive (Cortazar *et al*., 2015), which makes it suitable for limited spatial coverage but not for large-scale assessments. Methods based on remote sensing in contrast to those based on in situ plant chlorophyll-*a* estimation, are effective, non-destructive, and allow for assessments that cover a wider range of time and space (Homolová *et al*., 2013). Similarly, in recent decades, remote sensing data have been extensively utilized for the estimation of chlorophyll-*a* (Chusna and Chu, 2022).

Previous studies have successfully mapped chl-*a* at various spatial resolution. For instance, a research by Salman *et al*., (2013) has successfully utilized MODIS data in estimating chl-*a* at a much larger scale. However, although the success was realized with this instrument, there remained a challenge of finer scale quantification of chl-*a* due to its relevant for local agriculture and local carbon stock balance **(**Brewer *et al*., 2022**)**. This handicap is mainly due to the apparent low spatial resolution (~ 250m) found in MODIS instrument. To circumvent this, some studies utilized Landsat data with a spatial resolution of 30m to estimate chl-*a*. For example, Croft *et al*., (2019) have estimated chl-*a* with accuracy of R^2 = 0.81 and RMSE = 9.45µg/m². Additionally, a study by Zhou *et al*., (2020) also successfully estimated the chl-*a* in China, Shunyi District with Landsat data. The launch of Landsat-8 Landsat-9 satellite sensors since 2013 thus offered a continuation to this accurate quantification of vegetation biophysical and biochemical characteristics (Loveland and Irons, 2016). On the other hand, in 2015 the European Space Agency have launched a Sentinel-2 satellite with similar spectral configuration to Landsat-8 but with additional red-edge bands (Bramich *et al*., 2021). The Sentinel-2 data comes at a spatial resolution of 10-60 m and has shown to have successfully estimated chl-*a* in the retrieval of leaf area index (LAI) and chl-*a* content of a homogeneous agricultural potato crop using Sentinel-2 (Clevers *et al*., 2017). However, it remains to be seen whether the spatial resolution has any effect to the mapping of chl-*a* in a heterogeneous natural landscape.

A number of remote sensing approaches for monitoring vegetation biochemical characteristics include principal component analysis (PCA) and image texture analysis using gray level co-occurrence matrix (GLCM) (Lottering *et al*., 2021). Previous studies retrieved vegetation chlorophyll-*a* using various vegetation indices (Carmona *et al*., 2015). However, vegetation indices specifically the normal difference vegetation index are known to saturate at high biomass levels (Mutanga and Skidmore, 2004). Consequently, a number of remote sensing studies has present the ability of image texture analysis in retrieving vegetation chlorophyll-*a* based on high spatial resolution imagery effectively compared to vegetation indices (Lottering and Mutanga, 2012). For instance, mapping forest

aboveground biomass using a texture combination created from SPOT-6 pan-sharpened imagery in the reforested Buffelsdraai landfill site.

On the contrary, to the best of knowledge estimation of biochemical properties of plants with texture analysis more especially chlorophyll-*a* in heterogeneous landscape is rarely studied. Therefore, the study's objective was to investigate the suitable resolution for mapping chlorophyll-*a* between Landsat-8 OLI and Sentinel-2 MSI derived data. For this purpose, a statistical approach using regression models was adopted with the use of measured field data and texture features. This study was guided by the following objectives:

i. Comparing the best suitable spatial resolution for mapping plant chlorophyll-*a* between Landsat-8 OLI and Sentinel-2 MSI.
ii. Mapping of plant chlorophyll-*a* calibrated from heterogenous landscape through the best model.

3.2 Materials and methods

3.2.1 Description of the study area

This research was conducted in the Vhembe District Municipality (VDM), which is situated in the northern-most region of South Africa between 23º40’ S and 30º00’ E (Figure 2.1). The topography of the VDM landscape is varied and includes savannah, semi-arid regions, woodland vegetation, and croplands (Rosmarin, 2013), with different floral and faunal biodiversity. According to a community survey, the district has a total land area of 21 407 square km and a population of 1 393 949 people (Stats SA, 2016). Annual average rainfall of the district is approximately 500mm, and about 87.1% of which falls between October and March, primarily because of the modified orographic rain effect of the Drakensberg mountains connecting the Soutpansberg mountains (Vhembe District Municipality, 2017/18). The distribution of plants in the study area ranges from being sparsely populated in the west to being extremely densely populated with deciduous plants in the south and east. Patches of grassland and forest biomes are found in the savannah biome of the VDM commonly known as Bushveld (VDM, 2017/2018). The plant species that are commonly found in the VDM include *Adansonia digitata (*Baobob tree*)* with simple leaves that are much longer, unlike other *Adansonia* species, *Adansonia digitata* leaflets have varying sizes and stalk lengths. (Rutherford *et al*., 2006; Baum, D.A, 1995). Forest plant species include *Xymalos monospora* (Harv.) Baill with dark green leaves and *Kiggelaria africana* (L.) with yellow-green leaves beneath and green above (Mostert *et al*., 2008). Temperature in the Vhembe District Municipality is projected to rise by 1-3°C by the middle of the twenty-first century, and the South African National Climate Change Response found that rainfall in the Vhembe district will decline by 5-10%. (Department of Environmental Affairs and Tourism: DEAT, 2004).

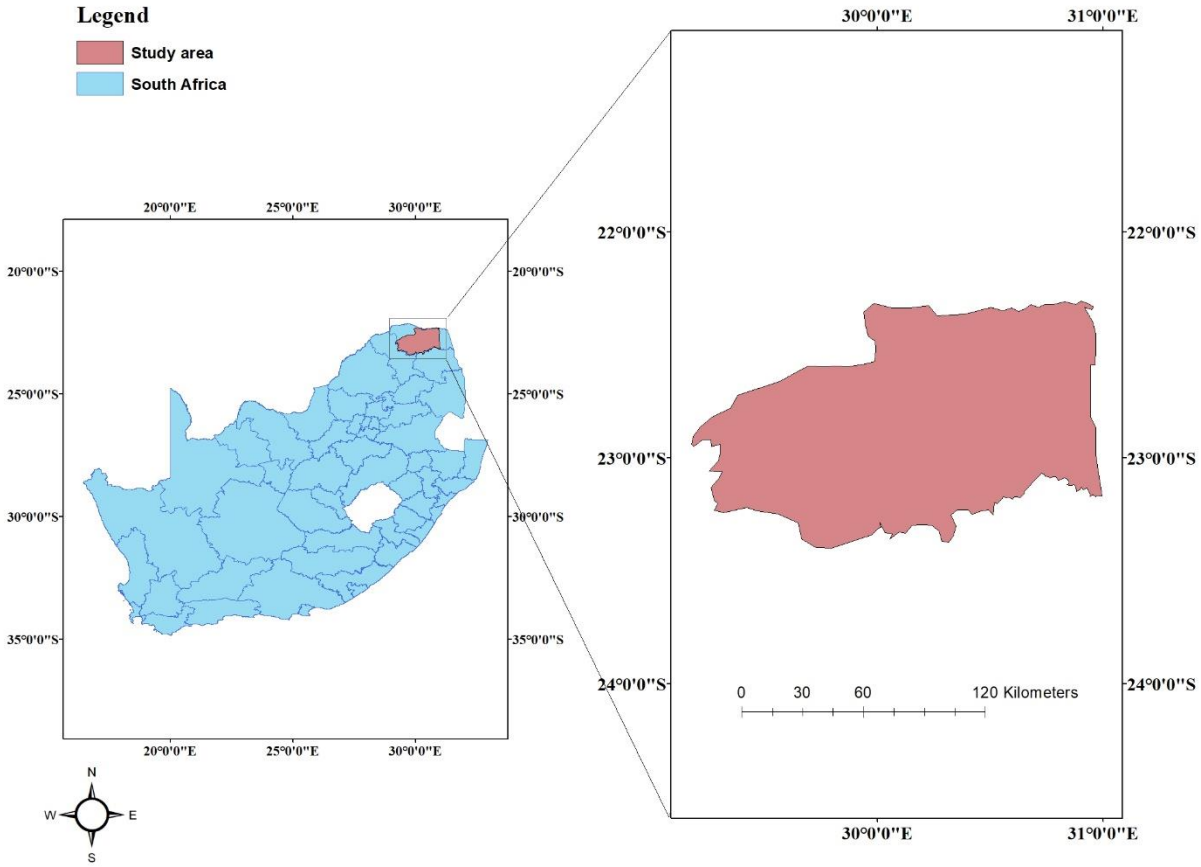


Figure 3.1 Location of the study area in the Limpopo province of South Africa.

3.2.2 Field data collections

Field data collection was carried out from the 112 sampling locations selected as random points (Figure 3.2). Sampling was done within an area of 30 m × 30 m, equivalent to a pixel size of Landsat-8 sensor in the Vhembe District Municipality, South Africa. At every sampling location, measurements were taken from a sun-lit leaves, then the chl-*a* content was recorded across five (5) leaves of individual tree species. The five chl-*a* were then averaged to derive a single value for that location. A hand-held standard Garmin eTrex 20™ Global Positioning System (GPS) at the same time determined sample positions. Given the time and cost constraints the field data were collected during August 2017 covering a total of 112 representative points (n = 112) in the Vhembe District Municipality using CCM-300 chlorophyll content meter. In this study, four (4) spatial resolutions of common multispectral datasets were compared.

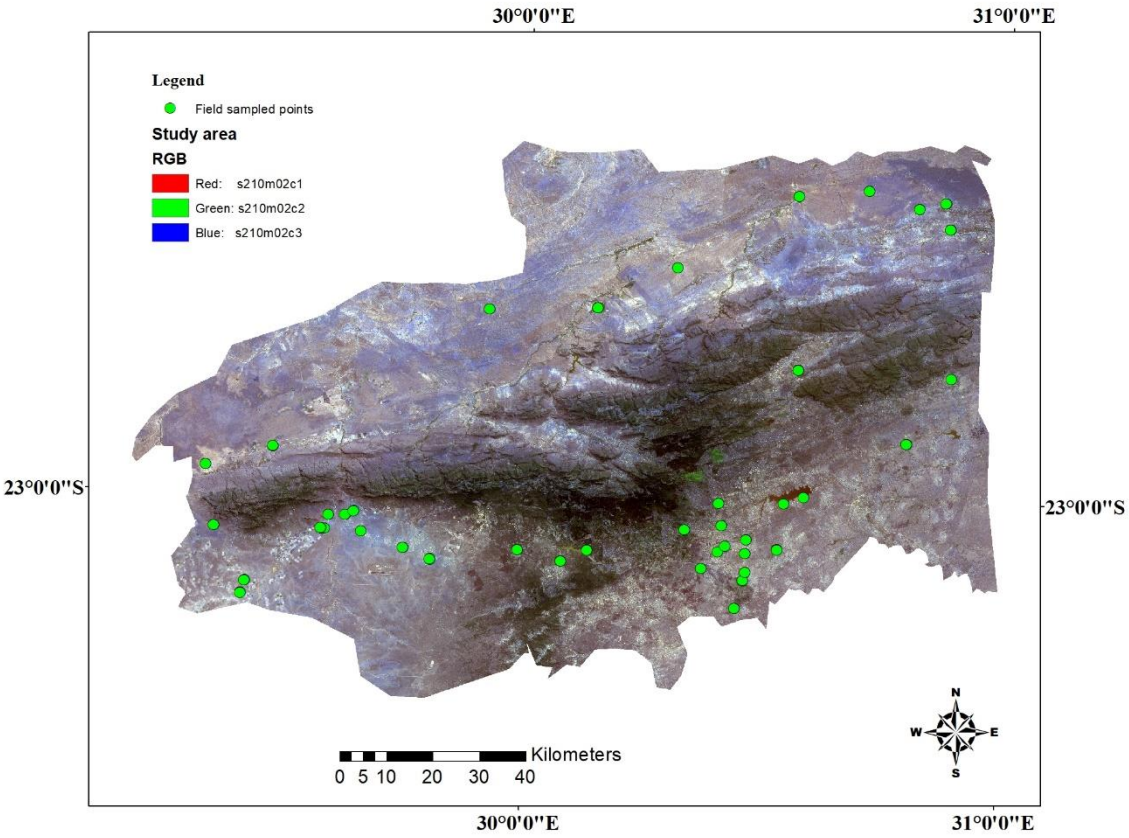


Figure 3.2 Sampling points collected in the Vhembe District Municipality.

3.2.3 Satellite data acquisition

In this study, Landsat-8 and Sentinel-2 multispectral images were used to estimate and map the spatial distribution of chlorophyll-*a* concentration in the Vhembe District Municipality.

3.2.4 Landsat-8 data acquisition

Landsat-8 Operational Land Imager (OLI) that covers the whole study area with less cloud cover were acquired, Landsat-8 images acquisition date was 20/08/2017 from USGS (United States Geological Surveys) https://earthexplorer.usgs.gov. Landsat consists of thermal infrared sensors (TIRS) and operational land imager (OLI). The visible near-infrared (VNIR) and shortwave infrared (SWIR) regions each contain nine spectral bands that make up the OLI sensor. For the objective of this chapter only seven (7) multispectral bands were used for the estimation of chlorophyll-*a*.

3.2.5 Sentinel-2 data acquisition

Multispectral imager (MSI) of Sentinel-2 were downloaded from USGS (Universal States Geological Surveys) https://earthexplorer.usgs.gov where the satellite passed over the study area on 30/08/2017. The Sentinel-2 data provided by the European Space Agency (ESA) comprising a total of 13 spectral bands ranging from the visible-near infrared at 10 m, red edge bands and shortwave infrared at 20 m and coastal, cirrus and water vapour at 60 m. However, for this study only eight (8) multispectral bands in the visible-near infrared, red edge region and shortwave infrared region were tested to estimate chl-*a* in the study.

3.3 Image pre-processing

3.3.1 Atmospheric correction

Atmospheric correction was performed on both Landsat-8 and Sentinel-2 imagery to accurately compensate for atmospheric effects i.e., distribution of aerosols, cloud cover and amount of water vapour in the atmosphere. The Quick Atmospheric Correction (QUAC) is a technique that atmospherically correct multispectral and hyperspectral imagery in the visible and near-infrared through shortwave infrared bands embedded in ENVI ™ 5.3 software (Exelis Visual Information Solutions, 2016). The QUAC module was used to reduce atmospheric effects from the downloaded Landsat-8 and Sentinel-2 multispectral images in ENVI version 5.3 software (Exelis Visual Information Solutions, 2016).

3.3.2 Mosaicking

Mosaicking is a method of combining images that can be acquired at various times, different viewing angles with overlapping parts to obtain a seamless, large-format, high-resolution image (Zhang *et al*., 2018). In order to obtain a one very large image that extends over the whole study, a total of three (3) images for Sentinel-2 and two (2) images for Landsat-8 were merged in Quantum Geographical Information System (QGIS) software (QGIS Development Team, 2016). The VDM shapefile (*.shp*) was used retrieve the final image of the research area.

3.3.3 Texture feature extraction

Image texture analysis is the process of classifying different areas of an image based on their texture content. Both occurrence and co-occurrence texture parameters are used in two different types of image textures (Lottering *et al*., 2021). The use of co-occurrence image texture parameters in identifying vegetation has been demonstrated in earlier studies and was effective (Hlatshwayo *et al*., 2019). For feature extraction, the well-known algorithm GLCM was used (Iqbal *et al*., 2021), established by Haralick, Shanmugan and Dinstein (1973) for processing remote sensing data. Firstly, during feature

extraction a grayscale version of the original image is created. Secondly, the relationship between brightness values and the centre pixel is used to extract spatial features from grayscale images, with each neighbourhood being defined by a window size. The relationship helps with the computation of different sets of texture information based on window size, direction, and gay-scale. Only four textural features, which are listed below, are taken into account in this study:

- Correlation (COR): It is a measurement of the linear dependencies between the image's gray tones and reflects the consistency of image texture (Iqbal *et al*., 2021).
- Dissimilarity (DIS): A linear measure of local variations in an image is called dissimilarity (Iqbal *et al*., 2021).
- Entropy (ENT) with values ranging from 0 to a log of the kernel size, statistically determines uncertainty and complexity of image texture (Iqbal *et al*., 2021).
- Homogeneity (HOM): It assesses how homogeneous an image's texture is, assuming that a high homogeneity value indicates the absence of intra-regional changes and locally homogeneous distribution in the image texture Mirjalili and Hardeberg, 2022). Homogeneity is more sensitive to the presence of near diagonal elements in the GLCM (Iqbal *et al*., 2021).

Each of the listed textural feature is computed using the Equations (5) to (8) (GLCM Equations, 2011; (Iqbal *et al*., 2021) :

$$\mathrm{COR} = \sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(i-\mu_j}{\sqrt{(\sigma_i{}^2)(\sigma_j{}^2)}} \right] \tag{5}$$

$$\mathrm{DIS} = \sum_{i,j=0}^{N-1} P_{i,j} |i-j| \tag{6}$$

$$\mathrm{ENT} = \sum_{i,j=0}^{N-1} P_{i,j} \left(-\ln P_{i,j}\right) \tag{7}$$

$$\mathrm{HOM} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{8}$$

where *N* denotes the number of gray levels, μ and σ are the average and variance of the pixel elements, and ln is the natural logarithm. *P (i, j)* is the normalized value of the gray scale at positions *i* and *j* of the kernel, with a sum of 1. (Iqbal *et al*., 2021).

The above co-occurrence texture parameters were computed and extracted individually for each band of both Landsat-8 and Sentinel-2.

3.4 Regression modelling

In this study, the stepwise multiple linear regression (SMLR) method was employed to predict and map plant chlorophyll-*a* (Neter *et al*., 1983). The prediction and mapping of chlorophyll-*a* in the Vhembe District Municipality was done by comparing the spectral reflectance measurements for chlorophyll-*a* from field measurements and remotely sensed data. The use of stepwise multiple linear regression was justified by the fact that it has been employed in other vegetation studies (Badreldin and Sanchez-Azofeifa, 2015; Ashourloo *et al*., 2022). SMLR can handle many potential predictor variables, and the model can be fine-tuned to select the best predictor variables from the available options. It is also found to be faster, robust and reliable than other automatic model selection methods. This simply demonstrate that the regression model allows each independent variables (spectral variables) of the satellite data to estimate the dependant variables (chlorophyll-*a*) (Vincent *et al*., 2004).

The field chlorophyll-*a* measurements from the 112 sampling points across the Vhembe District Municipality and the reflectance values from both satellite data were randomly split into two different data sets which is 60% training (n = 67) and 40% validation (n = 45). For both satellite data and field data, predictive models of chlorophyll-*a* was provided, which were then used to generate linear regression plots in Excel.

3.4.1 Stepwise multiple linear regression

This study employed stepwise multiple linear regression (SMLR) model with MASS package in R software (version 4.2.1) to derive the best final model with significant variables to estimate chl-*a* content using data collected from the field and spectral reflectance of Landsat-8 and Sentinel-2 bands (R Development Core Team, 2017). The multiple linear regression assessed the relationship between chl-*a* concentration and significant variables with spectral bands first, then GLCM features both from Landsat-8 and Sentinel-2. For modelling 67 sampling points of chl-*a* were used as response variable, while spectral bands and GLCM features from Landsat-8 seven bands (7) and Sentinel-2 eight bands (8) as predictor variables. In this regression technique, initially all the spectral bands and GLCM variables are added to the starting model, Aikaike's Information Criterion (AIC) which selects the model with the lowest AIC as the final model, was then used to make the decision through stepwise method.

The Akaike's information criterion (AIC), a tool for both model selection and estimation of prediction error (Stoica and Selen, 2004). AIC is a form of criterion that considers variables with lot of errors where in AIC is likely high. The study consisted of training datasets where stepwise multiple linear regression models were performed on each of the dataset, based on spatial resolution. The training of models was done on (i) Landsat-8 (15 m), then (ii) Landsat-8 (30 m), (iii) Sentinel-2 (10 m), and lastly (iv) Sentinel-2 (20 m).

The formula of multiple stepwise linear regression is given by equation (9):

$$Y\iota = \beta 0 + \beta 1\, X1\iota + \beta 2\, X2\iota + \ldots + \beta\kappa\, X\kappa\iota + \varepsilon\iota, \iota = 1, 2 \ldots N \qquad (9)$$

Where *Y* is the dependent variable, β0 is the intercept, βκ is the coefficient of Xκ, X1ι , X2ι …., Xκι are the independent or explanatory variables, ι index the number of observations in the sample N, and the term Ɛ is a random error term.

3.4.2 Model validation

Model validation is a measure of how many predictions there have been overall made using spectral reflectance of remotely sensed data when compared to the field measurements of the observed chlorophyll-*a*. This was done by correlating the R-squared and root mean square error (RMSE) values from Landsat-8 and Sentinel-2 with measurements from the field at various spatial resolutions. The performance of the regression models was then evaluated using the RMSE values (Mao *et al*., 2019). The average level of error is measured by the RMSE, which has values ranging from 0 to ∞, a model fits the data well when the RMSE has lower values. RMSE is helpful for evaluating how well a model can predict a variable based on actual data values.

Validation dataset which consisted of 40% (n = 45) of field data was used to validate the final models of each spatial resolution. The root mean square error (RMSE) and the mean absolute deviation (MAD) measured model validity using equations (10) and (11), as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - a_i)^2} \qquad (10)$$

$$\text{MAD} = \frac{1}{n}\sum_{i=n}^{n}(p_i - a_i) \qquad (11)$$

where n is the number of values in the validation dataset. p_i is the predicted value, a_i is the actual measured value.

3.5 Results

3.5.1 Descriptive statistics

Table 3.1 shows the descriptive statistics for measured chlorophyll-*a* in the field.

Table 3.1 Descriptive statistics of measured chlorophyll-*a* concentrations in the field.

Statistics	**Chlorophyll-*a* (µg/m²)**
Minimum	0.1
Mean	108.24
Standard deviation	90.04
Maximum	374
Sample size (*n*)	112

Stepwise linear regression model was used to compare in situ measured and predicted chlorophyll-*a* from Landsat-8 (15 m) spectral bands. Results are presented in Table 3.2.

Table 3.2 Results for stepwise linear model for Landsat-8 spectral bands at 15 m resolution.

Source	**Estimate**	**Std. Error**	**T Value**	**Pr(>|t|)**
(Intercept)	105.2	80.79	1.303	0.196
Band.1	0.1673	0.3023	0.554	0.581
Band.2	-0.1808	0.3339	-0.542	0.589
Band.3	0.000374	0.08646	0.004	0.997
Band.4	0.1429	0.02927	0.488	0.627

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.01, p-value: 0.80

Stepwise linear regression models were also used in validation datasets between field measured and predicted chlorophyll-*a* with Landsat-8 (30 m) spectral bands to see the performance of the model. Results are presented in Table 3.3.

Table 3.3 Results for stepwise linear model for Landsat-8 spectral bands at 30 m resolution.

Source	**Estimate**	**Std. Error**	**T Value**	**Pr(>|t|)**
Intercept)	39.348073	119.594533	0.329	0.7433
Band.1	0.751998	0.790115	0.952	0.3451
Band.2	-1.099510	0.853090	-1.289	0.2025
Band.3	0.263869	0.419603	0.629	0.5319
Band.4	0.266709	0.137681	1.937	0.0575 .
Band.5	0.013596	0.017326	0.785	0.4358
Band.6	0.009166	0.038165	0.240	0.8110

Band.7	-0 .073385	0.053681	-1.367	0.1768

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.14, p-value: 0.25

Stepwise linear regression model that was used to compare field measured and predicted chlorophyll-*a* from Sentinel-2 (10 m) spectral bands. Results are presented in Table 3.4.

Table 3.4 Results for stepwise linear model for Sentinel-2 spectral bands at 10 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	70.77971	76.53257	0.925	0.3589
Band.1	0.14420	0.26599	0.542	0.5898
Band.2	-0.31379	0.35386	-0.887	0.3789
Band.3	-0.05444	0.15221	-0.358	0.7219
Band.4	0.14941	0.18181	0.822	0.4146
Band.5	0.23218	0.34516	0.673	0.5038
Band.6	0.07427	0.30188	0.246	0.8065
Band.7	-0.41784	0.20614	-2.027	0.0473 *
Band.8	0.13906	0.08742	1.591	0.1171

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.10, p-value: 0.53

Stepwise linear regression model that was used to compare field measured and predicted chlorophyll-*a* from Sentinel-2 (20 m) spectral bands. Results are presented in Table 3.5.

Table 3.5 Results for stepwise linear model for Sentinel-2 spectral bands at 20 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
Intercept)	310.805418	142.505901	2.181	0.0333 *
Band.1	-0.516571	0.401664	-1.286	0.2035
Band.2	0.265597	0.276799	0.960	0.3413
Band.3	0.047920	0.170537	0.281	0.7797
Band.4	-0.052706	0.124596	-0.423	0.6739
Band.5	-0.003385	0.023646	-0.143	0.8867
Band.6	-0.082012	0.116977	-0.701	0.4860
Band.7	0.122192	0.122947	0.994	0.3244
Band.8	-0.037542	0.110588	-0.339	0.7355

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.09, p-value: 0.66

Stepwise linear regression model that were used to compare field measured and predicted chlorophyll-*a* from Landsat-8 (15 m) GLCM's features. Results are in Table 3.6.

Table 3.6 Results for stepwise linear model for Landsat-8 GLCM's features at 15 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	-1140.96	821.09	-1.390	0.1699
B4_COR	-42.73	25.99	-1.644	0.1055
B1_DIS	-1024.75	636.04	-1.611	0.1125
B2_DIS	1424.12	907.74	1.569	0.1220
B4_ENT	94.96	40.72	2.332	0.0231 *
B1_HOM	-2116.34	1299.17	-1.629	0.1086
B2_HOM	3016.77	1835.57	1.644	0.1056
B4_HOM	355.50	175.00	2.031	0.0467 *

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.16, *p*-value: 0.16

Stepwise linear regression model that was used to compare field measured and predicted chlorophyll-*a* from Landsat-8 (30 m) GLCM's features. Results are in Table 3.7.

Table 3.7 Results for stepwise linear model for Landsat-8 GLCM's features at 30 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	1214.94	863.49	1.407	0.165156
B5_COR	-37.67	18.64	-2.021	0.048201 *
B1_DIS	-908.77	477.44	-1.903	0.062322 .
B3_DIS	324.76	236.66	1.372	0.175653
B5_DIS	51.26	25.52	2.008	0.049629 *
B6_DIS	-161.34	61.40	-2.628	0.011161 *
B1_ENT	160.04	38.63	4.143	0.000121 ***
B2_ENT	-46.63	30.84	-1.512	0.136286
B4_ENT	-53.72	23.32	-2.304	0.025102 *
B5_ENT	50.00	27.82	1.797	0.077920 .
B1_HOM	-1538.02	1040.49	-1.478	0.145173
B3_HOM	896.31	507.64	1.766	0.083111 .
B6_HOM	-547.85	187.90	-2.916	0.005159 **

Significant codes: '***' 0.001 (*p*<0.001) '**' 0.01 (*p*<0.01) '*' 0.05 (*p*<0.05)

R^2: 0.55, *p*-value: 0.000005956

Stepwise linear regression model that were used to compare field measured and predicted chlorophyll-*a* with Sentinel-2 (10 m) GLCM's features. Results are in Table 3.8.

Table 3.8 Results for stepwise linear model for Sentinel-2 GLCM's features at 10 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	-1357.55	740.51	-1.833	0.07181 .
B4_COR	-101.44	36.80	-2.756	0.00777 **
B7_DIS	410.63	274.27	1.497	0.13968
B4_ENT	-85.51	33.06	-2.587	0.01218 *
B6_ENT	37.59	22.58	1.665	0.10128
B7_ENT	113.44	67.53	1.680	0.09829 .
B2_HOM	225.73	102.47	2.203	0.03151 *
B7_HOM	1334.57	736.62	1.812	0.07512 .

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.24, p-value: 0.01

Stepwise linear regression model that was used to compare field measured and predicted chlorophyll-*a* with Sentinel-2 (20 m) GLCM'S features. Results are presented in Table 3.9.

Table 3.9 Results for stepwise linear model for Sentinel-2 GLCM's features at 20 m resolution.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	-417.41	173.79	-2.402	0.019921 *
B2_COR	-72.39	27.94	-2.590	0.012410 *
B3_COR	215.14	109.69	1.961	0.055213 .
B4_COR	-214.19	108.81	-1.969	0.054343 .
B5_COR	57.36	26.41	2.171	0.034486 *
B7_COR	-54.66	20.27	-2.696	0.009437 **
B3_DIS	494.95	191.13	2.590	0.012432 *
B4_DIS	-526.98	168.99	-3.118	0.002963 **
B2_ENT	61.51	38.27	1.608	0.113987
B3_ENT	-293.51	91.33	-3.214	0.002253 **
B4_ENT	340.33	83.48	4.077	0.000157 ***
B6_ENT	46.79	17.06	2.742	0.008357 **
B1_HOM	264.28	90.29	2.927	0.005068 **
B2_HOM	471.53	189.82	2.484	0.016250 *
B5_HOM	-211.07	93.41	-2.260	0.028052 *

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.52, p-value: 0.00001

3.5.2 Comparison of field measured and estimated chlorophyll-*a* values

Graphs comparing field measures and estimated chlorophyll-*a* are presented in Figures 8-15. The data points demonstrate an excellent fit to the regression line. A poor fit might be caused by the small number of sampling points used for the satellite data prediction for the entire study area (VDM). Additionally, field data collection dates and remote sensing data not corresponding due to satellite overpassing the VDM on field data collection dates were affected by cloud cover rendering it not fit for this study. The models randomly used 67 sampling points as training points to train the model and the remaining 45 sampling points were used as validating points. Below are graphs 3.3-3.10 comparing in situ and predicted chlorophyll-*a* from spectral bands and GLCM's features from remotely sensed using multiple linear regression to statistically analyse data from both sensors.

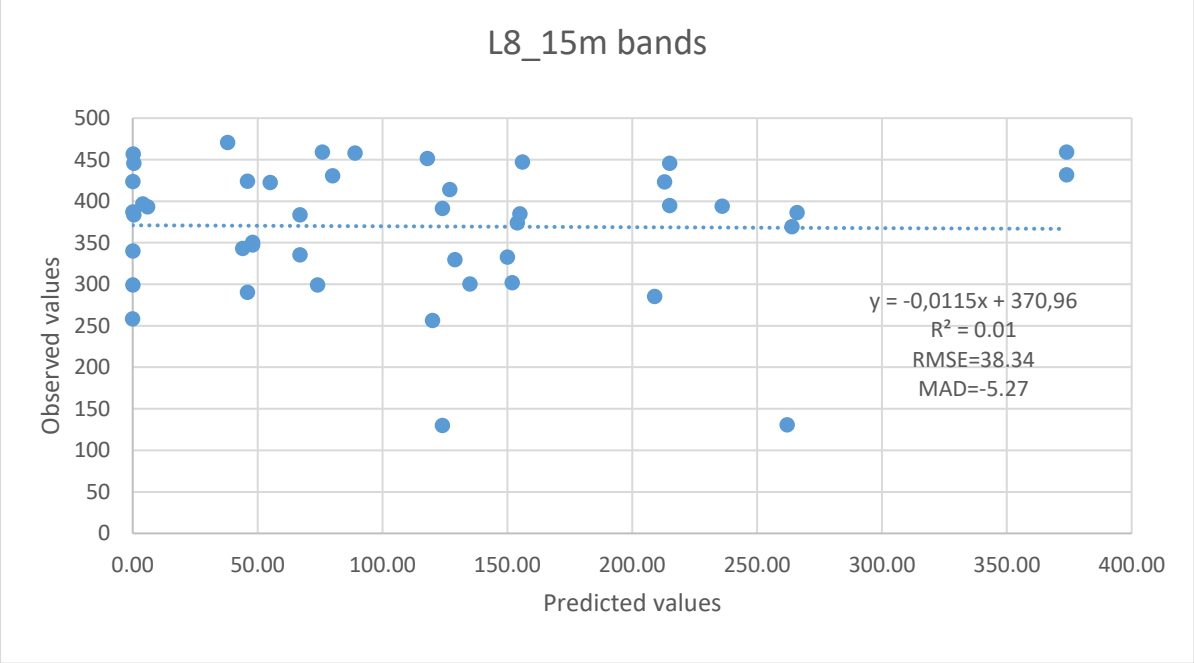


Figure 3.3 Graph comparing chlorophyll-*a* estimated and measured with Landsat-8 spectral bands at 15 m resolution.

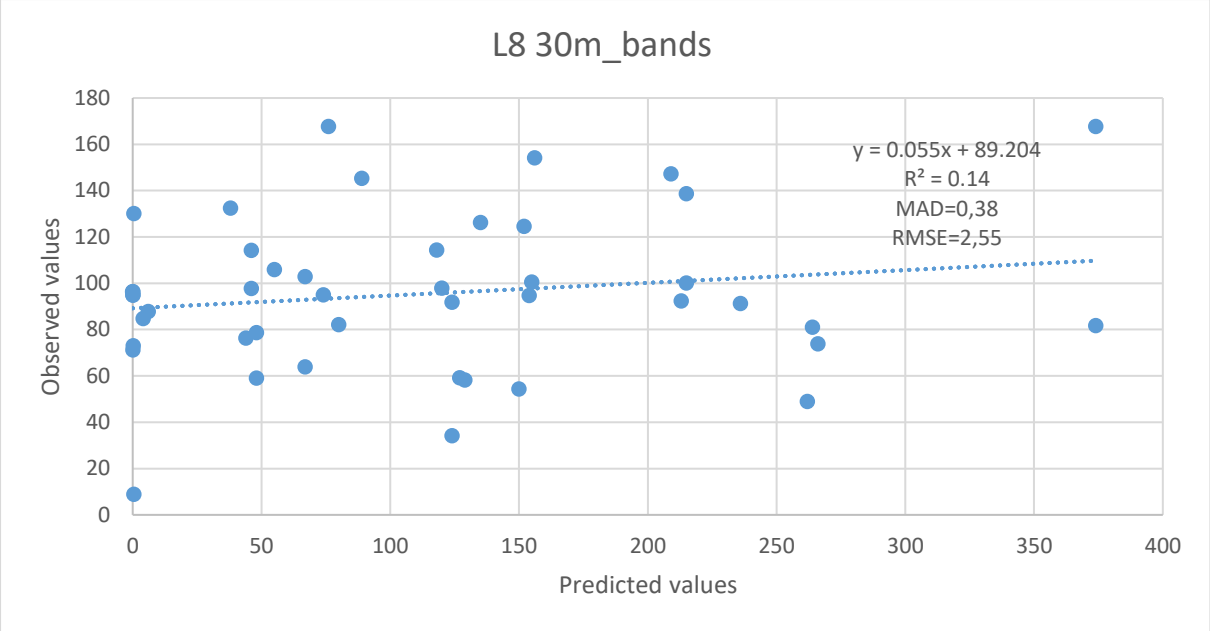


Figure 3.4 Graph comparing chlorophyll-*a* estimated and measured with Landsat-8 spectral bands at 30 m resolution.

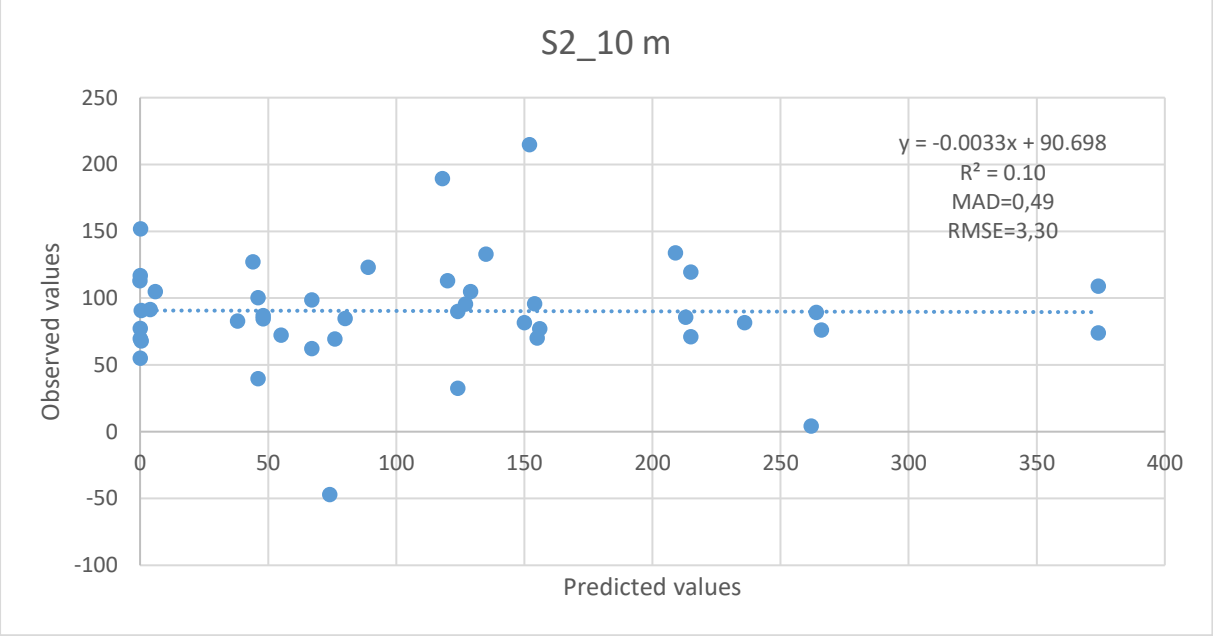


Figure 3.5 Graph comparing chlorophyll-*a* estimated and measured with Sentinel-2 spectral bands at 10 m resolution.

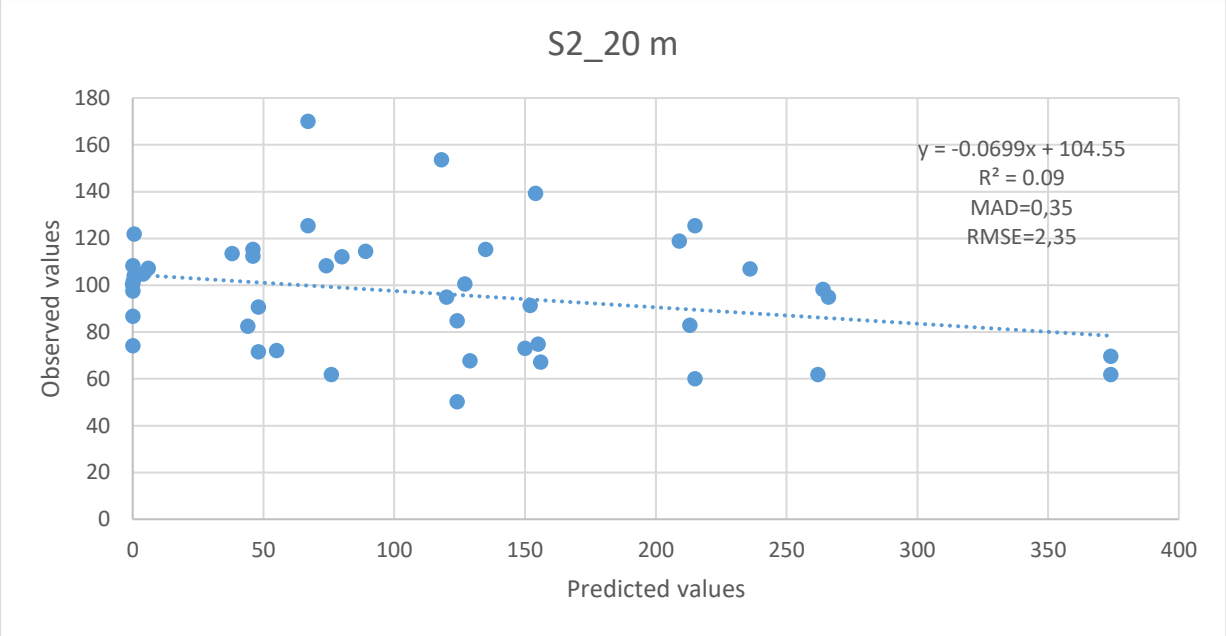


Figure 3.6 Graph comparing chlorophyll-*a* estimated and measured with Sentinel-2 spectral bands at 20 m resolution.

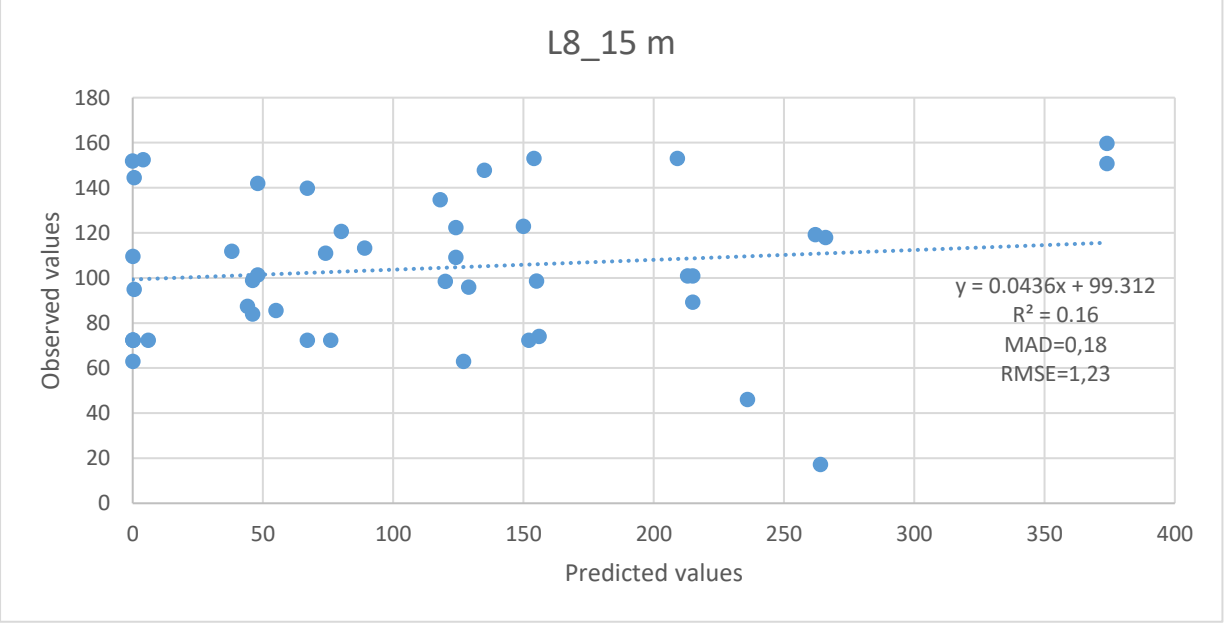


Figure 3.7 Graph comparing chlorophyll-*a* estimated and measures with Landsat-8 GLCM's features at 15 m resolution.

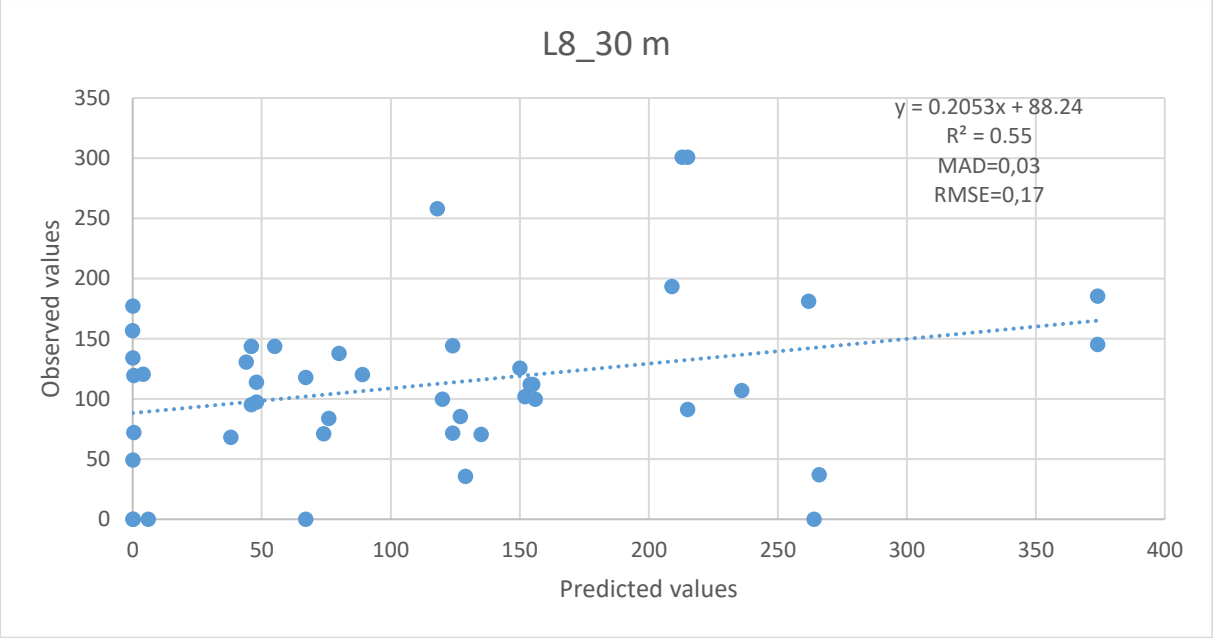


Figure 3.8 Graph comparing chlorophyll-*a* estimated and measured with Landsat-8 GLCM's features at 30 m resolution.

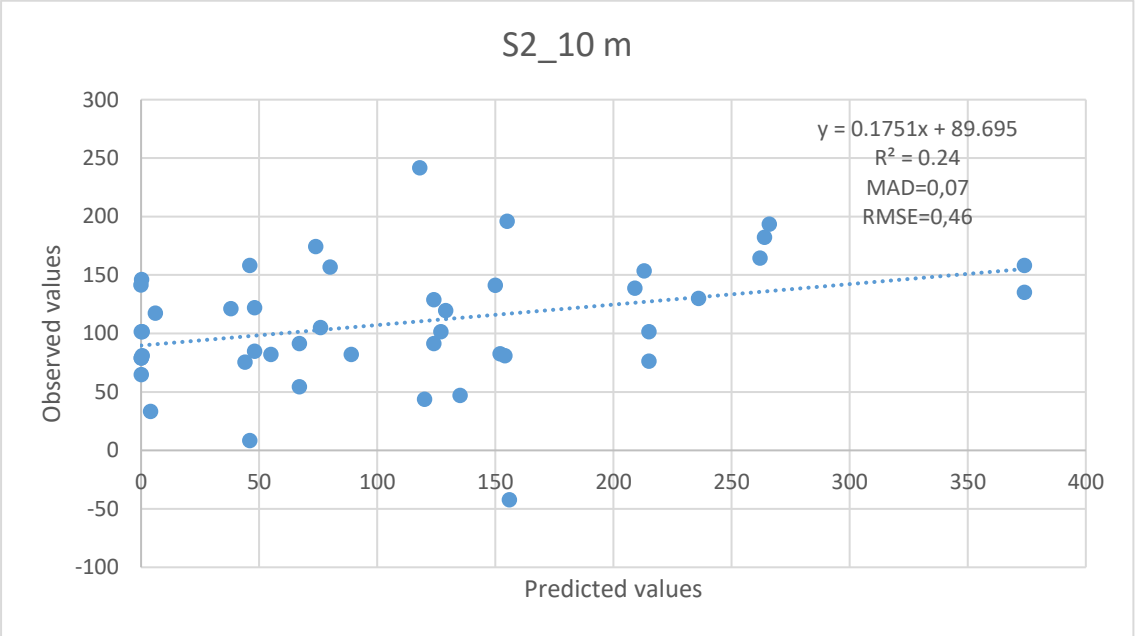


Figure 3.9 Graph comparing chlorophyll-*a* estimated and measured with Sentinel-2 GLCM's features at 10 m resolution.

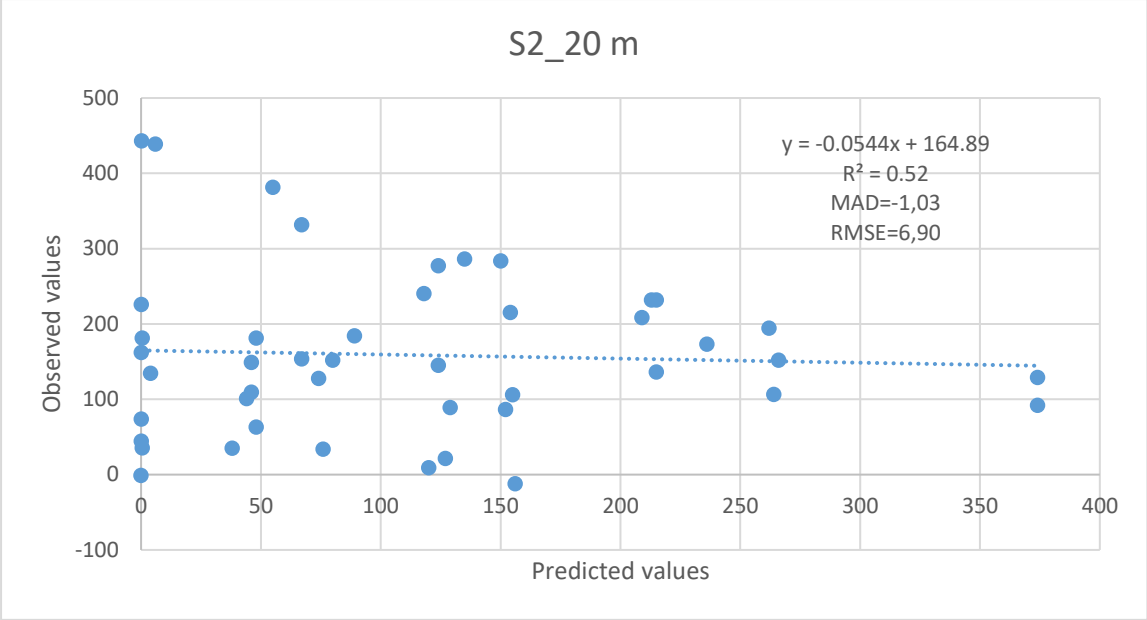


Figure 3.10 Graph comparing chlorophyll-*a* estimated and measured with Sentinel-2 GLCM's features at 20 m resolution.

3.5.3 Estimating chlorophyll-*a* using stepwise multiple linear regression models from remotely sensed data at various spatial resolution

Stepwise multiple linear regression Field measurements and predicted chlorophyll-*a* from models were compared. Landsat-8 and Sentinel-2 spectral bands at various spatial resolutions. The results are presented in Table 3.10

Table 3.10 Performance of stepwise linear regression models for prediction of chlorophyll-*a* concentrations using spectral bands from satellite data

Remote sensing dataset	**Stepwise regression**			
	R square	***p*-value**	**RMSE**	**MAD**
Sentinel-2 at 10 m	0.10	0.53	3.30	0.49
Sentinel-2 at 20 m	0.09	0.66	2.35	0.35
Landsat-8 at 15 m	0.01	0.80	1.23	0.18
Landsat-8 at 30m	0.14	0.25	0.17	0.03

Stepwise multiple linear regression models were used to compare field measured and predicted chlorophyll-*a* from Landsat-8 and Sentinel-2 GLCM features at various spatial resolutions. The results are presented in Table 3.11.

Table 3.11 Performance of stepwise linear regression models for prediction of chlorophyll-*a* concentrations using GLCM features from satellite data.

Remote sensing dataset	**Stepwise regression**			
	R square	***p*-value**	**RMSE**	**MAD**
Sentinel-2 at 10 m	0.24	0.01	0.46	0.07
Sentinel-2 at 20 m	0.52	0.00001	6.90	1.03
Landsat-8 at 15 m	0.16	0.16	1.23	0.18
Landsat-8 at 30m	0.55	0.000005956	0.17	0.03

3.6 Discussion

This study evaluated the feasibility of Landsat-8 and Sentinel-2 in predicting plant chlorophyll-*a* under a heterogeneous landscape in the Vhembe District Municipality by field data collection and stepwise linear regression. To date, a number of studies has demonstrated the prediction and mapping of chlorophyll-*a* using Landsat, Sentinel and other satellite images by field data collection and different regression models. Darvishzadeh *et al*. (2019) conducted a using the invertible forest reflectance model, Sentinel-2 and RapidEye data were used to map the leaf chlorophyll content in spruce stands, Croft *et al*. (2019) also retrieve field leaf chlorophyll-*a* content in agricultural crops for nitrogen management using Landsat-8 imagery. Thus, this study attempted to explore the optimal spatial resolution for mapping the distribution chlorophyll-*a*, integrating field data with the Landsat-8 and Sentinel-2 satellite imagery in a heterogeneous landscape. The findings of this research are significant for exploring the optimal spatial resolution for detecting and mapping chlorophyll-*a* across varying landscape. They enhance and assist the process of vegetation management strategies since there is a need for vegetation management in response to climate change.

The results obtained suggested that the performance of Landsat-8 at 30 m spatial resolution with R^2 = 0.55 (ρ = 0.000006), RMSE = 0.17 µg/m² and Sentinel-2 at 20 m spatial resolution with R^2 = 0.52 (ρ = 0.00001), RMSE = 6.90 µg/m², GLCM features were successful in estimating plant chlorophyll-*a*. On the other hand, Landsat-8 (15 m) with R^2 = 0.16 (ρ = 0.16), RMSE = 1.23 µg/m² with GLCM features showed very low performance and Sentinel-2 (10 m) with R^2 = 0.24 (ρ = 0.01), RMSE = 0.46 µg/m² showed slightly low performance in estimating chlorophyll-*a*. As for spectral bands stepwise linear models it was observed that spectral bands alone cannot effectively estimate chlorophyll-*a* due to models results of Landsat-8 (30 m) with R^2 = 0.14 (ρ = 0.25), RMSE = 0.17 µg/m² and Sentinel-2 (20 m), with R^2 = 0.09 (ρ = 0.66), RMSE =2.35 µg/m² Landsat-8 (15 m) with R^2 = 0.01 (ρ = 0.80), RMSE = 1.23 µg/m² and Sentinel-2 (10 m) with R^2 = 0.10 (ρ = 0.53, RMSE = 3.30 µg/m²).

The above results basically mean that GLCM of Landsat-8 at 30 m spatial resolution is optimal for chlorophyll-*a* mapping in a heterogeneous landscape, and the RMSE obtained in the optimal Landsat-

8 30 m resolution indicates that the study was inaccurate by estimating about 0.17 µg/m² of actual concentration of chlorophyll-*a*. In another study, Zhou *et al*. (2020) the ability of Landsat-8 OLI multispectral data was comprehensively investigated using different retrieval methods for assessing leaf chlorophyll for winter wheat. The feasibility for mapping chlorophyll-*a* was also observed within the Sentinel-2 medium resolutions bands (20 m) model performance. In the Bavarian forest national park, Germany Davishzadeh *et al.* (2019) successfully conducted a study with Sentinel-2 data on the chlorophyll content of leaves in spruce stands. The model for Sentinel-2 high resolution bands (10 m) also yielded significant results with a $p < 0.05$. For example, Phiri *et al*. (2020) in a study of Sentinel-2 data for land cover use mapping concluded that with high spatial resolution Sentinel-2 can achieve high accuracies.

Previous studies of leaf chlorophyll content and stress resistance relationship have indicated that estimation of plant chlorophyll-*a* provide crucial details about the presence of biotic stress factors as well as abiotic problems like light and drought (Agathokleous *et al*., 2020). Future research should focus more on mapping and monitoring chlorophyll-*a* from a heterogeneous landscape with various available moving windows as this study only explored a 3 × 3 moving window.

3.7 Conclusions

This study's goal was to investigate the optimal spatial resolution for estimating chlorophyll-*a* on a heterogeneous landscape in the Vhembe District Municipality using Landsat-8 and Sentinel-2 data. From this research, the following conclusions were made:

i. High predictive precision was attained with GLCM features from Landsat-8 at 30 m with $R^2 = 0.55$ ($\rho = 0.000006$), Sentinel-2 at 20 m with $R^2 = 0.52$ ($\rho = 0.00001$) and Sentinel-2 at 10 m final models compared to Landsat-8 15 m GLCM's final model.

ii. Spectral bands alone from both spatial resolution was not effective in estimating plant chlorophyll-*a* with very low predictive accuracy achieved.

In summary, this chapter provided an operational basis for modelling and mapping plant chlorophyll-*a* from a heterogeneous landscape using fine medium and high spatial resolution (30 m, 20 m and 10 m) of Landsat-8 and Sentinel-2 data. In this chapter the optimal spatial resolution for mapping plant chlorophyll-*a* have been successfully established using both multispectral data and field data. The establishment of the optimal spatial resolution was completed to support the study's aim and address objective 2. The following chapter presented the approach for mapping plant chlorophyll-*a* using multispectral data and the chlorophyll-*a* data.

3.8 Summary of study progress

	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5
	General Introduction	Assessing correlation of chl-*a* and satellite data	Exploring optimal spatial resolution of mapping chl-*a*	Mapping the concentration of chl-*a* across the heterogeneous landscape	
Findings	Gaps for GLCM mapping chl-*a*	GLCM outperforms standard reflectance values of spectral bands	GLCM at 30 m spatial resolution is optimum for chl-*a* mapping		

CHAPTER FOUR:

Mapping the concentration of chl-*a* across the heterogeneous landscape of the savannah biome with Landsat-8 at 30 m spatial resolution

Abstract

Chlorophyll-*a* (chl-*a*) mapping to mitigate the effects of climate change and danger posed by human activities requires precise information for a sustainable vegetation management. Various spatial resolutions and spectral regions allow for the use of remote sensing as a powerful source of crucial data, making it a crucial tool for biomass management and mapping of vegetation. Multispectral sensors have been extensively used to map, analyse, and monitor natural resources for conservation needs due to the growing availability of freely available data and software. The spatial mapping of chl-*a* distribution will significantly play a role in understanding vegetation dynamics. This chapter aimed to map chlorophyll-*a* from Landsat-8 at 30 m spatial resolution data across the heterogeneous landscape of the savannah biome in the Vhembe District Municipality (VDM) in Limpopo, South Africa. The results from this chapter indicate that gray level-co-occurrence matrix (GLCM) with Landsat-8 at 30 m spatial resolution have shown higher accuracy in mapping chlorophyll-*a* (R^2 = 0.55; RMSE = 0.17 µg/m² and MAD = 0.03%). This chapter highlighted the ability of Landsat-8 at 30 m resolution with GLM's features in predicting and mapping chlorophyll-*a* in a heterogeneous vegetation. The medium resolution chlorophyll-*a* map obtained from Landsat-8 might be part of the approach for vegetation monitoring strategies in a heterogeneous landscape.

Keywords: Chlorophyll-*a*; Gray level co-occurrence matrix; Landsat-8; Medium resolution

4.1 Introduction

Globally the economy, the biodiversity and the people's lives have been widely affected by deforestation. Deforestation is critical part of a heterogeneous landscape and ecosystems in the Vhembe District Municipality due to changes in land use, land cover, and climate, expansion of urban areas and reduction in forest management (Xavier *et al*., 2006). There is increasing evidence that deforestation is occurring and directly affect vegetation types and rainfall patterns (Lawrence and Vandecar, 2015). Plants consist of distinctive high evapotranspiration rates in comparison to other ecosystems, where it creates clouds and new precipitation before being recycled back into the atmosphere (Ellison *et al*, 2017). Additionally, plants provide essential ecosystem services such as habitat, gaseous exchange ensuring that there is enough oxygen (O_2) and regulate surface temperature. For plants to provide such ecosystem services this is highly dependent on its biochemical properties which is the green pigment mostly produced by plants i.e., chlorophyll-*a*.

Consequently, monitoring of plant chlorophyll-*a* is very important, necessitating large scale chlorophyll-*a* mapping and management (Kolström *et al*., 2011). Traditionally, vegetation mapping has been carried out with the use of field surveys, ancillary data analysis and image interpretation and chemical analysis to obtain leaf chlorophyll-*a*. However, this approach is known to be accurate but destructive, time-consuming, and costly. As a results it becomes a challenge using this approach to retrieve plant chlorophyll-*a* at a large heterogeneous landscape. Currently, data from remote sensing using active and passive sensors with low cost and in a short period of time and high spatial coverage are the main source of information for earth observation and proving recent vegetation maps and land use classification (Chaves *et al*., 2020).

Remote sensing has been known to be effective when it comes to monitoring large scale and real-time leaf parameters as compared to field surveys (Zhang and Zhou, 2019). With enormous studies on how remote sensing accurately retrieve vegetation parameters at a large scale at within a short space of time is highly related to the continuous advancement in remote sensing technology (Hourborg *et al*., 2015). Amongst this continuous advancement, some of the remote sensing techniques includes utilization of optical remote sensing, radar and thermal imaging. These Techniques for remote sensing can be used to effectively assess the physiological status of plant leaf such as nutritional status and plant productivity which assist in adopting relevant measures to protect the vegetation (Zhang and Zhou, 2019).

Previously, remote sensing has received much attention in terms of effectiveness and reliability in retrieving plant biophysical parameters for example chlorophyll and LAI (Clevers *et al*., 2017). However, the majority of remote sensing studies for mapping chlorophyll-*a* focused on a homogeneous landscape such as agricultural crops (Carmona *et a*l., 2015; Clevers *et al*., 2017). There are very limited studies that retrieved plant chlorophyll-*a* on a heterogeneous landscape. It is important to evaluate the

use of remote sensing for mapping plant chlorophyll-*a* at a heterogeneous landscape to understand the spatial distribution of chl-*a*. Thus, the current chapter used medium resolution satellite remote sensing to map the potential distribution of chlorophyll-*a* across a heterogeneous landscape of the Vhembe District Municipality area. The objective of this chapter was:

i. Mapping the distribution of chlorophyll-*a* across the heterogeneous landscape of the savannah biome.

4.2 Materials and methods

4.2.1 Description of the study area

Refer to Chapter 3 section **3.2.1**.

4.2.2 Field data collection

Chlorophyll measurements were carried out in the South African district municipality of Vhembe. Field data collection was done during the month August 2017, wherein a total of 112 sampling points (n = 112) were measured and recorded. Sampling points were selected using a random sampling method at the study area. The coordinates for each sampled points within the Vhembe District Municipality were captured with a portable device standard Garmin eTrex 20™ Global Positioning System (GPS). For plant chl-*a* measurements, a CCM-300 chlorophyll content meter was used. Only the sun-lit leaves were clipped, then the chl-*a* content was recorded across five (5) randomly sampled leaves of individual tree species. The five measured chl-*a* were then averaged to derive a single value for that location. Sampling was done within a quadrat of 30 m × 30 m, equivalent to a pixel size of Landsat-8 multispectral data.

4.2.3 Satellite data

Refer to Chapter 3 section **3.2.4**.

4.3 Image pre-processing

4.3.1 Atmospheric correction

Refer to Chapter 3 section **3.3.1**.

4.3.2 Mosaicking

Refer to Chapter 3 section **3.3.2**.

4.3.3 Texture feature extraction

Refer to Chapter 3 section **3.3.3**.

4.3.4 Model calibration

For this chapter, n = 112 samples were collected from the field and dataset was randomly split into 60% (n = 67) training data and 40% (n = 45) validating data in R statistical software. Training data set was used to train the model for chlorophyll-*a* concentration estimation in R statistical software. The study used stepwise multiple linear regression to determine variables that are significant i.e., band 1 entropy and band 4, band 5 and band 6 dissimilarity , band 6 homogeneity, and band 5 correlation, and

insignificant for the final model. Variables that were outliers for the respective models were eliminated, and models remained with variables that were significant.

4.3.5 Model validation

Model validation is a measure of how many predictions there have been overall made using spectral reflectance of remotely sensed data when compared to the field measurements of the observed chlorophyll-a. This was done by correlating the R-squared and root mean square error (RMSE) values from Landsat-8 with measurements from the field at various spatial resolutions. The performance of the regression models was then evaluated using the RMSE values (Mao *et al*., 2019). The average level of error is measured by the RMSE, which has values ranging from 0 to ∞, a model fits the data well when the RMSE has lower values. RMSE is helpful for evaluating how well a model can predict a variable based on actual data values.

Validation dataset which consisted of 40% (n = 45) of field data was used to validate the final models of each spatial resolution. The root mean square error (RMSE) and the mean absolute deviation (MAD) measured model validity using equations (13) and (14), as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_i - a_i)^2} \qquad (13)$$

$$\text{MAD} = \frac{1}{n}\sum_{i=n}^{n}(p_i - a_i) \qquad (14)$$

where n is the number of values in the validation dataset, a_i is the actual measured value, p_i is the predicted value.

4.4 Results

4.4.1 Stepwise regression results

Table 4.1 displays the outcomes of a stepwise linear regression, the final linear regression model included a total of twelve (12) variables that were all fairly significant. Entropy, homogeneity and correlation were strongly correlated with chlorophyll-*a* concentration ($p < 0.05$). Furthermore, model at 30 m medium resolution Landsat-8 have produce moderate correlation between predicted and observed values of plant chlorophyll-*a* with $R^2 = 0.55$.

Table 4.1 the final predictive model selected using stepwise multiple linear regression.

Source	Estimate	Std. Error	T Value	Pr(>|t|)
(Intercept)	1214.94	863.49	1.407	0.165156
B5_CO	-37.67	18.64	-2.021	0.048201 *
B1_DIS	-908.77	477.44	-1.903	0.062322 .
B3_DIS	324.76	236.66	1.372	0.175653
B5_DIS	51.26	25.52	2.008	0.049629 *
B6_DIS	-161.34	61.40	-2.628	0.011161 *
B1_ENT	160.04	38.63	4.143	0.000121 ***
B2_ENT	-46.63	30.84	-1.512	0.136286
B4_ENT	-53.72	23.32	-2.304	0.025102 *
B5_ENT	50.00	27.82	1.797	0.077920 .
B1_HOM	-1538.02	1040.49	-1.478	0.145173
B3_HOM	896.31	507.64	1.766	0.083111 .
B6_HOM	-547.85	187.90	-2.916	0.005159 **

Significant codes: '***' 0.001 (p<0.001) '**' 0.01 (p<0.01) '*' 0.05 (p<0.05)

R^2: 0.55, p-value: 0.000005956

Results showed that this model, gray level co-occurrence matrix from Landsat-8 with a spatial resolution of 30 m were the most significant GLCM features.

4.4.2 Performance of gray level co-occurrence matrix in mapping plant chlorophyll-*a*

Figure 4.1 shows the results of plant chlorophyll-*a* concentration models, which was constructed based on Landsat-8 30 m spatial resolution dataset. The model randomly selected 67 sampling points to train the model and used 45 as validation points, the Landsat-8 30 m model produced an accuracy of R^2 = 0.55, RMSE = 0.17 µg/m² and MAD = 0.03%. Based on the results, the study has selected the model for Landsat-8 with a spatial resolution of 30 m to map plant chlorophyll-*a* across the study area.

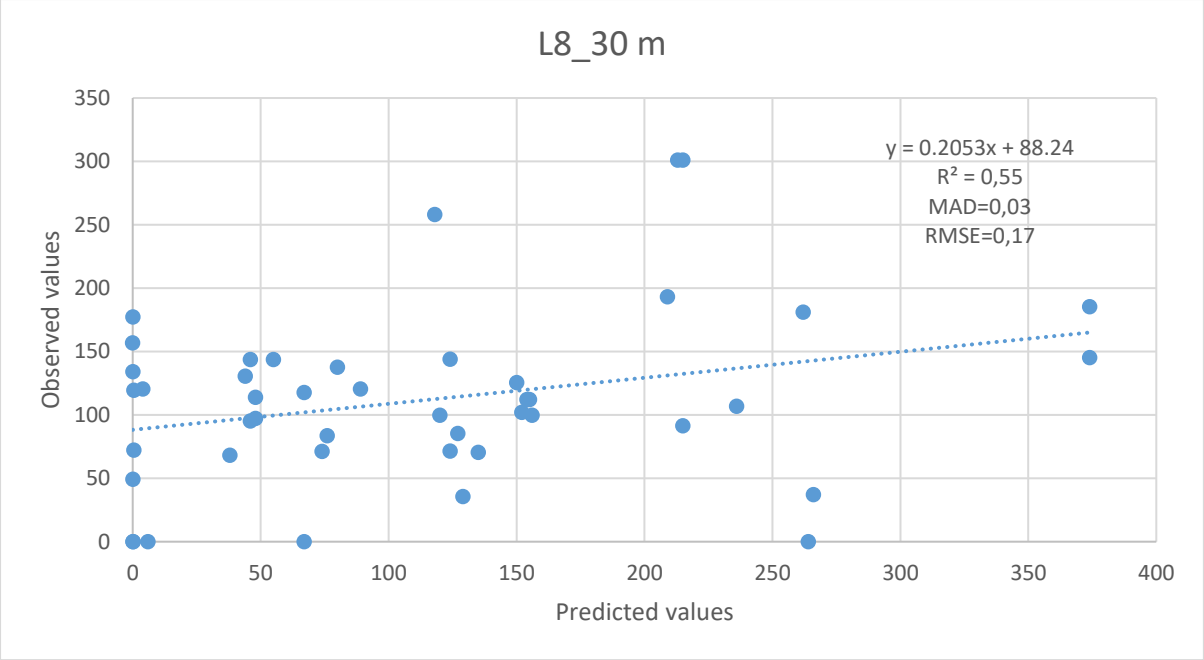


Figure 4.1 Graph comparing chlorophyll-*a* estimated and observed with Landsat-8 GLCM's features at 30 m resolution.

4.4.3 Plant chl-*a* mapping using multiple linear regression and GLCM features

Initially, in the mapping of plant chl-*a* across the study area of the Vhembe District Municipality two multispectral images of Landsat-8 at 30 m spatial resolution were merged to capture a single, large image that encompasses the entire study area. The regression model from stepwise regression was used to compute multiple linear regression equation using raster calculator in QGIS, which was then used on the image that had been mask and resulted in the map showing plant chl-*a* concentration across the Vhembe District Municipality. The spatial distribution of plant chl-*a* is unevenly distributed across the figure 4.2 below, which depicts the study area.

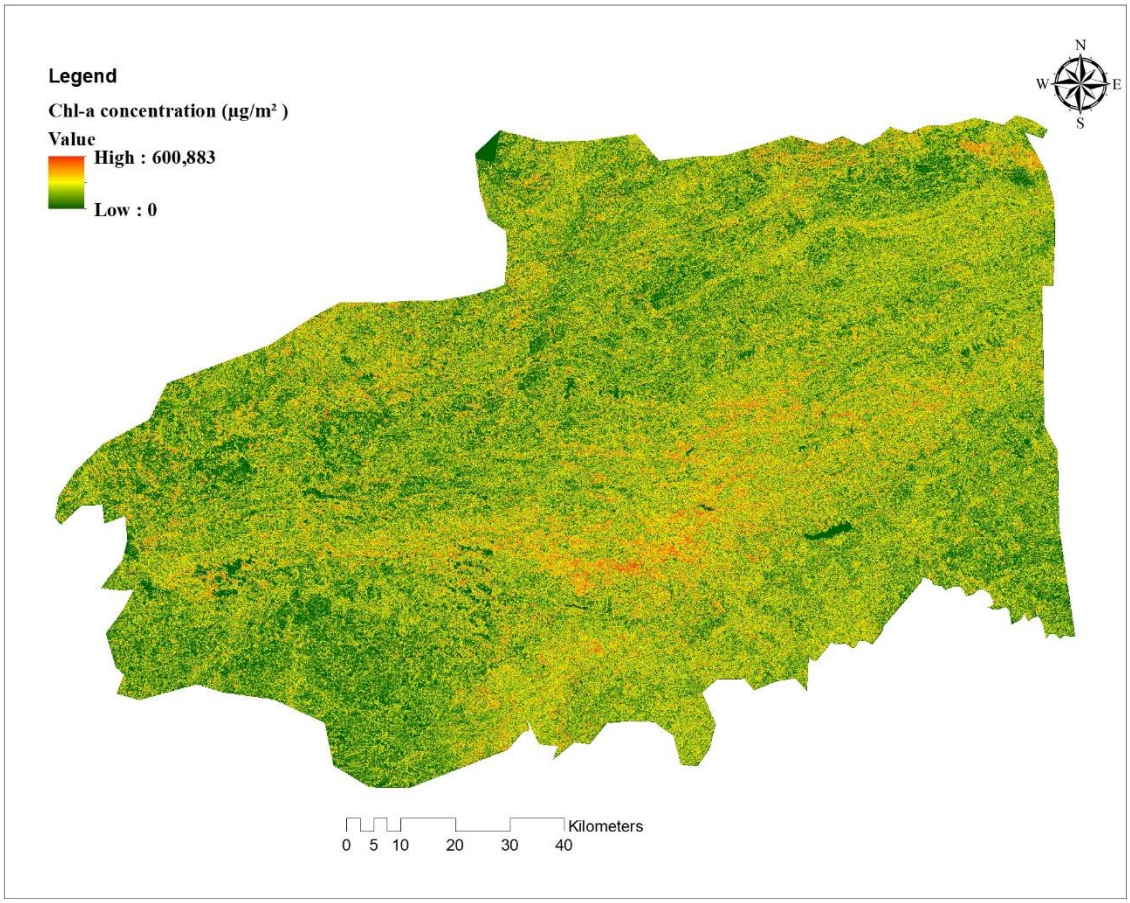


Figure 4.2 Plant chlorophyll-*a* concentration distribution map in the study area based on Landsat-8 at 30 m spatial resolution data and 12 GLCM features.

4.5 Discussion

Chlorophyll-*a* form a crucial part of plants and have been previously mapped using different approaches (Bramich *et al*., 2021; Carmona *et al*., 2015) but rarely in heterogeneous landscapes. Numerous studies concentrated on mapping chlorophyll from a homogeneous landscape for example in agricultural crops, and chlorophyll from water including freshwater lakes (Shah *et al*., 2019; Karimi *et al*., 2022). According to our knowledge, gray level co-occurrence matrix in texture analysis has not been fully used in mapping plant chlorophyll-*a* across a heterogeneous landscape. Findings from this study proves how feasible it would be to use medium spatial resolution Landsat-8 for mapping plant chlorophyll-*a* in a varying landscape.

The results of this chapter showed that band 1 entropy, band 6 homogeneity, and band 1 dissimilarity with Landsat-8 at 30 m spatial resolution possess the capacity for retrieval of plant chlorophyll-*a* from a heterogeneous landscape ($p < 0.05$). From the past vegetation studies, stepwise multiple linear regression have been proven in predicting and mapping chlorophyll using field data and multispectral data (Matus-Hernandez *et al*., 2018). The geographic distribution of chlorophyll-*a* across the Vhembe

District municipality were mapped using the equations derived from the intercept and coefficients of the final Landsat-8 model in Table 4.1, based on the analysis done using multiple linear regression model. Results shows the variability of chlorophyll-*a* distribution ranging from low concentrations in some parts, moderate, and high in other areas. To date, many studies have mapped chl-*a* using multispectral data and field data in either agricultural plants or freshwater and marine environments achieving significant results. As a results mapping of vegetation parameters from heterogeneous landscape is rarely studied.

4.6 Conclusion

Plant chlorophyll-*a*, were successfully mapped using Landsat-8 multispectral imagery at 30 m spatial resolution using multiple linear regression. From this research, the following conclusions were made:

i. The use of Landsat-8 multispectral data at 30 m spatial resolution increases prediction accuracies for mapping plant chlorophyll-*a* when compared to Sentinel-2 medium to high spatial resolution due to variations in the bandwidth of spectral channels.
ii. Gray level co-occurrence matrix features derived from Landsat-8 at 30 m resolution yielded higher accuracy than those of Sentinel-2 medium and high spatial resolution.
iii. The distribution of plant chlorophyll-*a* across the Vhembe District Municipality is unevenly distributed and this is due to varying plant species and height.

4.7 Summary of study progress

	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5
	General Introduction	Assessing correlation of chl-*a* and satellite data	Exploring optimal spatial resolution of mapping chl-*a*	Mapping the concentration of chl-*a* across the heterogeneous landscape	Conclusion and Synthesis
Findings	Gaps for GLCM mapping chl-*a*	GLCM outperforms standard reflectance values of spectral bands	GLCM at 30 m spatial resolution is optimum for chl-*a* estimation	GLCM for mapping of chl-*a* at 30 m spatial resolution	

CHAPTER FIVE:

Conclusion and Synthesis

5.1 Introduction

According to Liang. (2004), some vegetation quality studies have been successful using canopies as a source for remote sensing measurements because of their distinctive spectral signature, which can reveal important details about various vegetation parameters. The prediction and mapping of chlorophyll-*a* concentration in this study were successful, but not as successfully as in other studies on vegetation quality. This may have happened due to the gray level co-occurrence matrix (GLCM) approach from a heterogeneous landscape used in this study compared to other approaches such as image classification and image differencing.

5.2 Evaluation of objectives

The main objective of this study was to compare the spatial and spectral properties of Landsat-8 and Sentinel-2 data for estimating plant chlorophyll-a. To achieve this, the study had the following specific objectives:

i. To assess the correlation between chlorophyll-*a* and multispectral data available at various spatial resolution.

   The first objective of this study was to assess the correlation between chlorophyll-*a* and multispectral data at various spatial resolution at diverse landscape in the Vhembe District Municipality. This study used correlation analysis. The performance of the correlation was ranked from a value of -1 to +1. The higher the determination coefficient's value, the stronger the relationship is between the compared variables. The findings showed that a slightly higher and significant positive correlation between chlorophyll-*a* and GLCM's entropy with $R^2 = 0.39$ with Landsat-8 at a spatial resolution of 30 m. On the other, spectral bands alone yielded extremely low negative correlation with red and green band $R^2 = -0.11$ at 15 m spatial resolution.

ii. Exploring the optimal spatial resolution for mapping chl-*a* on a heterogeneous landscape using Landsat-8 and Sentinel-2 data.

   Among this study's objectives was to examine the optimal spatial resolution for mapping chl-*a* on the Vhembe District Municipality. This was done by selecting gray level co-occurrence matrix features over standard reflectance values of spectral bands. The selected GLCM's features was grouped into four categories (,L8 15m, L8 30m, S2 10 m, and S2 20 m), and

stepwise regression was performed for each spatial resolution using training dataset. Study concluded that GLCM's features at Landsat-8 with a spatial resolution of 30 m performed better with R^2 = 0.55, RMSE = 0.17 µg/m² and MAD = 0.03%, compared to other spatial resolutions models.

iii. Mapping the distribution of chl-*a* across the heterogeneous landscape of the savannah biome using Landsat-8.

The methodology used in chapter 3 was adopted for chapter 4 of this study. Stepwise linear regression was used on both spatial resolution GLCM's features in R statistical software using training dataset. With the use of stepwise regression, GLCM features that were not fit for the models were eliminated. The study compared the final models for each spatial resolution of GLCM features and established that the model for Landssat-8 at 30 m spatial resolution GLCM's features performed better (R^2 = 0.55, RMSE = 0.17 µg/m² and MAD = 0.03%, and was selected for plant chl-*a* mapping. Mapping of plant chl-*a* was done using multiple linear regression in QGIS and a produced map showing how chlorophyll-*a* is distributed across the study area. It is also established that plant chlorophyll-*a* concentration is unevenly distributed across the study area.

5.3 Conclusion

This study's primary objective was to compare the spatial and spectral properties of Landsat-8 and Sentine-2 data for estimating plant chlorophyll-*a* in the Vhembe District Municipality. This study has proven that Landsat-8 at 30 m medium spatial resolution can be used to predict and map plant chlorophyll-*a*. The study has reached the following conclusions:

i. Obvious correlation between chl-*a* and Band 6 entropy with R^2 = 0.39 at Landsat-8 30 m medium spatial resolution. Resulting to gray level co-occurrence matrix (GLCM) features outperforming standard reflectance values of spectral bands.

ii. The GLCM features used to estimate plant chlorophyll-*a* at 30 m spatial resolution have been found to be optimum and produced higher accuracy compared to other models with R^2 = 0.55, RMSE = 0.17 µg/m² and MAD = 0.03%, followed by Sentinel-2 at 20 m with R^2 = 0.52, RMSE = 6.90 µg/m² and MAD = 1.03%, and lastly, the Sentinel-2 at 10 m R^2 = 0.24 RMSE = 0.46 µg/m² and MAD = 0.07%.

iii. The use of Landsat-8 imagery, together with gray level co-occurrence matrix features in relation to biophysical and biochemical plant parameters have proven to be effective for predicting and

mapping plant chlorophyll-*a* at heterogeneous landscape. Landsat-8 at 30 m spatial resolution have outperformed other models with R^2 = 0.55, RMSE = 0.17 µg/m² and MAD = 0.03%.

5.4 Recommendation for further research

The use of Landsat-8 imagery with 30 m spatial resolution has been found to be effective in estimating and mapping chlorophyll-*a* in the Vhembe District Municipality, even in a landscape with varying characteristics. By utilizing Gray Level Co-occurrence Matrix (GLCM) features, higher accuracy rates were achieved in estimating chlorophyll-*a* compared to other spatial resolutions. However, using GLCM features at 10 m, 15 m, and 20 m spatial resolutions resulted in decreased estimation accuracies. One possible explanation for the lower performance accuracy compared to other studies may be due to the difference in dates between remote sensing data collection and field data collection which could have affected the regression model results. To improve the accuracy of these models in future studies, it is recommended that the field data collection coincides with the remote sensing data collection date.

5.5 Summary of the study progress

	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5
	General Introduction	Assessing correlation of chl-*a* and satellite data	Exploring optimal spatial resolution of mapping chl-*a*	Mapping the concentration of chl-*a* across the heterogeneous landscape	Conclusion and Synthesis
Findings	Gaps for GLCM mapping chl-*a*	GLCM outperforms standard reflectance values of spectral bands	GLCM at 30 m spatial resolution is optimum for chl-*a* estimation	GLCM for mapping of chl-*a* at 30 m spatial resolution	Summary of main findings of each chapter

References

Adam, E., Mutanga, O., Odindi, J., Abdel-Rahman, E.M., 2014. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, 2014, 35, 3440 – 3458**.**

Adede, C., Oboko, R., Wagacha, P.W., Atzberger, C., 2019. A mixed model approach to vegetation condition prediction using artificial neural networks (ANN): Case of Kenya's operational drought monitoring. https://doi.org/10.3390/rs11091099.

Agathokleous, E., Feng, Z., Penuelas, J., 2020. Chlorophyll hormesis: Are chlorophylls major components of stress biology in higher plants. *Science of the Total Environment*, Volume 726, 15 July 2020, 138637.

Ahmad, M., Lee, S., Haq, I., Mushtaq, Q., 2012. Hyperspectral Remote Sensing: Dimensional Reduction and End Member Extraction. *International Journal of Soft Computing and Engineering*, (IJSCE) ISSN: 2231 – 2307.

Al-doski, J., Mansorl, S.B., Shafri, H.Z.M, 2013. Image classification in remote sensing. *Journal of Environment and Earth Science*, ISSN 2224 – 3216 (Paper) ISSN 2225 – 0948 Vol. 3, No.10, 2013.

Ashourloo, D., Manafirfard, M., Behifar, M., Kohandel, M., 2022. Wheat yield prediction based on Sentinel-2, regression, and machine learning models in Hamedan, Iran Scientia Iranica D. *Computer Science & Engineering and Electrical Engineering*, (2022) 29(6), 3230 – 3243.

Badreldin, N., and Sanchez-Azofeifa, G.A., 2015. Estimating Forest Biomass Dynamics by Integrating Multi-Temporal Landsat Satellite Images with Ground and Airborne LiDAR Data in the Coal Valley Mine, Alberta, Canada. *Remote Sensing*, 7. 2832 – 2849.

Baum, D.A., 1995. A systematic Revision of Adansonia (Bombacaceae). *Annals of the Missouri Botanical Garden*, Volume.82, No.3 (1995), pp. 440 – 471. https://doi.org/10.2307/2399893.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., and Courchamp, F., 2012. Impacts of climate change on the future of biodiversity. *Ecology Letters*, (2012) 15: 365 – 377.

Bhargava, D., Vyas, S., Bansal, A., 2020. Comparative analysis of classification techniques for brain magnetic resonance imaging images. Advances in computational techniques for biomedical image analysis 2020, Pages 133-144. https://doi.org/10.1016/B978-0-12-820024-7.00007-4.

Botanical Garden, 1995, Vol.82, No.3, pp. 440 – 471.

Bramich, J, Bolch, J.S., Fischer, A., 2021. Improved red-edge chlorophyll detection for Sentinel 2.

Brewer, K., Clulow, A., Sibanda, M., Gokool, S., Naiken,V., Mabhaudhi., 2022. Predicting the Chlorophyll Content of Maize over Phenotyping as a Proxy for Crop Health in Smallholder Farming System. *Remote Sensing*, Volume 14. https://www.mdpi.com/2072-4292/14/3/518.

Buma, G.W. and Lee, S., 2020. Evaluation of Sentinel 2 and Landsat 8 images for estimating chlorophyll-*a* concentrations in Lake Chad, Africa. *Remote Sensing. 12*(15), 2437; https://doi.org/10.3390/rs12152437.

Carmona, F, Rivas, R, Diana,C, Fonnnegra, C, 2015. Vegetation Index to estimate chlorophyll content from multispectral remote data. *European Journal of Remote Sensing*, 48:319 – 326.

Carpenter, G.A., Gopal, S., Macomber, S., Martens, S., Woodcock, C.E., Franklin, J., 1999. A neural network method for efficient vegetation mapping. *Remote Sensing of Environment*, 70 (3), 326-338.

Carter, G. and Knapp, A., 2001. Leaf Optical Properties in Higher Plants: Linking Spectral Characteristics to Stress and Chlorophyll Concentration. *American Journal of Botany*, 88, 677 – 684.

Chaves, P., Ruokolainen, K.,Van donick, J., Kalliola, R., Rivero, E.G., Tuomisto, H., 2020. Mapping fluristics patterns of trees in Peruvian Amazonia using remote sensing and machine learning. *Remote sensing*, Volume 12/Issue 9/10.3390rs12091523.

Chen, J., and Chen, C, 2008. Correlation analysis between indices of tree leaf spectral reflectance and chlorophyll content. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B7.

Cheng, Y., Bhoot, V.N., Kumbier, K, 2021. A novel random forest approach to revealing interactions and controls on chlorophyll concentration and bacterial communities during coastal phytoplankton blooms. *Scientific Reports* 11, 19944 (2021). https://doi.org/10.1038/s41598-021-98110-9.

Cho, M.A., Mathieu, R, Asner, G.P., Naidoo, L., Aardt, J., Ramoelo, A., Debba, P., Wessels, K., Main, R., Smit, I.P.J., & Erasmus, B., 2012. Mapping tree species composition in South African savannas using an integrated airborne spectral and LiDAR system. *Remote Sensing of Environment* 125: 214 – 226.

Chusna, W.B., Chu, H., 2022.Estimating chlorophyll-*a* concentrations in tropical reservoirs from-ratio machine learning models. *Remote Sensing Applications: Society and Environment*, Volume 25, January2022, 100678.

Clevers, J.G.P.W and Koostra, L., 2012. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. Laboratory of Geo-information Science and Remote Sensing, PE&RC.

Clevers, J.G.P.W., Kooistra, L., van den Brande, M.M., 2017. Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sensing*. https://doi.org/10.3390/rs9050405.

Cornelissen, J.H.C., Lavorel, S., Garnier, E., Diaz, S., Buchmann, N., Gurvich, D.E., Reich, P.B., Steege, H., Morgan, H.D., Heijden, M.G.A., Pausas, J.G, Poorter, H, 2003. A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 2003, 51, 335 – 380.

Cortazar, B., Koudemir, H.C., Tseng, D., Feng, S.W, 2015. Quantification of plant chlorophyll content using google glass.

Costanza, R., d'Arge, R., De Groot, R. Farberk, S., Grasso, M., Hannon, B. Limburg. K., 1997. The value of the world's ecosystem services and natural capital *Natur*e, 387), Pages 253 – 260. https://doi.org/10.1038/387253a0.

Crick, J., 2016. How weather affects fall colors. Michigan State University Extension.

Croft, H., Jing, M.C., Luo, X., Bartlett, P., Chen, B., Staebler R.M., 2019. Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, Volume 23(9), Pages 3513 – 3524.

Damayanti, R., Rachma, N., Firmanda, D, and Hendrawan, Y, 2021. The prediction of chlorophyll content in African leaves (Vernonia amygdalina Del.) using Flatbed scanner and optimised Artificial Neural Network.

Darvishzadeh, R., Skidmore, A., Abdullah, H., Cherenet, E., Ali., A., Wang, T., Nieuwenhuis, W.,Vrieling, A., O'Connor, B., 2019. Mapping leaf chlorophyll content from Sentinel-2 and Rapid data in spruce stands using the invertible forest reflectance model. *International Journal of Applied Earth Observation and Geoinformation*, Volume 79, July 2019, Pages 58 – 70.

Darvishzadeh, R., Skidmore, A., Schlerf, M. and Atzberger, C., 2008. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote sensing of environment*, 112(5), Pages 2592 – 2604.

Delegido, J., Van Wittenberghe, S., Verrelst, J., Ortiz, V., Veroustraete, F., Valcke, R., Samson, R., Rivera, J.P., Tenjo, C., Moreno, J., 2014. Chlorophyll content mapping of urban vegetation in the city of Valencia based on the hyperspectral NAOC index. *Ecological Indicators*, 40, 34 – 42. https://doi.org/10.1016/j.ecolind.2014.01.002.

Department of Environmental Affairs and Tourism: DEAT, 2004. A national climate change response strategy for South Africa. Department of environmental affairs and tourism. Private BagX447, Pretoria0001. ⟨https://unfccc.int/files/meetings/seminar/application/pdf/sem_sup3_south_africa.pdf⟩.

DeRiggi, J., 2017. Remote sensing part 3: Identify healthy vegetation from space.

Duveiller, G., Hooker, J. & Cescatti, A., 2018. A dataset mapping the potential biophysical effects of vegetation cover change. *Scientific Data* **5**, 180014. https://doi.org/10.1038/sdata.2018.14.

Essays, UK. (November 2018). Causes and effects of Climate Change Essay. Retrieved from https://www.ukessays.com/essays/environmental-sciences/the-causes-and-effects-of-climate-changes-environmental-sciences-essay.php?vref=1.

Esteban, R., Moran, J. F., Becerril, J. M., & García-Plazaola, J. I. (2015). Versatility of carotenoids: An integrated view on diversity, evolution, functional roles, and environmental interactions. *Environmental and Experimental Botany*, Volume119, Pages 63 – 75.

Exelis Visual Information Solutions. (2016). Environment for Visualizing Images. Boulder, CO: Exelis Visual Information Solutions.

Ellison, D., Cindy E.M., Bruno, L., Douglas, S., Jane, C., Daniel, M., 2017. Trees, forests, and water: Cool insights for a hot world. *Global Environmental Change*, Volume 43, 2017, Pages 51 – 61, ISSN 0959-3780, https://doi.org/10.1016/j.gloenvcha.2017.01.002.

FAO. (2014). Soil and Water Conservation in Semi-Arid Areas. Natural Resources Management and Environment Department. http://www. fao.org/docrep/t0321et0321e-08.htm.

Fensham, R.J. and Fairfax, R.J., 2002). Aerial photography for assessing vegetation change: a review of applications and the relevance of findings for Australian vegetation history. *Australian Journal of Botany*, Volume 50(4), Pages 415 – 429.

Foundation Project. http://qgis.osgeo.org.

G. (2022, May 27). Landsat Program: 50+ Years Archive of Earth. GIS Geography. Retrieved June 6, 2022, from https://gisgeography.com/landsat/.

Garnier, E., Lavorel, S., Ansquer, P., Castro, H., Cruz, P., Dolezal, J., Eriksson, O., Fortunel, C., Freitas, H., Golodets, C., Grigulis, K., Jouany, C., Kazakou, E., Kigel, J., Kleyer, M., Lehsten, V., Leps, J., Meier, T., Pakeman, R., Papadimitriou, M., Papanastasis, V.P., Quested, H., Quetier, F., Robson, M., Roumet, C., Rusch, G., Skarpe, C., Sternberg, M., Theau, J.-P., Thebault, A., Vile, D., and Zarovali, M.P., 2007. Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European Sites. Annals of Botany. Volume 99, 967 – 985. https://doi.org/10.1093/aob/mcl215.

Gitelson, A. A., and Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160, 271 – 282. https://doi. org/10.1078/0176-1617-00887.

Gitelson, A. A., G. P. Keydan, and M. N. Merzlyak, 2006. Three band model for non-invasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters* 33: https://dx.doi.org/10.1029/2006GL026457.

Gitelson, A.A., Chivkuniva, B.O., Merzlyak, M.N., 2009. Non-destructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *American Journal of Botany*, Volume 96(10), Pages 1861 – 1868.

GLCM Equations, 2011. Gray Level Co-occurence Matrix equations. Available at https://scikitimage.org/docs/0.7.0/api/skimage.feature.texture.html#skimage.feature.texture.greycoprops (accessed 10 July 2020).

Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, 81, 416 – 426.

Haralick R.M., Shanmugam K., Dinstein I.H, 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, Volume: SMC-3(6): Pages 610 – 621 DOI 10.1109/TSMC.1973.4309314.

Hlatshwayo, S.T., Mutanga, O., Lottering, R. T., Kiala, Z. and Ismail, R., 2019. Mapping forest aboveground biomass in the reforested Buffelsdraai landfill site using texture combinations computed from SPOT-6 pan-sharpened imagery. *International Journal of Applied Earth Observation and Geoinformation*, Volume. 74, pp. 65 – 77, 2019/02/01/ 2019.

Houborg, R., Fisher, J., Skidmore, A., 2015. Advances in remote sensing of vegetation function and traits. *International Journal of Applied Earth Observation and Geoinformation*, 43. 10.1016/j.jag.2015.06.001.

Homolova, L., Malenovsky, Z., Clevers, J. G. P. W., Garcia-Santos, G. & Schaepman, M. E., 2013. Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity*, Volume 15, Pages 1 – 16.

Iqbal, N., Mumtaz, R., Shafi, U., Zaidi, S.M.H., 2021. Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*. 7:e536 DOI 10.7717/peerj-cs.536.

Jaramaz, D., Perovic, V., Belanovic, S., Saljnikov, E., Cakmak, D., Mrvic, V., Zivotic, L, 2013. The European Space Agency Sentinel-2 mission vegetation variables for Remote Sensing of Plant monitoring.

Karimi, B., Hashemi, S.H., and Aghighi, H., 2022. Performance of Sentinel-2 and Landsat-8 satellites in estimating Chlorophyll-*a* concentration in a shallow freshwater lake. 10.21203/rs.3.rs-1968542/v1.

Karthik, A and Shivakumar, B.R., 2017. Change detection using image differencing: A study over area surrounding Kumta, India. 10.1109/ICECCT.2017.8117851.

Kayitakire, F., Hamel, C., and Defourny, P., 2006. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. *Remote Sensing of Environment*, Volume 102, Pages 309 – 401. 10.1016/j.rse.2006.02.022.

Kolström, M., Lindner, M., Vilén, T., Maroschek, M., Seidl, R., Lexer, M.J., Netherer, S., Kremer, A., Delzon, S., Barbati, A., Marchetti, M., and Corona, P., 2011. Reviewing the Science and Implementation of Climate Change Adaptation Measures in European Forestry. *Forests*, Volume 2(4), Pages 961 – 982. https://doi.org/10.3390/f2040961.

Lawrence, D., & Vandecar, K.L., 2015. Effects of tropical deforestation on climate and agriculture. *Nature Climate Change*, Volume 5, Pages 27 – 36. https://doi.org/10.1038/nclimate2430.

Lemaire, G., Wilkins, R. and Hodgson, J., 2005. Challenges for grassland science: managing research priorities. *Agriculture, ecosystems & environment*, Volume 108(2), Pages 99 – 108. ISSN 0167-8809, https://doi.org/10.1016/j.agee.2005.01.003.

Liang, Y., Urano, D., Liao, KL. *et al*., 2017. A nondestructive method to estimate the chlorophyll content of *Arabidopsis* seedlings. *Plant Methods* 13, 26 (2017). https://doi.org/10.1186/s13007-017-0174-6.

Liang, S., 2004. Quantitative Remote Sensing of Land Surfaces. Wiley Praxis series in Remote Sensing. *Wiley & Sons*, Hoboken etc.

Lichtenthaler, H.K. and Buschmann, C. (2001) Chlorophylls and Carotenoids: Measurement and Characterization by UV-VIS Spectroscopy. *Current Protocols in Food Analytical Chemistry*. https://doi.org/10.1002/0471142913.faf0403s01.

Lichtenthaler, H. (1987) Chlorophyll and carotenoids: Pigments of photosynthetic biomembranes, *Methods of Enzymology*, Academic press, Volume 148, Pages 350 – 382. https://doi.org/10.1016/0076-6879(87)48036-1.

Liu, M., Yu, T., Gu, X., Sun, Z., Yang, J., Zhang, Z, Mi, X., Cao, W., and Li, J, 2020. The impact of spatial resolution on the classification of vegetation types in highly fragmented planting areas bases on unmanned aerial vehicle hypesrspectral images. *Remote sensing*, 12.146. 10.3390/rs12010146.

Lopatin, J., Dolos, K., Hernandez, H.J., Galleguillos, M., Fassnacht, F.E., 2015. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile.

Lottering, R, Mutanga, O, 2012. Forests in KwaZulu-Natal, South Africa, using texture measures and an artificial neural network. *Journal of Spatial Science*, Volume 57, Pages 153 – 173, 2012/12/01.

Lottering, R, Peerbhay, K, Lottering, S, 2021. Discriminating invasive *Solanum Mauritianum* using image texture and sparse PLS discriminant analysis. *International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Mauritius, 2021, Pages 1 – 5, doi: 10.1109/ICECCME52200.2021.9590904.

Lottering, R., Mutanga, O., and Peerbhay, K., 2018. Detecting and mapping levels of Gonipterus scutellatus-induced vegetation defoliation and leaf area index using spatially optimized vegetation indices. *Geocarto International*, 33:3, 277 – 292, DOI: 10.1080/10106049.2016.1250823.

Loveland, T.R., and Irons, J.R., 2016. Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 185(11) DOI:10.1016/j.rse.2016.07.033.

Marcinkowska A.,Zagajewski B.,Ochtyra A.,Jarocińska A.,Raczko E.,Kupková L.,Stych P. and Meuleman K., 2014. Mapping vegetation communities of the Karkonosze National Park using APEX hyperspectral data and Support Vector Machines. *Miscellanea Geographica*, Volume18 (2), Pages 23-29. https://doi.org/10.2478/mgrsd-2014-0007.

Mao, H., Meng, J., Ji, F., Zhang, Q., Fang, H., 2019. Comparison of Machine Learning Regression Algorithms for Cotton Leaf Area Index Retrieval Using Sentinel-2 Spectral Bands. *Applied Sciences*, 9, 1459. https://doi.org/10.3390/app9071459.

Matus-Hernandez, M.A., Hernandez-Saavedra, N.Y., Martinez-Rincon, O., 2018. Predictive performance of regression models to estimate chlorophyll-a concentration based on Landsat imagery. *PLOS ONE*, https://doi.org/10.1371/journal.pone.0205682.

Meddens, A.J., Hicke, J.A., Vierling, L.A., 2011. Evaluating the potential of multispectral imagery to map multiple stages of tree mortality. *Remote Sensing of Environment*. Volume 115(7), Pages 1632 – 1642. https://doi.org/10.1016/j.rse.2011.02.018.

Microsoft Corporation, 2013. Microsoft Excel, Available at: https://office.microsoft.com/excel.

Millennium Ecosystem Assessment (MA) Ecosystems and Human Well-Being: Synthesis Island Press, Washington, DC (2005).

Mirjalili, F., and Hardeberg, J.Y., 2022. On the Quantification of Visual Texture Complexity. J. Imaging 2022, 8, 248. https://doi.org/10.3390/jimaging8090248.

Mohammadpour, P., Viegas, D.X., Viegas, C., 2022. Vegetation Mapping with Random Forest Using Sentinel-2 and GLCM Texture Feature. A Case Study for Lousã Region, Portugal. *Remote Sensing*. 2022, 14, 4585. https:// doi.org/10.3390/rs14184585.

Mostert TH, Bredenkamp GJ, Klopper HL, Verwey C, Mostert RE, Hahn N. 2008. Major vegetation types of the Soutpansberg conservancy and the Blouberg nature reserve, South Africa. Koedoe 50: 32 – 48.

Mridha, N., Chakraborty, D., Biswal, A., Mitran, T, (2021). Retrieval of Crop Biophysical Parameters Using Remote Sensing. In: Mitran, T., Meena, R.S., Chakraborty, A. (eds) Geospatial Technologies for Crops and Soils. Springer, Singapore. https://doi.org/10.1007/978-981-15-6864-0_3.

Mucina, L., and Rutherford, M.C., 2006. The vegetation of South Africa, Lesotho, and Swaziland. Strelitzia 19, South African National Biodiversity Institute, Pretoria.

Mutanga, O, Skidmore, A. K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *International Journal of Remote Sensing*, Volume 25, Pages 3999 – 4014.

Nanda, A., Mohapatra, B.B., Mahapatra, A., Mahapatra, A., 2021. Multiple comparison test by Tukey's honestly significant difference (HSD): Do the confident level control type I error. International Journal of Statistics and Applied Mathematics 2021; Volume 6(1): Pages 59 – 65 https://doi.org/10.22271/maths.2021.v6.i1a.636.

Neter, J., Wasserman, W. and Kutner, M.H., 1983. Applied linear regression models.

Nofrizal, A.Y., Sonobe, R., Yamashita, H., Ikka, T., Morita, A., 2021. Estimation of chlorophyll content in radish leaves using hyperspectral remote sensing data and machine learning algorithms, Proc. SPIE 11856, *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXIII*, 1185609 (12 September 2021); https://doi.org/10.1117/12.2600072.

Peng, Y, Fan M, Wang Q, Lan W, Long Y., 2018. Best hyperspectral indices for assessing leaf chlorophyll content in a degraded temperate vegetation. Ecology and Evolution, Volume 8(14): Pages 7068 – 7078. https://doi.org/10.1002/ece3.4229.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V.V., Murayama, Y., and Ranagalage, M., 2020. Sentinel-2 data for land cover/use mapping: A review. *Remote Sensing*, Volume 12(14):10.3390/rs12142291.

Project, G.C. International Geosphere-Biosphere Programme; IGBP: Stockholm, Sweden, 2010; Volume 21, Page 36.

QGIS Development Team. (2016). QGIS Geographic Information System. Open-Source Geospatial Foundation Project. http://qgis.osgeo.org.

R Development Core Team (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rascher, U., Nichol, C.J., Small, C., Hendricks, L., 2007. Monitoring spatio-temporal dynamics of photosynthesis with a portable hyperspectral imaging system. *Photogrammetric Engineering & Remote Sensing*, Volume. 73, No. 1, January 2007, Pages 045 – 056.

Rosmarin, J., 2013. Vhembe biospehere reserve. Retrieved August 1, 2014, from Environmental Affairs.www.environmental.gov.za/?q=content/projects_programmes/manand_thebiosphere_reserves/list/vhembe.

Rubner, Y., Puzicha, J., Tomasi, C., and Buhmann, M., 2001. Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding*, Volume 84, Pages 25 – 43.

Rutherford M.C., Mucina L., Lötter M.C., Bredenkamp G.J., Smit C., Scott-Shaw R., Jacobus, H.L., Hoare D.B., Goodman P.S., Bezuidenhout H., Scott L., Ellis F., Powrie 107 L.W., Siebert F., Mostert T.H., Henning B.J., Venter C.E., Camp K.G.T, Siebert S.J., Matthews W.S., Burrows J.E., Dobson L., van Rooyen N., Schmidt E., Winter P.J.D., Johann du Preez P., Ward R.A., Williamson S., and Hurte P.J.H., 2006. Savanna biome. *South African National Biodiversity Institute*, Pages 438 – 538.

Salman, M.A., Fendereski, F., Hosseini, S.A., Fazli, H., 2012. A MODIS-based estimation of chlororphyll-*a* using ANN model and in-situ measurement in the southern Caspian Sea. *Indian Journal of Geo-Marine Sciences*, Volume 42 (7), November 2013, Pages 924 – 928.

Shah, S.H., Angel, Y., Houborg, R., Ali, S., McCabe, M.F, 2019. A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sensing*, Volume 11(8), 920; https://doi.org/10.3390/rs11080920.

Simon, D. and Helliwell, S, 1998. Extraction and quantification of chlorophyll a from freshwater green algae, *Water Research*, Volume 32(7), Pages 2220 – 2223.https://doi.org/10.1016/S0043-1354(97)00452-1.

Singhal, G., Bansoda, B., Mathew, L., Goswami, J., Choudhury, B.U., Raju, P.L.N., 2019. Chlorophyll estimation using multi-spectral unmanned aerial system based on machine learning technique. *Remote sensing applications: Society and Environment*, Volume 15 (2019). ISSN 2352 – 9385. https://doi.org/10.1016/j.rsase.2019.100235.

Skidmore, A. K., Pettorelli, N., Coops, N. C., Geller, G. N., Hansen, M., Lucas, R., Wegmann, M., 2015. Environmental science: Agree on biodiversity metrics to track from space. *Nature*, 523, 403 – 405. https://doi.org/10.1038/523403a.

Small, C., 2021. Grand challenges in Remote Sensing Image Analysis and Classification. *Frontiers in Remote Sensing*, 1:605220. https://doi.org/10.3389/frsen.2020.605220.

Solovchenko, A., Chivkunova, O. and Merzlyak, M., 2014. Photosynthetic pigments: chemical structure, biological function, and ecology.

Sonobe, R., Yamashita, H., Mihara, H., Morita, A., Ikka, T., 2021. Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms, *International Journal of Remote Sensing*, Volume 42:4, Pages 1311 – 1329, DOI: 10.1080/01431161.2020.1826065.

Sothe, C., Almeida, C.M., Liesenberg, V., Schimalski, M.B., 2017. Evaluating Sentinel-2 and Landsat-8 Data to Map Sucessional Forest Stages in a Subtropical Forest in Southern Brazil. *Remote Sensing*, Volume 9, 838. https://doi.org/10.3390/rs9080838.

Stoica, P. and Selen, Y. (2004). Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, Volume 21(4), Pages 36 – 47.

Tang, L. and Shao, G., 2015. Drone remote sensing for forestry research and practices. *Journal of Forestry Research*. Volume 26, Pages 791 – 797. DOI: 10.1007/s11676-015-0088-y.

Thiemann, S and Kaufmann, H, 2000. Determination of chlorophyll content and trophic state of lakes using spectrometer and IRS-1C satellite data in the Mecklenburg Lake District, Germany. *Remote sensing of environment*. Volume 73(2), Pages 227 – 235. https://doi.org/10.1016/S0034-4257(00)00097-3.

Tuttle, E.M., Jensen, R.R, Formica, V.A., and Gonser, R.A, 2006. Using remote sensing image texture to study habitat use patterns: a case study using the polymorphic white-throated sparrow (Zonotrichia albicollis), *Global Ecology and Biogeography*, Volume. 15, Pages 349 – 357.

Verrelst, J., Camps-Valls, G., Munoz-Mari, J., Rivera, J.P., Veroustraete, F., Clevers, G.P.W., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 108, Pages 273 – 290, http://dx.doi.org/10.1016/j.isprsjprs.2015.05.005.

Vhembe District Municipality, 2017. Integrated Development Plan 2017/2018. Vhembe District Municipality, Thohoyandou, South Africa.

Vincent, R.K., Qin, X., McKay, R.M.L., Miner, J., Czajkowski, K., Savino, J. and Bridgeman, T. 2004. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment*, Volume 89, Pages 381 – 392.

Wang, H.; Zhao, Y.; Pu, R.; Zhang, Z. Mapping Robinia *pseudoacacia* forest health conditions by using combined spectral, spatial, and textural information extracted from IKONOS imagery and random forest classifier. *Remote Sensing*, Volume 7, Pages 9020 – 9044.

Wang, M., Zheng,Y., Huang, C., Meng, R., Pang, Y., Jia, W., Zhou, L, Huang, Z., Fang, L., Zhao, F., 2022. Assessing Landsat 8 and Sentinel 2 spectral-temporal features for mapping tree species of northern plantation forests in Heilongjiang province, China. *Forest Ecosystems*, Volume 9 (202) 100032.

Xavier, A., Rudorff, B., Shimabukuro, Y., Berka, L., and Moreira, M., 2006. Multi-temporal analysis of MODIS data to classify sugarcane crop. *International Journal of Remote Sensing*, Volume 27, Pages 755 – 768. 10.1080/01431160500296735.

Xie, Y., Sha, Z., and Yu, M, 2008. Remote sensing imagery in vegetation mapping: A review. *Journal of Plant Ecology*, Volume 1(1), Pages 9 – 23.

Xu, M., Liu, R., Chen, J., Liu, Y., Shang, R., Ju, W., Wu, C., Huang, W., 2019. Retrieving leaf chlorophyll content using a matrix-based vegetation index combination approach. *Remote Sensing of Environment*, Volume 224, Pages 60 – 73.

Xu, H., Wang, J., Qu, Y., Hu, L., Tang, Y., Zhou, Z., Xu, X., Zhou, Y., 2022. Estimating leaf Chlorophyll content of Moso Bamboo based on unmanned Aerial Vehicle Visible Images. *Remote Sensing*, Volume 14, 2864. https://doi.org/10.3390/rs14122864.

Xue, J. and Su, B., 2017. Significant remote sensing vegetation indices: A Review of developments and applications. *Journal of Sensors*, Volume 2017, Article ID 1353691, 17 pages, 2017. https://doi.org/10.1155/2017/1353691.

Yuan, X., King, D., and Vlcek, 1991. Sugar maple decline assessment based on spectral and textural analysis of multispectral aerial videography, *Remote Sensing of Environment*, Volume. 37, Pages 47 – 54.

Zarco-Tejada, P.J., Hornero, A., Beck, P.S.A., Kattenborn, T., Kempeneers, P., Hernandez-Clemente, R., 2019. Chlorophyll content estimation in an open-canopy conifer forest with Sentinel-2A and hypespectral imagery in the context of forest decline. *Remote Sensing of Environment*, Volume 15 (223), Pages 320 – 335. doi: 10.1016/j.rse.2019.01.031.

Zhang, W., Li, X., Yu, J., Kumar, M., Mao, Y., 2018. Remote sensing image mosaic technology based on SURF algorithm in agriculture. *EURASIP Journal on Image and Video Processing*, Volume 85 (2018). https://doi.org/10.1186/s13640-018-0323-5.

Zhang, Y., Chen, J.M., Miller, J.R, and Noland, T.L, 2008. Leaf chlorophyll content retrieval from airborne hyperspectral imagery. *Remote Sensing of Environment*, Volume 112, Pages 3234 – 3247.

Zheng, H., Du, P., Chen, J., Xia, J., Li, E., Xu, Z., Li, X., and Yokoya, N., 2017. Performance evaluation of downscaling sentinel-2 imagery for Land Use and Land Cover classification by spectral-spatial features. *Remote Sensing*, Volume 9, 1274.

Zhou, X, Zhang, J, Chen, D, Huang, Y, Kong, W, Yuan, L, Ye, H, Huang, W, 2020. Assessment of leaf chlorophyll content models for winter wheat using Landsat 8 multispectral remote sensing data. *Remote Sensing*, Volume 12, 2574; doi: 10.3390/rs12162574.