University of the Witwatersrand School of Geography, Archaeology & Environmental Studies

CAPABILITY OF MULTI-REMOTE SENSING SATELLITE DATA IN DETECTING AND MONITORING CYANOBACTERIA AND ALGAL BLOOMS IN THE VAAL DAM, SOUTH AFRICA.

by

Altayeb Adam Alsafi Obaid
Student number: 2126105

Supervisors

Prof. Elhadi M I Adam
Prof. Khalid Adem Ali

Dissertation submitted to the Faculty of Science, University of Witwatersrand in fulfilment of the academic requirement for the degree Doctor of Philosophy in Remote Sensing

Johannesburg, South Africa

Abstract

Vaal Dam is a large dam in South Africa. It is the primary source of potable water for the metropolitan and industrial areas of Gauteng province and other surrounding areas. The dam's surface area is about 320 km^2. It's the second biggest dam in South Africa in terms of surface area, and it drains a catchment area of approximately 38,000 km^2. The dam's total capacity is about 2.603×10^6 m^3 (Haarhoff and Tempelhoff, 2007). The dam catchment area holds various anthropogenic activities, including major agricultural activities, mining, and some industrial activities (Obaid et al., 2023, Du Plessis, 2017), as well as many formal and informal settlements. The dam water is strongly affected by such activities, releasing chemical, physical, and biological contaminants and dissolved urban effluents, most of which enrich the nutrients that reach the dam water in some way.

Water resources assessment and monitoring are crucial practices due to their direct contribution to the effective use of such resources. They require precise information about the water quantity and quality. Monitoring of inland water resources has been conducted using in-situ sampling and in-vitro measurement of the water quality constituents. However, these methods have limitations such as high cost, labor-intensive limited spatial and temporal coverage, and time consumption. Over the last few years, remote sensing has been examined for water quality monitoring as a cost-effective system. This research has tested satellite remote sensing to detect some water quality parameters in the Vaal Dam of South Africa. The main objective of this research is to examine the recent generation multispectral satellite sensors, Sentinel-2 MSI, and Landsat-8 OLI data to detect and assess chlorophyll-a and cyanobacteria in the Vaal Dam, South Africa to be used as a cost-effective monitoring tool.

To achieve the objective, the research first aimed to understand how the spatial and temporal dynamics of land use, and land cover (LULC) impact algal growth in the dam reservoir. Land use land cover classification was conducted in the catchment area of the Vaal Dam using a pixel-based classification method. Landsat data for the period from 1986 to 2021 were classified using a random forest (RF) classifier in seven-year intervals (1986, 1993, 2000, 2007, 2014, and 2021). Applying the RF classifier revealed that overall classification accuracies (OA) ranged from 87% in the 2014 classified image to 95% in the 2007 image. The change-detection analysis revealed the continuous increase of the settlement class owing to the continuous population growth. A lot of anthropogenic activities associated with population growth have been recognized to release

contaminants into the surrounding environment and might end up reaching the water resources causing significant deterioration. As a result, Vaal Dam encounters significant nutrient input from multiple sources within its catchment. This situation raised the frequency of the Harmful Algal Blooms (HABs) within the dam reservoir during recent years.

The study also performed a time series analysis for the potential nutrients expected to be the enhancing factors for algal blooms in the Vaal Dam. Using chlorophyll−a (Chl−a) as a proxy of HABs, along with the concentrations of potential nutrients, statistical measures, and water quality data were applied to understand the trend of selected water quality parameters. These parameters were: Chl−a, total phosphorus (TP), nitrate and nitrite nitrogen (NO_3NO_2_N), organic nitrogen (KJEL_N), ammonia nitrogen (NH_4_N), dissolved oxygen (DO) and the water temperature. The results reveal that the HAB productivity in the Vaal Dam is influenced by the levels of TP and KJEL_N, which exhibited a significant correlation with Chl−a concentrations. From the Long-term analysis of Chl−a and its driving factors, some very high values of Chl−a concentrations and its driving factors TP and KJEL_N were recorded in erratic individual dates which suggested some nutrients rich in wastes find their way to the dam. Another important notice was that the average Chl-a concentration significantly increased during the period of the study (1986 to 2023) it increased from 4.75 μg/L in the first decade (1990–2000) to 10.51 μg/L in the second decade (2000–2010) and reaching 16.7 μg/L in the last decade (2010–2020). Additionally, Chl−a data extracted from Landsat-8 satellite images was utilized to visualize the spatial distribution of HABs in the reservoir. The satellite data analysis during the last decade revealed that the spatial dynamics of HABs are influenced by the dam's geometry and the levels of discharge from its two feeding rivers, with higher concentrations observed in meandering areas of the reservoir, and within zones of restricted water circulation. These spatial distribution patterns of HABs are associated with spatial variations of algal species in term of domination through the seasons of the year.

The research also examined the utility of remote sensing techniques for mapping algal blooms using the current generation Sentinel-2 and Landsat-8 data. The effectiveness of some band ratio indices in the blue-green and red-near infrared wavelengths was tested. The results suggested that the blue-green band ratio of Landsat-8 [$R_{rs}(560)/R_{rs}(443)$], and red/NIR of Sentinel-2 [$R_{rs}(705)/R_{rs}(665)$] were found to be the best indices for Chl-a retrieval in the Vaal Dam. Results for the Landsat OLI dataset showed $R^2 = 0.89$; RMSE = 0.36 μg/L, $P < 0.05$, and the Sentinel MSI dataset revealed $R^2 = 0.75$; RMSE = 0.48 μg/L, $P < 0.05$ which is a high degree of accuracy.

As the potential toxicity comes from the cyanobacterial bloom, the study examines different models to assess and map cyanobacteria concentration in the dam reservoir. Sentinel-2 and in-situ hyperspectral data have been used. None of the Sentinel-2 band ratios showed a significant correlation with the laboratory-measured values of the cyanobacteria. The in-situ measured Hyperspectra showed strong correlations between the band ratios $R_{rs}(705)/R_{rs}(655)$ and $R_{rs}(705)/R_{rs}(620)$, and the measured cyanobacteria ($R^2 = 0.96$ and $R^2 = 0.95$ respectively).

Chlorophyll−a concentration was retrieved using band ratio indices in the red-NIR region. The strongest correlation was found between the retrieved Chl−a of band ratio Rrs(705)/Rrs(665) and the laboratory-measured Chl−a concentrations for both reflectance datasets. This correlation resulted in an R^2 value of 0.78 for Sentinel-2 reflectance data and an R^2 value of 0.93 for in-situ hyperspectral data.

A Semi-analytical algorithm for estimating the Chl−a and phycocyanin (PC) pigments has also been examined. The algorithm uses the ratio of the calculated Chl−a absorption at 665 and phycocyanin absorption at 620 nm to their specific absorption coefficients $a^*_{Chl}(655)$ and $a^*_{PC}(620)$ to estimate the concentration of Chl−a and phycocyanin respectively. It resulted in a strong correlation with measured chlorophyll-a, $R^2 = 0.95$. The algorithm also strongly correlated with measured cyanobacteria using the absorption to specific absorption ratio at 620 nm ($R^2 = 0.97$). However, the estimated values of cyanobacteria using a Semi-analytical algorithm resulted in cyanobacterial concentration values a little bit higher compared to the measured ones, hence, some factors used by the model need to be adjusted to the Vaal Dam site for better estimations.

This research revealed that using band ratio indices of Landsat-8 and Sentinel-2 data are valuable tools for mapping chlorophyll-a in the Vaal Dam, a key indicator of phytoplankton biomass. Furthermore, using the semi-analytical algorithm with hyperspectral data is key for estimating the cyanobacteria concentration in the dam water. Models developed in this research will significantly improve near-real-time and long-term chlorophyll-a monitoring of the Vaal Dam. It will effectively help researchers and environmental agencies monitor changes in algal biomass of the dam water to address public health issues related to water quality. It helps to identify areas of high nutrient input and assess the effectiveness of water quality management strategies.

It is of prime importance that the developments within the catchment of the Vaal Dam be carefully considered as it is one of the primary sources of dam water. The research recommends implementing the existing regulatory policies for effluent dispersal within the catchment to protect

ecosystem functioning and water resources from further deterioration in their quality. It also recommends regular monitoring to detect real-time changes in HABs using satellite remote sensing.

Keywords: Vaal Dam; remote sensing; image classification; random forest; time series; chlorophyll−a; harmful algal blooms; water quality; South Africa

Declaration

I declare that this work is my original work and has not been previously submitted to obtain any academic qualification. Data and information obtained from published and unpublished work of others have been acknowledged in the text and a list of references is herein provided.

Signature: ____  ____

Date: _______27/03/2024____

Declaration – Plagiarism

I, Altayeb Obad, declare that this thesis titled "CAPABILITY OF MULTI-REMOTE SENSING SATELLITE DATA IN DETECTING AND MONITORING CYANOBACTERIA AND ALGAL BLOOMS IN THE VAAL DAM, SOUTH AFRICA." and the work presented in it is my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the University of the Witwatersrand, Johannesburg.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. Apart from such quotations, this thesis is entirely my work.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signature: ____  ____

Declaration – Publications and manuscripts

I **Obaid, A.;** Adam, E.; Ali, K.A. (2023) Land Use and Land Cover Change in the Vaal Dam Catchment, South Africa: A Study Based on Remote Sensing and Time Series Analysis. *Geomatics* 2023, *3*, 205-220. DOI: https://doi.org/10.3390/geomatics3010011.

II **Obaid, A.;** Adam, E.; Ali, K.A., Abiye, T. A. (2024) Time series analysis of water quality factors enhancing harmful algal blooms (HABs): a study integrating in-situ and satellite data, Vaal Dam, South Africa. Water 2024, 16, 764. DOI: https://doi.org/10.3390/w16050764.

III **Obaid, A.**; Ali, K.; Abiye, T.; Adam, E. (2021) Assessing the utility of using current generation high-resolution satellites (Sentinel 2 and Landsat 8) to monitor large water supply dam in South Africa. *Remote Sensing Applications: Society and Environment* 2021, *22*, 100521. DOI: https://doi.org/10.1016/j.rsase.2021.100521.

VI **Obaid, A.;** Adam, E.; Ali, K.A., Abiye, T. A., Ortez, J. (in preparation) Using Sentinel-2 satellite data and In-situ Hyperspectra for Cyanobacteria Assessment in the Vaal Dam, South Africa.

Dedication

To my mother for her prayers and constant support during my entire educational career,

To the soul of my father who always believed on my success during his life,

Acknowledgements

I would like to express my gratitude and acknowledgement to all institutions and individuals who contribute to this research. I am grateful to the University of the Witwatersrand, Johannesburg, South Africa. For giving me the Post-graduate Merit Award and providing me with the research facilities during my PhD study.

I am also would like to send my thanks to the University of Alfashir, Sudan for giving me the chance to come and study at Wits, and for their valuable support.

I wish to express my sincere gratitude to Prof. Elhadi M I Adam for his unwavering academic supervision, mentoring, and steadfast support throughout every phase of this research endeavor. His guidance and encouragement were instrumental in completing this study. Additionally, I extend my thanks to Prof. Khalid A Ali for his invaluable contribution, inspiration, and guidance, which significantly enriched the research process. Thank you both for your exceptional support and guidance.

I am grateful to my family for their help and encouragement. My mother, brothers, and sisters many thanks to all of you for your continuous support, my success is your reward to all of you, my beloved family.

Finally, I am grateful to my colleagues, and PhD candidate's fellow, many thanks to all of you for providing me with a comfortable collaborative study environment with special thanks to my PhD lab mates for the family sense that you provide. Thank you to all and to everyone who contributed to this research in some way.

Table of Contents

List of Figures

List of tables

List of abbreviations

RS	=	Remote Sensing
AC	=	Atmospheric Correction
AERONET-OC	=	Ocean Color component of the Aerosol Robotic Network
CDOM	=	Colored Dissolved Organic Matter
DWAF	=	Department of Water Affairs and Forestry
HABs	=	Harmful Algal Blooms
LULC	=	Land-Use/Land-Cover
L_w	=	Water Leaving Radiance
MLA	=	Machine Learning Algorithms
MSI	=	Multi-Spectral Instrument
OLI	=	Operational Land Imager
NIR	=	Near Infra-Red
SWIR	=	Short Wave Infra-Red
VDR	=	Vaal Dam Reservoir
VNIR	=	Visible Near Infra-Red
WQ	=	Water Quality
WQPs	=	Water Quality Parameters
Chl−a	=	Chlorophyll−a (pigment)
PC	=	Phycocyanin (pigment)
ASD	=	Analytical Spectral Devices
SeaWiFS	=	Sea-viewing Wide Field-of-view Sensor
MERIS	=	Medium Resolution Imaging Spectrometer
MODIS	=	Moderate Resolution Imaging Spectroradiometer
OLCI	=	Ocean and Land Cover Instrument
TM	=	Thematic Mapper
ETM+	=	Enhanced Thematic Mapper
MSS	=	Multispectral Scanner
HRG	=	High-Resolution Geometric Sensor
ISI	=	International Scientific Indexing
RF	=	Random Forest

WMA	=	Water Management Area
SANLC	=	South African National Land Cover
OOB	=	Out-of-Bag
OA	=	Overall Accuracy
UA	=	User Accuracy
PA	=	Producer Accuracy
CART	=	Classification and Regression Trees
KJEL_N	=	Organic nitrogen
SLR	=	Stepwise Logistic regression
NH_4_N	=	Ammonia Nitrogen
NO_3NO_2_N	=	Nitrate and Nitrite Nitrogen
DO	=	Dissolved Oxygen
USGS	=	United States Geological Survey
OC	=	Ocean Colour
SNR	=	Signal to Noice Ratio
RMSE	=	Root Mean Squared Error
SDD	=	Secchi Disk Depth
SANAS	=	South African National Accreditation System
C2RCC	=	Case 2 Regional Coast Color
WHO	=	World Health Organization
SNAP	=	Sentinel Application Platform

Chapter One

General Introduction

1.1 Background

Survival on the earth requires water, and even though there is a lot of water on the planet, less than 3% of the total water resources is represented by freshwater. Among this small percentage, less than 1% is in the form of the earth's liquid surface freshwater which includes rivers, lakes, marshes and the seasonal or ephemeral transient waters (Revenga and Kura, 2003). The remaining freshwater is in the form of subsurface groundwater and water being frozen in ice caps. This small available fraction of freshwater on earth supports a high level of biodiversity of habitats including human living. Historically, all major human civilisation and settlements are associated with river systems (Bănăduc et al., 2022, Revenga and Kura, 2003). Now, over 50% of the world's population are living very close to the surface freshwater bodies, closer than 3 km compared to just around 10% who are living further than 10 km away (Kummu et al., 2011). Humans rely on freshwater resources for various important uses and as population growth and water-use demands increase, human activities on the land have enormous negative effects on the freshwater resources by allowing pollution to reach the rivers and been transported to be accumulated in reservoirs, lakes, and other wetlands or to be transported to the sea. However, activities such as agriculture, forestry, mining, urbanization and power generation coupled with adverse natural climatic conditions such as highly variable rainfall and excessive evaporation have created in some areas a highly polluted river system (Gyedu-Ababio and Van Wyk, 2004, Braune and Roger, 1987). Some of such pollutants loaded into the water resources act as nutrients enhancing phytoplankton growth in freshwater bodies. Among such nutrients, phosphorus (P) and nitrogen (N) are considered the most enhancing factors to algal blooms (Davis et al., 2015, Ding et al., 2018), their high loads leading to excessive plant biomass growth which leads to as known as eutrophication. Eutrophication is a process by which lakes, rivers and coastal water become increasingly enriched with nutrients leading to excessive plant growth in water bodies (Griffin et al., 2014, Van Ginkel, 2011). It is one of the most prevalent inland water quality problems which substantially impairs beneficial uses of water (Brooks et al., 2016, Chislock et al., 2013, Smith, 2003).

South African freshwater bodies are no exceptions, eutrophication is at the centre of water quality challenges confronting water authorities, it is a major threat to surface freshwater in the country and a matter of public and scientific concern (Griffin et al., 2014, Malahlela et al., 2018). Algal and cyanobacterial blooms are common, they are associated with many undesired effects including foul odours and tastes, and toxicity alterations. These aspects will limit the functionality of such freshwater resources and pose a risk to human health (Griffin et al., 2014, Van Ginkel, 2011).

In Gauteng province, domestic water supply interdepends on an embedded large dam known as Vaal Dam. The dam water seems strongly affected by the interaction of chemical, biological, and physical elements which raises the turbidity and makes it optically complex water. Within its catchment, several anthropogenic activities were identified as primary contributors to the water quality deterioration such as diffusion of dissolved urban effluents, mining wastes, agriculture, and industrial discharges, most of which are nutrient-rich. Such activities resulted in raising some nutrients loading the rivers feeding the dam, riverine habitat degradation in some sub-catchments, and frequent recent blooms evident in the Vaal Dam. Most of such contaminants have their origin in the upstream areas, such as those of the Vaal and Wilge rivers (Gyedu-Ababio and Van Wyk, 2004). The tributaries of the Vaal River may be most important in terms of draining some highly industrialized areas of Gauteng and Mpumalanga. Studies have reported that most of the significant water quality challenges emerge from biological materials such as solid faecal materials and chemical materials such as those from gold mining & industrial pollution (Malahlela et al., 2018). Besides anthropogenic activities, natural processes also have effects on the water quality in the study area. Weather, climate, seasonal trends, underlying geology, and hydrology have natural impact on water quality (Khatri and Tyagi, 2015).

As a result, the dam reservoir experienced frequent algal blooms, and it became more noticeable recently which warned the authorities and scientists to pay more attention to such important issues. To support sustainability and to better manage this vital resource, water resource managers need enhanced capabilities of near real-time monitoring of HABs to understand the state conditions and better protect, manage, and control its blooms in the dam water. The conventional methods for monitoring HABs using in-situ biogeochemical measurements, and sample collection give accurate estimations of chlorophyll-a concentration at the points of sampling, but it is limited in spatial and temporal coverage, it is time-consuming and expensive to cover the spatial dynamics of the HABs. Moreover, the traditional approach does not allow for adequate water quality monitoring in surveillance programs which is crucial to ensuring public safety (Kim et al., 2016).

1.2 The inland freshwater quality monitoring

The degradation of inland freshwater bodies, such as lakes and reservoirs, has heightened the need to assess and monitor their quality for better understanding of their status, which is essential for human and aquatic health. However, the most common monitoring and assessment practices involve direct measurements of contaminant concentrations in the water through sampling and laboratory estimation of water quality parameters. Data obtained from this method does not provide a synoptic coverage of the entire body of water; rather, it allows for accurate measurements only at discrete points and at the time of sampling. Consequently, it does not reflect the water quality status between measurement intervals, potentially leading to inadequate detection of intermittent contamination events (Debén et al., 2015). In many countries, alongside traditional methods, there is a growing adoption of near real-time monitoring and detection of contaminants. This involves deploying a network of sensors strategically positioned to cover targeted sources. These sensors promptly transmit real-time measurements to a centralized system, providing immediate insight into the water's condition. Nevertheless, it is important to highlight that certain monitoring systems can be expensive (Ahmed et al., 2020) and require specific facilities for implementation, which may not be accessible to some economically disadvantaged nations globally.

Hydro-morphological and biological measures are considered for integrated contamination diagnosis. Another method has been studied for routine monitoring practices to detect bio-available metals in the water using non-biological accumulators such as Diffusive Gradients in Thin film (Diviš et al., 2007). For decades, some different organisms have been used as contamination bio-monitors for contamination such as bryophytes, algae, and mussels, it allows for to detection of sporadic contamination and the contamination source's location precisely. This method is not yet been used publicly (Debén et al., 2015).

Recently, Remote sensing techniques have increasingly been used worldwide to monitor aquatic systems and provide an instantaneous synoptic overview once the data is calibrated using in-situ measurements, Therefore, the intensive focus in current studies is on advancing research using remote sensing capabilities for monitoring inland water quality.

1.3 Remote sensing for inland freshwater quality monitoring

Remote sensing for water quality monitoring depends on the reflection, absorption, and transmission characteristics of the electromagnetic spectrum in different wavelengths to differentiate between the water properties. Some water quality parameters such as algal blooms have proxies to be measured by remote sensing to indicate their abundance in the targeted water body. Chl−a is a universal pigment normally used as a proxy for phytoplankton through its alteration of the irradiance light by scuttering and absorption and changing the spectral signature of the reflected radiance from the water surface (Franz et al., 2014, Gholizadeh et al., 2016). The electromagnetic spectrum signature within the ranges of visible to near-infrared bands has been used for Chl−a concentrations modelling (Ali et al., 2022a), Chl−a absorbs the spectrum at two peaks; approximately at 443 nm and 686 nm (Richardson, 1996). Chl−a concentration in water indicates the abundance of phytoplankton biomass, but it does not provide an accurate estimation of cyanobacterial bloom (Khan et al., 2021). Hence, in the case of cyanobacterial dominance, the estimation of phycocyanin pigment is required. It is the proxy for the cyanobacterial occurrence in the water. The peak absorption of the phycocyanin pigment is around 615 nm (Simis et al., 2005). In some cases, multispecies of phytoplankton will be present together, which will affect the absorption peaks of phycocyanin by overlapping with Chl−a and the absorption peak might range between 615 and 650 nm (Khan et al., 2021, Kutser et al., 2005).

Remotely sensed data can be acquired through spaceborne sensors, airborne sensors, and ground-based data collection using, for example, handheld analytical spectral devices (ASD) spectrometers. A suitable sensor choice for a particular project should be made regarding the spectral, spatial, and temporal resolutions to achieve certain objectives of the project (Ibrahim and Monbaliu, 2011). Spaceborne data acquired from commonly used sensors including the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), The Medium Resolution Imaging Spectrometer (MERIS), Moderate Resolution Imaging Spectroradiometer (MODIS), and the Ocean and Land Cover Instrument (OLCI) onboard Sentinel-3 have been used for monitoring chlorophyll-a in Oceans. In many studies, they have been used for detecting HABs in inland water (Qin et al., 2015, Matthews et al., 2010, Shen et al., 2017). Still, there are challenges associated with their data type such as their insufficient spatial resolution. Luckily, many other spaceborne sensors are available at low data costs, and some have free access to their data and are much finer in their spatial resolution. Such sensors can capture high-quality remote sensing reflectance data for the land surface and inland water bodies, which can successfully be used to retrieve chlorophyll-a concentrations.

To retrieve the concentration of Chl−a from remote sensing reflectance (R_{rs}) of satellite data captured from inland waters, algorithms have been developed using the red-infrared spectrum band combination (Pahlevan et al., 2017a) instead of blue-green algorithms which were successfully used for retrieving Chl−a concentration in the open ocean (Case I) water (Ali et al., 2016, Hieronymi et al., 2016). This is because various constituents in inland (Case II) waters absorb light in the blue region which might create uncertainty when using it to detect chlorophyll-a concentration in such productive waters (Gons, 1999, Gitelson et al., 2007, Franz et al., 2014).

1.4 Multispectral sensors for inland freshwater quality monitoring

The multispectral land surface optical sensors are characterised by higher spatial resolution (10 to 30 meters) in contrast to coarser spatial resolution ocean colour sensors (Sensors have been used to monitor chlorophyll-a levels in oceans). The finer spatial resolution of the land surface sensors such as; Landsat-8 (OLI), Landsat-7 (ETM+), Landsat-5 (TM), Landsat-4 (MSS), Sentinel-2 A & B (MSI), SPOT5 (HRG) as well as some commercial sensors allows them to provide useful information for estimating the characteristics of the earth's surface over large areas (Kloiber et al., 2002, Wang et al., 2004, Tyler et al., 2006, Watanabe et al., 2015, Kutser et al., 2016, Doxaran et al., 2002, Dvornikov et al., 2018, Hellweger et al., 2007). Some of the mentioned-above sensors can detect the spatial patterns of water quality in inland water bodies such as small lakes, dam reservoirs, rivers, …etc., and can be used to measure and monitor chlorophyll−a, watercolour, and suspended sediments (Ibrahim and Monbaliu, 2011, Shang and Zhu, 2019, Eugenio et al., 2020). But their coarser temporal resolution (around one to two weeks) limits their measurement to only twice or three times a month depending upon revisit time and cloud cover conditions compared to the ocean colour sensors with a shorter revisit time of one day (Topp et al., 2020).

Recently launched satellite generations including Landsat-8 and Sentinel-2 with their high-resolution sensors can detect the spatial variations of the optically active constituent reflectance from open water surfaces such as Chl−a (Watanabe et al., 2017, Buma and Lee, 2020). The Operational Land Imager (OLI) sensor has a relatively high spatial resolution. It can monitor the Algal bloom dynamic in large open water bodies (Markogianni et al., 2020, Cao et al., 2020). The sensor has a high radiometric resolution (12-bit) compared to its previous generation which has the capability to distinguish very small differences in the optics of the water column as a function of water quality parameters (Pahlevan et al., 2014). An evaluation has been performed for OLI R_{rs} products at sites within the Ocean Colour component of the Aerosol Robotic Network (AERONET-OC), it showed that the near-infrared and short wave infrared (NIR-SWIR) band combinations yield the most robust and stable R_{rs} retrievals in moderately turbid waters (Pahlevan et al., 2017c). Multispectral Instrument (MSI) is very similar to OLI sensor in design and requirements with some differences in band number, nominal band centres, spatial resolutions, and revisit times (Pahlevan et al., 2017b). The sensor has 13 bands of different spatial resolutions: 10 m, 20 m, and 60 m, see Table 4.1. MSI data is also promising in driving high-quality products for retrieving inherent optical properties of the water column in coastal and inland areas such as chlorophyll-a (Pahlevan et al., 2017b, Ansper and Alikas, 2018). MSI has a finer spatial resolution (10 meters) which might detect the spatial variations in narrower estuaries than OLI.

OLI and MSI sensors have been more suitable sensors for inland water applications in recent years (Dörnhöfer and Oppelt, 2016, Niroumand-Jadidi et al., 2019). Many studies have successfully retrieved chlorophyll-a from OLI and MSI data under various atmospheric conditions and for different aquatic systems using various approaches, ranging from simple regression analysis between satellite and in-situ data to semi-analytical models and machine-learning approaches (Zhang et al., 2020, Laneve et al., 2021, Wang and Chen, 2024, Ouma et al., 2020).

1.5 Challenges in remotely mapping inland freshwater quality parameters

Satellite observations have been effectively employed for marine remote sensing using sensors that measure ocean colour in open sea waters (Case I), primarily dominated by phytoplankton. However, the application of these sensors in monitoring inland water bodies (Case II) has seen limited advancement. They encounter various challenges, including coarse spatial resolutions that render them unsuitable for remote sensing applications in inland waters such as lakes, reservoirs, and rivers. However, higher spatial resolution sensors like those in the Landsat series have been utilized. These sensors offer sufficient spatial resolution for numerous inland water bodies. Their original design is geared toward land applications, and their radiometric sensitivity, spectral resolution, and coverage may not be optimal for many inland water monitoring applications (Palmer et al., 2015).

Additionally, many unresolved problems added more challenges to inland waters compared to ocean colour, such as atmospheric correction issues due to the optical complexity of inland waters, adjacency effects, high concentrations of phytoplankton biomass which in some cases reach up to 350 mg m^{-3} (Gitelson et al., 1993), presence of the suspended sediments, presence of coloured dissolved organic matter (CDOM), tripton, (Balasubramanian et al., 2020, Malahlela et al., 2018, Pahlevan et al., 2020, Dekker et al., 1992, Resolution, 1993), and the high variability of the optical properties even within the same water body. All these issues have complicated the development of the algorithms for inland waters making the identification of the HABs proxies not easy, due to absorption overlapping with the above-mentioned components. For example, studies indicated that sediments affect phycocyanin peak absorption while CDOM affects Chl−a peak absorption mainly in shortwave lengths (Dekker, 1993, Gholizadeh et al., 2016). Another significant challenge in inland water remote sensing research is that it is often viewed as a local or national issue, with limited regional focus. In contrast, ocean colour remote sensing products, developed through well-supported multinational projects, receive greater attention and importance (Palmer et al., 2015).

1.6 Research problem statement

Vaal Dam is the most important dam in South Africa; it is the leading water supplier to the economic heartland of the country, around Johannesburg and Pretoria; the dam serves potable water to over 13 million in Gauteng Province and the areas around the dam. The dam has also been used for recreational, water sports, and fishing activities. It's world-renowned for freshwater carp and catfish fishing; it has also been in the Guinness Book of Records as the "*Most Yachts in an Inland Yacht Race in the World*".

The dam's water quality is under threat due to intensive anthropogenic activities in its catchment. These activities release contaminants and nutrient-rich effluents into the water courses, significantly impacting the dam's water quality. The situation is exacerbated by the failure of waste treatment systems in some areas, leading to the release of raw sewage into the water courses such as the overwhelmed sewage flood in the Deneysville town at the Vaal Dam bank 2015 (https://mg.co.za/article/2015-07-23-sewage-in-gautengs-drinking-water/), and seepage of sewage from holding ponds at the Bethal wastewater treatment works into the Vaal River catchment system 2022 (https://www.groundup.org.za/article/sewage-seepsinto-vaal-dam-as-mpumalanga-water-treatment-plants-fail/). This has resulted in a severe deterioration of water quality, reaching alarming levels (Mnguni, 2022).

The dam water is highly turbid (Sakuno et al., 2018b) and experiences frequent harmful algal blooms. Among the species, cyanobacteria exist during whole seasons of the year, with excessive growth during the summer season; it dominates over other species during the summer bloom between February and March (Chinyama et al., 2016). Cyanobacteria are known to produce cyanotoxins that may affect the consumer's health. Monitoring of HABs in the dam reservoir has been conducted using the traditional methods of collecting water samples and sending them to the laboratories to analyse them; the sample results have sometimes been delayed from a few hours to a couple of days depending on many factors (Swanepoel et al., 2016), this method is labouring, time-consuming, and costly and it can't cover the spatial variations of HABs. More effective management and monitoring of HABs is necessary. Using remote sensing techniques can overcome the limitations of spatial coverage and temporal resolution to map the HAB dynamics accurately (Kerrigan and Ali, 2020). Only two studies so far were found to evaluate Chlorophyll-a concentrations in the Vaal Dam using OLI and MSI remote sensing data utilising stepwise logistic regression (SLR) and band ratio indices, none of them tried to estimate the phycocyanin pigment, the proxy of the cyanobacteria. This research examined multispectral and hyperspectral data with different indices and algorithms to detect chlorophyll-a and phycocyanin concentrations in the Vaal Dam water.

1.7 Research aim and objectives

The main goal of this research is to test the potential use of satellite remote sensing techniques in mapping and monitoring harmful algal blooms in the Vaal Dam Reservoir, South Africa.

The specific objectives are:

1. To quantify the land use/land cover changes of the Vaal Dam Catchment over the last four decades to assess the impacts of such changes on the water quality of the Vaal Dam.
2. To perform long-term time series analysis of some chosen water quality parameters serving as potential nutrients of algal growth in the dam water, and to explore the annual spatial distribution of summer blooms during the last decade using satellite-based time series analysis.
3. To examine the utility of using high-resolution Landsat-8 and Sentinel-2 multispectral data to map and assess the concentration of harmful algal blooms in the Vaal Dam Reservoir.
4. To examine the applicability of some existing algorithms, indices and band ratios for detecting the concentration of cyanobacteria in the dam reservoir, and to adjust some regional values if applicable to improve the remote sensing retrievals.

1.8 Scope of the research

Many previous studies have utilised various methodologies to assess aquatic and water quality issues in inland water bodies using remote sensing capabilities. Different aquatic systems have been studied in various atmospheric conditions. The knowledge acquired behind them enhances our understanding of challenges that should be resolved. This research represents a further effort to identify immediate and cost-effective solutions for monitoring and addressing water quality concerns in the Vaal Dam. The study aimed to assess the use of satellite remote sensing techniques in monitoring harmful algal blooms (HABs) in the Vaal Dam, South Africa. Landsat-8 OLI and Sentinel-2 MSI satellite data have proven to be invaluable tools in monitoring harmful algal blooms in inland waters. They provide high-resolution multispectral imagery that allows for the detection and mapping of algal blooms over the water body. Mapping HABs and assessing their spatial and temporal distribution (growth patterns) is crucial to water supply authorities, scientists in the field of the aquatic and environmental sciences, and other stakeholders to draft and implement programs that can assist in controlling HABs for timely decision-making and mitigation efforts leading to sustainable monitoring and management. This can be achieved by analysing the spectral signatures over the spatial coverage of the water body and the satellite's revisit times to enable near real-time detection of algal blooms.

1.9 The thesis outline

This research consists of six chapters, of which chapters 2 to 4 have already been published in International Scientific Indexing (ISI) rated peer-reviewed journals. Chapter 5 is being prepared for publication in ISI-rated peer-reviewed journals. The chapters were presented in a standalone scientific article format with aims, objectives, results, discussion, and conclusion. The conclusions in each chapter address the main aim of this research study. Little or no changes were made to the chapters that have been published. Therefore, the previously mentioned approach involves some reiterations in various sections which is unavoidable in some of the chapters. However, this drawback is insignificant considering that each chapter got a peer review as a standalone article after submission to international journals, and it can be read independently without losing its meaning. This research is organized as follows:

Chapter 1 - Introduces the general introduction of the study. It highlights the general background for the study area, the use of multispectral remote sensing sensors in mapping harmful algal blooms, and the challenges encountered in satellite-based HABs mapping. The research objectives were also outlined in this chapter.

Chapter 2 - Explores the LULC classification in the Vaal Dam catchment area using Landsat-8 data. This aimed to show the changes in the patterns of LULC for the period between 1986 to 2021 and highlight the anthropogenic activities within the catchment. A Random Forest (RF) classifier was used to identify the LULC changes within the study area. A high level of classification accuracy was achieved in this chapter by using such an algorithm, showing its potential accuracy for mapping, and detecting the changes in LULC patterns using OLI multispectral data. The following chapter evaluates the historical water quality data available in the Department of Water and Sanitation website.

Chapter 3 - Assesses archived water quality data available to the public, accessed on the Department of Water and Sanitation website. In this chapter, the researcher evaluated some water quality parameters that serve as nutrients enhancing algal blooms in the Vaal Dam Reservoir. The chapter also tested the utility of applying a previously used algorithm for mapping chlorophyll-a in different dating OLI data to perform satellite-based time series analysis to map the annual changes in the spatial distribution of HABs during the last decade. The applicability of the current generation of high-resolution satellite data was tested in the mapping of HABs in the Vaal Dam and it is covered in the following chapter.

Chapter 4 - Examines the capability of using current-generation high-resolution satellite imagery for mapping chlorophyll−a in an embedded water supply dam in South Africa. The study tested the utility of using Sentinel-2 and Landsat-8 data to map the spatial distribution of HABs in the Vaal Dam. The use of the ocean colour algorithms and band ratio indices were examined for chlorophyll−a detection. The results showed that the blue-green ocean color algorithm for OLI data, and red to near-infrared band ratio for MSI data were vital in the mapping of the chlorophyll−a in the study area. The testing of different algorithms for detecting the phycocyanin pigment was discussed in the next chapter.

Chapter 5 - Evaluate the ability to use some adjusted algorithms on multispectral imagery in mapping HABs and counting the cyanobacteria within the dam reservoir. The chapter examined the effectiveness of such algorithms in mapping cyanobacterial blooms in the study area. The study indicated the effectiveness of seasonality in the domination of blue-green algae upon other species, which is important to the inter-seasonal management of HABs in the study area.

Chapter 6 - Summarises the outcomes of this research. It outlines the main contributions of this research to the detecting, mapping, and monitoring of algal and cyanobacterial blooms. In this chapter, the limitations of the current study are outlined, and it also gives an outlook into other research aspects that can be considered in further studies. This study seeks to assist water supply authorities with knowledge of the data and techniques used in HAB mapping. This is vital for minimizing the cyanotoxin hazards which promote supplying safe water to consumers.

Chapter Two

Land Use and Land Cover Change in the Vaal Dam Catchment, South Africa: A Study Based on Remote Sensing and Time Series Analysis

This chapter is based on **Obaid, A.;** Adam, E.; Ali, K.A. (2023) Land Use and Land Cover Change in the Vaal Dam Catchment, South Africa: A Study Based on Remote Sensing and Time Series Analysis. *Geomatics* 2023, *3*, 205-220. Available at: https://doi.org/10.3390/geomatics3010011.

Abstract: Understanding long-term land use/land cover (LULC) change patterns is vital to implementing policies for effective environmental management practices and sustainable land use. This study assessed patterns of change in LULC in the Vaal Dam Catchment area, one of the most critically important areas in South Africa since it contributes a vast portion of water to the Vaal Dam Reservoir. The reservoir has been used to supply water to about 13 million inhabitants in Gauteng province and its surrounding areas. Multi-temporal Landsat imagery series were used to map LULC changes between 1986 and 2021. The LULC classification was performed by applying the random forest (RF) algorithm to the Landsat data. The change-detection analysis showed grassland being the dominant land cover type (ranging from 52% to 57% of the study area) during the entire period. The second most dominant land cover type was agricultural land, which included cleared fields, while cultivated land covered around 41% of the study area. Other land use types covering small portions of the study area included settlements, mining activities, water bodies and woody vegetation. Time series analysis showed patterns of increasing and decreasing changes for all land cover types, except in the settlement class, which showed continuous increase owing to population growth. From the study results, the settlement class increased considerably for 1986–1993, 1993–2000, 2000–2007, 2007–2014 and 2014–2021 by 712.64 ha (0.02%), 10245.94 ha (0.26%), 3736.62 ha (0.1%), 1872.09 ha (0.05%) and 3801.06 ha (0.1%), respectively. This study highlights the importance of using remote sensing techniques in detecting LULC changes in this vitally important catchment. The natural landscapes in many areas within the catchment are significantly altered through settlements, mining, and agriculture increasing nutrient runoff into the tributaries and rivers feeding the dam. This process fuels algal growth in the reservoir. Therefore, it is crucial to carefully manage land use practices in this catchment to control nutrient runoff and mitigate the potential for algal blooms in the dam reservoir.

1.1 Introduction

Land resources have been used for social, material, and cultural human demands, which leads to significant changes in LULC patterns (Birhane et al., 2019, Twisa and Buchroithner, 2019). Such changes have consequently been associated with various impacts and effects on different fields at multiple scales; local, regional, and global (Deng et al., 2013, Tang et al., 2021), including surface energy balance through its effect on the weather and climate at local, regional (Prijith et al., 2021, Dutta et al., 2022) and global scales, such as how changing the tropical rainforest areas impacts the global climate (Snyder et al., 2004). LULC changes affect the hydrological cycle by altering the hydrological response of watersheds in terms of surface runoff, decreasing groundwater recharge, water quality and pollutant transfers. These factors affect the dispersion of non-point water pollutants and direct them into freshwater bodies (Abdulkareem et al., 2018). Furthermore, changes in LULC affect biodiversity and aquatic systems (Twisa and Buchroithner, 2019, Tang et al., 2021, Du Plessis, 2017, Hua, 2017) and the receiving ecosystem service values by affecting their structure and functioning (Wang et al., 2017, Gong et al., 2022). Therefore, understanding the spatial and temporal variations of LULC on a watershed level is critical for effective monitoring, planning and management of the resources and ecosystems (Hua, 2017). Accurate information on LULC changes can also contribute to other applications, including the assessment of damage and deforestation, the monitoring disasters and measurement of the expansion of urban areas. It also assists with land use management and planning (Attri et al., 2015). However, obtaining such information can be achieved through performing long-term time series analysis of the LULC changes (Birhane et al., 2019, Attri et al., 2015).

The Vaal Dam Catchment forms a vast part of the Upper Vaal Water Management Area (WMA). It is part of the Vaal drainage system in South Africa [(Ilunga, 2017). Extensive gold and coal mining activities have taken place within the Upper Vaal catchment (DWAF, 2004). In the Vaal Dam Catchment area, the range of land use includes major agricultural activities (encompassing mainly cattle grazing and dry land cultivation), gold and coal mining and some industrial activity (Abdulkareem et al., 2018, Du Plessis, 2017). In the past, most human activities within the Vaal Dam Catchment area were dependent on or related to agriculture (Chutter, 1970, Archibald, 1971, Gyedu-Ababio and Van Wyk, 2004). The most extensively cultivated areas within the Vaal Dam Catchment have been near the Vaal and the lower-lying valleys of the Wilge River, with stock farming in the hilly parts (Chutter, 1970). With the building for Sasol of the synthetic fuel complex and the commencement of coal mining in late 1970 (alongside other economic developments), the land use character of the catchment has changed substantially. The towns related to these developments grew significantly, mainly in the Waterval sub-catchment in the northern part of the Vaal Dam Catchment (Gyedu-Ababio and Van Wyk, 2004). These economic developments have contributed to the ongoing expansion of settlement areas in the catchment area (Hua, 2017). The southern and southeastern part of the Vaal Dam Catchment is contained in the Wilge River Sub-catchment, which is dominated by agricultural land (consisting of non-cultivated arable land, cultivated farms, livestock pastures and some human settlement areas) (Gong et al., 2022). As human activity has intensified in recent decades, ecosystems within the catchment have been degraded (Gyedu-Ababio and Van Wyk, 2004). As a result of these activities, the discharge of treated effluents from the mine dewatering and urban areas within some areas of the Vaal Dam Catchment returns into the river system and causes significant impacts on river water quality (DWAF, 2004).

A remote sensing (RS) approach is particularly effective for characterising the LULC changes for large areas such as the Vaal Dam Catchment. Satellite RS has been used as a cost-effective method in mapping and developing a clear understanding of LULC changes. Various satellites with moderate to high spatial resolution and temporal coverage are available that can be used to study long-term changes in LULC (Owojori and Xie, 2005). Many studies have used satellites such as the Moderate Resolution Imaging Spectroradiometer (MODIS) (Clark et al., 2010, Thenkabail et al., 2005, Wardlow and Egbert, 2008) and Landsat series data to study LULC at regional scales in different regions in the world (Manandhar et al., 2009). MODIS coverage started in 2000 with a spatial resolution of one kilometre and a revisit time of one day. Starting in 1972, Landsat has had a longer coverage time; since then, a new era in earth observations has begun, using moderate (60-metre) spatial resolution imagery (Yuan, 2008, Kenea et al., 2021). After 1982, Landsat sensors started to acquire data in higher (30-metre) spatial resolution but with a lower (16-day) temporal resolution. However, since LULC change is not noticeable over short periods (such as hours and days), Landsat is better at detecting LULC change owing to its higher spatial resolution and its comprehensive archive compared to the coarse resolution satellite data such as NASA's MODIS. The Landsat archive is very valuable for estimating area changes over time; it allows the LULC changes to be thoroughly assessed and statistically quantified (Owojori and Xie, 2005). Another advantage of using Landsat data is the consistency of configurations of the various generation sensors (TM, ETM+ and OLI) within visible to shortwave infrared bands, as well as their spatial resolution (30 m), which allows us to use a continuous data set starting in 1984 (Yuan, 2008). Obtaining information on LULC change based on RS has been used in many parts of the world to address various environmental challenges (Gong et al., 2022, Owojori and Xie, 2005, Yuan, 2008). RS data and field observations can be combined to accomplish LULC classification and change detection. RS data provides faster processing and is more cost-effective than traditional methods (Diallo et al., 2009). Many studies conducted within the Vaal Dam Catchment area investigating water quality issues, few of them have highlighted the impact of LULC owing to human activities on surrounding water quality (Gyedu-Ababio and Van Wyk, 2004, Verheul, 2012). However, the studies conducted to date on LULC have lacked spatial and temporal resolution for accurately characterising the impact of LULC on the large-scale functioning of the ecosystem. The application of geospatial techniques can provide robust methods for studying the impacts of LULC on large catchment systems such as the Vaal Dam Catchment (Chughtai et al., 2021). The Vaal Dam Catchment is a primary water source for the Vaal Dam Reservoir beside the Lesotho highland water project, from which Rand Water supplies potable water to Gauteng province (Plessl et al., 2019). Many issues of water quality and ecosystem deterioration have been discussed in the literature for this crucial area (du Plessis et al., 2018); understanding the LULC patterns and evaluating their effects on ecosystems and water quality are essential. To understand their dynamics, there is a need to first carry out LULC classification and then to detect respective changes in LULC. This will provide critical information that can be used for effective and sustainable environmental restoration and management of resources to avoid and minimise further deterioration in water quality and eco/aquatic systems. This research aims to accurately characterise the LULC change in the Vaal Dam Catchment area over recent decades (1986 to 2021) using NASA's Landsat data. It is anticipated that this could contribute towards evaluating requirements for better management of catchment.

2.2 Materials and Methods

2.2.1 Study Area

The Vaal Dam Catchment is located between 26.27_ and 28.77_ S, and 28.00_ and 30.31_ E, in the central plateau of South Africa. The annual rainfall generally ranges between 600 and 800 mm/y and occurs mostly between October and March during the summer season. In the south and southeastern part of the catchment, annual rainfall can reach up to 1500 mm/y (Chutter, 1970, Obaid et al., 2021). The catchment has an area of about 38,000 km^2, with an altitude ranging between 1300 and 1850 m.a.s.l. The Vaal Dam Catchment consists of two main sub-catchments drained by two major river systems, the Vaal River and Wilge River catchments (Figure 2.1). The two river systems are the main arteries for the Vaal Dam Catchment that supplies fresh water to the Vaal Dam Reservoir, from which water is supplied to Gauteng province and its surrounding areas. The southern and southeastern parts of the catchment are mountainous areas, and they are the highest parts of the catchments. In this part, the Wilge River rises from the northern slopes of Mont-aux-Sources in the Drakensberg Mountain range, whereas the northern and northeastern parts of the catchment are relatively flat areas.

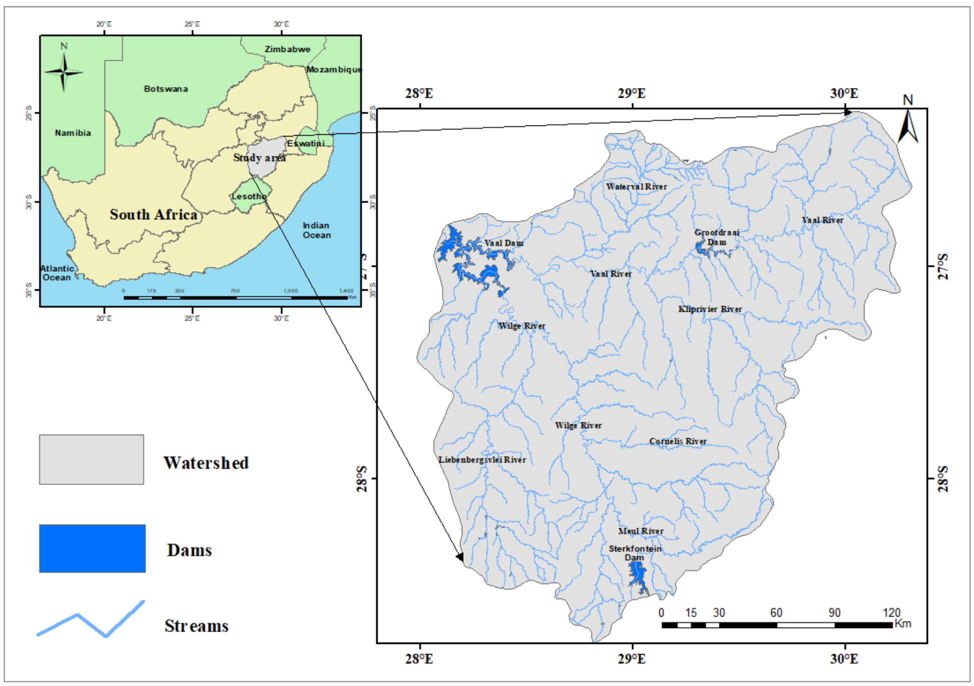


Figure 2.1 The location map of the study area.

The study area was chosen since the Vaal Dam is one of the most important water sources in South Africa and since there are considerable concerns regarding water quality in the catchment over recent decades. Further, there is interest from the water authorities in understanding the pattern of LULC changes in the catchment since this information will assist them in better understanding potential threats to the water quality of the dam.

2.2.2 Identification of LULC Types

Historical Google Earth images and available South African National Land Cover (SANLC) classification schemes were used as guides for the identification of the major LULC types in the study area (SANLC data sets for 1990, 2013–2014 and the latest data (SANLC- 2018). The data sets can be obtained from the Department of Forestry, Fisheries and the Environment website (https://egis.environment.gov.za/sa_national_land_cover_datasets) accessed on 15 October 2020.

The SANLC-2018 data set uses 20-metre resolution Sentinel-2 imagery. The study made use of seven major LULC classes, including agriculture, cleared fields, grassland, mining, settlements, water bodies and woody vegetation (Table 2.1).

Table 2.1 Land cover classes used in this study.

Class	Description
Agriculture	Crop areas that were green at the time of imagery
Cleared fields	Cleared fields, bare lands and shaded agricultural fields (greenhouses)
Grassland	Sparsely wooded and natural grassland areas plus fellow land and old fields
Mining	Ash piles, dumps and slime dams accumulated from coal and gold mining
Settlements	Residential, industrial, commercial and mixed urban build-up areas
Water bodies	Rivers, open water, lakes, ponds and reservoirs
Woody vegetation	Natural, planted forests and orchards

2.2.3 Data and Methods

2.2.3.1 Data Acquisition and Pre-Processing

Level 1 imageries were acquired from the USGS Earth Explorer platform (http://earthexplorer.usgs.gov/) accessed on 13 June 2021. A total of 36 cloud-free or good-quality images of Landsat 5 (TM), Landsat 7 (ETM+) and Landsat 8 (OLI) were obtained for the years 1986, 1993, 2000, 2007, 2014 and 2021 to produce 7-year-interval LULC maps. For each year, six scenes were required to cover the span of the Vaal Dam Catchment, and their respective paths/rows were specified (Table 2.2).

Table 2.2 The Landsat satellite images used in the study.

Year	Satellite	Sensor	Path/Row	Resolution (m)	Acquisition Date
1986	Landsat 5	TM	169/078, 169/079, 169/080	30	20.05.1986
	Landsat 5	TM	170/078, 170/079, 170/080	30	11.05.1986
1993	Landsat 5	TM	169/078, 169/079, 169/080	30	24.06.1993
	Landsat 5	TM	170/078,170/079, 170/080	30	30.05.1993
2000	Landsat 7	ETM+	169/078, 169/079, 169/080	30	22.08.2000
	Landsat 7	ETM+	170/078, 170/079, 170/080	30	28.07.2000
2007	Landsat 5	TM	169/078, 169/079, 169/080	30	06.05.2007
	Landsat 5	TM	170/078, 170/079, 170/080	30	29.05.2007
2014	Landsat 8	OLI	169/078, 169/079, 169/080	30	17.05.2014
	Landsat 8	OLI	170/078, 170/079, 170/080	30	24.05.2014
2021	Landsat 8	OLI	169/078, 169/079, 169/080	30	04. 05.2021
	Landsat 8	OLI	170/078, 170/079, 170/079	30	11.05.2021

Radiometric calibration was applied to each scene, followed by atmospheric correction using Fast Line-of-sight Atmospheric Analysis of Spectral Hypercube (FLAASH) module in ENVI (v5.4). The scenes of each year were mosaicked together using a seamless mosaic in ENVI to produce a single image for each year before cropping them to the study area boundaries using a shape file that had been prepared for this purpose by digitising the watershed boundary following the water divide based on stream network of the watershed.

The gap-fill module attached to ArcGIS was used to fill the gaps from Landsat 7 ETM+ scan line corrector failure. Images were then ready for LULC classification.

2.2.3.2 Reference Data

The reference dataset was obtained through visual interpretation of the pre-processed Landsat data sets. Well-distributed training areas were selected to ensure adequate representation of each class within the entire image extension. Point features with class labels for all classes were digitised on all the mosaicked images using ArcGIS. The point features of each year’s image were then loaded on historical high-resolution Google Earth Pro imagery (http://earth.google.com/) accessed on 23 November 2021 and the temporal data was used to assess consistency. Points that fell in the area of inconsistent LULC were excluded from the analysis. Then, pixel values were extracted using the feature extraction tool in ArcGIS. An approximately equal number (100) of sample points was obtained for all classes, and, for the poorly separated classes, more representative point samples were added. The extracted pixel values of each class were randomly divided into training and validation data sets, 70% as the training data set used to train the classifier and 30% as the validation data set to test the classification accuracy. Table 2.3 shows the number of extracted reference data sets of each year used in the classification.

Table 2.3 The reference data numbers used in this study. Training data (Tr), test data (Te) and total number of data points (To).

LULC Class	1986			1993			2000			2007			2014			2021		
	Tr	Te	To	Tr	Te	To	Tr	Te	To	Tr	Te	To	Tr	Te	To	Tr	Te	To
Agriculture	67	28	95	70	30	100	70	30	100	63	27	90	70	30	100	70	30	100
Cleared fields	182	78	260	108	46	154	192	81	273	132	56	188	160	68	228	140	60	200
Grassland	70	30	100	70	30	100	72	30	102	70	30	100	70	30	100	70	30	100
Mining	70	30	100	70	30	100	56	24	80	50	21	71	70	30	100	70	30	100
Settlements	70	30	100	70	30	100	70	29	99	70	30	100	70	30	100	70	30	100
Water bodies	87	37	124	70	30	100	70	30	100	70	30	100	70	30	100	70	30	100
Woody vegetation	70	30	100	70	30	100	67	28	95	70	30	100	70	30	100	70	30	100

2.2.3.3 Image Classification

In this study, the random forest (RF) model was used as the LULC classifier. The model is an ensemble-based classification algorithm suggested by Breiman in 2001 to improve the performance of classification and regression trees (CART) (Xu et al., 2018). It combines a large set of decision trees. The algorithm uses bagging and random selection techniques to build several binary classification trees (ntree) by using bootstrap samples with replacements driven from the original training data sets (each bootstrap sample produces a tree, and ntree is grown from bootstrap samples) (Adam et al., 2014). Then, each predicted classification tree contributes a single vote to assign the most frequent class for the input data. The classifier output is determined by the majority of tree votes. If a separate test data set is not available, the out-of-bag (OOB) method can be used; around one-third of the samples are left out randomly for each newly generated training data set, called OOB samples. The OOB samples are used to measure the variable importance and estimate the misclassification errors (Rodriguez-Galiano et al., 2012, Feng et al., 2015). In RF classification algorithm, two parameters need to be defined, mtry (the number of variables to split at each node) and ntree (the number of trees to grow). A given number of input variables (mtry) at each node are randomly chosen from a random feature subset and the best split is then calculated using only this subset of input features. However, to ensure lower similarity between individual trees and thus a low bias, the tree is allowed to grow fully without pruning (Talukdar et al., 2020). RF parameters (mtry and ntree) have to be optimised to improve the classification accuracy (Feng et al., 2015, Probst et al., 2019). Using the algorithm, the default number of decision trees (ntree) is set to 500, while the default value of the variables number (mtry) corresponds to the square root of the total number of predictor variables (spectral bands) used in the study (Breiman, 2001, Akar and Güngör, 2012). Based on several studies, the RF algorithm is regarded as a robust machine-learning LULC classifier with higher performance (Rodriguez-Galiano et al., 2012, Talukdar et al., 2020, Piao et al., 2021, Ali et al., 2022b, Shetty et al., 2021). RF has many advantages over other machine learning classifiers, in that it is less sensitive to outliers, noise and overtraining; it can also handle large data sets and gives estimations of the important variables in the classification, along with internal generalisation error (OOB error) estimates, in an unbiased fashion (Rodriguez-Galiano et al., 2012, Piao et al., 2021, Hütt et al., 2016).

In this study, the classifications were separately performed using stacked images containing bands from 1 to 6 of TM/ETM+ images, and bands from 1 to 7 of OLI images. The mtry and ntree were kept as default values, the mtry was set to the number of variables which were the number of bands of each stacked image (6 for TM and ETM+ images and 7 for OLI images) and ntree was set to 500, and a repeated 10-fold cross-validation was used to obtain the parameters of RF optimisation using the training data set only. The RF algorithm was run using the caret package in R statistical software, and the best combination of mtry and ntree was obtained based on the lowest OOB error (Table 2.4).

Table 2.4 The combination of mtry and ntree used to train the model for image classification.

Year	mtry	ntree	Best Performance
1986	2	2500	0.08
1993	5	500	0.07
2000	5	500	0.08
2007	4	5500	0.06
2014	2	500	0.09
2021	4	1500	0.07

2.2.3.4 Accuracy Assessment

To evaluate the quality of the thematic LULC maps developed from Landsat images using RF classifier, accuracy measures were calculated for each year using an independent set of data obtained from the reference data.

A confusion matrix was generated for each classified image using the test data set to quantify the classification accuracy of the RF performance based on the kappa index, overall accuracy (OA), producer accuracy (PA) and user accuracy (UA). The kappa index is a measure of how the classification results compare to values assigned by chance. Its values are between 0 and 1. If the kappa coefficient equals 0, it means no agreement between the classified image and the reference image. If the kappa coefficient equals 1, then the classified image and the reference image are identical. Thus, the higher the kappa coefficient is, the more accurate the classification. OA represents the percentages between the total of pixels correctly classified for all classes and the total number of pixels used in the data set. While PA is the number of correctly identified pixels divided by the total number of pixels in the reference image, UA is the number of the correctly identified pixels of a class, divided by the total number of pixels of the class in the classified image.

The flowchart below (Figure 2.2) outlines the steps of the classification method and change detection.

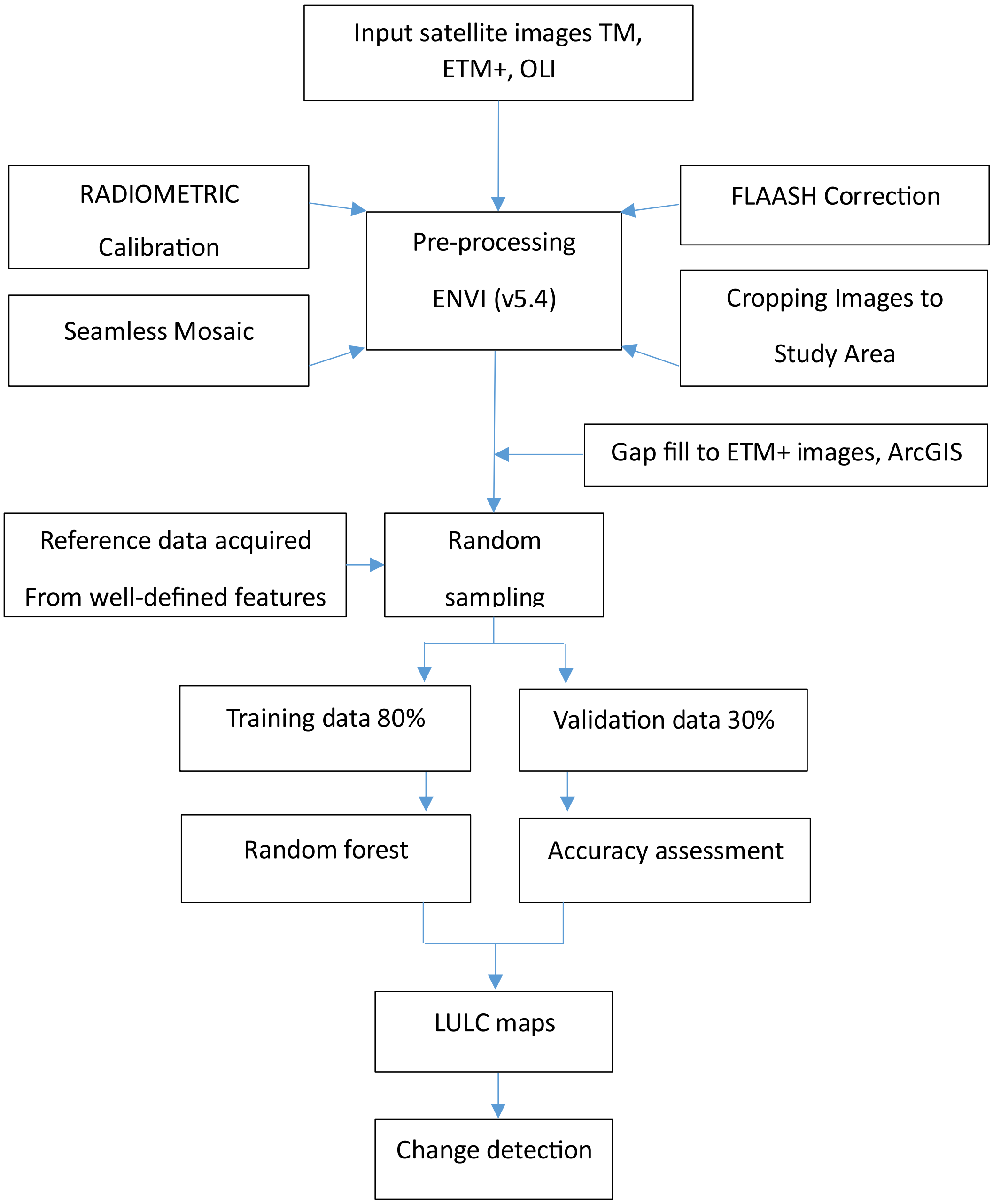


Figure 2.2 The methodology flow chart for the study to outline the classification and change detection steps.

2.3 Results

2.3.1 Spatial Pattern of LULC Classes (1986 to 2021)

Seven LULC classes were extracted from the processed images of the study area. Referring to Figure 3 below, the spatial distribution of the LULC classes from 1986 to 2021 was as follows: the grassland land cover type was generally dominant in the study area throughout the study period.

This was followed by the lands used for crop cultivation activities, with land uses including cleared fields and agricultural land cover types. The grasslands covered between 52.0 per cent in 2007 to 56.9 per cent of the total study area in 2021, followed by cleared fields ranging between 28.6 in 2021 and 41.8 per cent in 1986. Agriculture was 1.41 per cent in 1986 and ranged to 10.0 per cent in 2021. The remaining classes covered small portions of the total area. Mining ranged between 0.14 in 1986 and 0.39 in 2021, settlements varied between 0.87 in 1986 and 1.43 in 2021, while water bodies ranged from 0.65 in 1986 to 2.10 in 2007, and woody vegetation was at 2.66 in 1986 and decreased to 0.29 in 2000.

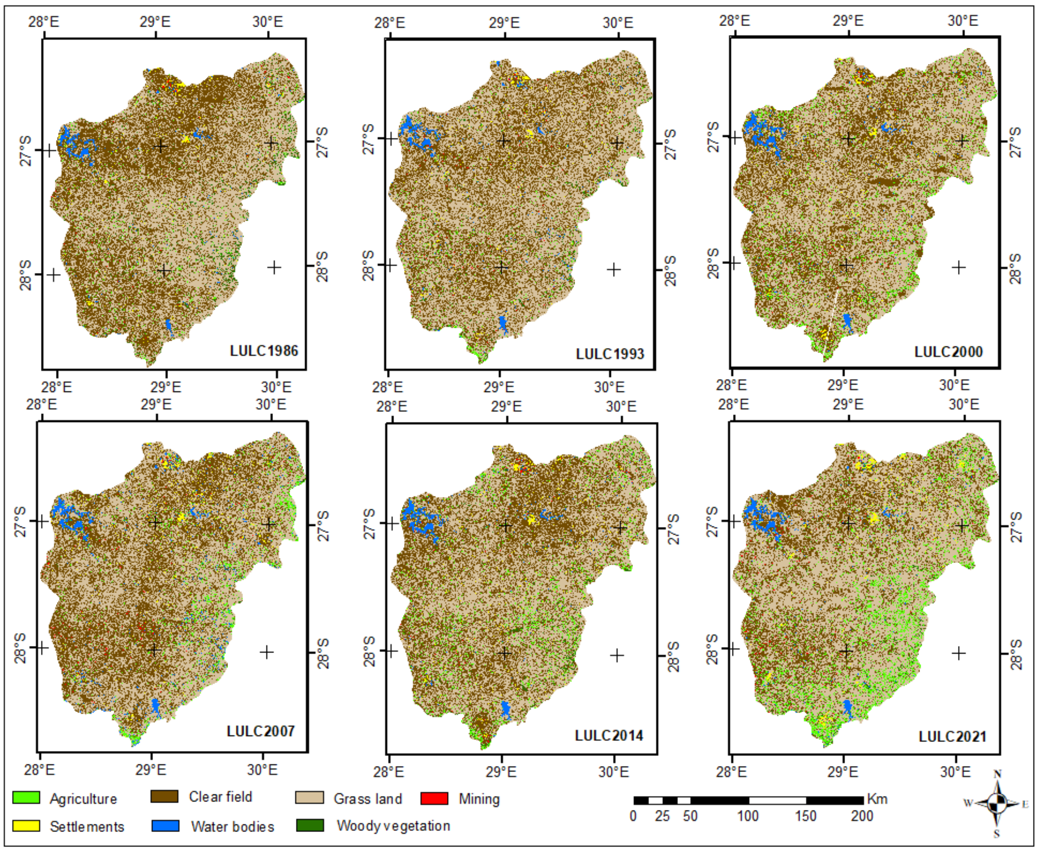


Figure 2.3 LULC patterns in the Vaal Dam Catchment from 1986 to 2021 in 7-year intervals.

The LULC classifications based on six different images were obtained based on the best combination of *mtry* and *ntree* parameters in RF. Table 2.4 above shows the selected combinations of the two input parameters used to train the classifier for each year according to their best performance. Table 2.5 summarises the class areas in hectares LULC patterns for 1986,1993,2000,2007,2014 and 2021.

Table 2.5 The LULC class areas in hectares and their relative proportions, based on the total area.

Year	1986		1993		2000		2007		2014		2021	
LULC Area (Hectare)	**Area**	**%**	**Area**	**%**	**Area**	**%**	**Area**	**%**	**Area**	**%**	**Area**	**%**
Agriculture	54,605.88	1.41	111,290.7	2.88	176,796.5	4.58	133,960.9	3.46	166,057.5	4.29	387,083.0	9.98
Cleared fields	1,618,476	41.75	1,442,757	37.28	1,411,707	36.5	1,566,830	40.48	1,480,856	38.3	1,107,549.3	28.6
Grasslands	2,035,750	52.52	2,186,781	56.51	2,151,995	55.7	2,014,450	52.04	2,058,324	53.2	2,206,309.3	56.9
Mining	5545.08	0.14	10,179	0.26	8687.07	0.22	12,219.12	0.32	7,996.32	0.21	14,996.07	0.39
Settlements	33,539.67	0.87	34,252.31	0.89	44,498.25	1.15	48,234.87	1.25	50,107.77	1.29	55,361.52	1.43
Water bodies	25,198.56	0.65	34,653.87	0.9	58,830.57	1.52	81,455.49	2.10	50,914.08	1.32	55,778.85	1.44
Woody vegetation	103,116.5	2.66	49,653.36	1.28	11,214	0.29	13,881.78	0.36	56,395.35	1.46	52,673.22	1.36
Total	3,876,232	100	3,869,568	100	3,863,728	100	3,871,032	100	3,870,651	100	3,879,751	100

2.3.2 The Accuracy Assessment

The accuracy assessment for RF classifier was performed to evaluate the predicted performance of the trained model using the test data set; this was determined to be 30% of each data set. A confusion matrix for each classified image was then driven. The results of LULC classification indicated that the overall LULC classification accuracies (OA) for the six different date-classified images ranged from 87% in the 2014 classified image to 95% in the 2007 image, with kappa agreement indices ranging between 0.79 in 2014 and 0.92 for 2007. The user accuracy (UA) and producer accuracy (PA) are shown in Table 2.6. The performance of the internal classifier was good, with low OOB errors ranging between 3.75% and 8.98%. All classes had high user and producer accuracies ranging between 80 and 100, except for the settlement classes in 1986 and 2014; the UAs were 79% and 66%, respectively, and, in 2014 and 2021, the PAs were 59% and 63%, respectively. However, the nature of some classes and the different date-stacked images from neighbouring paths made it challenging to separate features in some classes owing to their similarity in their spectral signatures. For example, there was confusion between settlements and the shaded fields (greenhouses), with both containing spectrally similar features. This is also the case between settlement areas surrounded by trees and the woody vegetation class. The kappa values of the six classification results are sufficient for the study area because they satisfy the minimum 85% accuracy.

Table 2.6 The producer accuracies (PAs) and user accuracies (UAs) of the LULC classifications for the Vaal Dam.

	1986		1993		2000		2007		2014		2021	
LULC	**PA%**	**UA%**	**PA%**	**UA%**	**PA%**	**UA%**	**PA%**	**UA%**	**PA%**	**UA%**	**PA%**	**UA%**
Agriculture	89	93	87	93	90	93	93	88	93	100	93	90
Cleared fields	86	91	91	91	96	92	80	89	85	80	94	84
Grassland	93	93	93	88	93	93	100	100	97	88	91	91
Mining	93	80	90	100	96	92	97	97	80	89	88	88
Settlement	87	79	97	97	86	96	97	100	59	66	63	86
Water	96	100	100	100	100	100	100	100	100	100	100	100
Woody vegetation	93	97	100	91	93	96	100	100	100	97	93	97

The RF classifier generally performs well in obtaining the seven determined classes in the Vaal Dam Catchment, as shown in the OA above. These results are the bases for the subsequent analysis of LULC change detections.

2.3.3 LULC Changes (1986 to 2021)

The ratio of each LULC class to the total area within the catchment varies per LULC class, with both increasing and decreasing trends being seen. Table 2.7 shows the pattern of LULC changes during the period studied (1986 to 2021).

The agriculture class showed an increasing pattern, except for the area change between 2000 and 2007. At the same time, cleared fields showed a decreasing pattern except between 2000 and 2007. The sum of the agriculture and cleared field areas indicates that the areas used in crop cultivation activities covered around 41 per cent +/- 3% of the total study area.

The grassland areas show different patterns of increasing and decreasing in the size of their relative area. In general, they did not show a significant change, and grassland consistently remained the dominant class with little variation in extent in the study area.

Table 2.7 Area and percentage changes of the land use land cover (LULC) in the Vaal Dam Catchment for 1986, 1993, 2000, 2007, 2014 and 2021 images.

Year	1986–1993		1993–2000		2000–2007		2007–2014		2014–2021	
LULC Change (in hectares)	**Area**	**%**	**Area**	**%**	**Area**	**%**	**Area**	**%**	**Area**	**%**
Agriculture	56,684.82	+1.46	65,505.8	+1.69	−42,835.6	−1.11	32,096.6	+0.83	221,025.5	+5.71
Cleared Fields	−175,719	−4.53	−31,050	−0.80	155,123	+4.01	−85,974	−2.22	−423307	−10.94
Grassland	151,031	+3.90	−34,786	−0.90	−137,545	−3.56	+43,874	+1.13	197,985.6	+5.11
Mining	4633.92	+0.12	−1491.93	−0.04	3532.05	+0.09	−4222.8	−0.11	6999.75	+0.18
Settlements	712.64	+0.02	10,245.94	+0.26	3736.62	+0.10	+1872.9	+0.05	5253.75	+0.14
Water Bodies	9455.31	+0.24	24,176.7	+0.62	22624.92	+0.59	−30,541.41	−0.79	4864.77	+0.13
Woody Vegetation	−53,463.14	−1.38	−38,439.36	−0.99	2667.78	+0.07	+42,513.57	+1.10	−3722.13	−0.10

The settlement area showed expansion in the Vaal Dam Catchment from 1986 to 2021 Figure 4.

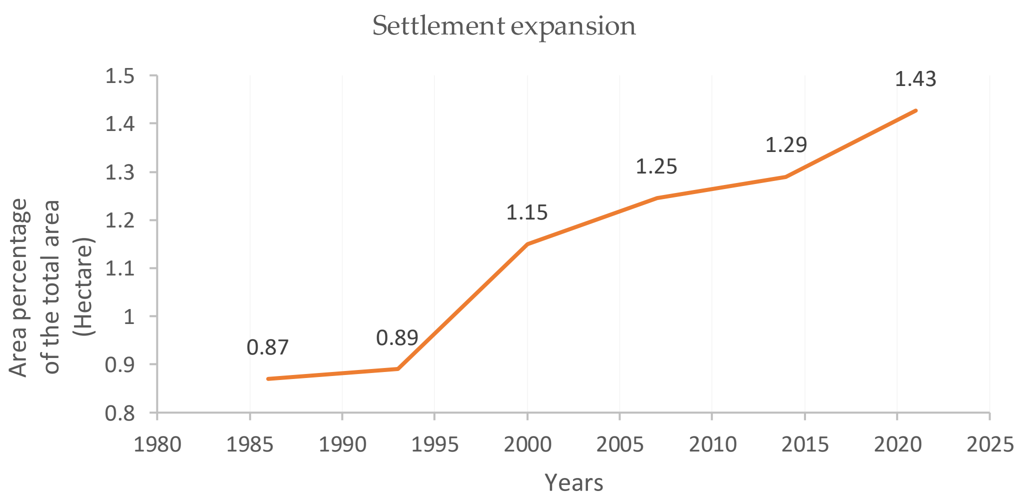


Figure 2.4 Settlement area changes between 1986 and 2021, as a percentage of the total area.

Figure 2.5 shows a sample of settlement expansion in eMbalenhle township from 1986 to 2021. There are many settlements within the catchment, with differing rates of expansion during the study period.

Regarding the remaining classes (mining, water body and woody vegetation), there were different trends in area changes; for example, there was a noticeable increase in the mining area, except for the periods 1993–2000 and 2007–2014. The change detection did not show any trends for the water body class, but it showed that the woody vegetation areas decreased from 1986 to 2000 and expanded from 2000 to 2014 before it decreased again in 2021 (see Figure 2.6 and Table 2.7). The woody vegetation class was particularly noticeable in the mountainous parts on the east and southeastern boundaries of the catchment.

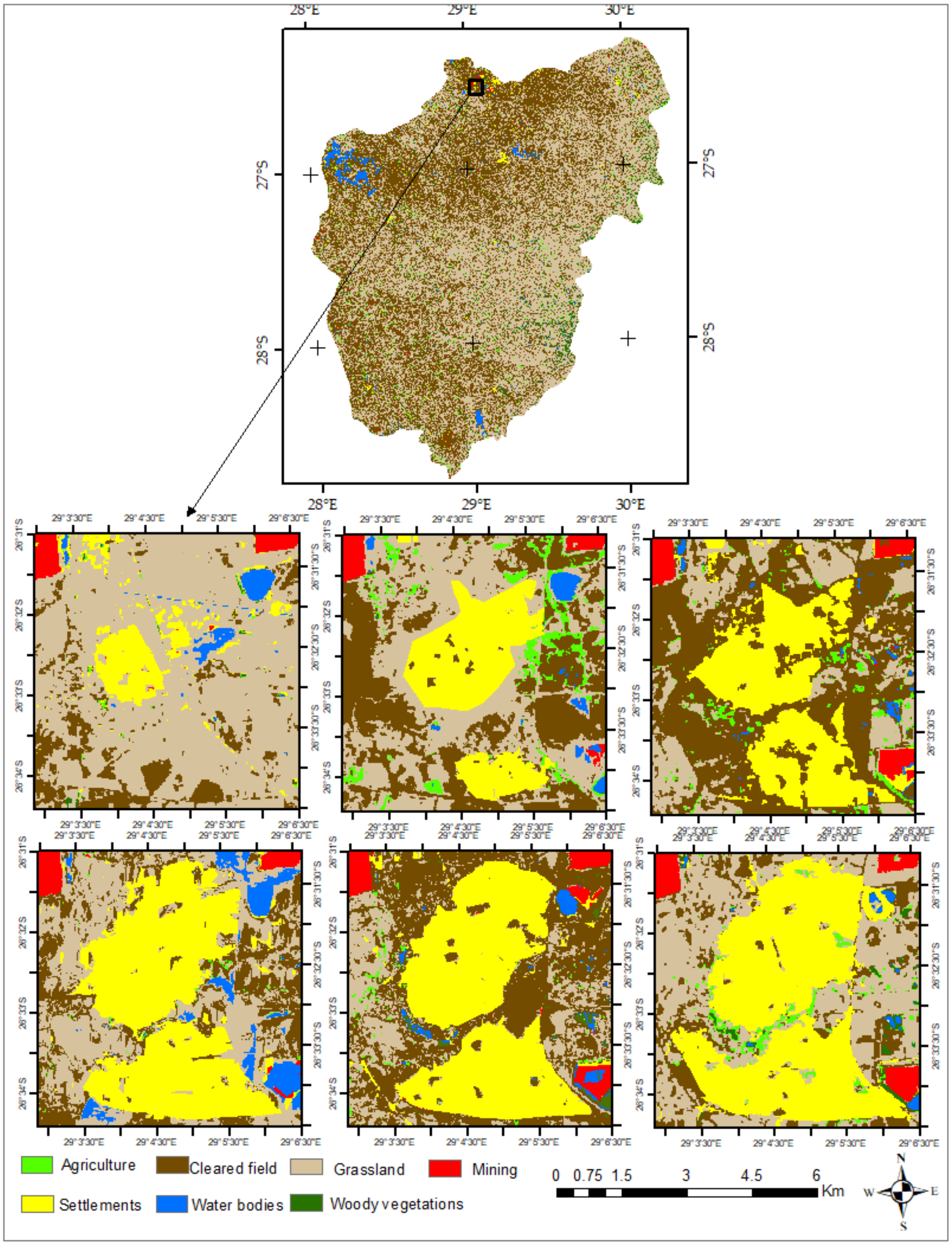


Figure 2.5 A sample of settlement expansion for eMbalenhle township in the northern part of the Vaal Dam Catchment during the period 1986–2021.

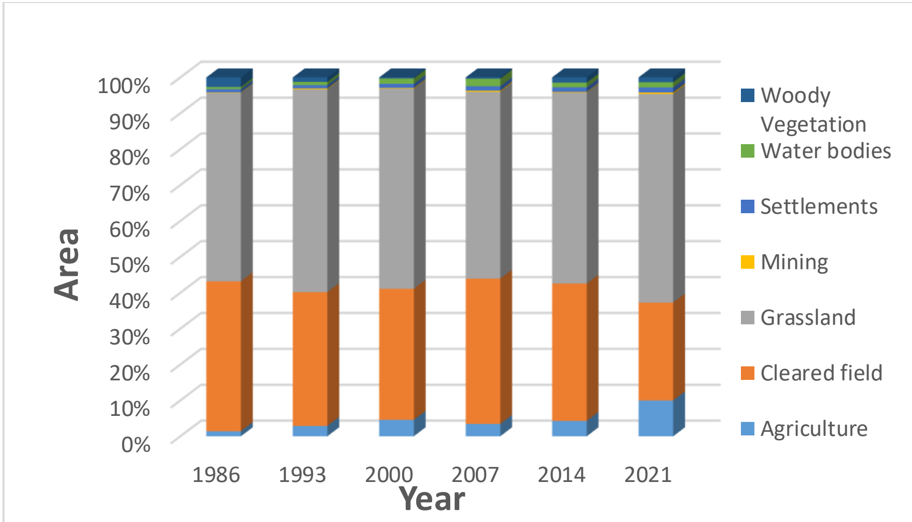


Figure 2.6 Land use land cover (LULC) change for 1986, 1993, 2000, 2007, 2014 and 2021 in the Vaal Dam Catchment.

2.4 Discussion

The LULC classification results showed the domination of the grassland class (including the land used for cattle and sheep grazing farms and pastures) in the study area, followed by the land used for agricultural activities (that is, the summation of agriculture and cleared field classes). The changing pattern of these two categories showed different rates of increase and decrease, but, contrary to what was expected, no significant ongoing increase was noticed for land used for agricultural activities. Although the settlement area class consists of only a relatively small proportion of the total study area, it showed continuous expansion from 1986 to 2021. In contrast, the mining land cover class showed varying patterns of increase and decrease, and varying patterns were also noticed for the remaining classes.

The study successfully achieved its aim, in that it mapped the changes and patterns in LULC and associated dynamics for this strategically important area of South Africa over the 35 years ending in 2021. The most important land cover change within the catchment was the expansion of settlement areas related to the economic activities within the area in recent decades. Many gold and coal mines (Chutter, 1970), as well as synthetic fuel complexes, were constructed and operated in the late 1970s (Gyedu-Ababio and Van Wyk, 2004). This economic growth significantly increased the expansion of towns associated with these developments (see Figure 2.4 and Table

2.7). Figure 2.5 focuses on eMbalenhle township and is a sample of this settlement expansion. The latter township was established in 1978 to increase accommodation needs for the people working for the synthetic fuel manufacturing plant at Sasol. The expansion of many settlement areas within the catchment owing to population growth and their need to grow more food, the construction of houses and industries, etc., have contributed both directly and indirectly to environmental, ecosystem and water quality degradation. In most of the previous studies conducted, water quality and ecosystem degradation have been identified as the most concerning problems within the catchment (Chutter, 1970, Gyedu-Ababio and Van Wyk, 2004, Verheul, 2012, du Plessis et al., 2018).

The relative area used for agricultural activities showed no significant changes, even though it was expected to expand during the study period to meet population growth and their increasing needs for food. In fact, according to Biggs (2002), privately owned farms occupy around 68% of South Africa's land area (Biggs and Scholes, 2002). This indicates that the extent of privately owned farmland remains relatively stable (Biggs and Scholes, 2002), and only limited areas are available for agricultural expansion (Xu et al., 2018). Nevertheless, increasing productivity using modern techniques and soil fertilisers may have been applied to increase the productivity per unit hectare (Biggs and Scholes, 2002).

The results of LULC classification using RF classifier (Dahinden and Ethz, 2011) showed good accuracy in mapping the Vaal Dam Catchment. It has provided much more useful details for the study area than those already published on global land cover patterns (https://lcviewer.vito.be/download,http://maps.elie.ucl.ac.be/CCI/viewer/download.php and http:// www.globallandcover.com/home) accessed on 8 January 2023. The settlement class was clearly detected with all classification schemes; they showed high similarity between them when comparing same-year maps. The remaining classes appeared more generalised in the published global schemes while they were well-detected in this study. Results of this comparison confirm that the method used in this study gives reliable, highly accurate LULC classification results and can be adopted in other regions. However, this methodology has some limitations in some land cover categories, such as the mining class. The detected mining class in this study represents rock and ash dumps and piles from gold and coal mining; only highly reflective materials of those dumps and piles were detected, while many of the mining sites contained small dams and materials for which the colours could not be detected. They may thus have been misclassified. Furthermore, acid mine drainage within or around the study area from abandoned and active mine sites potentially causes severe water quality and environmental issues but this could not necessarily be detected in this study (Clark et al., 2010, Wardlow and Egbert, 2008). With the Landsat data (30-m spatial resolution) methodology applied, it is not easy to detect the narrow surface drainage ditches carrying acid mine effluences. The same limitation can be considered regarding mapping the uncontrolled seepage and flooding of sewage in some settlement areas within the catchment. This situation poses risks to public health, ecosystems and water quality and has received extensive coverage by media channels in recent years (https://www.groundup.org.za/article/sewage-seeps-into-vaal-dam-as-mpumalanga-water-treatment-plants-fail/) accessed 27 September 2022. Another concerning issue is that the water-body areas detected included wastewater storage dams (such as Leewpan near to eMbalenhle) (Gyedu-Ababio and Van Wyk, 2004). Consideration should be given to the possibility of such dams contaminating nearby water resources.

The chosen study area is characterised by the relative homogeneity of its dominant land cover classes; this made it possible to map the area with RF classifier using only a small but representative training data sample (see Section 2.3.3). As shown by Ebrahimy et al. (2021), the RF approach is successful in land cover mapping with limited reference sample data (Ebrahimy et al., 2021).

2.5 Conclusions

This study used Landsat data series to assess LULC changes in the Vaal Dam Catchment area over 7-year intervals for the period 1986–2021. A random forest classifier method consisting of 500 trees was used owing to its advantages over most of the other classifiers in LULC detecting. It was used to classify six mosaicked images covering the study area in seven land cover classes: namely, agriculture, cleared fields, grassland, mining, settlements, water bodies, and woody vegetation. The results of the classification reveal the following percentage composition for the total study area for the period investigated: grasslands covered between 52.0% and 56.9%, cleared fields ranged between 28.6% to 41.8%, agriculture covered 1.4% to 10.0%, mining ranged between 0.14% and 0.39%, settlements covered between 0.87% and 1.43%, water bodies ranged between 0.65% and 2.10% and woody vegetation ranged between 0.29% and 2.66%. The results showed varying patterns of change in LULC (both increases and decreases) observed in most of the classes except for the settlements class. The latter showed a clear increasing trend from 1986 to 2021 resulting from economic development within the area in the last few decades. The OA of the classification of the various LULC types ranged between 91% and 97%. The RF models reported an average OOB error that ranges from 3.75% to 8.98%. The area was dominated by the grassland class for the study period, followed by cultivation land use (which includes the agriculture and cleared field classes). The generated maps provide spatial and temporal patterns of land cover and the changes for the periods studied. An ongoing rapid increase in population growth will have even more significant effects on the region's environment and economic spheres. Therefore, it is of prime importance that these developments be carefully considered in this important catchment as it is one of the primary sources of water for the Vaal Dam, which supplies more than 13 million people in the Gauteng and Mpumalanga provinces. Such studies can support efforts to protect ecosystem functioning and water resources from further deterioration in water quality.

Chapter Three

Time series analysis of water quality factors enhancing harmful algal blooms (HABs): a study integrating in-situ and satellite data, Vaal Dam, South Africa

This chapter is based on **Altayeb A. Obaid,** Elhadi M. Adam, K. Adem Ali and Tamiru A. Abiye. Time Series Analysis of Water Quality Factors Enhancing Harmful Algal Blooms (HABs): A Study Integrating In-Situ and Satellite Data, Vaal Dam, South Africa. *Water* 2024, *16, 764.* Available at: https://doi.org/10.3390/w16050764

Abstract: The Vaal Dam catchment, which is the source of potable water for Gauteng province, is characterized by diverse human activities, and the dam encounters significant nutrient input from multiple sources within its catchment. As a result, there has been a rise in Harmful Algal Blooms (HABs) within the reservoir of the dam. In this study, we employed time series analysis on nutrient data to explore the relationship between HABs, using chlorophyll−a (Chl−a) as a proxy, and nutrient levels. Additionally, Chl−a data extracted from Landsat-8 satellite images was utilized to visualize the spatial distribution of HABs in the reservoir. Our findings revealed that HAB productivity in the Vaal Dam is influenced by the levels of total phosphorus (TP) and organic nitrogen (KJEL_N), which exhibited a positive correlation with chlorophyll-a (Chl−a) concentration. Long-term analysis of Chl−a in-situ data (1986–2022) collected at a specific point within the reservoir showed an average concentration of 11.25 µg/L. However, on certain stochastic dates, Chl−a concentration spiked to very high values, reaching a maximum of 452.8 µg/L, coinciding with elevated records of TP and KJEL_N concentrations on those dates, indicating their effect on productivity levels. The decadal time series and trend analysis demonstrated an increasing trend in Chl−a productivity over the studied period, rising from 4.75 µg/L in the first decade (1990–2000) to 10.51 µg/L in the second decade (2000–2010), and reaching 16.7 µg/L in the last decade (2010–2020). The rising averages of the decadal values were associated with increasing decadal averages of its driving factors, TP from 0.1043 to 0.1096 to 0.1119 mg/L for the three decades, respectively, and KJEL_N from 0.80 mg/L in the first decade to 1.14 mg/L in the last decade. Satellite data analysis during the last decade revealed that the spatial dynamics of HABs are influenced by the dam's geometry and the levels of discharge from its two feeding rivers, with higher concentrations observed in meandering areas of the reservoir and within zones of restricted water circulation.

3.1 Introduction

Globally, freshwater resources have been put under increasing pressure due to the rapid increase in the world population and their improvement of living standards (Makarigakis and Jimenez-Cisneros, 2019). It is indicated that an estimation of more than 83% of land surfaces that are surrounding freshwater systems have been significantly influenced by the footprint of human beings as a response to anthropogenic activities (Makarigakis and Jimenez-Cisneros, 2019, Sanderson et al., 2002). Contaminants from ineffective waste management, pesticides and fertilizers from agricultural areas, pollution from urban, industrial, and domestic wastewater can often be released into ground and surface water (Taheri Tizro et al., 2014), and even the landfills or slag heap disposals may release pollutants seeping into nearby water resources (Oberholster and Ashton, 2008, Oberholster et al., 2008). The inland freshwater bodies are more vulnerable to such problems of pollution and contamination (Ho and Goethals, 2019) resulting in the decline of quality and availability. Among the loaded pollutants, some serve as nutrients for the enhancing growth of algae. Such nutrients can be loaded from point or non-point sources, point sources, for example, wastewater treatment plants can easily be recognized and given more attention (Adu and Kumarasamy, 2018, Carey and Migliaccio, 2009). Non-point sources including urban areas, cultivated lands, natural forests, and pastures are more difficult to recognize and attention should be given to them according to the amounts of nutrients they can potentially release. For example, it was found that the estimated loads of nitrogen, phosphorus, and suspended sediments from urban areas and cultivated land are 10 to 100 times greater than idle and forested lands in the Great Lakes catchments of USA and Canada (Sonzogni et al., 1980).

Eutrophication is a serious freshwater problem caused by the excessive growth of harmful algal blooms (HABs). In recent years, the factors enhancing such blooms have become interesting research subjects. Many studies have been conducted in this field, and most of them were related to the HABs to the loading of nutrients such as nitrogen and phosphorus into the water bodies (Chaffin et al., 2021, Glibert et al., 2014).

In Africa, societies are highly dependent on their natural inland freshwater resources creating more pressure on them resulting in significant changes in their water quality (Robarts and Zohary, 2018). South Africa has over 4000 freshwater storage dams, among which around 700 are public dams controlled and managed by the governments used for domestic, irrigation, and industrial water supply (Hara and Backeberg, 2014), they are underpinning the economic and social development of the country. Despite such a huge number of dams, South Africa is a water-stressed and scarce country facing critical challenges due to poor water management practices, inadequate infrastructure, and a relentless surge in water demands (Ewerts, 2010). Owing to the fact that most South African dams are located downstream of metropolitan and urban areas, they have become more enriched with nutrients (Oberholster et al., 2008, Ali et al., 2022a). The climatic conditions of South Africa associated with nutrient loads resulted in extensive and widespread eutrophication and cyanobacterial blooms in the inland freshwater bodies (Matthews and Bernard, 2015, Mudaly and Van der Laan, 2020). This will continue to increase the cost of using these valuable resources. Nitrogen and phosphorus are considered the leading factors in accelerating the lakes and reservoirs' eutrophication (Beaulac and Reckhow, 1982). The Water Act of 1956, Section 21(1) (a) and the Department of Water Affairs in 1980 set exceptional standards of the phosphate concentration for effluent discharged to some mentioned rivers or their tributaries, it promulgated that the phosphate concentration should not be higher than 1 mg/L (Ewerts, 2010, Walmsley, 2002). An investigation was conducted by CSIR and the Department of Water Affairs from 1985 to 1988 to predict the impact of the phosphorus standard (1 mg/L). Based on their findings, the eutrophication control aim was set to maintain the mean concentration of chlorophyll-a within receiving water bodies at the level that no conditions of severe nuisance would occur more than 20% of the time. To achieve this aim, it is required to keep the phosphorus concentration level in the reservoirs water less than 0.14 mg/L (Van Ginkel, 2011).

Vaal Dam is one of multiple water bodies affected by eutrophication and cyanobacterial blooms in South Africa. It was constructed by the Department of Irrigation of the national government and functionally completed in October 1936 at 20 km downstream of the confluence of the Vaal and Wilge rivers (Haarhoff and Tempelhoff, 2007). It is the second biggest dam by surface area in South Africa (about 320 km^2). The dam has undergone several rises to increase its storage capacity which ended up with a total capacity of approximately 2.603×10^6 m^3 (Haarhoff and Tempelhoff, 2007). The dam catchment area is approximately 38,000 km2 holding various human activities including major agricultural activities (crop cultivation and cattle grazing), mining, and some industrial activities (Obaid et al., 2023, Du Plessis, 2017) as well as many formal and informal settlements. In many areas within the Vaal Dam catchment, mine dewatering and urban/industrial treated effluent discharges find their way to the streams causing serious water quality issues (Obaid et al., 2023). Such activities have direct or indirect effects on the dam water quality in terms of nutrient loading which enhances the growth of HABs. A study in 2013 stated that the Vaal Dam Reservoir was classified as a mesotrophic water body according to the South African Department of Water Affairs Classification System, the mean concentration of chlorophyll-a was 14.8 µg/L, the mean total phosphorus concentration was 0.077 mg/L, and the time percentage where chlorophyll-a exceeds 30 µg/L was found was 17% (Swanepoel et al., 2016).

HABs have major effects on water quality and their aquatic system function; therefore, monitoring of their distribution in space and time is very important for water resources managers to address the issues related to it. Moreover, it is very important to address the question of which factors enhance and control the HABs in the Vaal Dam Reservoir. Historically, many studies were conducted to assess different methods to detect HABs and cyanobacteria in the Vaal Dam Reservoir, for example, an attempt was made to help the managers of the drinking water treatment facility with advanced prediction of Microcystis sp. concentration in the Vaal Dam water, by building a model using physical, chemical, and biological water quality records between 2000 and 2012. The model showed a promising result in estimating Microcystis sp. in 7 days in advance (Swanepoel et al., 2016). The composition of the phytoplankton in the reservoir was examined over the course of a year, it was found that Anabaena and Microcystis are the most abundant cyanobacteria that dominate the phytoplankton on the dam reservoir during summertime; during February and March, the study reviled that they are not necessary dominating throughout the year (Chinyama et al., 2016). Few studies were conducted in the Vaal Dam using remote sensing data (Landsat 8 OLI and Sentinel 2 MSI), the two sensors have relatively high spatial resolution and a capability to distinguish very small differences in the optics of the water column as a function of water quality parameters such as Chl−a. Sakuno, Y et al., (2018) have obtained a strong correlation between the satellite estimation using unified red-to-near-infrared two-band and three-band models, and the in-situ measurements [see ref. (Sakuno et al., 2018b)]. Another study tested the red-to-near-infrared (NIR-red) bands by means of stepwise logistic regression (SLR) to classify Chl−a concentrations using Landsat 8 data for 2014–2016, the SLR gave overall accuracy ranged between 65 to 80% [see ref. (Malahlela et al., 2018)]. A comparison of Landsat 8 and Sentinel 2 was conducted in mapping water quality at Vaal Dam, Sentinel 2 generated better results than Landsat 8 for estimating both chlorophyll-a and turbidity with R^2 = 0.86 compared to 0.68 for Chl−a, and 0.59 and 0.50 for turbidity (Sakuno et al., 2018b, Malahlela et al., 2018, Bande et al., 2018). The latest remote sensing-based article published on the Vaal Dam was (Obaid et al., 2021) applied to high-resolution sensors that successfully used both the blue-green and red-infrared OC algorithms in estimating Chl‒a concentrations (Ali et al., 2022a, Obaid et al., 2021).

Blue-green algorithms are typically utilized to estimate chlorophyll-a concentration in Case I waters, where water quality is predominantly influenced by phytoplankton (Gyedu-Ababio and Van Wyk, 2004, du Plessis et al., 2015). However, in Case II waters, where the optical signature is influenced by various water quality constituents, employing blue–green models can result in inaccuracies when estimating chlorophyll-a concentration. The successful use of the blue-green OC algorithm in the Vaal Dam, a Case II water body, may be attributed to the strong signal of HABs arising from the high biomass concentration in the reservoir water column. While previous studies have focused on developing methods to detect HABs and cyanobacterial blooms in the reservoir, none have attempted to uncover the factors contributing to such blooms in terms of nutrient loading and environmental factors.

This study aims to explore the correlation between HABs, nutrients, and environmental factors during peak blooms over recent decades. Historical water quality records from the dam reservoir will be analysed to investigate this relationship by comparing nutrient levels with measured chlorophyll-a (Chl‒a) concentrations, thereby identifying which nutrients contribute to algal bloom enhancement. Additionally, the study seeks to assess the spatial distribution of HABs using Landsat-8 satellite images captured in late summer (April) between 2013 and 2023. This analysis will facilitate the identification of the frequently and most affected areas of the dam reservoir. Leveraging remote sensing techniques provides a synoptic view of HABs' spatial and temporal distribution, enhancing our understanding of their dynamic variations. This approach has the potential to enable effective monitoring of HABs, utilizing high-resolution data for more precise studies.

3.2 The Study Area

The Vaal Dam holds water used to supply potable water to the Gauteng metropolitans and its surrounding areas. The Vaal Dam Catchment extends within Free State, Mpumalanga, and Gauteng provinces and drains by the Vaal and Wilge river systems, (Figure 3.1). The main rivers consist of many tributaries draining different areas in terms of land use land cover types and human activities.

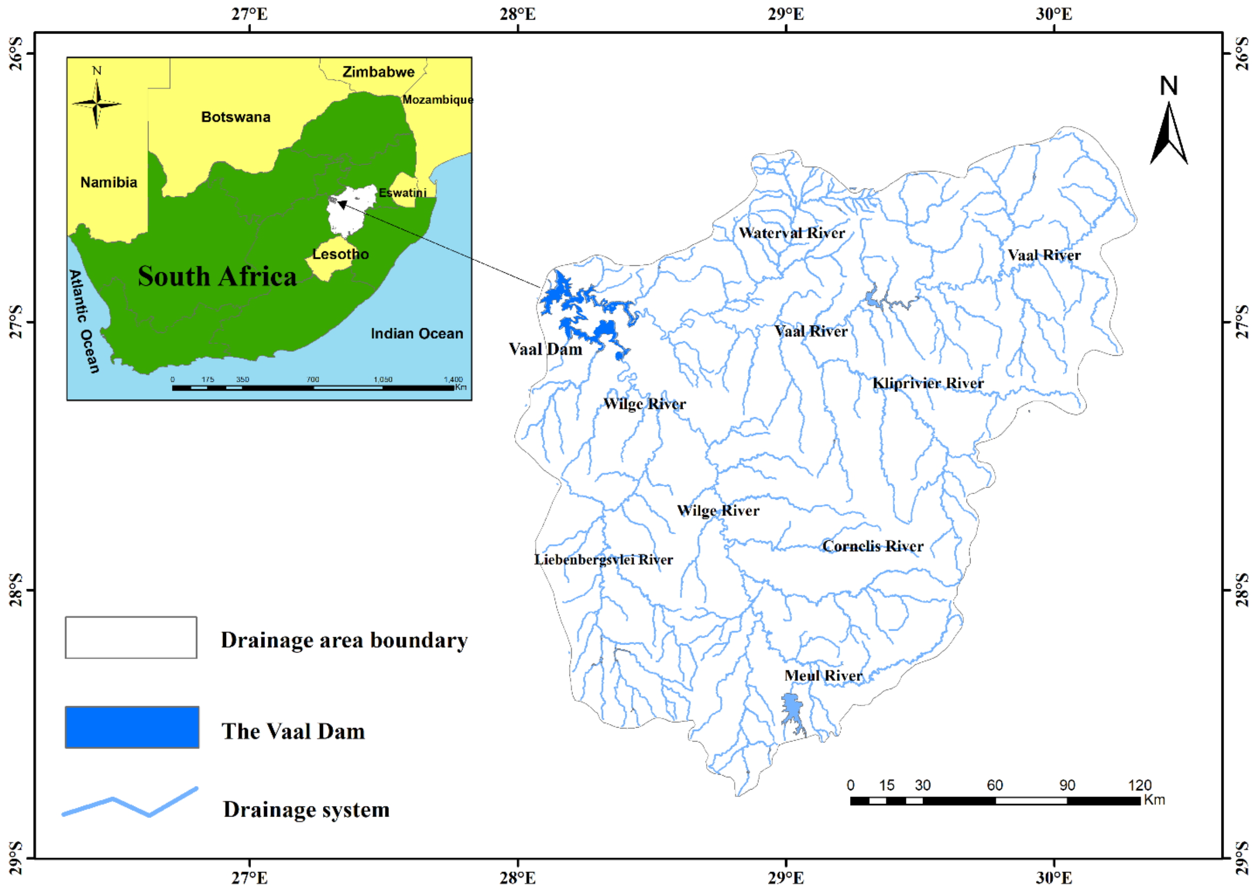


Figure 3.1 The location map of the Vaal Dam.

The Waterval River contributes approximately 111 × 106 m3 of water annually to the Vaal River. It drains very active areas holding intense human activities such as agriculture, industry, mining, and urban and rural settlements, mainly in the upper reach of the Waterval River, these activities have been shown to be responsible for the deterioration of the water quality and ecological integrity of the river system (Gyedu-Ababio and Van Wyk, 2004). Another active area with mining and industrial activities within the Vaal River site is the Grootdraai Dam catchment, it is found that the greatest area of concern was the region's closeness to the downstream of the urban, industrial, mining, and cultivated lands cover (du Plessis et al., 2015). These two active areas in terms of land use activities put the Vaal River under more concern about its water quality issues.

The Wilge River system drains areas dominated by agriculture and grasslands, it contributes a great deal of water to the dam from its catchment and from the Lesotho Highland Water Project (Matete and Hassan, 2006). The Tugela–Vaal Water Project also contributes a good portion of water to the Vaal Dam via the Wilge and Nuwejaarspruit Rivers (Haarhoff and Tempelhoff, 2007).

3.3 Materials and Methods

3.3.1 The water quality data

Historical water quality was obtained from the Department of Water and Sanitation (http://www.dwa.gov.za/iwqs/wms/data/WMA08_reg_WMS_nobor.htm), accessed on 2 January 2023. The data contains chemical and physicochemical water quality parameters from the Vaal Dam Reservoir. Daily water quality data include Chl‒a, Total Phosphorus (TP), Dissolved Oxygen (DO), KJEL Nitrogen (KJEL_N), Ammonia Nitrogen (NH_4_N), Nitrate and Nitrite Nitrogen (NO_3NO_2_N), and water Temperature (Temp) have been downloaded. Some gaps of data missing data existed in some chosen parameters, the records were poor with a big gap noticed between 2000 and 2010 for all-targeted water quality parameters except Chl‒a and TP. A good record was only found for the period between 2010 to 2019. Thus, the distribution, time series, and decadal trend plots as well as regressions between the variables were generated using the available data only, no gap-fill approach was applied. The continuous DO records are only available from 2010 to 2019.

3.3.2 The satellite data

Cloud-free Landsat-8 OLI, Level 1 data were downloaded from the United States Geological Survey (USGS) Earth Explorer website (https://earthexplorer.usgs.gov/), ac- cessed on 1 March 2023. One image was acquired for each summer season in a year, and a total of eleven images were downloaded covering the period between 2013 and 2023 to understand the spatial dynamics of HABs in the reservoir, specifically, the images acquired in April if available otherwise any summer month. The images were processed using a level-2 generator (l2gen) of the NASA SeaDAS software version 8.3, which is the standard ocean colour processing software of NASA. To perform the atmospheric correction, l2gen uses an iterative process to estimate the aerosol radiance, it uses the near-infrared and short-wave infrared to overcome the limitation of dark pixel assumption, and it has proven to give a better result in turbid waters (Ilori et al., 2019). Landsat 8 OLI bands 865 and 2201 were used in this study for atmospheric correction and then normalized water-leaving radiance (nLw) and remote sensing reflectance (R_{rs}) have been retrieved in l2gen. All necessary masks have been applied using SeaDAS v.8.3 standard masking to retrieve only the open water body images.

3.3.3 Data analysis

Historical in-situ water quality data has been analysed using statistical methods in R v4.3. The long-term average and the standard deviation (SD) for any single variable have been calculated from the whole data. The SD values have been corrected using a correction factor to obtain the known SD in textbooks. Multiplot types including frequency scatter plots, time series plots, and decadal trend plots have been generated. The variable data have been indexed and matched using the index and match function in an Excel sheet. Then, the above-mentioned plots have been generated. The erratic high values have been excluded from the frequency distribution plots to explore the most frequent values while all data have been included in long-term time series (1986 to 2022), decadal time series, and decadal trend plots. The analysis was performed to understand the changes in the studied parameters. The correlation between these variables has also been evaluated, scatter plots of Chl−a against temperature and the potential nutrients, and between the variables that are well-correlated with Chl−a were plotted to reveal the values that show excessive HAB conditions were chosen for detailed analysis alongside their concurrent nutrient concentrations. Since the data were collected over only one point within this large dam reservoir, more understanding of the spatial distribution of HABs was conducted using satellite remote sensing data to explore the spatial change of productivity levels. A simple two-band ratio of blue-green based ocean colour algorithm was applied (O'Reilly and Werdell, 2019) for Vaal Dam water modelling productivity. Level 2 Landsat-8 products obtained from level−2 generator (l2gen) of NASA SeaDAS v8.3 were used for chlorophyll-a retrieval. The ratio of ($R_{rs}560/R_{rs}443$) from our previous work (Obaid et al., 2021) is used in this study, it was successfully used to retrieve Chl−a from Landsat-8 data in the Vaal Dam.

3.4 Results

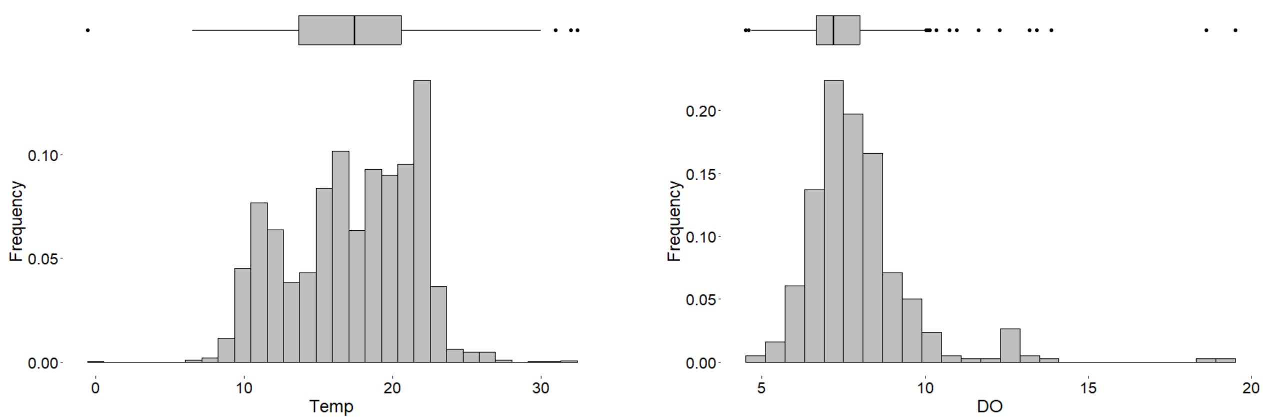
3.4.1 Time series of targeted parameters

Time series of the Chl−a, TP, DO, KJEL_N, NH_4_N, NO_3NO_2_N, and temperature have been plotted from 1986 to 2022. The analysis of the data shows that the concentration levels for the targeted water quality parameters in the Vaal Dam ranged from very low values to extremely high recorded values on some stochastic dates. Table 3.1 summarizes the availability of the data coverage periods and their minimum and maximum recorded values as well as their averages which are calculated from the whole available data. The averages tend to lower values which explains that the extremely high values do not always occur.

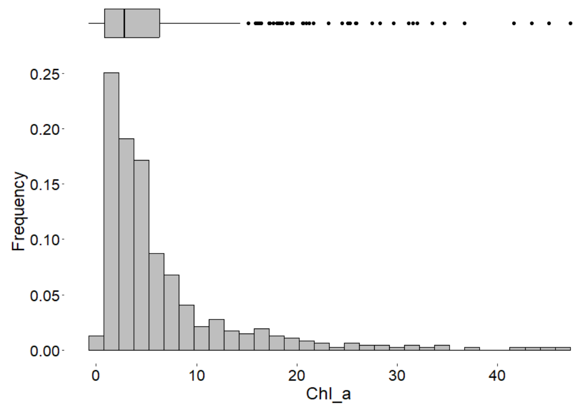
Table 3.1 Summary of the availability periods and the characteristics of the targeted WQ parameters.

Parameter	**Data availability**	**Minimum Value**	**Maximum Value**	**Average**	**Standard Deviation**
Chl–a	1986 – 2022	0.5 µg/L	452.8 µg/L	11.25 µg/L	34.72
TP	1986 – 2022	0.01 mg/L	1.4 mg/L	0.113 mg/L	0.121
DO	2010 – 2018	4.98 mg/L	19.31 mg/L	7.9 mg/L	1.615
KJEL_N	1986 – 2018	0.05 mg/L	18.058 mg/L	0.94 mg/L	1.352
NH_4_N	1977 – 2018	0.02 mg/L	0.28 mg/L	0.046 mg/L	0.029
NO_3NO_2_N	1968 – 2018	0.02 mg/L	0.921 mg/L	0.27 mg/L	0.201
Temperature	1968 – 2018			18°C	4.160

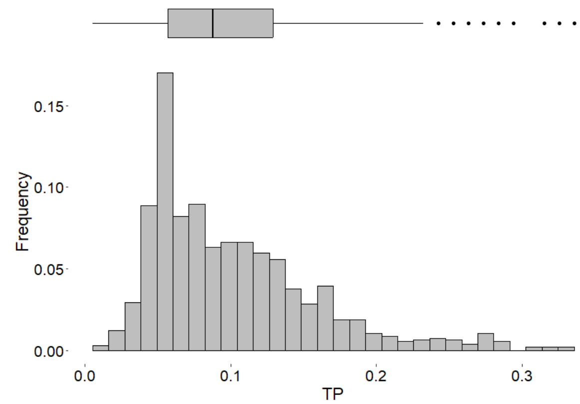
The frequency distribution plots (Fig 3.2) show that most of the Chl−a data are centred near the low-range values, below its average. A similar situation for HN_4_N and TP which most of the values fall below 0.1mg/L and 0.15mg/L respectively. Temperature and NO_3NO_2_N values are distributed stochastically with the most values falling below 20°C and 0.5mg/L, respectively. KJEL_N and DO data values have relatively followed the normal distribution and most of the data cantred around their averages.



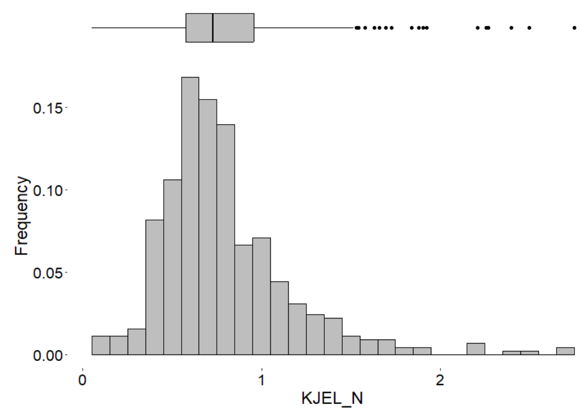
(**a**) (**b**)



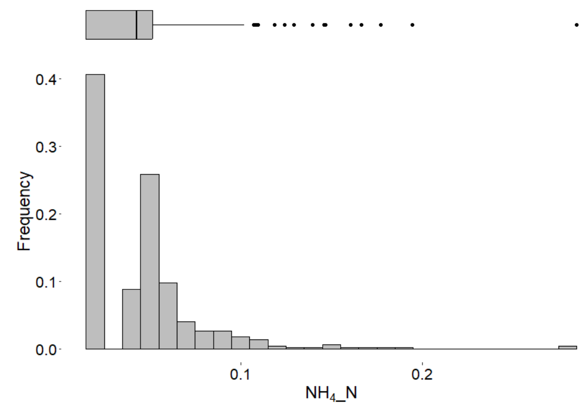
(c)



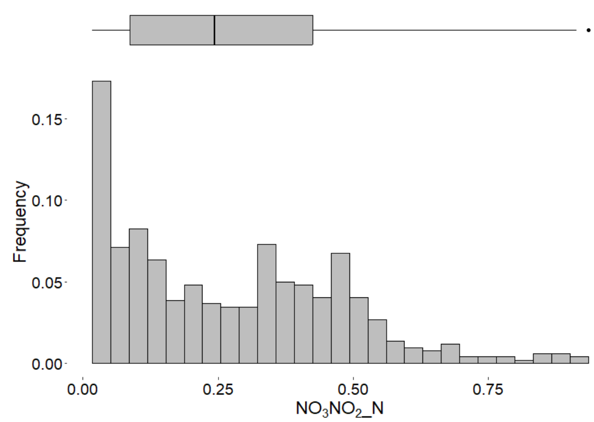
(d)



(e)



(f)



(g)

Figure 3.2 Frequency distribution plots of targeted water quality parameters in the Vaal Dam: (a) Temperature; (b) Dissolved oxygen; (c) Chlorophyll-a; (d) Total phosphorus; (e) Organic nitrogen; (f) Ammonia; and (g) Nitrate and Nitrite. The extreme values falling outside the whiskers of the boxplots are represented by black dots.

The time series of the temperature showed seasonal variations which have a strong influence on DO levels, the DO concentrations are decreasing during summer months when the temperature and Chl–a are relatively high (Figures 3.3 and 3.4c). The Chl–a concentration time series recorded many peak values (Figure 3.3) on some erratic dates. The time series plots of the TP and KJEL_N also show some recorded high values corresponding to the same dates of Chl–a high values, reaching up to 1.15 mg/L and 18 mg/L for TP and KJEL_N, respectively, mainly during the period between 2015 and 2017 (Figure 3.4c) and between 2004 and 2008 for phosphorus where it jumped up to 0.75 mg/L (Figure 3.4b). DO, NO_3NO_2_N and NH_4_N did not show a clear correlation with the Chl–a time series. But in general, the time series of NO_3NO_2_N showed low values during the high levels of Chl-a concentrations, and it appears to show a periodic cycle (1.5-year) during the last decade (Figures 3.3 & 3.4c).

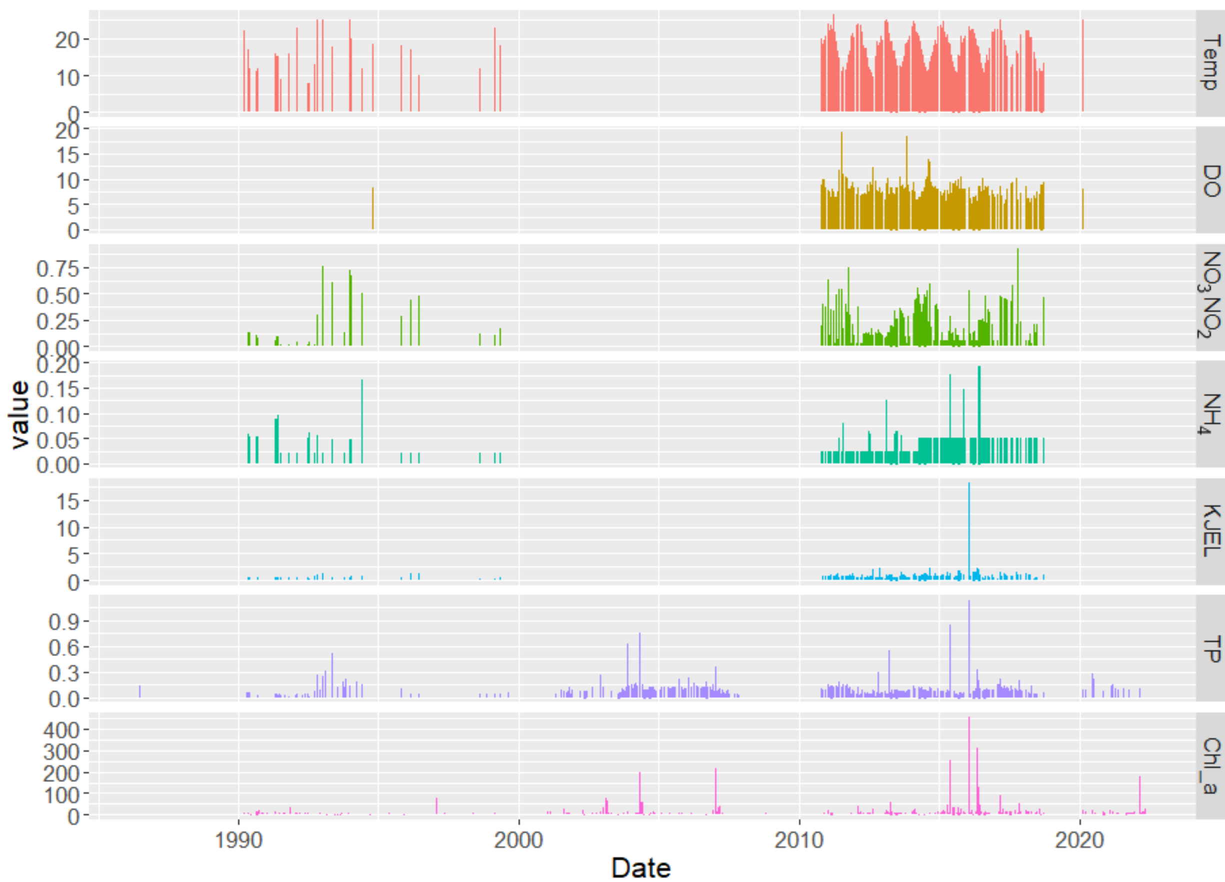
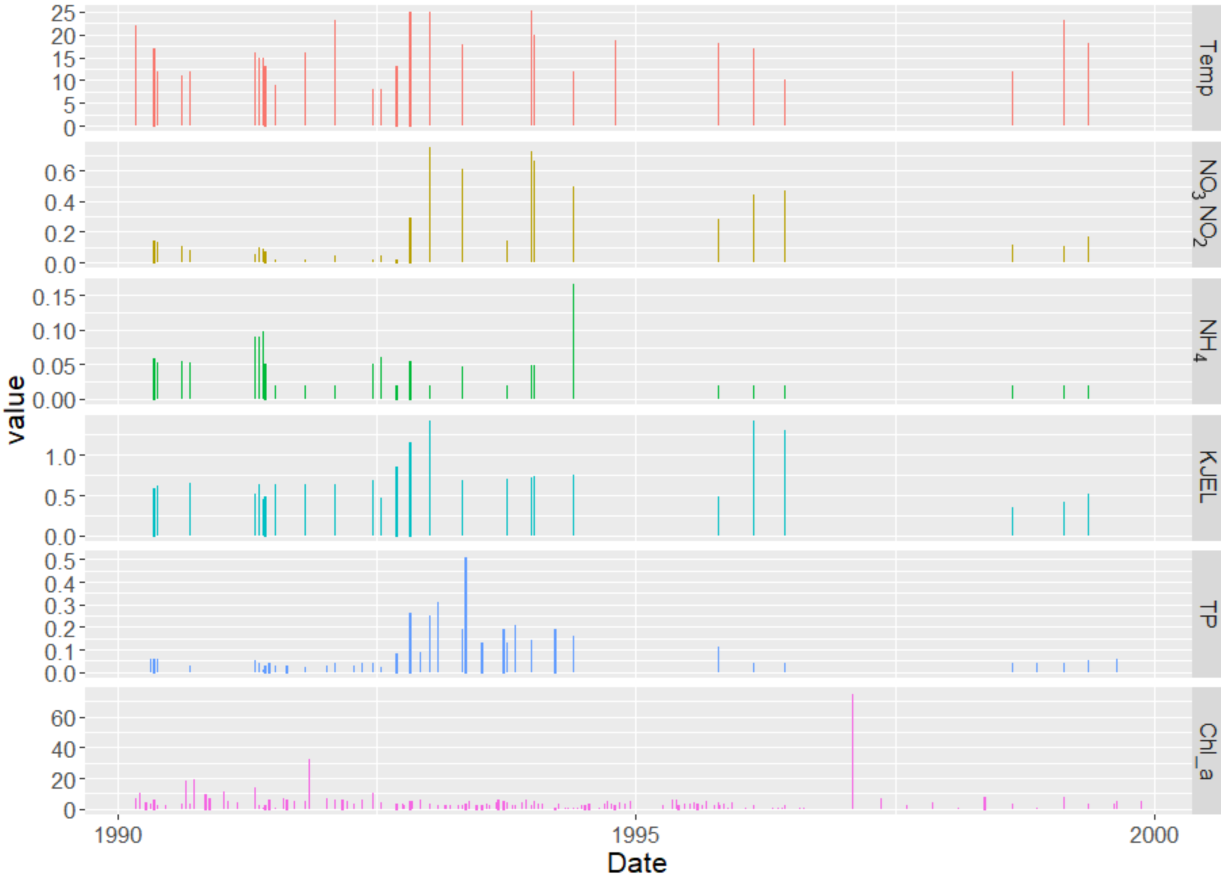


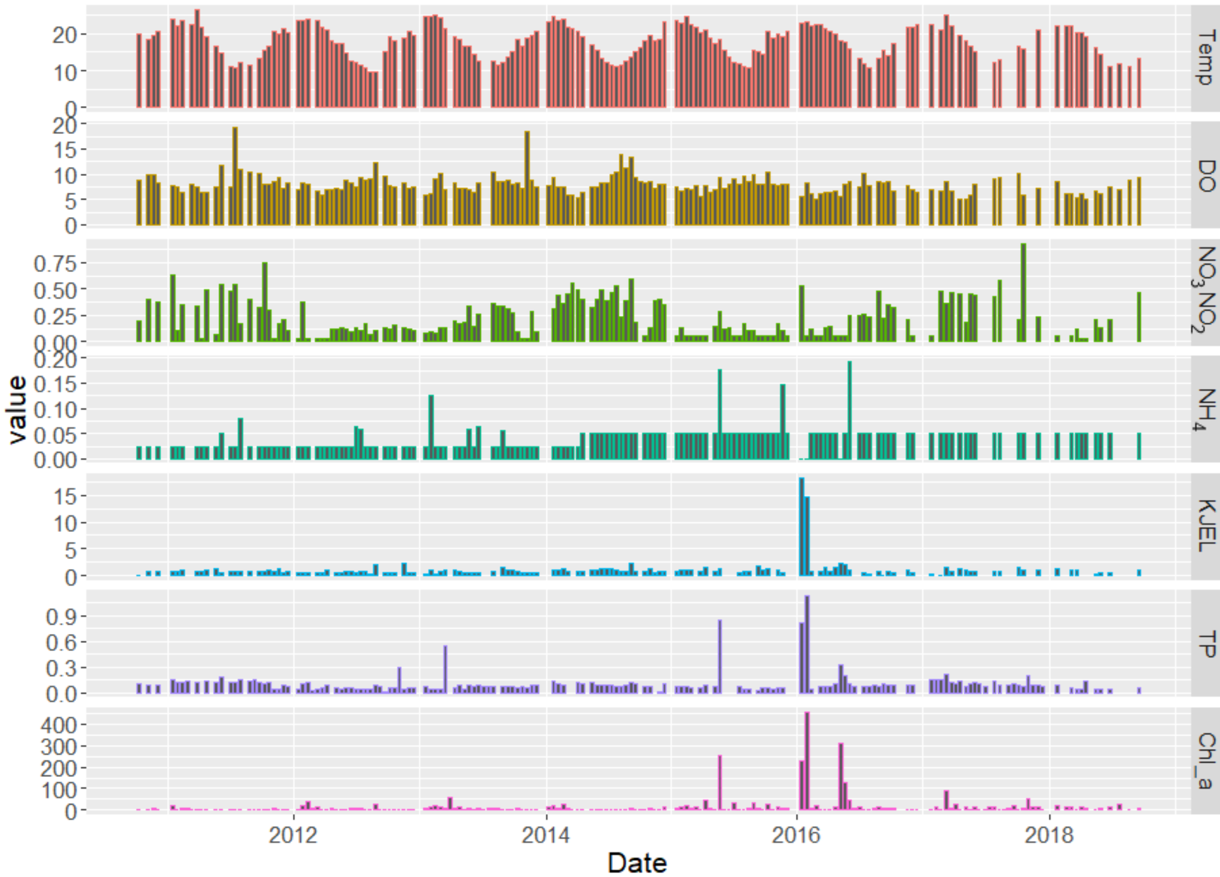
Figure 3.3 Time series of the targeted water quality parameters in the Vaal Dam.


value
Temp
NO_3NO_2
NH_4
KJEL
TP
Chl_a
Date
1990
1995
2000

(a)


value
TP
Chl_a
Date
2000
2002
2004
2006
2008

(b)



(c)

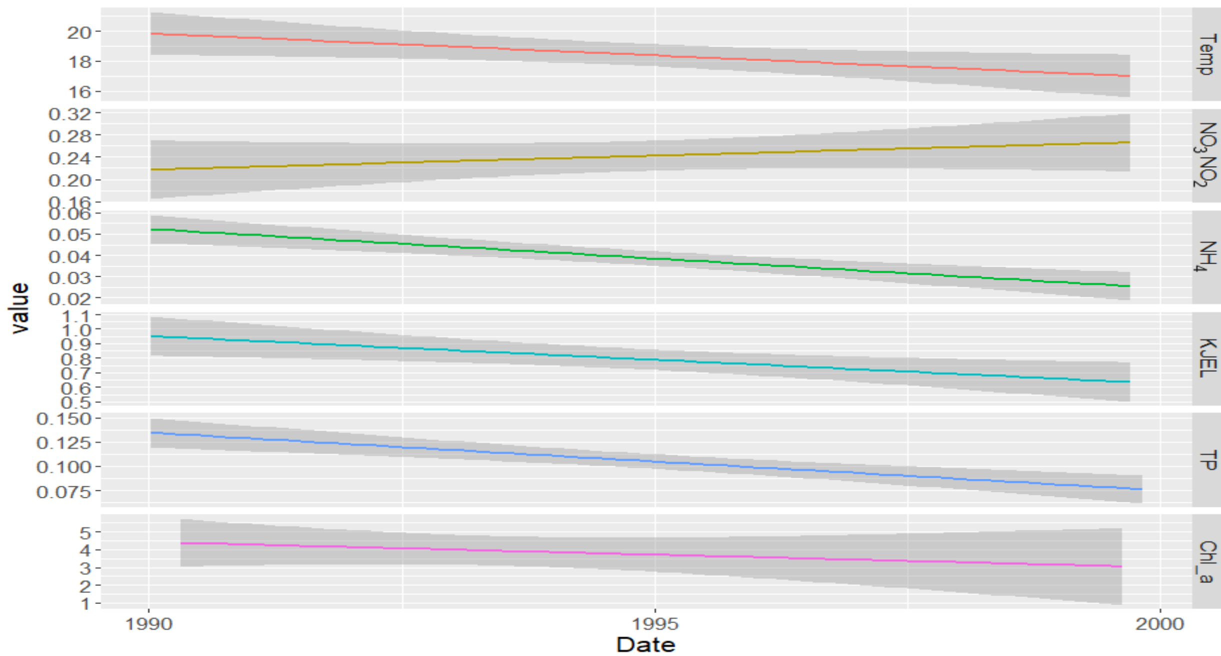
Figure 3.4 Decadal time series of the targeted water quality parameters in the Vaal Dam (**a**) first decade 1990 – 2000, (**b**) second decade 2000 – 2010 and (**c**) third decade 210 – 2020.

Figure 3.4 shows the decadal time series for the targeted parameters. In the period between 1990 to 2000, there were very poor records (Figure 3.4a) which show relatively low values of chlorophyll-a concentrations except on specific dates. TP and KJEL_N show some high values between 1993 and 1995 but there are no correlated peaks with Chl‒a during this period due to the poorness of the records and no matchup in their measuring dates. Between 2000 and 2010 only records for Chl‒a and TP were available, some high values were recorded on specific dates (Figure 3.4b). The best records were found between 2010 and 2018 which covered all targeted parameters (Figure 3.4c), extremely high values were noticed for some of the targeted parameters during this decade. There are increasing decadal averages of Chl‒a, TP, and KJEL_N, while there were decreasing decadal averages of NO_3NO_2_N and no significant change in the NH_4_N decadal average (Table 3.2).

Table 3.2 The decadal averages of the studded water quality parameters.

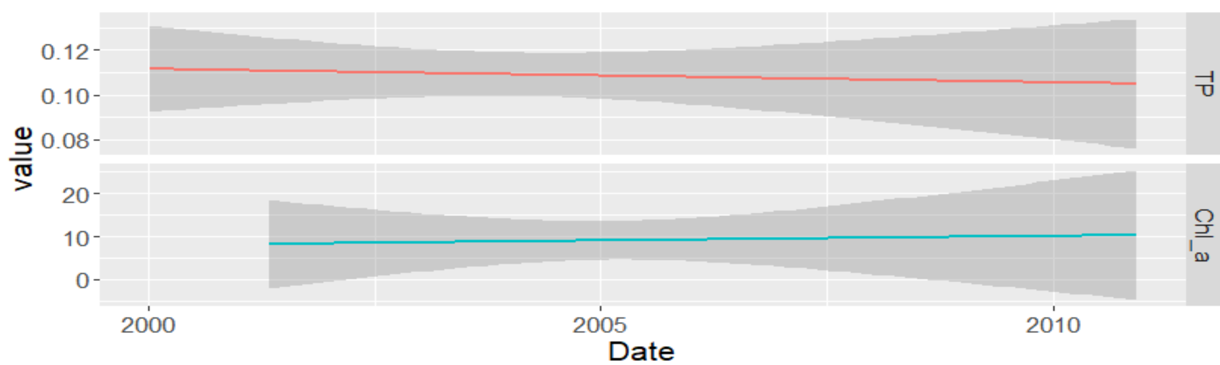
Decades **WQ Parameter**	**1st decade (1990-2000)**	**2nd decade (2000-2010)**	**3rd decade (2010-2020)**
Chl-a (µg/l)	4.75	10.51	16.7
TP (mg/l)	0.1043	0.1096	0.1119
KJEL_N (mg/l)	0.8	-	1.14
NO_3NO_2_N (mg/l)	0.246	-	0.225
NH_4_N (mg/l)	0.04	-	0.043
DO (mg/l)	-	-	7.93
Temp. (˚C)	17.9	-	18

From the decadal trend analysis (Figure 3.5a–c), Chl−a shows a decreasing trend during the first decade from 1990 to 2000 with low trend average values ranging from ~5.5 to ~4 µg/L and it remains constant during the second decade from 2000 to 2010 with higher mean trend average value, then it showed an upward trend during the last decade from 2010 to 2020, which raised from ~7 to 30 µg/L. TP average trend levels showed the same pattern as Chl−a, it showed a downward trend from around 0.137 mg/L to around 0.075 mg/L during the first decade and showed a nearly constant trend during the second decade with a trend average value of ~0.11 mg/L before it started showing an up- ward trend during the last decade, where the trend average value increased from around 0.11 to 0.125 mg/L. KJEL_N and NH_4_N followed the same pattern as Chl−a and TP, they showed a downward trend during the first decade from ~0.95 mg/L to ~0.65 mg/L and from around 0.053 mg/L to 0.025 mg/L, respectively, and they showed an upward trend during the last decade, KJEL_N from 0.75 to 1.75 mg/L and NH_4_N from ~0.028 to 0.06 mg/L. DO showed a decreasing trend during the last decade which decreased from ~9 to ~7 mg/L. The NO_3NO_2_N concentration showed an upward average trend during the first decade from ~0.22 to ~0.26 mg/L while remaining nearly constant with an average trend value of around 0.225 mg/L in the last decade. The average trend of temperature shows a slight decrease from ~20 to ~17 ◦C from 1990 to 2000 and a nearly constant average trend value of ~18 ◦C from 2010 to 2020.



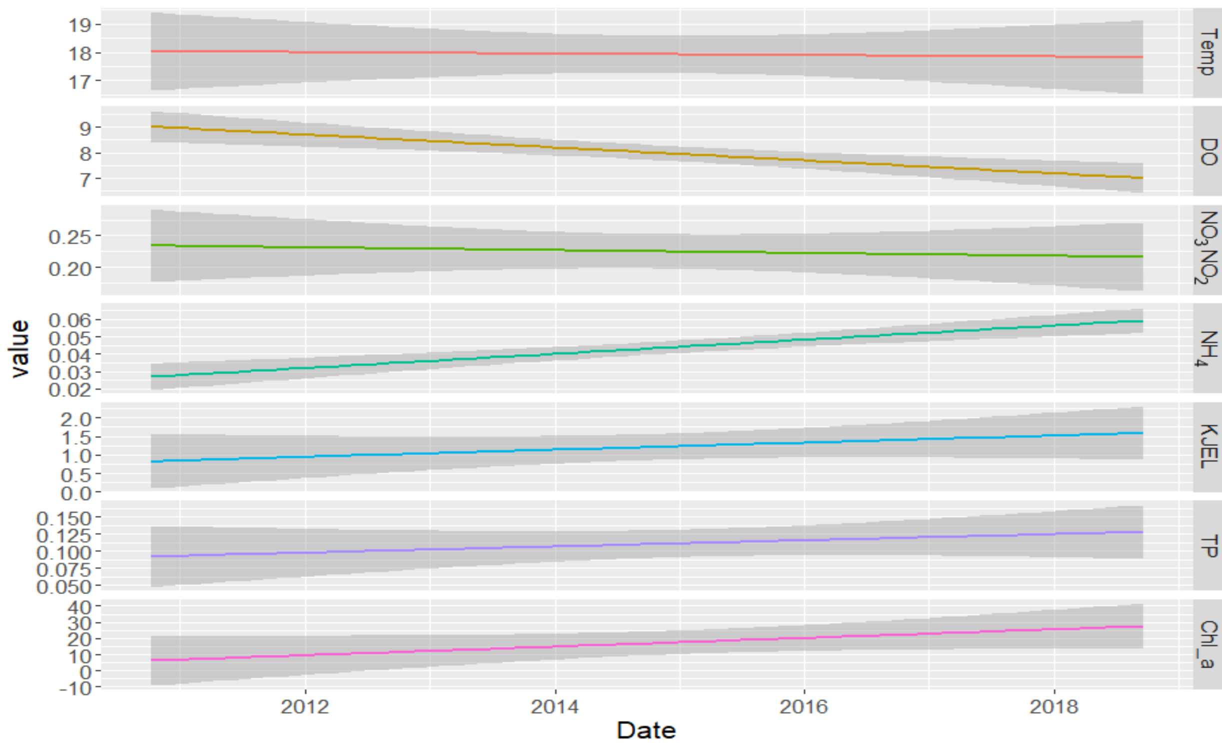
value
Date
1990
1995
2000
Temp
NO3NO2
NH4
KJEL
TP
Chl_a

(a)



value
Date
2000
2005
2010
TP
Chl_a

(b)



(c)

Figure 3.5 Decadal trends of Chl−a, TP, DO, NO3NO2_N, NH4_N, KJEL_N, and Temperature in the Vaal Dam. (a–c) showing the decadal trends for the first, second, and third decades respectively.

The regression analysis between the chosen variables showed positive correlations between Chl−a and TP; Chl−a and temperature; Chl−a and KJEL_N, also showed positive correlations between KJEL_N and TP, and KJEL_N and temperature (Figure 3.6) while s showing negative correlations between Chl−a and DO, Chl−a and NO_3NO_2_N and between KJEL_N and DO but no correlation between Chl−a and NH_4_N (Figure 3.6). The satellite-based time series analysis showed the spatial and temporal variability and the HABs dynamics across the dam. The satellite data were normalized to compare the variation between those mentioned on different dates. In general, the images showed high levels of Chl−a concentration which is linked to the summer blooms and the load of nutrients to the system with the summer storm runoff, for example, the image dated 7 February 2016, was captured 10 days after a very high record of Chl−a concentration at the measuring point (452 µg/L). Generally high productivities were seen near the Vaal and Wilge Rivers discharge areas in most of the images, also high productivity levels were seen where the reservoir is shallow and meandering (Figure 3.7).

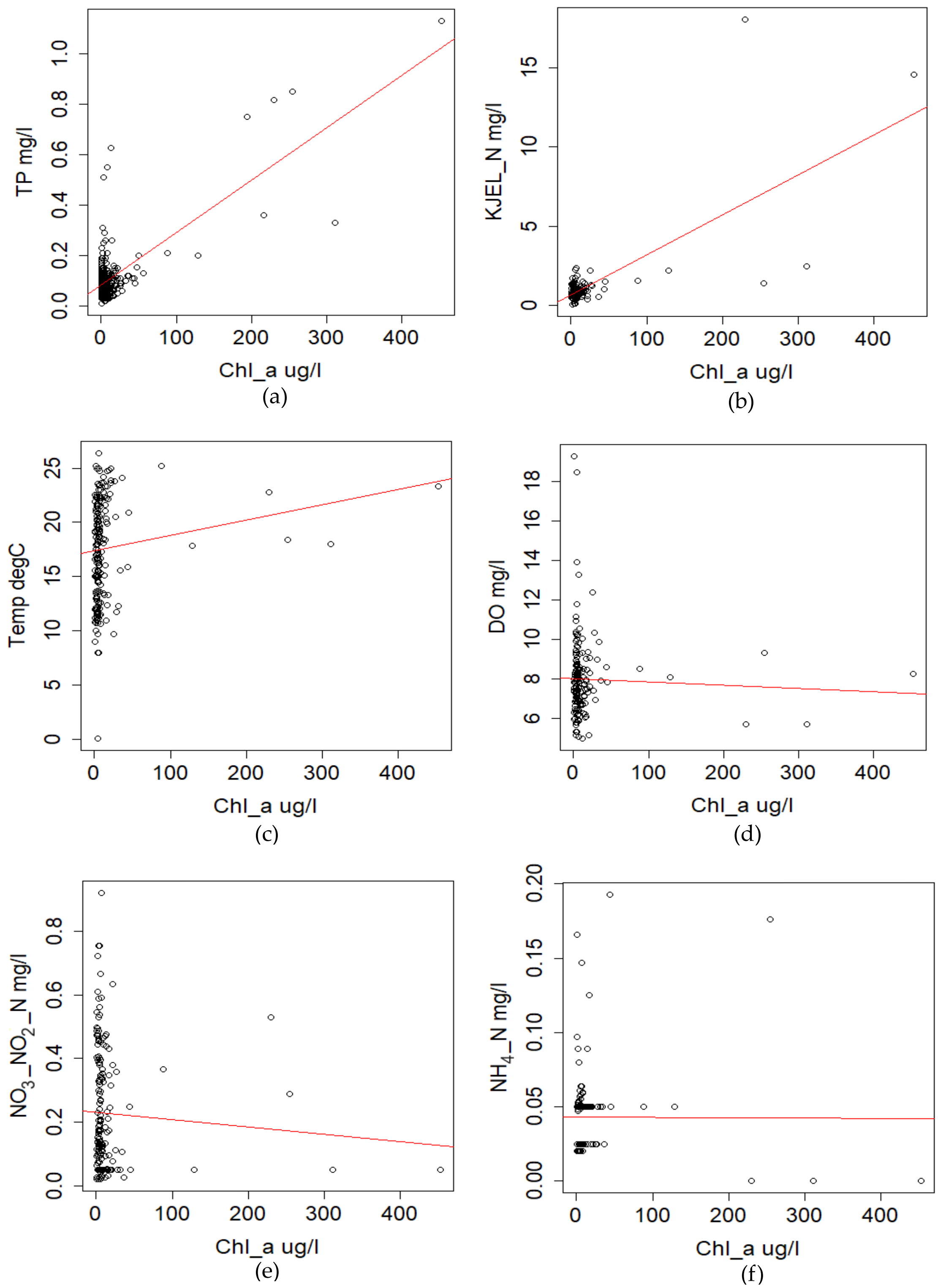
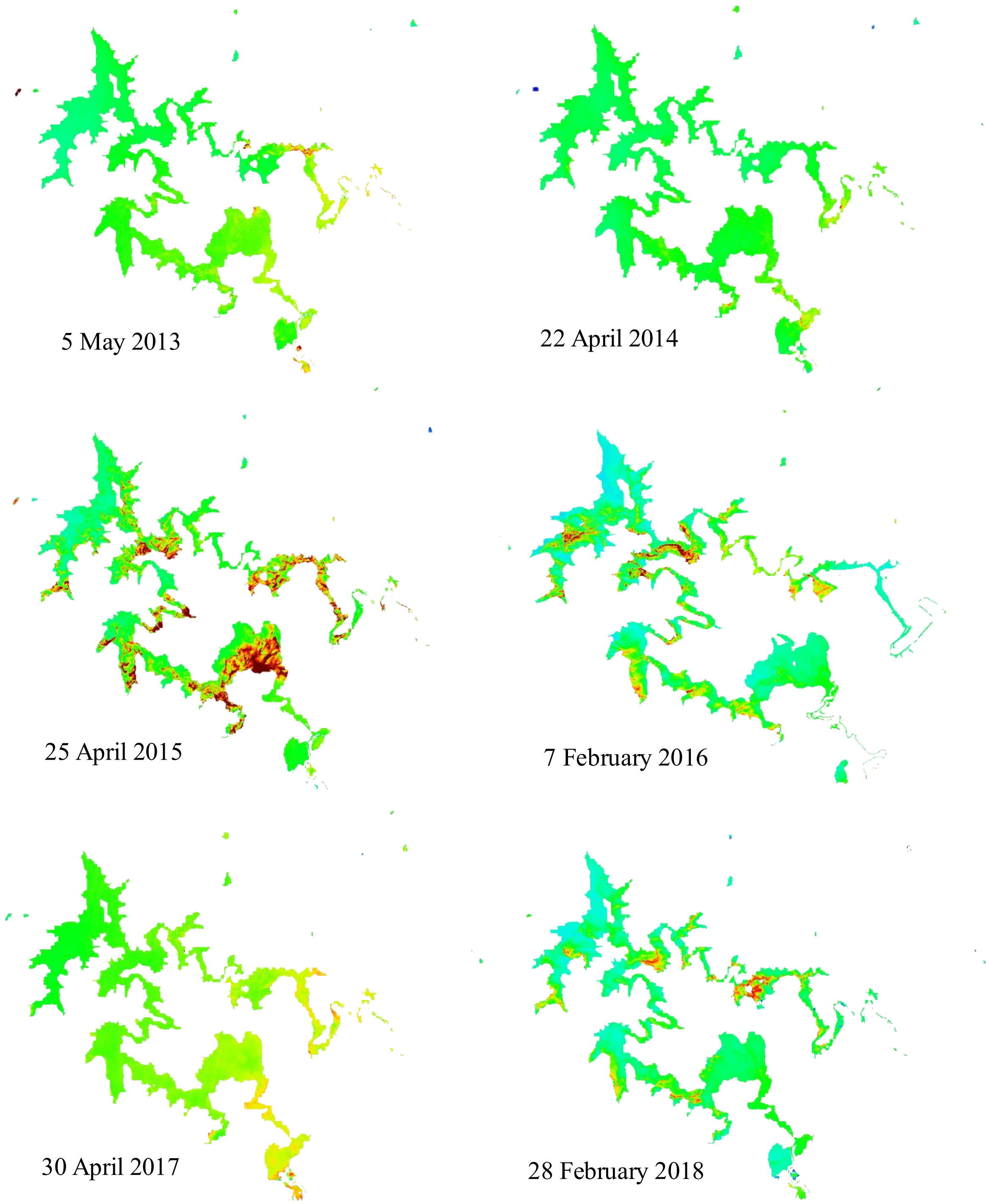


Figure 3.6 Scatter plots between Chl–a and various water quality parameters; (a) Chl–a vs. TP; (b) Chl–a vs. KJEL_N; (c) Chl–a vs. temperature; (d) Chl–a vs. DO; (e) Chl–a vs. NO_3NO_2; and (f) Chl–a vs. NH_4. The trend lines (red) showing the relationship between the two variables composing the scatter plot.


5 May 2013
22 April 2014
25 April 2015
7 February 2016
30 April 2017
28 February 2018

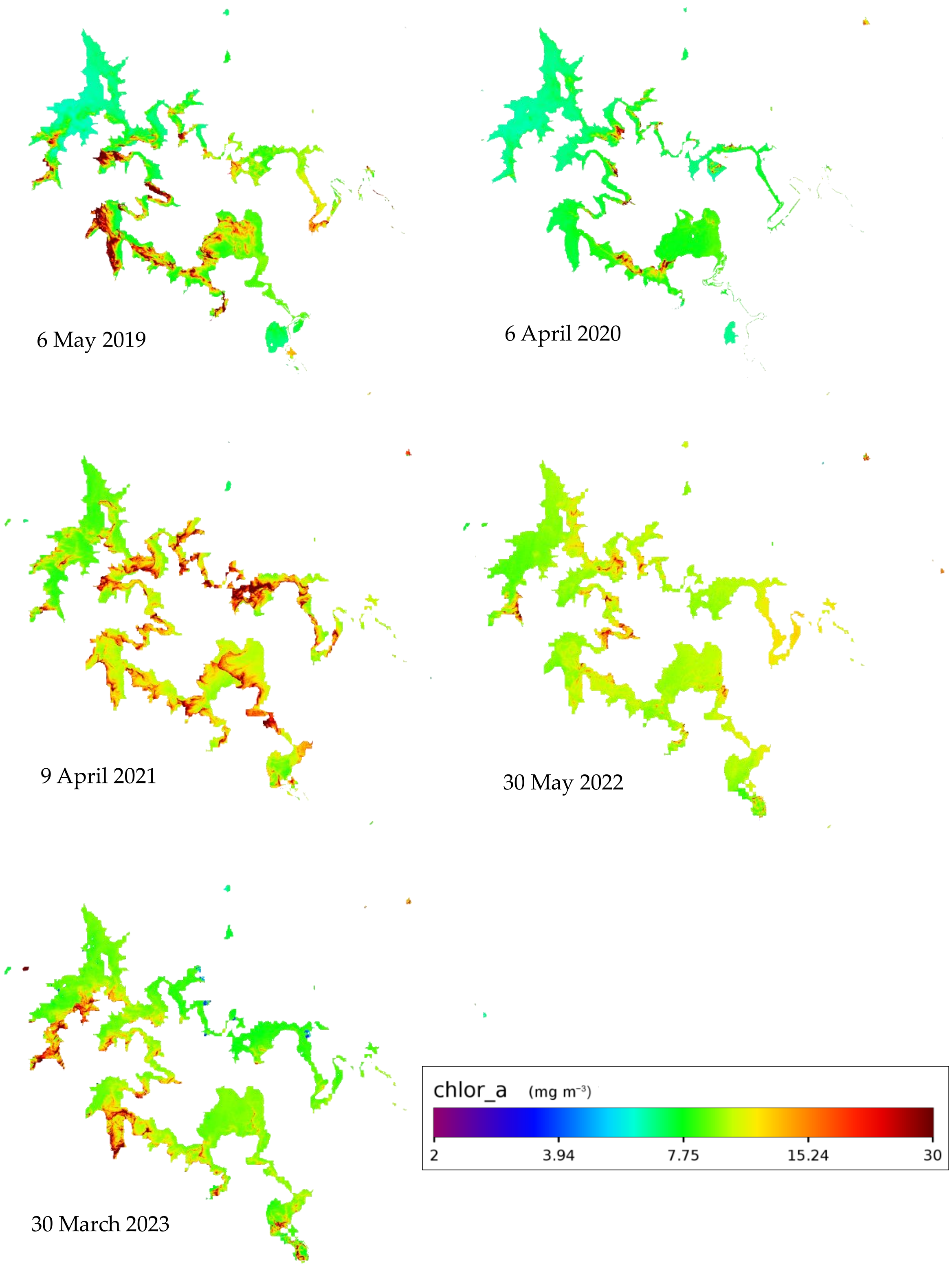


Figure 3.7 HABs productivity in the Vaal Dam between 2013 and 2023 from Landsat.

3.5 Discussion

The data have been explored to uncover the important water quality properties. The perusal of the graphs reveals some extremely high values as well as some trends, it also reveals the seasonality of temperature and dissolved oxygen. The distribution frequency plot of chlorophyll−a (Figure 3.2) after excluding the very high values, shows that most of the data is cantered near its lower range, below 10 µg/L. This illustrates that most of the time, the chlorophyll−a concentration in this measurement point remains below its average value. Algal blooms are spatially and temporarily varying, it might happen that an excess bloom might occur at the location of the measuring point during the time of sampling. This might explain the extremely high values of Chl−a concentration. The distribution plots of TP, NH_4_N, and KJEL_N after removing the very high values, also show that most of the data are cantered relatively near their lower values in the reservoir water. Therefore, the extremely high to very high values probably indicate that some types of pollution ended up reaching the reservoir which apparently increases the capacity of the reservoir water to support high rates of biomass productivity during such high nutrient load times.

The productivity in the open water systems is a function of multi biophysical factors, such as light, temperature, nutrients, etc. In many aquatic systems, some limiting nutrients (mainly TP and N) are considered the main drivers of HABs (Welch and Cooke, 2005, Gholizadeh et al., 2016). However, the time series plot (Figure 3.3) showed that TP levels between 1990 and 2022 are relatively low except on some specific dates, a corresponding increase in Chl−a and KJEL_N concentrations at such specific dates were detected, which jumped to extremely high values. These are clearer when we look closely at the decadal time series plots in Figure 3.4b,c. Thus, this explains that the dam water HABs productivity is primarily driven by TP and KJEL_N when their concentrations suddenly jumped to high values. The water temperature time series in Figures 3 and 4a,c shows no significant change, and its decadal trends show a slightly decreasing trend for the first decade (1990 to 2000) in Figure 5a while showing no increasing or decreasing trend during the last decade (2010–2020) in Figure 5c. The average temperatures for the first and last decades were 17.94 and 18.04 ◦C, respectively, which suggests a slight increase in the average water temperature between the first and last decades. This situation of increasing the decadal average of Chl−a during the study period while the temperature remains with no significant change suggests that the productivity in the Vaal Dam is not limited to the changes in the temperature during the study period, but its value ranges are ideal for algal growth.

Chl−a decadal trends (Figure 3.5a–c) showed a slight decrease in the first decade (1990 to 2000) and remained constant within the second decade (2000 to 2010) before it increased noticeably during the last decade (2010 to 2020). The average decadal concentrations were 4.75, 10.51, and 16.7 μg/L, respectively. Such an increasing trend of the Chl−a during the last decade alongside the increasing decadal average through the study period raises a significant warning to the situation of algal blooms in the dam. It will become a serious challenge facing the authorities to control this rising impact of HABs in the future. The significant increases in the Chl−a decadal average might be affected by the erratic extremely high values that occurred at individual dates during the second and third decades, for example, when excluding all Chl−a values above 50 μg/L, the decadal averages will decrease to 6.26 and 10.38 μg/L for the second and third decades, respectively. It still has high jumps on the decadal averages through the last three decades.

TP, KJEL_N, and NH_4_N decadal trends followed the same trend behaviour of the Chl−a in the first and last decades, where they were associated with general upward and downward trends of Chl−a. The average decadal concentrations of TP were 0.1043, 0.1096, and 0.1119 for the first, second, and third decades, respectively. These concentrations show that the TP concentrations were low except on the above-mentioned individual dates where they jump to high levels greater than the hypertrophic TP threshold level which is 0.25 mg/L (Ali et al., 2022a). For KJEL_N, the decadal average concentrations were 0.8033 and 1.1417 mg/L for the first and last decade, respectively, and NH_4_N were 0.0397 and 0.0431 mg/L, respectively. However, the corresponding behaviour of the Chl−a, TP, and KJEL_N decadal trends alongside the erratic high Chl−a values recorded at the same individual dates of very high TP and KJEL_N records, support that they are the driving factors of the algal blooms within the Vaal Dam.

On the other hand, the NO_3NO_2_N concentrations showed a negative trend compared to the Chl−a decadal trends during the first and last decade and DO also showed a negative trend during the last decade. The NO_3NO_2 decadal average concentrations were 0.2455 and 0.2248 mg/L, respectively, while the average concentration of DO was 7.928 mg/L. The trend of the DO is well understandable because dissolved oxygen is usually consumed during excessive algal blooms which have been known through the last few decades (Welch, 1969), but this situation might not be applicable to the behaviour of NO_3NO_2_N trends. The analysis of the results suggests that it has no direct relation with the Chl−a trend, but a paper (Kharbush et al., 2023) focused on sources and forms of the most important nitrogen substrate for blooms in eutrophic Lake Erie suggested that the NO_3^- was the most important source of N except in late blooming stages where phytoplankton relays on recycled N derived from dissolved organic nitrogen. Their study further showed that the NO_3^- depletion was related to the consumption by phytoplankton during its blooms showing a negative relationship. In this study, NO_3NO_2_N also showed a negative correlation with Chl‒a, but the assumption of consumption by the HABs needs more detailed investigations. The regression between the targeted WQ parameters showed a strong correlation between Chl‒a and TP, Chl‒a and KJEL_N in great agreement with the trend analysis. It also showed a positive correlation between Chl‒a and temperature while there was no correlation between Chl‒a and NH_4_N which explains that NH_4_N is not a driving factor for HAB in the Vaal Dam. Like the decadal trend plots, the regression showed a negative correlation between the Chl‒a, DO and NO_3NO_2_N.

The need for more effective monitoring of the temporal and spatial distribution of HABs in the Vaal Dam Reservoir has led to the use of satellite remote sensing techniques. The satellite data gives a synoptic view of the spatial distribution and the temporal frequency of HABs. Thus, it is an effective tool for the detection, mapping, and monitoring of the blooms. The spatial distribution of HABs in the reservoir has been analysed using Chl‒a concentrations derived from satellite data. Despite the presence of only one ground station for validation, the use of remote sensing data is justified based on the robustness of the employed algorithms. The accuracy of the retrieved data has been confirmed through rigorous evaluation of these algorithms. This assessment included correlating the in-situ data point with 16 distinct satellite images obtained across varying dates, seasons, and weather conditions. From each image, a singular value corresponding to the measuring point coordinates was extracted. The resulting R^2 value indicates a reasonable fit, thereby affirming the reliability of the algorithms for chlorophyll retrieval in the Vaal Dam. Additionally, the observed accuracy of the retrieved data underscores the effectiveness of this approach, despite the limited number of in-situ data points utilized for evaluation. Notably, observations indicate temporal variations in Chl‒a concentrations, including instances of elevated levels of Chl‒a concentrations in 2015 and the following summers. Such high concentrations are strongly correlated with the in-situ measured values of Chl‒a, TP, and KJEL_N during April and May 2015, and January 2016. High productivity was also noticed in images of 2019, 2021, 2022, and 2023. These high productivities may relate to the periods where TP and KJEL_N levels increase due to the increase in human activities such as the discharge of partially treated or untreated wastewater, or through runoff from urban and agricultural areas around the dam. For example, the media reported some waste found its way to the reservoir on 23 July 2015, a report mentioned that an uncontrolled sewage discharge overwhelmed the Deneysville town which is located on the Vaal Dam bank, just next to the dams wall (https://mg.co.za/article/2015-07-23-sewage-in-gautengs-drinking-water/), accessed 7 March 2023.

The satellite data showed high productivity in the Vaal Dam for the period from 2019 up to 2023 while there are no historical records available on the website during this time, which put some warnings of increasing nutrient loading to the dam reservoir recently. The geometry of the dam and the fluxes of the Vaal and Wilge rivers draining into it are directly linked to the pattern of HAB spatial dynamics, the shallow parts of the dam as well as the meandering areas where the velocity of the water decreases always exhibit relatively high blooms compared with the rest of the dam. Satellite monitoring is useful in detecting the frequency of the blooms. In this study, we processed one image each summer during the last decade, but there is a potential to monitor the spatial and temporal dynamic of the blooms every 16 days using Landsat−8 images only, or shorter intervals if using a harmony of Landsat-8 and Sentinel−2 a and b satellite data.

3.6 Conclusions

This study revealed that the Vaal Dam productivity is a function of TP and KJEL_N levels. The analyzed data between 1986 and 2018 showed a positive correlation of TP and KJEL_N with chlorophyll-a concentrations. They are the HABs drivers alongside the seasonal effect of surface water temperatures which enhances the productivity levels. When TP and KJEL_N levels are low, the Chl−a concentrations are usually below the threshold level of the eutrophic conditions of 7 μg/L. Chlorophyll-a retrieved from Landsat data showed the temporal and spatial dynamics of HABs over the past decade. Time series analysis of the data has shown the effect of sudden rises of the driving parameters on productivity levels, the productivity increases with an increased flux of TP and KJEL_N. The relatively high concentrations were observed where the reservoir is shallow and meandering and where restricted water circulation occurs due to the low level of mixing and near Vaal and Wilge rivers discharge points. According to the location of the in-situ measurement point, which is in the Vaal River stream near the dam wall, the measured Chl−a concentration might be much less than the concentrations on many other parts of the dam reservoir because the HABs can be washed out by high flow speed. Thus, one measuring point for this large dam is not enough to reveal the overall situation of the chlorophyll-a concentration on the dam, mainly when considering the spatial variations. Therefore, more water quality monitoring points are needed and to be well distributed among the reservoir for a better understanding of the spatial variations and to give valuable information to be used for satellite monitoring algorithm building and adjustments. However, the inherent limitations of relying on a solitary sample point have been recognised. Therefore, to enhance the reliability and generalizability of the findings, this approach has been emphasised to serve as an initial exploration to provide the general trend of the water quality. However, it also highlights the need for further investigation and data collection to strengthen the robustness of our findings. The results obtained from this time series analysis revealed a high level of anthropogenic impacts on the Vaal Dam which is being used to provide drinking water for Gauteng province, and for agricultural and industrial uses in the region. The periodic deterioration of the quality of the dam water over some stochastic dates was aggravated by the discharge of nutrient-rich, poor-quality effluents from the wastewater treatment works in the dam catchment.

Chapter Four

Assessing the utility of using current generation high-resolution satellites (Sentinel 2 and Landsat 8) to monitor large water supply dam in South Africa

This chapter is based on **Obaid, A.**; Ali, K.; Abiye, T.; Adam, E.(2021) Assessing the utility of using current generation high-resolution satellites (Sentinel 2 and Landsat 8) to monitor large water supply dam in South Africa. *Remote Sensing Applications: Society and Environment* 2021, *22*, 100521. Available at: https://doi.org/10.1016/j.rsase.2021.100521.

Abstract: The Vaal Dam is one of the largest impounded water resources in South Africa with a capacity of 2330 × 10^6 m^3 and a surface area of 321.07 km^2. The Dam is part of the Vaal River system with a catchment area of about 37 100 km^2 and is designed to supply water to the large metropolitan city of Johannesburg in the north and the surrounding areas in the Gauteng Province serving approximately 13 million people. The Dam is biologically productive and experiences frequent algal blooms, especially during the summer months (February to April). Recent studies have indicated the presence of toxic cyanobacteria genera such as Anabaena, Lyngbya, and Microcystis and these can be of great concern for ecosystem health as well as biodiversity. Currently, ground-based methods are employed to assess and monitor the water quality of the Dam. This method is costly and lacks the temporal and spatial resolution. Therefore, more efficient observation tools are required for effective monitoring. Satellite remote sensing is an efficient tool that can allow accurate and timely detection of harmful algal blooms (HABs) such as the toxic cyanobacteria genera and other algal types. This study focuses on assessing the utility of using current generation Satellites (Sentinel-2 and Landsat-8) to characterize the water quality and estimate concentrations of chlorophyll−a as a proxy for HABs and the total suspended matter (TSM). The optimal correlation between in-situ and chlorophyll−*a* retrieved from the blue-green band ratio of Landsat 8 [R_{rs} (560)/R_{rs} (443)], and red/NIR of Sentinel-2 [R_{rs} (705)/R_{rs} (665)] were found to be the best indices for Chl−a retrieval in the Vaal Dam. Results for Landsat OLI dataset (R^2 = 0.89; RMSE = 0.36 µg/L, $P < 0.05$) and Sentinel MSI dataset (R^2 = 0.75; RMSE = 0.48 µg/L, $P < 0.05$) show high degree of accuracy. Models developed in this study will significantly improve near-real-time and long-term water quality monitoring of the Vaal Dam and will effectively help to address public health issues related to water quality.

4.1 Introduction

Eutrophication has been a recurring problem in many freshwater bodies across South Africa (Matthews et al., 2010, Van Ginkel, 2004). Despite legislation that restricts the discharge of nutrients from wastewater treatment plants, eutrophication has continued to increase in recent years. Over the past decades, the Vaal Dam reservoir has been under increasing and considerable pressure due to a variety of interlinked humanmade and natural pollution. This is leading to the deterioration of water quality, putting this important resource at considerable risk. Terrestrially derived inorganic and organic wastes; the interaction of chemical, biological and physical elements gives rise to the turbidity observed in the dam reservoir water. The influx of nutrients is increasing primary productivity in the dam and reaches its maximum during the summer period. The dam reservoir is a shallow system with an average depth of 22.5 m and is located in a subtropical zone with warm temperatures during the summer season with noticeably warm extremes increasing during the last five decades (Kruger and Sekele, 2013). These situations are further enhancing productivity (Chinyama et al., 2016). Some of the observed algal blooms are cyanobacteria species that produce toxins (e.g., *Microcystis aeruginosa*) and most of the animal poisonings in South Africa since 1927 have been linked to ingestion of these algal species (Harding, 2008, Mamba et al., 2007, Scott, 1991). The Dam serves as important resource in the region through agriculture, industries, and for drinking water supply. To better manage this system and support the sustainability of the Vaal Dam water resources, it is important to accurately observe and monitor the water quality.

Obtaining these observations through conventional in-situ measurement methods is challenging and does not provide the spatial resolution required for continuous assessment of water quality in open freshwater bodies (Ali et al., 2016, Gholizadeh et al., 2016). With appropriate validation, satellite-based remote sensing can provide complementary information to provide high spatial and temporal resolution water quality data and greatly improves our understanding of variations in water quality across the reservoir (Ali et al., 2016, Gholizadeh et al., 2016, Wang et al., 2009). Multispectral sensors that detect variations in reflectance from the water surface have proven promising for water quality assessment (Ali et al., 2014a, Ali et al., 2016). Measured reflectance at multiple wavelengths from the water surface carries spectral signatures that characterize various water quality parameters (WQP) such as chlorophyll−a, turbidity and coloured dissolved organic matter (CDOM) (Morel and Prieur, 1977, Gitelson et al., 2007, Gitelson et al., 2008, Gholizadeh et al., 2016, Sòria-Perpinyà et al., 2021). Phytoplankton pigment chlorophyll−a is an optically active constituent of water that interacts with light and drives variations in ocean colour (OC) through their absorption and scattering properties; it changes the energy spectrum of reflected solar radiation from water bodies (Franz et al., 2014, Gholizadeh et al., 2016). These interactions in the portion of visible and near-infrared bands of the solar spectrum can be used to obtain robust correlations between water column reflectance and chlorophyll−a concentration.

Many bio-optical algorithms have been developed to obtain information on the biochemical constituents of the upper euphotic zone using the upwelling radiance from the water column. These algorithms relate measurements of remote sensing reflectance $R_{rs}(\lambda)$ to chlorophyll−a to provide phytoplankton biomass approximation (Franz et al., 2014). In the open sea (Case-1 waters), many OC algorithms depend on blue and green bands have been developed and used for detecting chlorophyll−a concentrations which dominate the resulting spectral signals; while the inland and coastal complex (Case-2) waters are outside the standard conditions used for OC algorithm designed for Case-1 (Ali et al., 2016, Hieronymi et al., 2016). In the coastal and inland waters, multiple WQPs affect remote sensing reflectance, $R_{rs}(\lambda)$, including CDOM, suspended sediments, tripton and the presence of harmful algal bloom pigments. (Ali et al., 2014a, Doerffer and Fischer, 1994, Malahlela et al., 2018, Sakuno et al., 2018a). The various WQPs competingly absorb light in the blue spectral region and may add to uncertainties when using this region to detect specifically chlorophyll−a concentration (Franz et al., 2014, Gons, 1999, Gons et al., 2002, Gitelson et al., 2007). On the other hand, Rrs products have been evaluated in Case-2 waters using data from the Aerosol Robotic Network (AERONET-OC) and showed that the near-infrared and shortwave-infrared (NIR-SWIR) band combinations generate robust Rrs retrievals in optically complex turbid waters (Pahlevan et al., 2017b).

New generation satellite sensors such as the Sentinel 2 MSI and Landsat 8 OLI are equipped with spectral bands that are optimal for measurements of WQPs in turbid waters such as the Vaal Dam. These sensors have high spectral, spatio-temporal resolution and therefore, allow for better characterization of the spectral signals associated with the various WQPs. The goal of this study was to evaluate existing visible/NIR empirical models using reflectance data from Sentinel 2 MSI and Landsat 8 OLI sensors and their application for monitoring bio-mass productivity in the Vaal

Dam reservoir. The algorithms use spectral bands in the visible and near-infrared regions that are sensitive to chlorophyll−a pigment attenuation feature. Therefore, this study tries to accurately monitor changes in water quality in the Vaal dam through the implementation of spectral data using the new generation satellites.

4.2 Material and Methods

4.2.1 Study area

The Vaal Dam is located in the northeastern region of South Africa (Fig. 4.1) and is characterized by a strong seasonal cycle with a diurnal cycle of precipitation during summer (Rouault et al., 2013). The driest months are the winter months (May–July) and the most rainfall is received in January as shown in Fig. 4.2 Average annual temperature and rainfall amount are 15.7 ◦C and 681 mm, respectively (Fig. 4.2). The dam is the most important infrastructure in the Vaal River system that provides water to the economic heartland of South Africa. It lies about 56 km south of Johannesburg within the eastern part of the Vaal River drainage region. The Vaal Dam catchment is the most important supplier of water to the Vaal River water supply system (Braune and Roger, 1987). The Dam supplies water to the metropolitan cities of Pretoria and Johannesburg besides small towns within which about 40% of South Africa's population resides and which accounts for more than 50% of the gross national product. The Rand Water Board manages the Dam. Some of this water originates through the Lesotho Highlands Water Project, and it pumps into the Clarens ultimately flowing into the Vaal River system (Grobler et al., 1987, Naidoo, 2017). The surface area of the Vaal Dam Reservoir (VDR) is approximately 321 km^2 and the adjoining catchments area of about 38 000 km2; it is a shallow reservoir reaching its maximum depth at the location of the maximum height of the dam wall, its average depth is 22.5 m. The total capacity of the dam reservoir is around 2540 million m^3 (Bande et al., 2018, Chinyama et al., 2016, Sakuno et al., 2018a). The VDR characterized by highly turbid water and the average annual rainfall of the Vaal Dam catchment is about 700 mm (Sakuno et al., 2018a). The Vaal Dam Catchment includes gold and coal mining, the Sasol fuel processing plants, agricultural activities, cattle grazing on natural pastures, dry-land cultivation and urbanization areas (Braune and Roger, 1987).

4.2.2 Data and field measurements

Field surveys were conducted in the VDR on March 19, April 27 during the summer season and on 15 and 26 August in the winter season of 2019. A total of 243 observation data from 49 stations were used in this study. The sample locations were stored as waypoints in a handheld GPS to verify reoccupation of the same site during each visit, chlorophyll−a, turbidity and other WQPs were measured using a submersible bio-optical sonde (Hydrolab Data Sonde 5x instruments with chlorophyll−a sensor: range = 0.03–500 mg m^{-3}, accuracy = ±3%, resolution = 0.01 mg m^{-3}). Ten to fifteen water samples were collected from the various sites across the VDR during each field campaign for analysis in the laboratory to measure multiple WQPs, including nutrients, chlorophyll−a and other conventional water quality parameters. These measurements were used for model calibration and validation. For chlorophyll−a, the samples were filtered using 0.45 μm glass fibre filters (GF/F) and stored in a cooler at – 20 ◦C to prevent degradation. The chlorophyll−a concentration was measured fluorometrically after the extraction process was conducted within 24 hours of collection following the US EPA 445 protocol (Arar and Collins, 1997). 10 mL of 90% acetone buffered with $MgCO_3$ was used in extracting the chlorophyll-a pigment and then macerated. The extraction process took place for 24 hours in a standard refrigerator (~4 ◦C). Chlorophyll−a concentrations, which were corrected for degradation products such as pheophytin, were measured fluorometrically using a benchtop Fluorometer (Turner Designs, Inc.) fitted with a daylight white lamp and chlorophyll optical kit (340–500 nm excitation filter and emission filter ≥ 665 nm). A three-point calibration was conducted for the optical kit using liquid chlorophyll−a standard acquired from Turner Design, Inc. Previously collected chlorophyll−a data in 2016, 2017 and 2018 were also incorporated in data analysis and a total of 243 observation data points that were collocated with cloud-free satellite pixel data were used for the modelling.

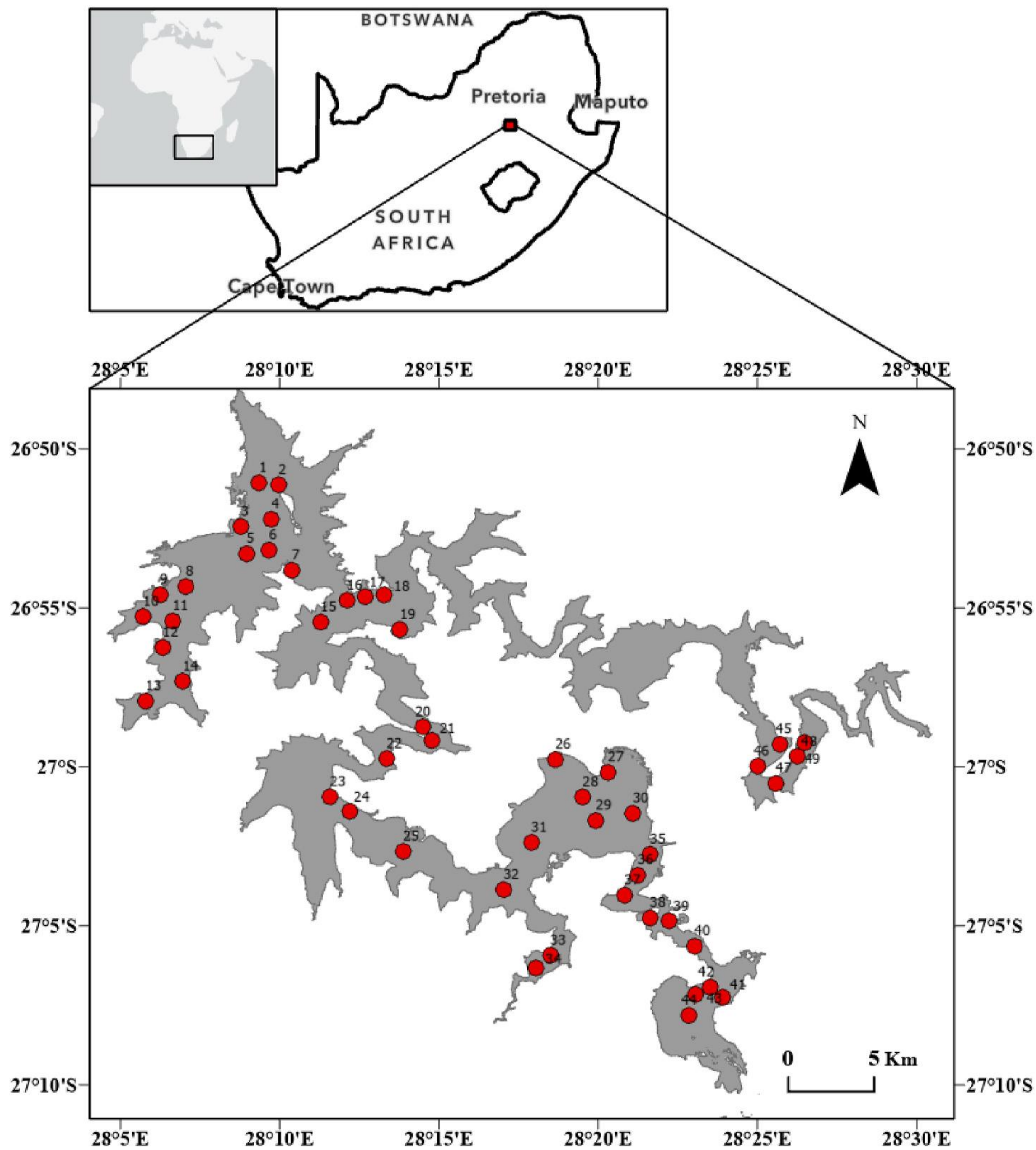


Fig. 4.1 Location map of the Vaal dam and sampling positions.

4.2.3 Satellite imagery and modelling

Two current-generation high-resolution sensors from NASA (Landsat 8) and ESA (Sentinel 2) were employed in this study to characterize the water quality of the Vaal Dam (Table 4.1). Landsat 8 Operational Land Imager (OLI) and Sentinel 2 Multispectral Imager (MSI) level 1 images that were acquired within ~±3 days of the field campaign were obtained from the USGS Earth Explorer data portal. The level 2 generator (l2gen) algorithm from NASA's standard OC processing software, the SeaWiFS Data Analysis System (SEADAS; version 7.5.1) was used to process the satellite data, including atmospheric correction process and retrieval of the remote sensing reflectance (R_{rs}) (Franz et al., 2014, Ilori et al., 2019, Pahlevan et al., 2017a). Atmospheric correction was carried out on images from both satellites. In the Vaal Dam, the assumption of near-zero reflectance at NIR wavelength cannot be made due to its turbidity and increased scattering effects from non-algal components. Therefore, the commonly used dark object subtraction method cannot be applied (Schalles, 2006, Shi and Wang, 2009). Iterative-based atmospheric correction methods that use NIR (Gordon and Wang, 1994) and SWIR bands (Wang et al., 2009) have been developed and applied in optically complex waters (Moses et al., 2009). For this analysis, an iterative-based atmospheric correction model was applied using a 2-band NIR and SWIR iterative model with non-negligible water leaving radiance (Pahlevan et al., 2017b). Only satellite images with 20% or less cloud cover were downloaded, and additional masking was also applied for land and cloud pixels.

4.3 Bio-optical modelling

Many of the existing OC models are based on the relationship that exists between remote sensing reflectance R_{rs} and the inherent optical properties of waters – the backscattering coefficient ($bb(\lambda)$) and the absorption coefficient ($a(\lambda)$) expressed as:

$$Rrs \approx \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{1}$$

For chlorophyll−a, the attenuation (absorption and scattering) features of the pigment are used for model development. However, in waters where multiple WQPs exist, the performance of the chlorophyll−a model will depend on the extent to which the pigment attenuation features from non-algal components may be removed to independently extract the chlorophyll−a signature (Ali et al., 2014a). In the Vaal Dam, in-water constituents such as CDOM and total suspended matter (TSM) attenuate the chlorophyll−a signal and influence the reflectance in the visible spectral region masking chlorophyll−a generated signals (Bukata et al., 1995). Isolating the spectral signatures that are a function of chlorophyll a, independent of the other WQPs requires analysis of various spectral regions. In the Vaal Dam, the spectral signatures of the various WQPs overlap, and therefore the differentiation of the attenuation features requires the selection of prime bands(Ali et al., 2014b). Multiple band ratios (e.q. 2) from the Visible and NIR bands of the level 2 Landsat and Sentinel products were evaluated for chlorophyll−a retrieval. Linear-based empirical modelling was performed based on ratios of spectral regions $R(\lambda)x$ that represent maximum absorption and minimum absorption due to chlorophyll−a and the observed chlorophyll−a concentration in Vaal Dam.

$$Chl\ a \approx \frac{R(\lambda)_1}{R(\lambda)_2} \tag{2}$$

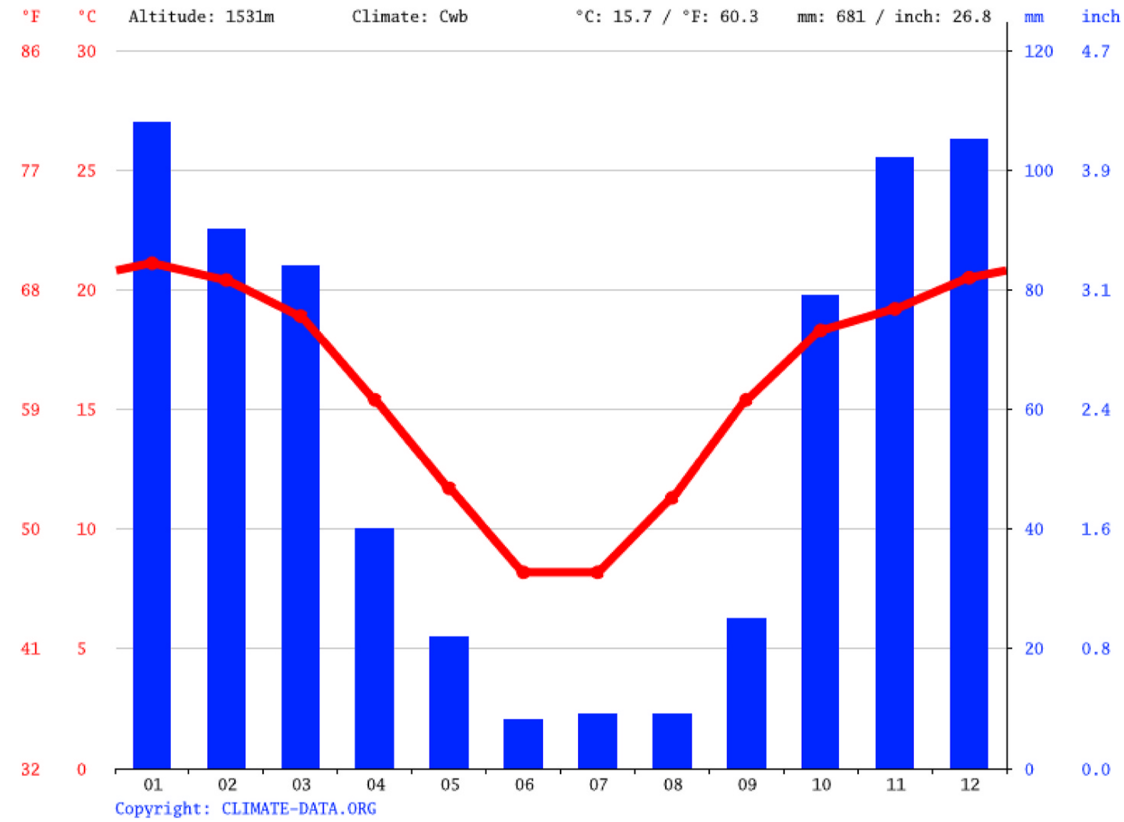


Fig. 4.2 Vaal Dam region monthly average temperature and rainfall measured at the Vaal community centre weather station (*South Africa Weather Services)*.

Table 4.1 Specifications of Landsat-8 OLI and Sentinel-2 MSI sensors.

L8 Band		CW (µm)	Wavelength Lower-Upper	Res. (m)		S2 Band		CW (µm)	Wavelength Lower-Upper	Res. (m)
1	C/A	0.443	0.435–0.451	30	C/A	1	C/A	0.443	0.421–0.457	60
2	Blue	0.482	0.452–0.512	30	Blue	2	Blue	0.494	0.439–0.535	10
3	Green	0.561	0.533–0.590	30	Green	3	Green	0.560	0.537–0.582	10
4	Red	0.655	0.636–0.673	30	Red	4	Red	0.665	0.646–0.685	10
						5	VRE	0.704	0.694–0.714	20
						6	VRE	0.740	0.731–0.749	20
						7	VRE	0.781	0.768–0.796	20
						8	NIR	0.834	0.767–0.908	10
5	NIR	0.865	0.851–0.879	30	NIR	8a	NIR	0.864	0.848–0.881	20

						9	WV	0.944	0.931–0.958	60
6	Cirrus	1.373	1.363–1.384	30	Cirrus	10	Cirrus	1.375	1.338–1.414	60
7	SWIR	1.609	1.567–1.651	30	SWIR	11	SWIR	1.612	1.539–1.681	20
8	SWIR	2.201	2.107–2.294	30	SWIR	12	SWIR	2.194	2.072–2.312	20
9	Pan	0.590	0.503–0.676	15	Pan					
10	TIRS	10.895	10.60–11.19	100[a]	TIRS					
11	TIRS	12.005	11.50–12.51	100[a]	TIRS					

[a] Acquired at 100 m resolution and resampled to 30 m.

4.3.1 Validation

The relatively limited collocated observation-satellite data (243 for Landsat and 90 for Sentinel) does not accurately allow for the conventional single holdout calibration validation procedure. This is because the holdout cross-validation estimator will have a higher variance and therefore will show large variation in the performance estimates. In this study, the leave-one-out cross-validation procedure (Kohavi, 1995) was used as a resampling method for training the model iteratively on all the data leaving one data out and a prediction is made for that point. This will bring a more robust assessment regarding the efficiency of the model without the limitations of the holdout method. Validation of the model was evaluated using the coefficient of determination (R^2), and root mean square error (RMSE) expressed as:

$$R^2 = \left(\frac{Cov(o,s)}{SDo\ X\ SDs}\right) \tag{3}$$

Cov(*o, s*) is the covariance of observation (O) and satellite (S), and SDo and SDs are the standard deviation of the two variables.

$$RMSE = \sqrt{\frac{1}{N}\sum_{1}^{N}(y_o - y_s)^2} \tag{4}$$

Where, *N* is the number of samples; *y0* is the predicted value of *chlorophyll-a*; and *ys* is the measured value of *chlorophyll-a* for each sample.

4.4 Results

4.4.1 Chlorophyll–a

Chlorophyll−a concentrations observed in the Vaal Dam ranged from 3.9 to 13 µg/L with an average concentration of 5.4 µg/L. Rivers flow at maximum during the summer (December to February) and significant amounts of both organic and inorganic materials are discharged into Vaal Dam. This increases the concentrations of phytoplankton and results in higher chlorophyll−a concentrations, especially at stations located near the inlets of the feeding rivers (Wilge River and Vaal River) (Sites 40–49, Fig. 4.3). The Vaal River drainage basin includes one of the most active sub-catchments in the area (Waterval catchment); characterized by urban, and complex industrial and agricultural activities (Steward, 2015). These activities coupled with direct solid waste disposal generate a substantial amount of effluent polluting the Vaal River system that eventually discharges near the southeastern end of VDR. Nutrient fluxes via polluted rivers lead to eutrophication near the VDR inlets. This condition in turn results in a large algal bloom near the discharge zone of the VDR leading to the observed variability of chlorophyll−a between the inlet region and the northern parts of the Vaal Dam, where trophic levels are relatively low. At the end of the rainy season (December–April), the river discharge decreases and the waters between the southern and northern sections of Vaal Dam mix reducing the chlorophyll−a variability in the reservoir. Based on observed data average summer chlorophyll−a concentration was 5.5 µg/L and decreased to 4.9 µg/L in the winter. To understand seasonal variations, Fig. 4.4, shows the average concentrations measured from each season's field campaign. Concentrations during summer were consistently higher except at a few sites. During this period, higher rainfall coupled with higher temperatures causes greater biological production and greater phytoplankton biomass.

Sites closest to the shore of the reservoir undergo the greatest exposure to land-based sources of pollution and runoff. This effect was more pronounced in narrow regions of the reservoir where water circulation may be limited. Effects in near shore and narrow regions of the reservoir are seen by the relatively higher chlorophyll−a concentration (Fig. 4.3). The variability in the reservoir volume and shape besides the availability of nutrient sources particularly treated sewage disposal and agricultural runoff causes variability in the chlorophyll−a concentration.

4.5 Remote sensing

The spectral plot from Landsat and Sentinel represents the optical properties of the water in the Vaal Dam (Fig. 4.5). The reflectance pattern in the visible and NIR range shows the optical signature of the multiple WQPs. The blue-green spectral range shows spectral patterns typically observed in turbid waters (Ali et al., 2014a, Gitelson et al., 2007, Kutser et al., 2013) with significant absorption due to chlorophyll−a, CDOM and other associated WQPs. Absorption features linked to chlorophyll−a at known wavelength regions (440, 490, 560, 670 and 700 nm) are represented depending on the band placement of each satellite. Both sensors show absorption troughs in the 440–490 nm region. Spectral plots from both sensors show relatively high reflectance near 560 nm (Fig. 4.5). This peak is attributed to minimum absorption by chlorophyll−a combined with scattering effects from non-algal particles. Sentinel 2 captures the chlorophyll−a fluorescence peak at the red/NIR edge (near 700 nm) which is a function of in-elastic scattering (Figure 4.5B). This peak is also observed in previous studies (e.g., Babin et al., 1996; Gower et al., 2004; Gower and King, 2007; Huot et al., 2007). The absorption feature near 665 nm represents the chlorophyll−a absorption feature. The absorption troughs near 665 nm and the reflectance peaks at 705 nm are more noticeable in the Sentinel-2 spectral plot from sites in the open waters of Vaal Dam.

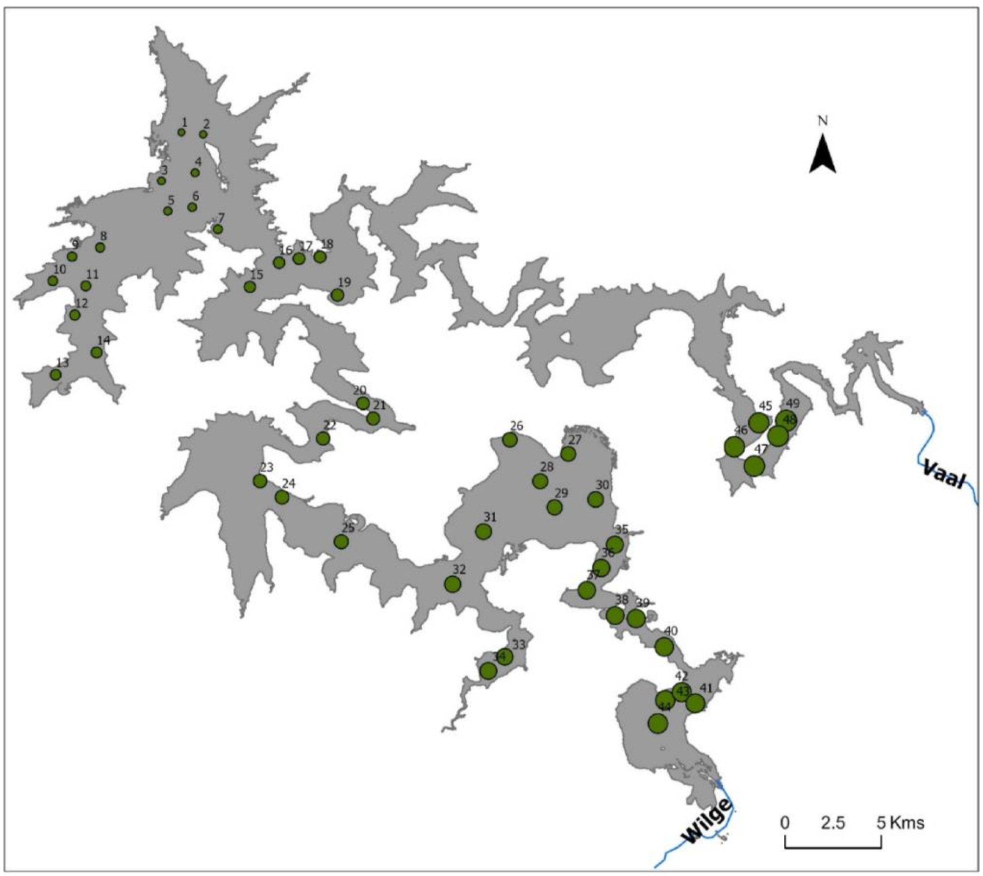


Fig. 4.3 Graduated symbol map showing average Chlorophyll−a variability across Vaal Dam, with the Wilge and Vaal Rivers draining into the dam.

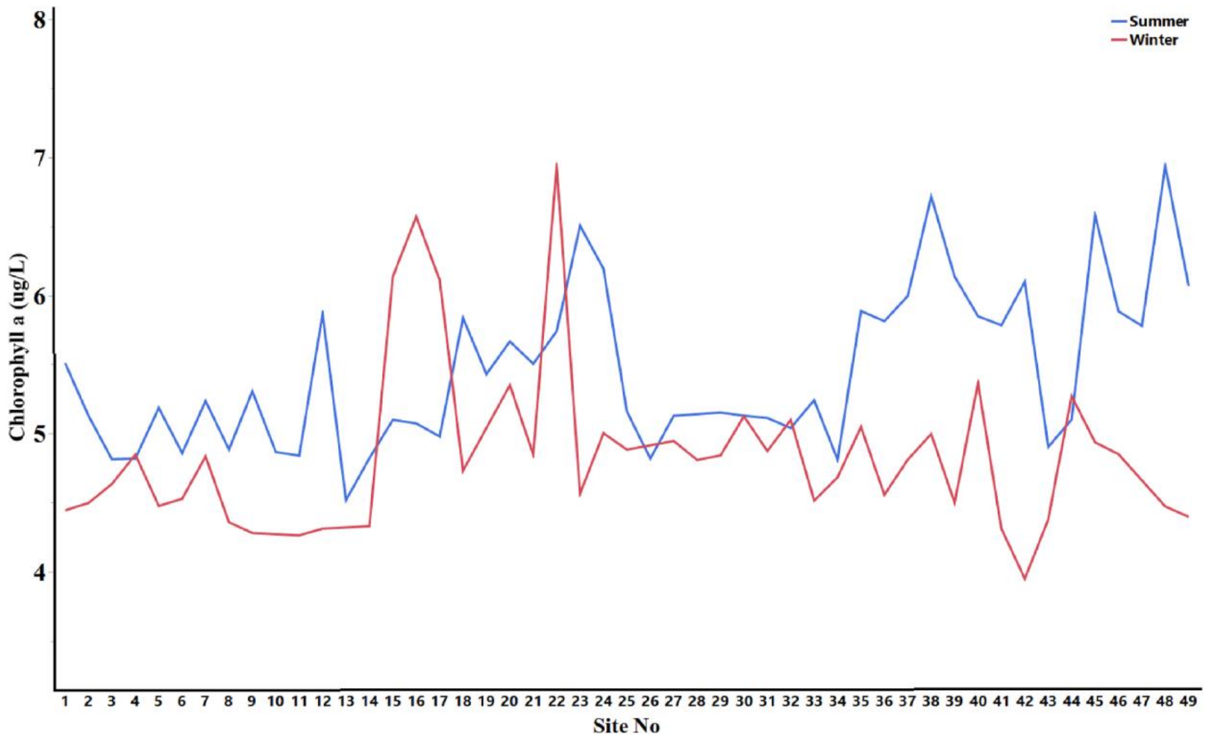


Fig. 4.4 The seasonal average in the Chlorophyll−a concentration (μg/L) by sample site.

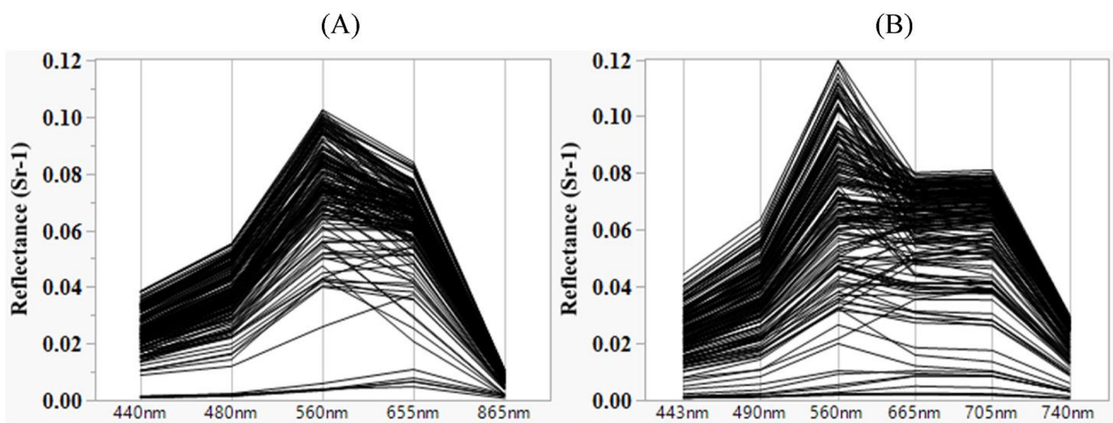
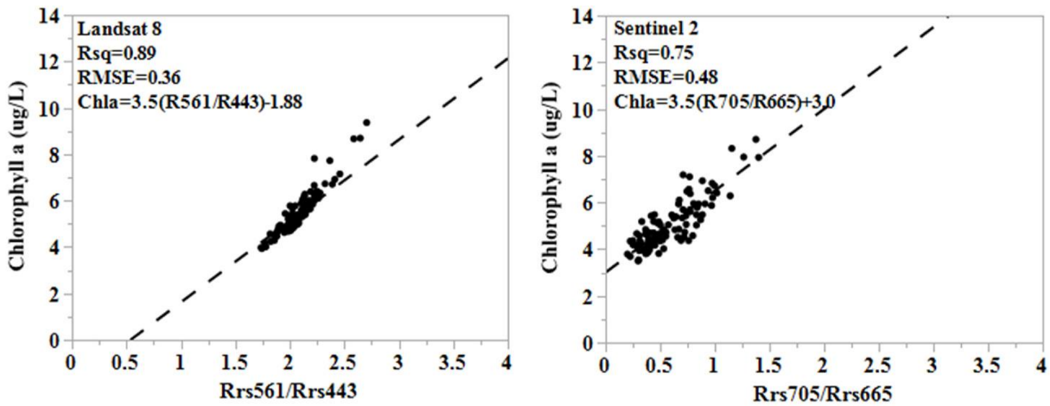


Fig. 4.5 Water leaving reflectance measured in Vaal dam from A) Landsat 8, B) Sentinel 2.



A B

Fig. 4.6 Optimal correlation between *in situ* and retrieved chlorophyll−a from (A) Landsat 8 OLI [R_{rs} (*560*)/R_{rs} (440)] and (B) Sentinel 2 MSI [R_{rs} (*705*)/R_{rs} (665)].

4.6 Bio-optical modelling

Algorithms that use band ratios in the blue/green and red/NIR spectral ranges are applied to the satellite-based remote sensing data that were collected during the overpass. Band ratios such as R440/R560, R490/R560 and R700/R665 are commonly used to estimate chlorophyll−a in large open water bodies (Ali et al., 2014a, Gordon and Morel, 1983, Gitelson and Yacobi, 2004, Moses et al., 2009, Menken et al., 2006, O'Reilly et al., 1998). These band combinations were evaluated to select the best model that describes the observed chlorophyll−a variability in the Vaal Dam. Based on the spectral characteristics of the Vaal Dam reflectance data set, each band ratio was optimized appropriately to enhance the predictive capacity of the associated algorithm. Since band placements of the Landsat 8 and Sentinel 2 sensors vary slightly, different band combinations were applied. For Landsat 8 regression between measured chlorophyll−a and blue-green ratio ($R_{rs}560/R_{rs}443$) gave the optimum results with R^2 = 0.89 and RMSE = 0.36 μg/L. This algorithm was optimized using data from both seasons, and the result is illustrated in Fig. 4.6A. The algorithm-predicted values lie along the 1:1 line with a linear relationship with the field measurements (Fig. 4.6A). The application of the blue-green model performed poorly when applied to Sentinel 2 data (R^2 = 0.3). This can be attributed to the higher signal interference from non-algal constituents and thus low signal-to-noise ratio (SNR) resulting from the relatively finer spatial resolution of the sensor, 10 m for Sentinel 2 and 30 m for Landsat 8. However, unlike Landsat 8, Sentinel has a red band (665 nm) and an NIR band (708) which are less sensitive to absorption due to associated WQPs. Regression analysis between Sentinel-based red/NIR band ratio ($R_{rs}705/Rrs665$) model and chlorophyll−a concentration in the Vaal Dam produced $R^2 = 0.75$ and RMSE = 0.48 µg/L (Fig. 4.6B). The NIR/red band is more sensitive to chlorophyll−a in-water constituents and therefore is strongly correlated with the measured chlorophyll−a concentration following the 1:1 correlation line. Using the respective optimal models for each sensor, chlorophyll-a concentration maps were produced using the 2019 Landsat and Sentinel image (Fig. 4.7). The Map clearly shows higher chlorophyll−a concentration near the inlet of the two rivers in the south. High values are also shown near narrow neck regions of the reservoir.

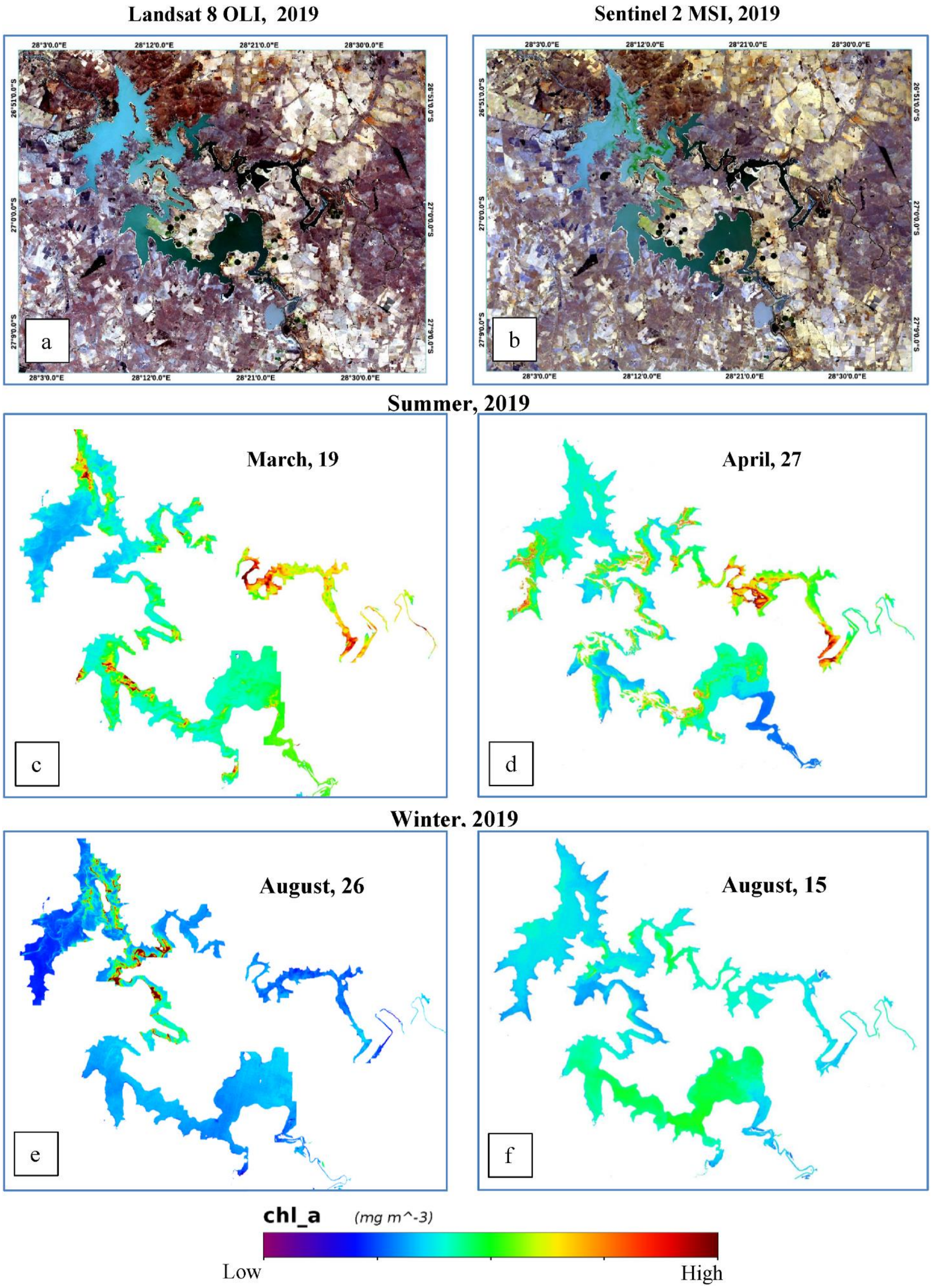


Fig. 4.7 Predicted distribution of chlorophyll–a within the Vaal Dam using Landsat-8 and Sentinel-2. Column 1 is for Landsat 8 products and Column b represents Sentinel 2 products.

4.7 Discussion

The blue-green and red/NIR band ration-based chlorophyll−a OC models illustrate the potential of using high-resolution satellite data (Landsat-8 and Sentinel-2) in the Vaal Dam. The application of band ratios removes systematic errors and minimizes atmospheric effects (Binding et al., 2008, Gilerson et al., 2010). However, scattering effects associated with other WQPs can limit the efficiency of the model. For the Landsat 8 model, the performance of the blue-green model may be limited due to the absorption effects from accessory pigments and CDOM, and also due to backscattering of non-algal organic materials that do not necessarily covary with chlorophyll−a (Bellacicco et al., 2018). The effects of non-algal organic materials are relatively minimal in the longer wavelength (red – NIR) regions and thus allow for optimal use of the red/NIR based band ratio model using Sentinel 2 data. However, the absorption of water also increases exponentially with wavelength (Babin et al., 1996) and this leads to low reflectance values that become more sensitive to effects related to the atmosphere and other random errors, hence lowering the SNR (Sokoletsky et al., 2011). Several studies have shown that OC algorithms come with certain uncertainties. The models are not global and calibration work has always been difficult due to variations in physical factors (Ali et al., 2014a; (Ali et al., 2014a, Kutser et al., 2005).

Despite promising results, the generation of a more stable global model that can accurately retrieve the various WQPs independently poses several challenges. Although satellite sensors can instantly model the optical properties of large water bodies, the calibration and validation data collected in the field cannot sample multiple water points instantaneously. Therefore, temporal variability between the sampling period and the acquisition of satellite data may be large, especially in dynamic turbid waters such as the Vaal Dam. Wind action over the waters can cause mixing, significantly changing the optical and hydrochemical properties of the water column. The small difference in the time frame of data collection may reduce the correlation factor between the retrieved constituents, hence degrading model accuracy. Although in this study, relatively high spatial resolution satellite data were used compared to previous studies (300 m–1 km, pixel resolution), the spatial variability of WQPs may still be large (Gitelson et al., 2008, Gons et al., 2002, Gordon and Morel, 1983, Gordon and Wang, 1994, Moses et al., 2009, O'Reilly et al., 1998). The Landsat and Sentinel data acquired at pixel resolution may be too coarse to correlate with the field-based point measurements and therefore, collocated field data may not represent conditions averaged over the satellite's pixel, limiting model performance. Finally, the effects due to atmospheric scattering can be significant and the performance of satellite-based OC models depends on the accuracy of the applied atmospheric correction. Many atmospheric correction methods including the iterative NIR bands procedure (Gordon and Wang, 1994), the SWIR bands procedure (Wang and Shi, 2005, Wang and Shi, 2007), Bright pixel atmospheric correction (Aiken and Moore, 2000) and the Case 2 regional processer (Doerffer and Schiller, 2007) have been tested to correct for path radiance. No single model has produced absolute correction of path radiance effects and this remains to be a challenge within the remote sensing community.

4.8 Conclusion

The study has illustrated the potential of remote sensing algorithms utilizing high-resolution satellite data to estimate chlorophyll−a concentration in biologically productive in-land open water bodies. The blue-green-based band ratio performed reasonably well when applied to the Landsat 8 data producing $R^2 = 0.89$ and RMSE = 0.36 µg/L. The performance of this model is attributed to the combined effect of the high SNR of chlorophyll−a in the Vaal Dam and the averaging effect of the other associated WQP due to relatively coarse Landsat pixel resolution. Application of the higher resolution Sentinel 2 data gave poor results when using the blue-green region. This may have lowered the SNR. For Sentinel 2 the red/NIR-based model performed well with $R^2 = 0.75$ and RMSE = 0.48 µg/L. This model is less sensitive to non-algal organic matter and more sensitive to chlorophyll−a signal. Landsat 8 lacks the narrow chlorophyll−a sensitive band (665 and 700 nm) in the NR region. The increased spatial and spectral resolution of the new generation sensors along with the incorporation of chlorophyll−a sensitive band significantly improved the chlorophyll−a OC model allowing to effectively monitor productivity in open water bodies such as the Vaal Dam.

Chapter Five

Using Sentinel-2 satellite data and In-situ Hyperspectra for Cyanobacterial Assessment in the Vaal Dam, South Africa

This chapter is based on **Obaid, A.;** Adam, E.; Ali, K.A., Abiye, T. A., Ortez, J. (in Preparation) Using Sentinel-2 satellite data and In-situ Hyperspectra for Cyanobacteria Assessment in the Vaal Dam, South Africa

Abstract: The Vaal Dam, the primary water supplier to the Gauteng province, has experienced frequent algal blooms in recent years. The algal community in the dam water varies spatially and temporally. During the summer months, the bloom is mainly dominated by cyanobacteria, while in the rest of the seasons, an abundance of other species is experienced alongside the cyanobacteria. This study aims to examine Sentiel-2 and in-situ hyperspectral data and different models, including a semi-analytical algorithm introduced by Simis et al. (2005), to detect cyanobacteria concentration in the Vaal Dam. The Chlorophyll–a (Chl–a) concentration has successfully been retrieved by both data sets using band ratio indices in the red-NIR region; the best correlation obtained between band ratio $R_{rs}(705)/R_{rs}(665)$ and the laboratory-measured Chl–a concentrations with $R^2 = 0.78$ for Sentine-2 data, and $R^2 = 0.93$ of in-situ measured hyperspectral data. A semi-analytical algorithm has also been examined, resulting in a strong correlation with measured chlorophyll-a, $R^2 = 0.95$. The cyanobacterial proxy is phycocyanin pigment, with its peak absorption at approximately 620 nm. This wavelength is not available in Sentinel-2 band composition. However, the Sentinel-2 available bands in the range of red to near-infrared wavelengths have been tested to estimate cyanobacteria but resulted in non-significant correlations with laboratory-measured cyanobacteria. The estimated cyanobacteria from in-situ measured hyper-spectra showed strong correlations with laboratory-measured cyanobacteria, $R^2 = 0.96$ for the band ratio $R_{rs}(705)/R_{rs}(655)$ and $R^2 = 0.95$ for the band ratio $R_{rs}(705)/R_{rs}(620)$. The semi-analytical algorithm also successfully estimated cyanobacteria with a strong correlation to measured values, $R^2 = 0.97$. but its estimated values are slightly higher than the measured ones. Hence, some factors used by the model need to be adjusted for the Vaal Dam site.

However, Sentinel-2 bands in the range of red to NIR can successfully be used for Chl-a monitoring in the Vaal Dam which can be indications of cyanobacterial blooms during its domination in summertime. For the rest of the year seasons, the study recommends using the Semi-analytical algorithm with hyperspectral remote sensing for cyanobacterial assessment in highly productive inland waters such as Vaal Dam. The study also recommends examining various satellite data containing a band centred at a wavelength around 620 nm if they have adequate spatial resolutions.

5.1 Introduction

Inland freshwaters are vulnerable to climate change and human activities; they are exposed to several threats from accumulative effects of nutrient enrichment, organic and inorganic pollution, and global warming (Shi et al., 2019, Winslow et al., 2018). The human nutrient enrichment waters (mainly with nitrogen and phosphorus) over the past several centuries, associated with the development in agricultural, industrial, and urban fields, have further accelerated the rates of harmful algal blooms (HABs) and cyanobacteria primary productivity (Paerl and Paul, 2012). When it is in chain equilibrium, the microscopic planktonic algae are a critical component of the aquatic food chain. Still, if it grows out of control as harmful blooms, it will cause several environmental and human health problems (Wu et al., 2019). The most important consequence is the production of biotoxin by some species (Blue-green algae), which has detrimental effects on human health (Davidson et al., 2011). Many studies indicated that the blue-green algae or cyanobacterial blooms increased globally in magnitude, frequency, and duration (Huisman et al., 2018, Carey et al., 2012).

From the early 2000s up to 2016, only 21 out of all African countries were found to have research information regarding cyanobacteria, and 11 of them recorded toxicity related to cyanobacteria. The dominant genus was *Microcystis*, and the most noticeable finding was that the more toxic blooms related to higher nitrogen concentrations in the water column (Ndlela et al., 2016).

Among the African countries, South Africa has more documented reports on cyanobacterial toxic blooms, with the dominant species found to be *Microcystis aeruginosa* (Oberholster et al., 2009a, Oberholster et al., 2009b, Oberholster et al., 2005, Oberholster and Botha, 2007, Oberholster, 2009, Nchabeleng et al., 2014), *Anabaena* (Oberholster et al., 2015) and *Oscillatoria* (Ballot et al., 2013). This increasingly problematic situation requires a consistent monitoring of inland freshwater bodies. One of the essential critical parameters of HABs, which can be used as the leading indicator of the phytoplankton biomass and then the eutrophic status of the water, is chlorophyll-a pigment (Ansper and Alikas, 2018).

Recently, there has been increasing demand for an effective way to determine the occurrence of blue-green algae in water bodies related to cyanotoxin production. The pigment related to this toxin-productive type of algae is the phycocyanin pigment. There are many ways to measure chlorophyll-a concentrations and thus monitor algal blooms in the water bodies; the traditional in-situ monitoring through collecting water samples and analysing them in the laboratory is perfect and very effective in counting the different species of algae in the water samples, the only issues are; labour-intensive method and needs well-trained crew for sampling and saving the samples in similar conditions till reaching the laboratory, consumes time and money mainly if been used in regular basis for water quality monitoring and it lack of spatial coverage (Ansper and Alikas, 2018, Nazeer et al., 2020).

Remote sensing is another method that allows people to observe water bodies regularly. It can estimate the water quality parameter in the non-accessible part of water bodies, and it is beneficial for temporal and spatial coverage of the water bodies (Ansper and Alikas, 2018). Remote sensing is a valuable tool for deriving cyanobacterial and phytoplankton pigments by measuring the number of various radiation wavelengths reflecting from the water surface using spaceborne or airborne sensors, the reflected solar radiation energy spectrum from the water body interacted with its optically active constituents, which form the majority of surface water important parameters (Haji Gholizadeh et al., 2016). Remote sensing for water quality is suitable for regional studies because it covers large geographic areas with each scene. Still, it can be used effectively in relatively more minor water bodies. Currently, some satellite images are free, and some are low-cost. Many of them are very useful in water quality multi-temporal studies to reveal the history of HABs changing in such water bodies and, thus, for HAB monitoring. Operational Land Imager (OLI) and Multispectral Imager (MSI) on board Landsat-8 and Sentinel-2 provide high spatial-resolution data valuable for aquatic studies. The data are accessible on USGS Earth Explorer and European Copernicus websites.

Vaal Dam is one of the many water bodies that frequently experience HAB blooms in South Africa. According to a previous study, managers of the drinking water treatment works (DWTW) rely on the chemical, physical and biological results of monthly water samples grabbed from the Dam reservoir for their decisions regarding water treatments and management (Swanepoel et al., 2016) which are effective in identifying severity and conditions of the blooms and other water quality parameters in each single sampling point. However, it takes some time to get the results back from the laboratory; this might vary from a few hours to a couple of days; at the same time, it’s not possible with in-situ sampling to cover the spatial distribution of the blooms throughout the reservoir area. Therefore, different types of estimating and monitoring approaches must be checked to overcome the limitations of the existing methods in the Vaal Dam. Some previous studies have been conducted on the Vaal Dam water using remote sensing approaches, and all of them successfully retrieved Chl−a concentrations from Landsat-8 and Sentinel-2 satellite data, none of them examined the estimation of the specific phycocyanin pigment as a proxy of cyanobacterial biomass. This study attempts to fill this gap by testing some remote sensing-based algal and cyanobacterial estimation approaches using regionally tuned band ratio indices, which can be implemented for routine monitoring of Chl–a (Sakuno et al., 2018a, Obaid et al., 2021), and the semi-analytical algorithm developed by Simis et al. (2005) which has consistently demonstrated superior performance compared to other existing PC retrieval algorithms for various turbid inland lakes and reservoirs (Lyu et al., 2013). This can contribute towards increasing understanding and interpretation of the spatial and temporal dynamics of algae and cyanobacteria blooms in the Vaal Dam. This is significant to improve the capability of the existing methods and previously achieved approaches for monitoring the reservoir water quality.

The objectives of this study were i) to examine the applicability of Sentinel-2 available bands data for detecting the concentration of cyanobacteria in the dam reservoir, ii) to examine the potential use of the semi-analytical algorithm in monitoring cyanobacteria in the Vaal Dam water using hyperspectral data.

5.2 Study area.

Vaal Dam lies approximately 56 km to the south of Johannesburg (Sakuno et al., 2018b), 77 km from OR Tambo International Airport (https://en.wikipedia.org/wiki/Vaal_Dam) accessed July 6 2023. At latitude of 26.88444867 S, longitude of 28.12059553 E and elevation of approximately 1480 meters above mean sea level (Chutter, 1970). The main arteries for the Vaal Dam catchment, which supplies fresh water to the Vaal Dam Reservoir, are the Vaal and Wilge River systems. These two river systems consist of many tributaries draining different areas in terms of land use, land cover types, and human activities, from which return flows enriched with nutrients and pollutants find their way to such streams that feed the Vaal Dam Reservoir. In its total capacity, the reservoir surface area will reach 320 km^2, and its maximum depth will reach 47 m; the water is turbid. The South African Department of Water Affairs (DWA) classified it as mesotrophic according to its mean total phosphate and mean chlorophyll-a concentration, which is 0.077 mg/l and 14.8 ug/l, respectively (Swanepoel et al., 2016). The reservoir lies in a subtropical highland (Highveld region of South Africa), characterised by a dry winter; its most annual rainfall occurs between October and April in the summer (Landman et al., 2017). The average annual rainfall is around 681 mm, and the mean average yearly temperature is 15.7 °C (Obaid et al., 2021). The reservoir is used to supply Zuikerbosh Water Treatment Plant which supplies potable water to the metropolitan areas in Gauteng Province. Around 80% of Zuikerbosh's raw water comes from the Vaal Dam Reservoir (Chinyama et al., 2016).

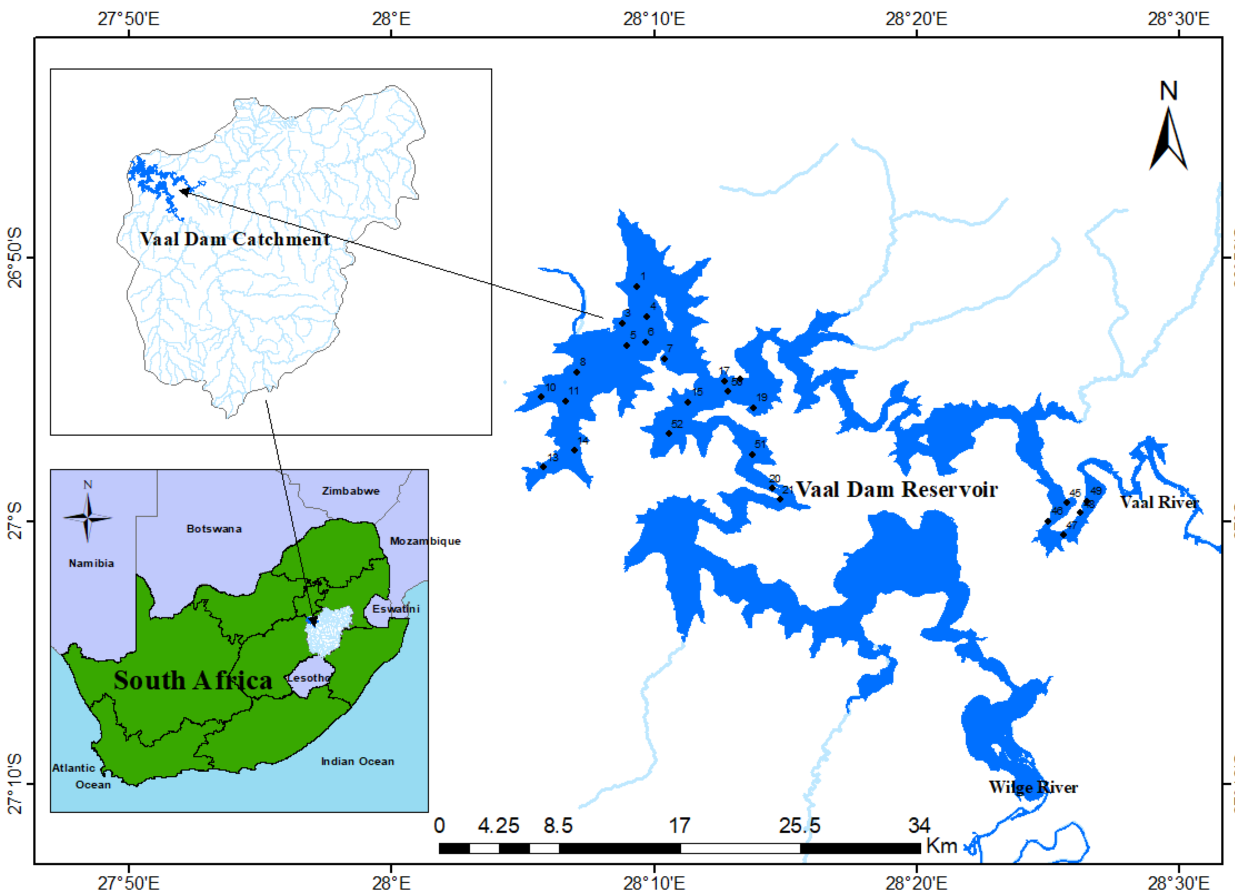


Figure 5.1 The location map of the study area including the sampling points.

5.3 The Data Used

5.3.1 In-situ Water Samples

In situ sampling campaigns have been carried out between June 21 and June 27, 2024. The sampling sites have been pre-determined to span the most critical areas within the reservoir. GPS (Garmin Montana 650) has been used to access the measuring points (see Figure 5.1).

The Secchi disk depth (SDD) was collected to estimate the water clearness and water samples were collected, saved in triplicate one-litre bottles, preserved in darkness, and refrigerated. Seventeen water samples were collected during the sampling campaigns on 21, 25, 26 and 27 June 2023.

Then, the water samples were sent to the South African National Accreditation System (SANAS), Testing Laboratory T0625, upon arrival for analysis.

5.3.2 In-situ Probes Data

Water quality parameters were measured at each sampling point using the Eureka Water Probe (Hydrolab) model; Manta+40 measures multi-water quality parameters, including chlorophyll-a, turbidity, temperature, pH, conductivity, dissolved oxygen, and it can add more sensors. Also, a bbe FluoroProbe has been used; the Flouraprobe instrument measures the concentration of algal types, including green algae, blue-green algae (cyanobacteria), diatoms and cryptophyta. The probe data has been recorded on laptop computers.

5.3.3 In-situ Spectral Data

Above-water spectral measurements were adopted using ASD FieldSpec HandHeld 2 hand-held spectroradiometer (Wavelength Range: 325 – 1075 nm, Spectral resolution <3.0 nm @ 700 nm). The sensor was connected to a Dell laptop computer equipped with the required software via a fibre connector to initiate spectral scanning of the ASD, visualise the output spectral data and store it. The sky conditions were evident during the sampling within the field campaigns. At each sampling point, total upwelling radiance, L_t (W m^{-2} sr^{-1}) was recorded, ASD was pointed at a nadir-viewing angle above the water surface, and the instrument was attached to an extendable pole to eliminate the boat influence. The measured total upwelling radiance was then corrected to remove the influence of the sky radiance reflected by the water surface (L_{sky}). Downwelling radiance L_d(W m^{-2} sr^{-1}) was also collected at each sampling point (Randolph et al., 2008). At each point, 30 valid scans were collected, consisting of an average of eight measures. The radiometer acquired a measurement scan in about 136 to 544 milliseconds, depending on the radiation strength, and a spectral curve was displayed for each scan and saved to the computer. Any erroneous scans were treated as invalid and replaced with valid scans.

5.3.4 Satellite Data

Sentinel-2 Multispectral Imager (MSI) level 1C images were acquired from ESA Copernicus Open Access Hub server (https://www.copernicus.eu/en/access-data/conventional-data-access-hubs), two images were downloaded; Sentinel-2 A taken on June 20th and Sentinel-2 B taken on June 25th, 2023, they were found to match the day of the field data acquisition or (+/– 1 day).

Table 5.1 Sentinel-2 specifications including band number, band central wavelength (nm), bandwidth (nm), band range, and band spatial resolution (m)

Band Number	Band Centers (nm)	Bandwidth (nm)	Band Range	Spatial Resolution (m)
1	443	20	VIS	60
2	490	65		10
3	560	35		10
4	665	30		10
5	705	15	NIR	20
6	740	15		20
7	783	20		20
8a	842	115		10
8b	865	20		20
9	945	20		60
10	1375	30	SWIR	60
11	1610	90		20
12	2190	180		20

5.4 Data Processing and Analysis

5.4.1 Sample testing and data processing

Chlorophyll-a and microcystin concentrations in water samples were tested in the SANAS Testing Laboratory T0625 using the spectrophotometric method, and the results were recorded in micrograms per litre. Moreover, the algae and cyanobacteria cell counts were analysed in the laboratory using the membrane filter method, and results were provided as a count of cells per millilitre. Then, they were used to validate the estimated chlorophyll-a and phycocyanin pigments.

5.4.2 In-situ Spectral Data Processing

Invalid reflectance scans have been removed from the data sets. Then, normalised data was arranged by the sampling station and saved on a spreadsheet. An average reflectance value was calculated for eight scans above the water surface for every wavelength and sorted by sample station and date.

Following the suggestion of the NASA Ocean Optics protocols 2003, Chapter 3, Method 2, the water-leaving radiance (L_w) was calculated from the surface and sky radiance measurements using the equation below.

$$L_w(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o) = F_L(\lambda)[L_{sfc}(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o) - \rho L_{sky}(\lambda,\theta',\phi' \in \Omega'_{FOV};\theta_o)$$

Where $F_L(\lambda)$ is the instrument's unknown radiance response calibration factor, and $L_{sfc}(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o)$ and $L_{sky}(\lambda,\theta',\phi' \in \Omega'_{FOV};\theta_o)$ are the radiometer's measured responses. ρ is the reflectance factor.

$$\rho = \frac{L_{sky}(\theta,\phi)(fully\ shaded)}{L_{sky}(\theta',\phi')}$$

The incident spectral irradiance $E_s(\lambda;\theta_o)$ was calculated using the following equation.

$$E_s(\lambda;\theta_o) = \frac{\pi}{R_g(\lambda,\theta'',\phi'' \in \Omega_{FOV};\theta_o,\phi_o)} \times F_L(\lambda)L_g(\lambda,\theta'',\phi'' \in \Omega_{FOV};\theta_o,\phi_o)$$

Where $L_g(\lambda,\theta'',\phi'' \in \Omega_{FOV};\theta_o,\phi_o)$ is the sensor response signal when the plaque is viewed at angles (θ'',ϕ'') with the sun at (θ_o,ϕ_o), and $R_g(\lambda,\theta'',\phi'' \in \Omega_{FOV};\theta_o,\phi_o)$ is the plaques bidirectional reflectance function (BRDF) for that sun and viewing geometry this study, the sensor was pointing vertically to view the centre of the white plaque normal to its surface, and BRDF were determined for illumination angles between normal and 90° at 5° increments.

The $R_{RS}(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o)$ then calculated using the following equation.

$$R_{RS}(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o) = \frac{L_w(\lambda,\theta,\phi \in \Omega_{FOV};\theta_o)}{E_s(\lambda;\theta_o)}$$

5.4.3 Satellite Data Processing

The Sentinel-2 images have been processed with the Sentinel Application Platform (SNAP). SNAP is free software for processing and analysing satellite data. It includes tools to process Sentinel mission and other satellite data. The software consists of several methods of atmospheric correction, one of them, Case 2 Regional Coast Color (C2RCC), developed for aquatic subjects. It has a version designed for turbid waters, the Case-2 Extreme Cases (C2X) process, which was adopted to process the continental Case-2 waters (Jeffrey and Humphrey, 1975).

In this study, the C2X–COMPLEX processor was chosen to perform the atmospheric correction due to the turbidity of the water at the studded site. Additionally, SNAP has many built-in algorithms, each with different biophysical variables, which can allow the development of new algorithms (Toming et al., 2016), The data were resampled to 10 m spatial resolution for further processing, all bands need to have the exact spatial resolution, which is necessary to run the atmospheric correction. After resampling, the images were subset to the area of interest before facilitating the atmospheric correction process. Values of maximum transparency, the concentration of Chl–a, and total suspended sediments have been calculated as automatic results during the atmospheric correction process. Then, the band pixel values were extracted from each sample point.

5.4.4 Calibration

Using the remote sensing reflectance values, various indices from different wavelengths in the regions of red/near infra-red, and blue/green have been calculated and correlated with the measured Chl–a and cyanobacterial concentrations using regression analysis to find the best fit to the measured data. Sentinel-2 band centres have been used for both multispectral and hyperspectral data; two of such indices were successfully used in previous studies and have been tested in this study as well (Table 5.2).

In addition, a semi-empirical bio-optical model for inland turbid waters adopted from *Simis et al. (2005)* has also been tested for Chl–a and PC pigments using in-situ hyperspectral data. The model is based on the relationship between the subsurface irradiance reflectance (R(0−,λ)) and the absorption coefficient and backscattering of the water column's optically active constituents. The model uses Chl–a peak absorption at 665 nm and PC peak absorption at 620 nm to estimate their concentrations in the water column. Equations 1 & 2 were introduced to estimate Chl–a and PC absorption coefficients.

$$a_{Chl}(655) = \left(\left\{\left[\frac{R(709)}{R(665)}\right][a_w(709) + b_b]\right\} - b_b - a_w(665)\right)\gamma^{-1} \quad \text{....(1)}$$

$$a_{PC}(620) = \left(\left\{\left[\frac{R(709)}{R(620)}\right][a_w(709) + b_b]\right\} - b_b - a_w(620)\right)\delta^{-1} - [(\varepsilon)a_{chl}(665)] \quad \text{....(2)}$$

Where:

$a_{chl}(665)$ = The chlorophyll-a absorption at 665 nm

$a_{PC}(620)$ = The phycocyanin absorption at 620 nm

R(λ) = The reflectance value of the specified wavelength

$a_w(\lambda)$ = The coefficients of the pure water absorption at specified locations ($a_w(709) = 0.727\ m^{-1}$, $a_w(665) = 0.401\ m^{-1}$, $a_w(620) = 0.281\ m^{-1}$, $a_w(778) = 2.71\ m^{-1}$)

γ = 0.68, a constant to estimate chlorophyll-a absorption,

δ = 0.84, correction factor to estimated phycocyanin absorption,

ε = 0.24; the proportion of absorption at 620 nm attributed to chlorophyll–a,

γ , δ and ε values have been derived from Lake Loosdrecht data and bee used in Simis et al., 2005 model. (Randolph et al., 2008)

The value of b_b can be retrieved from the reflectance at 778 nm using the following equation.

$$b_b = [a_w(778)\alpha R(778)][\gamma' - \alpha R9778)]^{-1} \text{ ....(3)}$$

This study adopted pure water absorption and other constant values, as well as Chl–a and PC-specific absorption coefficients at 665 and 620 nm, respectively, from Simis et al., 2005.

The ratio of the optically active constituent absorption coefficient, $a_i(\lambda)$, to its specific absorption coefficient, $a_i^*(\lambda)$, yields the estimate of pigment concentration. (Gons, 1999; Simis et al., 2005) Therefore, chlorophyll-a and phycocyanin concentrations can be determined from:

$$\mathrm{Chl} = \frac{a_{\mathrm{Chl}}(665)}{a^*_{\mathrm{Chl}}(665)}$$

$$\mathrm{PC} = \frac{a_{\mathrm{PC}}(620)}{a^*_{\mathrm{PC}}(620)}$$

Where:

$a_{\mathrm{Chl}}(665)$ Chlorophyll-a absorption at 665 nm, estimated by Eq. (1)

$a^*_{\mathrm{Chl}}(665)$ 0.0153 m^2 mg Chl^{-1}, The average specific absorption value of the chlorophyll-a at 665 nm for several lakes and reservoirs in Spain and The Netherlands (taken from Simis et al., 2005 and 2006)(Randolph et al., 2008)

$a_{PC}(620)$ Phycocyanin absorption at 620 nm, estimated by Eq. (2)

$a^*_{PC}(620)$ 0.0070 m^2 mg PC^{-1}, The average specific absorption value of the phycocyanin at 620 nm for several lakes and reservoirs in Spain and The Netherlands (Randolph et al., 2008)

Table 5.2 Empirical indices can be used as reference for this study and their corresponding values. N: samples number, R^2 coefficient of determination, RMSE: root mean squire error.

Sensor	Atmospheric Correction	Bands Relation	N	R^2	RMSE	Data Range
S2-MSI	Sen2cor	R740/R665	21	0.84	141	10–1287 mg/m^3
S2-MSI	C2RCC	R705/R665	44	0.75	0.48	3.9–13 mg/m^3

After the normality test of the reflectance data was verified, they were correlated with the Chl–a and PC concentrations using linear regression analysis, and the adjustment was performed using the best ratio. The determination coefficients (R^2) were obtained for all indices and models tested.

5.4.5 Validation

The validation processes were performed for all indices and models to find the best fit for this specific study case. Some adjustments were made to relate the estimated Chl–a and PC values to the observed ones. The index or a model with the best R^2 and root mean square error (RMSE) was adopted to apply to this case study.

5.5 Results

5.5.1 Summary of algal types in the Vaal Dam water

The laboratory analysis showed that around six algal species were detected in the water samples. Among such species, only Microcystis sp. has been detected in all water samples in highly variable cell counts ranging from 39 to 440554 cells/ml. The WHO classification for cyanobacteria states that a cell density below 200 cells/ml is categorised as a surveillance level, 2000 cells/mL is alert level one, and 100,000 cells/ml is alert level two (Randolph et al., 2008). Depending on such classification, the spatial variation of alertness in the Vaal Dam water was evident in June 2023. Among 17 sampling sites assessed, five exhibited cyanobacterial concentrations exceeding 2000 cells/ml, with two sites displaying exceptionally high levels of 217447 and 440554 cells/ml. Cyanobacteria exists in most dam sites as part of the algal group, consisting of 40% of the total algal biomass, and this ratio raised to 48% when excluding sites 45 – 49, which are dominated by diatoms and no significant abundance of cyanobacteria there. Cyanobacteria only dominated two sampling sites, points 13 and 14 (Figure 5.2).

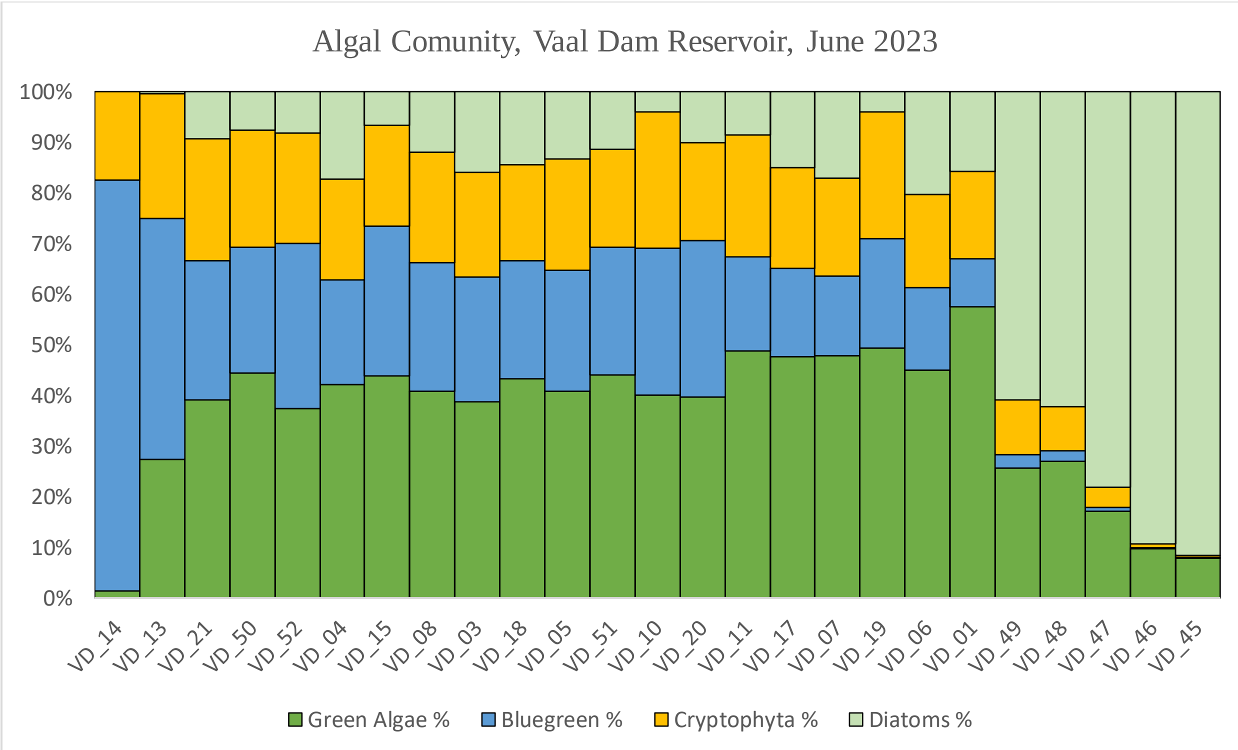


Figure 5.2, Shows the composition of the algal community in the Vaal Dam Reservoir in June 2023, the community includes green algae, blue-green algae, cryptophyta and diatoms.

The fluoroprobe dataset showed four groups of algae in the Vaal Dam, including green algae, blue-green algae diatoms and cryptophyta, as yellow substances. The results showed significant spatial variations in the algal community compositions throughout the dam reservoir (Figure 5.2 and Table 5.3); the blue-green algae showed an extensive range from 0.01 to 45.97 µg/L, while the green algae showed a lower range, between 0.84 and 5.89 µg/L. In the eastern part of the reservoir (points 45 – 49), the blue-green algae are in deficient concentrations (0.01 to 0.25 µg/L) at the time of the measurements. In the rest of the reservoir, its concentration varies, reaching its domination over other species in points 14 and 13 (55.89 and 9.94 µg/L respectively). A positive correlation (R^2=0.89, n=17) between the laboratory cyanobacterial cells/mL and the in-situ cyanobacteria measured by Fluoroprobe µg/L, and approximately a similar correlation was obtained from the relationship between the laboratory-measured chlorophyll–a concentrations and the fluoroprobe total concentration of algae (R^2=0.88, n=17) see Figures 5.3 A and B. The mean ratio of total cyanobacteria to total algae is 0.4, indicating the proportion of total algal biomass attributed to cyanobacteria. Measurements by Hydrolab sonde revealed a reasonable range of Chl-a concentrations and high turbidity values see Table 5.3.

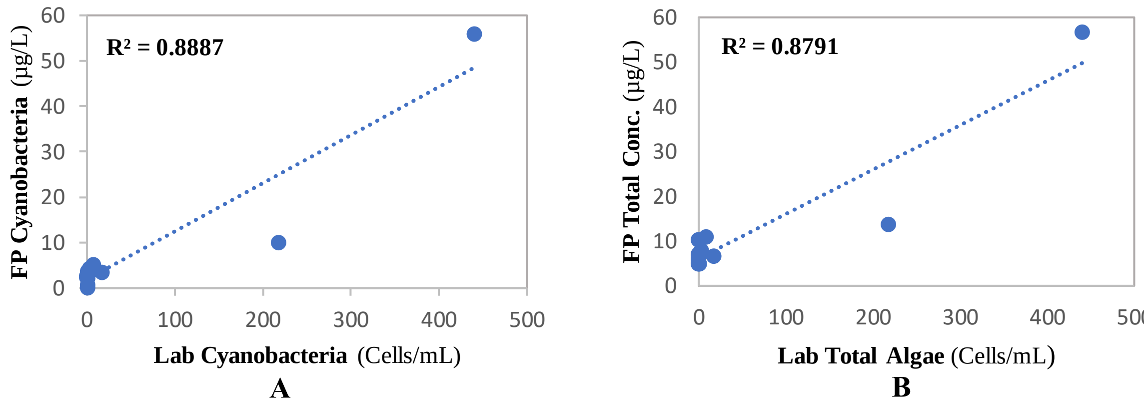


Figure 5.3 The relationship between the laboratory total counts and in-situ fluoroprobe total concentrations of algae and cyanobacteria, **A.** Cyanobacteria and **B.** Total algae concentration.

Table 5.3 Statistical summary of some water quality parameters in the Vaal Dam Reservoir including Secchi Depth (cm), Chlorophyll-a (µg/L), Turbidity (FNU), Green Algae (µg/L), Bluegreen (µg/L), Diatoms (µg/L), Cryptophyta (µg/L), Total concentration of algae (µg/L), and the ratio of Bluegreen to Green Algae.

WQ Parameter	n	Maximum	Minimum	Mean	SD
Secchi Depth (cm)	25	83	21	37	14.1
Chl-a Laboratory	17	70	1.2	18.55	23.94
Chl-a Hydrolab	24	15.33	2.37	4.14	2.75
Turbidity (FNU)	24	77.51	7.06	40.7	18.16
Green Algae (µg/L)	25	5.89	0.84	2.96	1.03
Bluegreen (µg/L)	25	45.97	0.01	3.28	8.99
Diatoms (µg/L)	25	33.27	0.00	3.78	7.92
Cryptophyta (µg/L)	25	9.92	0.15	1.72	1.84
Total conc. (µg/L)	25	56.72	4.89	11.74	11.72
Total cyano/Total algae	25	0.99	0.01	0.40	0.22

5.5.2 The Vaal Dam water spectra

The spatial variations of the dam water highly influence the intrinsic optical properties. Significant variations in the spectra related to different dam parts are noticed. Three different spectral patterns are detected; the first pattern is where the diatom dominates the water at the most eastern part of the dam (points 45 – 49, the blue circle, Figure 5.4); the most noticeable feature is the chlorophyll-a pigment absorption at approximately 675 nm and the peak reflectance at around 575 nm (Figure 5.5C). The second pattern includes three points (13, 14 and 50, black circles, Figure 5.4) where the cyanobacteria are dominated over other algal species in such points and the phycocyanin pigment absorption feature is noticed at around 625 nm, and the peak reflectance is at around 709 nm (Figure 5.5D). The third pattern is the remaining sampling sites where the cyanobacteria exist alongside other algal species, and the chlorophyll-a absorption feature at 675 is noticeable, and the peak reflectance is at around 660 nm (Figure 5.5E). At the same time, there is no clear absorption feature of phycocyanin at 620 nm.

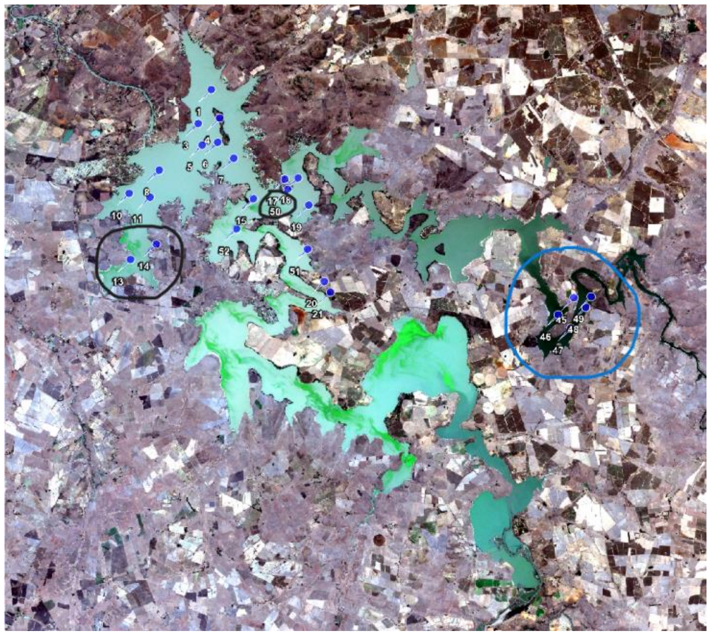
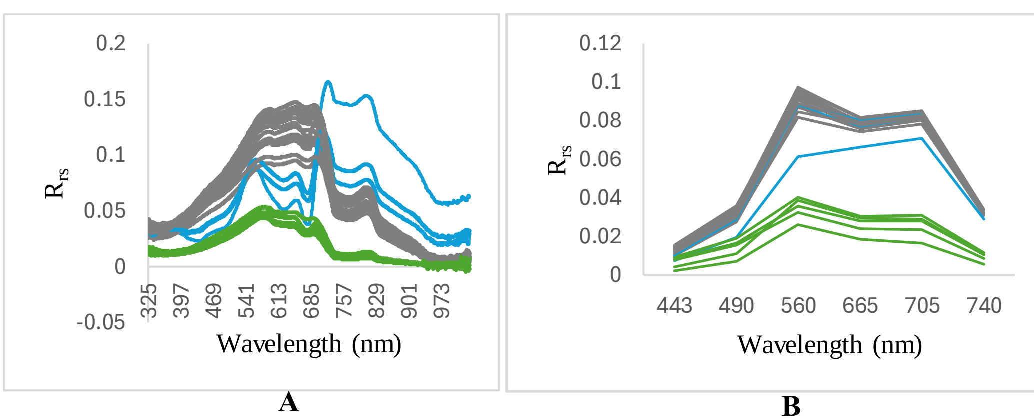
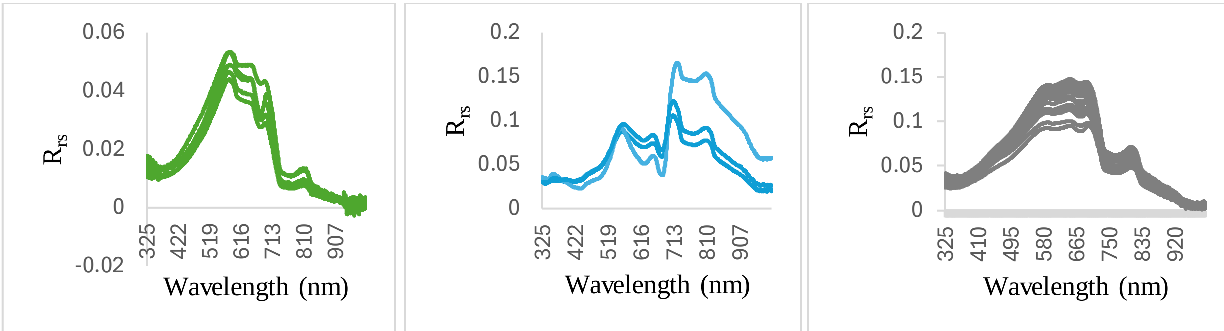


Figure 5.4 RGB satellite image (Bands 4-3-2) view showing the bloom of the algal community in the Vaal Dam Reservoir (June 25th, 2023), Cercles showing different inherent optical properties sites.

The Sentinel-2 spectra showed the same behaviour as the hyperspectral data; it showed lower reflectance values for the sampling points 45 – 49, while the cyanobacterial features do not appear due to the lack of the PC peak absorption wavelength in the Sentinel-2 bands (Figure 5.5B).



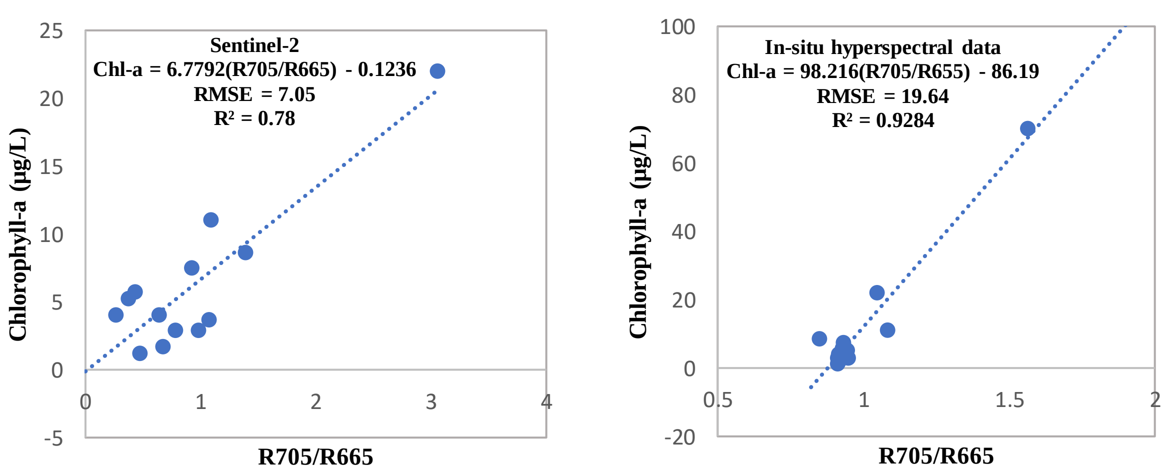


C **D** **E**

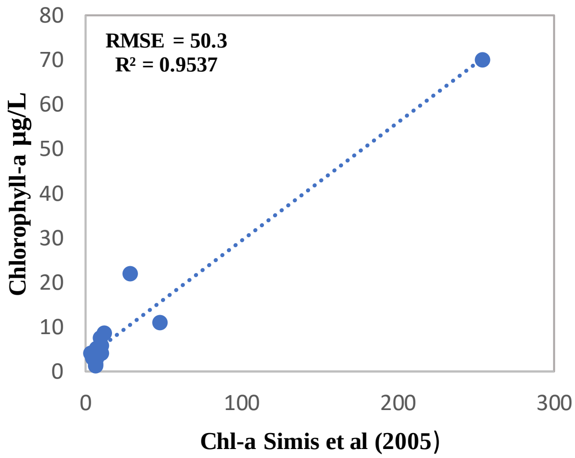
Figure 5.5 Remote sensing spectral data of the Vaal Dam, A. In-situ hyperspectral, B. Sentinel-2 MSI multispectral, C. Multi algal species spectra, D. Diatom dominated sites spectra, and E. Cyanobacteria dominated sites spectra.

5.5.3 Bio-optical models for Chlorophyll–a

Band ratios in red/NIR spectral ranges have been examined on remote sensing reflectance from sentinel-2 multispectral satellite images and in-situ hyperspectral measurements, which are more sensitive to Chl–a concentrations in the water; the data were collected during the satellite overpass while the hyperspectral data were collected during the sampling campaigns. Blue/green and red/NIR band ratios were examined using the Sentinel-2 band centres for multispectral and hyperspectral data. The regression analysis showed a weak correlation between blue-green ratios and Chl–a for Sentinel-2 data and gave the best results between the measured Chl–a and red-NIR ratio (705/665), $R^2 = 0.78$, n = 14, and RMSE = 7.05 µg/L for Sentinel-2 data (Figure 5.6A). And it gives R^2 of 0.92, n = 14, RMSE =19.64 µg/L using hyperspectral data (Figure 5.6B).



A **B**



C

Figure 5.6 The best correlation between measured and retrieved chlorophyll-a from **A.** Sentinel-2 MSI Rrs705/Rrs665, **B.** In-situ measured reflectance Rrs705/Rrs665 and **C.** Simis et al (2005) model.

A semi-analytical algorithm for retrieval chlorophyll–a depending on the ratio between its absorption at 665 nm ($a_{Chl}(665)$) to its specific absorption ($a^*_{Chl}(665)$) was applied to the in-situ hyperspectral data measured in the Vaal Dam reservoir, the algorithm gives a solid relationship (R^2=0.95, n=15, RMSE = 50.3) with the laboratory-measured chlorophyll-a (µg/L) and the model estimated values (Figure 5.6C).

5.5.4 Bio-optical models for Phycocyanin

The phycocyanin pigment serves as an indicator of cyanobacterial biomass in the water, it absorbs light at a wavelength of approximately 620 nm. Since Sentinel-2 MSI lacks a band specifically centred around this wavelength, numerous studies have attempted to estimate phycocyanin concentration using the available bands of MSI. Previous studies have explored various indices, some of which involve subtracting the remote sensing reflectance at $R_{rs}(665)$ from the reflectance in the red edge bands, such as $R_{rs}(740) - R(665)$ and $R_{rs}(705) - R_{rs}(665)$. However, these two methods cannot be applied in this study due to the complexity of the water in the Vaal Dam, where the reflectance at 665 nm is higher than that at 705 nm and 740 nm. Therefore, two indices (Table 5.2) were examined alongside Simi's algorithm for cyanobacterial PC retrieval. The regression analysis of the band ratios mentioned above yielded a strong correlation between both of them and the cyanobacteria measured in the laboratory (cells/mL), with a slightly higher ratio obtained from the relationship of $R_{rs}(705)/R_{rs}(665)$ with the measured cyanobacteria (cells/ml), R^2 = 0.96, n = 17, RMSE = 118 cells/ml (Figure 5.7A).

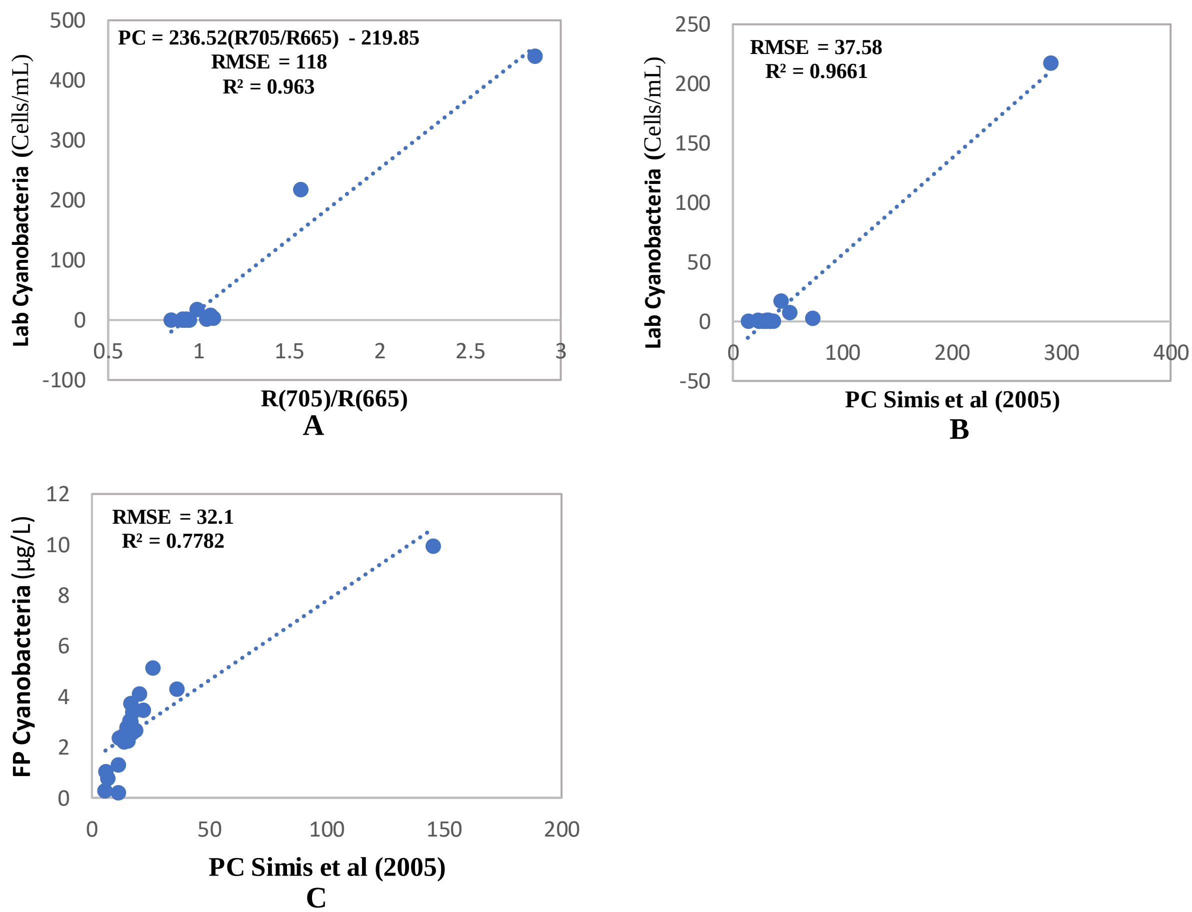


Figure 5.7 The best correlation between the measured and retrieved cyanobacteria, **A.** Laboratory measured and the estimated from the reflectance ratio $R_{rs}705/R_{rs}665$, **B.** Laboratory measured and the estimated from the Simis et al (2005) model, and (C) Fluoroprobe in-situ measured and Simis et al (2005) model.

Application of the reflectance band–ratio algorithm for retrieval of phycocyanin using the ratio of its absorption at 620 nm to its specific absorption ($a^{*}_{PC}(620)$) results in a strong correlation between the Fluoroprobe measured cyanobacteria (µg/L) and the estimated phycocyanin pigment (µg/L) using the model, it gave $R^2 = 0.78$, n = 24, and RMSE = 32.1 (Figure 5.7C) and showed a solid relationship with laboratory-measured cyanobacteria (Cells/ml) yielded R^2 values of 0.97, n=16, and RMSE = 67.3 Cells/ml (Figure 5.7B). A single outlier in the estimated phycocyanin exists; it gives a negative value of phycocyanin and has been excluded from the correlation. The most important notice is the reflectance band–ratio algorithm overestimated phycocyanin concentration in the Vaal Dam compared to the chlorophyll-a model, which gave values very similar to the measured ones.

Applying the reflectance band–ratio algorithm for phycocyanin retrieval from $a_{PC}(620)$ using Vaal Dam Reservoir data yielded an overestimation of the phycocyanin pigment concentration despite having a high correlation with in-situ data.

5.6 Discussion

Figure 5.2, and Table 5.3 from the results above reveal that there are significant spatial variations in the algal community compositions throughout the dam reservoir; the most eastern side of the reservoir at the inlet of Vaal River is characterised by the domination of diatoms with low concentrations of green algae and very low or absence of blue-green algae. The water in such a site contains less suspended matter and it is more transparent than the remaining parts of the reservoir, with an average Secchi-disk depth of 60 cm compared to 30.7 cm in the remaining parts of the reservoir, see Figure 5.4. The algal biomass composition is highly varying spatially; in most parts of the reservoir, the algal community is composed of many species abounding together, including Microcystis and Cryptophyta, the two forms of the cyanobacteria. Only three sampling sites were found to be dominated by cyanobacteria.

Moreover, the dam water was in highly turbid condition during the field data collection; this situation strongly affected the water's optical properties within the reservoir. The in-situ measured spectra agreed with these significant variations; they exhibited some productive inland water's distinct spectral characteristics. According to the literature, the unimodal peak of the reflectance spectra between 400 and 600 nm is attributed to absorption by Chl–a, which causes low reflectance between 400–500 nm with a strong effect of CDOM and other non-algal constituents (Menken et al., 2006, Shi and Wang, 2009). In the turbid eutrophic waters, three reflectance peaks could be identified; at around 570, 700 and 815 nm with occasional reflectance trough at around 625 nm associated with phycocyanin pigment absorption, and another well-identifiable trough at 675 nm which is attributed to strong absorption by Chl–a (Sun et al., 2013). The water reflectance in this study showed the same behaviours (see Figure 5.5A).

Three pattern groups can be noticed in the Vaal Dam in-situ spectra, associated with the abovementioned spatial variations of the dam water. The first pattern group, including the spectra of points 45 to 49 (Figure 5.5C, Green), which is located in the eastern part of the reservoir, their reflectance values, in general, are much lower than the spectral values from the other parts of the reservoir, this is due to high absorption in this part comparing to the different parts which affected by the highly turbid waters causing more spectral scuttering and reflectance. In such a pattern, the peak reflectance is around 575 nm. The reflectance dropped at around 675 nm indicating Chl–a an absorption, there are no clear apparent absorption features at 625 nm compared to what appears in waters dominated by cyanobacteria.

The second group, including points 13, 14 and 50, which are dominated by cyanobacteria (see figure 5.5D, Blue), their spectral shape is typically similar to the reflectance from waters dominated by cyanobacteria, showed high reflectance at approximately 560 nm and three low reflectance values, strictly between 400 – 500 nm, a reflectance trough at 625 nm indicating to the phycocyanin absorption, and a reflectance trough at 675 nm indicating to the common Chl–a absorption in inland turbid waters. Moreover, the literature concluded that a visible scattering peak at 709 nm is considered a cyanobacterial primary reflectance spectrum characteristic in inland waters. In the cases of extreme domination of cyanobacteria, the scuttering peak at 709 nm will shift toward longer wavelengths (Matthews and Odermatt, 2015, Shi et al., 2019, Duan et al., 2010, Gower et al., 2005, Vishnu Prasanth et al., 2022). In this study, the three sampling points dominated by cyanobacteria showed the same behaviour: a high scuttering peak at 711, 713, and 724 nm, which supports the excessive abundance of cyanobacteria in such sampling points.

Finally, the third pattern group represents the remaining sampling sites (Figure 5.5E, Grey), which contain multi-algal species; its pattern showed high reflectance values compared to the first pattern and their peak value was a bit different from the first and second patterns; the peak shifted to around 653 nm, this value is a peak reflectance feature of phycocyanin pigment (Cheng et al., 2013) which suggests the effect of cyanobacteria in this spectral pattern beside the impact of the high turbidity and mixture of algal species.

The proxy of the cyanobacteria PC pigment is determined by secondary spectral characteristics from the reflectance spectra, with maximum absorption around 620 nm and a reflectance peak near 640 nm. Unfortunately, Sentinel-2 MSI has no band centred around these wavelengths where PC affects the reflectance features. Another challenge is the simultaneous increase or decrease in the red reflectance between 600 and 700 nm due to other constituents' absorption and reflectance; this will affect the stability of the satellite-derived PC reflectance characteristics from multispectral images. When the Chl–a concentration exceeds 100 μg/L, and in the case of surface scum in cyanobacterial bloom, the Chl–a reflectance peak at 753 nm exceeds the scuttering peak at 709 nm. The reflectance peak at 753 nm is used as an indicator for surface scum conditions of the cyanobacterial community (Matthews and Odermatt, 2015).

Although Sentinel-2 multispectral data is valuable for chlorophyll-a detections in the water column, using it in the inland aquatic environment still has some challenges because of the significant bio-optical complexity of such waters. For example, other water quality constituents such as CDOM, detritus, and tripton significantly affect the Chl–a signal reflectance in the blue-green regions (Mishra and Mishra, 2012). In this study, the blue/green ratios seem strongly overlapped by non-algal spectral attenuators making it very difficult to isolate the spectral signature of Chl–a from the signature of other WQPs.

Red to near-infrared spectral regions perform much better to estimate chlorophyll-a in highly productive waters, benefitting from the advantage of the impact of phytoplankton on the optical properties of water in these spectral regions (Tran et al., 2023). However, for Sentinel-2 data, other issues related to the impact of the atmosphere create some attenuations resulting in lower R^2 values in contrast to in-situ hyperspectral reflectance when correlating to in-situ data. This was achieved in this study, the sentinel-2 red/NIR band ratio performed very well in estimating Chl–a with R^2 values a bit lower than those from in-situ hyperspectral data. Another advantage of hyperspectral data over sentinel-2 multispectral data is it captures data across hundreds of narrow spectral bands, allowing for more precise identification of chlorophyll-a and phycocyanin estimations.

The average specific absorption coefficient of phycocyanin at 620 nm ($a^*_{PC}(620)$) has been calculated from various lakes and reservoirs in The Netherlands and Spain; details can be found in Simis et al. (2006). This average specific coefficient of phycocyanin can be used to estimate phycocyanin concentrations in reservoirs and lakes. The estimation of phycocyanin concentration can be obtained by applying the ratio of phycocyanin absorption at 620 nm ($a_{PC}(620)$) to its specific absorption ($a^*_{PC}(620)$) ratio. This study examined this ratio using the average values of $a^*_{PC}(620)$ reported by Simis et al. (2005). Results obtained in this study are strongly correlated with the laboratory-measured cyanobacteria (Figure 5.5.6 B), but it seems to overestimate the concentrations of the PC pigment which agrees with the algorithm presenter (Simis et al., 2005). Several possible explanations were reported that the algorithm always produces an overestimation at low concentrations of PC values. Overestimation at this site is probably attributed to increased absorption by chlorophyll accessory pigments such as chlorophyll–c, with its peak absorption at 580 and 630 nm, known to be present in diatoms. Its peak absorption at 580 and 630 nm is present in diatoms found in relatively high concentrations of the Vaal Dam water.

5.7 Performance of the Semi-analytical algorithm for phycocyanin estimation in the Vaal Dam.

The linear least-squares regression obtained a strong relationship between the estimated phycocyanin using the absorption algorithm and laboratory-measured cyanobacterial concentrations; the regression gives $R^2 = 0.97$. Due to the nature of the optical complexity of the dam water, several potential challenges exist in the estimation of the phycocyanin pigments from Vaal Dam water using this Simis et al. (2005) model including analysis of the specific absorption coefficients of phycocyanin, the back scuttering estimation issues due to the presence of non-algal particles, and issues in the presence of various algal types alongside cyanobacteria in the water.

The Chl–a and PC-specific absorption coefficients used to retrieve the pigment concentrations in this study were adopted from previous research, the average values of the particular coefficients measured by Simis et al. (2005 and 2006) from diverse lakes and reservoirs in the Netherlands and Spain (Randolph et al., 2008). Thus, the conditions in the Vaal Dam reservoir site are expected to be different, even though their application to Vaal Dam Reservoir data was generally successful, strongly correlating with the measured Chl–a and cyanobacteria concentrations. The model retrieved Chl–a values that are strongly comparable with measured values, which reveals that the value of chlorophyll–a specific absorption coefficient $a^*_{Chl}(665)$ is optimal for use in the Vaal Dam water conditions. At the Vaal Dam site, the cyanobacterial concentrations during the field campaigns varied, and the measurements showed deficient concentrations of cyanobacteria except at three sampling points. The lakes and reservoirs that derive the pigment-specific coefficients $a^*_{PC}(620)$ are characterised by high concentrations of dominant cyanobacteria. Thus, according to Simis et al. (2005), the particular PC absorption became higher in low concentration values of PC relative to Chl-a compared to waters dominated by cyanobacteria. Therefore, applying such $a^*_{PC}(620)$ values to the sites exhibiting low cyanobacteria would overestimate phycocyanin concentration (Simis et al., 2005 and 2006). This might be the reason for the phycocyanin over-estimation in the Vaal Dam Reservoir data, where the model estimated phycocyanin concentrations ranged between 11–51 µg/L and reached 290 and 412 µg/L in points 13 and 50, respectively. Also, the absorption by other pigments, such as chlorophylls b & c, might elevate $a_{PC}(620)$ absorption values and lead to overestimations of the PC concentration.

According to a study by Chinyama et al. (2016) on the algal species in the Vaal Dam, the cyanobacterial domination in the dam water was suggested to be a function of seasonality; during the summer months (February and March), the cyanobacteria dominating other species while in the rest of the year, it will face phytoplankton compositing and species competition (Chinyama et al., 2016). This situation could be due to the overestimation of phycocyanin concentrations in the reservoir sites during June 2023, where the cyanobacteria did not dominate the phytoplankton species all over the reservoir.

The estimation of phycocyanin pigment using the semi-empirical model proposed by Simis et al. (2005) is assumed to function best in systems dominated by cyanobacteria, and its estimation accuracy decreases with the decrease of phycocyanin to chlorophyll-a ratio. A threshold has been specified by Simis et al. (2005), after which the error in phycocyanin estimation is increased, and the water with PC: Chl–a ratio less than 0.4 is facing increasing estimation errors. This study supports the assumption that the cyanobacteria to total algae ratio was 0.4 and the phycocyanin overestimation occurred.

5.8 Conclusion

The Vaal Dam algal community is significantly varying, different algal types were identified through the spatial extension of the dam. It also varies depending on the weather conditions through the year's seasons. Around six algal species were identified from the water samples collected between June 21st and June 27th, 2023. Some species were found in deficient concentrations and limited to one or two sampling sites; noticeably, only *Microcystis sp*. is found in the 17 analysed samples alongside other species. In-situ measurements taken by fluoroprobe showed the composition of the algal community within the dam water. The cyanobacteria composed over 40% of the total algal community during the sampling period, and it will rise to nearly 50% when excluding the samples located in the most eastern part of the reservoir at the inlet of Vaal River, diatoms are dominated by such site. The Sentinel-2 satellite data performed very well in detecting chlorophyll-a (R^2 = 0.78); unfortunately, it has no band centred at 620 nm wavelength, and it was not possible to check the phycocyanin pigment, the proxy of the cyanobacteria which its peak absorptions wavelengths at 620 nm. In-situ hyperspectral data performed very well with both tested methods, band ratio indices and Semi-analytical algorithm (R^2 = 0.96 and 0.97, respectively). The reflectance band–ratio algorithm estimates cyanobacteria by using the ratio of phycocyanin absorption at 620 nm [$a_{PC}(620)$] to its specific absorption coefficient $a^*_{PC}(620)$.

In-situ data has been used to validate chlorophyll-a and phycocyanin pigments. The validation showed an accurate estimation of Chl–a using both, band-ratio and semi-analytical algorithms. Regarding PC, a slight overestimation is noticed in the pigment concentration retrieved from the semi-analytical algorithm. The overestimation is because of the non-domination of the cyanobacterial concentration at this time of the year. The multispectral satellite remote sensing estimation of cyanobacteria in inland turbid waters remains a tough task. However, the study concludes that the Semi-analytical algorithm can effectively assess and monitor the cyanobacterial bloom in the Vaal Dam reservoir if the values of the absorption coefficient and the specific PC absorption coefficient were adjusted to the Vaal Dam conditions. The study recommends the use of hyperspectral remote sensing data for PC estimation in this site. Sentinel-2 data can still be used successfully for mapping Chl–a which can be used for estimating the cyanobacteria biomass during its dominance over other species in summertime.

Chapter 6

Conclusions and Recommendations

6.1 Introduction

Algae species occur naturally in water resources; they play a vital role by forming the energy base of the food chain for all organisms in aquatic environments (Stevenson, 2014). Cyanobacteria is one of the algal species known for its ability to produce cyanotoxin (Rastogi et al., 2014) which affects the health of humans, animals, and aquatic inhabitants.

The Vaal Dam Reservoir experienced frequent algal blooms during the past years, and among the algal species identified in the dam water, cyanobacteria were abundant. As the Vaal Dam is a primary source of potable water to metropolitan areas in Gauteng province and its surrounding areas, it's essential to find an effective method for instant detecting and estimating the cyanobacteria concentration in the dam water to help the authorities of water treatment works and other related authorities to deal with this critical issue. Recently, significant advances have occurred in remote sensing applications in various fields of science. With the recently launched multispectral high-resolution sensors, remote sensing techniques became one of the most potentially essential tools for multiple fields of studies, benefiting from its low cost and valuable spatial and temporal coverage.

This thesis examined remote sensing techniques for mapping and detecting cyanobacteria and algal blooms in the Vaal Dam reservoir; high-resolution Landsat-8 and Sentinel-2 multispectral data have been used to achieve the research objectives. Regarding detecting cyanobacteria, it isn't easy to precisely measure its concentration in its nondominant cases. Many studies have tried to estimate its concentrations. Still, it is more difficult to use OLI and MSI sensor data. They lack bands centred near the peak absorption wavelength of the phycocyanin pigment, the proxy of cyanobacteria. The following sections conclude the findings of the research.

6.2 Land Use and Land Cover Change in the Vaal Dam Catchment, South Africa: A Study Based on Remote Sensing and Time Series Analysis.

The chapter aimed to detect the land use land cover changes within the Vaal Dam catchment area to explore the potential effects of anthropogenic human activities and natural processes on the water quality; the study revealed the change detection during the last four decades and their possible impacts to the quality of water in the dam reservoir. Landsat series data has been used to assess LULC changes for the period between 1986 and 2021 in 7-year intervals. The random forest classifier was used successfully; and classified the study area into seven land cover classes. The results of the classification reveal that the grassland class dominated the Vaal Dam catchment through all the classified seven-year interval stacked images comprising more than 50% of the total catchment area. The grasslands are known as one of the most land cover classes with less contribution to water resources contamination unless they contain intensive cattle farming. In the Vaal Dam catchment area, the grassland class included cattle farming and the natural pastures, which are expected to release some types of nutrient contaminants that may find their way to the water courses and then accumulate in the large water bodies such as Vaal Dam Reservoir. The area experienced increasing economic development through the last few decades, increasing the effluent released to the surrounding environment; this is also reflected in the expansion of the settlement areas within the catchment due to the rise in the populations associated with such economic development. Such rapid increase in population growth will have even more associated significant effects on the region's aquatic environment and economic spheres. Therefore, it is of prime importance that these developments be carefully considered to control the contaminants better, not allowing them to be released into the watercourses. The dam water responding to such LULC changes, and its associated nutrient releases are checked in the following chapter.

6.3 Time series analysis of water quality factors enhancing harmful algal blooms (HABs): a study integrating in-situ and satellite data, Vaal Dam, South Africa.

The chapter aimed to explore the correlations between the HABs and the potential nutrients enhancing their growth in the Vaal Dam Reservoir to identify the limiting factors of its blooms. Some water quality parameters serve as nutrients enhancing algal and cyanobacterial growth in freshwater bodies; it's well known in the literature that besides the optimal weather conditions, total phosphorus and some forms of nitrogen are the driving factors for algal blooms in many water systems around the world. Historical water quality data are available and freely accessible on the South African Water and Sanitation website, from which some selected water quality parameters have been obtained, including total phosphorus, dissolved oxygen, nitrate, nitrite nitrogen, organic nitrogen, ammonia nitrogen and the water temperature. The obtained data were then examined to check if they were the potential drivers for the algal blooms of the Vaal Dam water. The time series analysis of the targeted data suggested that Vaal Dam HABs productivity is a function of TP and KJEL_N levels. A positive correlation has been detected between Chl–a and TP Chl–a and KJEL_N. Thus, these factors are the main drivers enhancing the growth of algae when weather conditions, such as surface water temperatures, are ideal for algal productivity levels. The time series showed some erratic dates on which concentrations of chlorophyll-a jumped to very high recorded levels; these high records were accompanied by high measurement levels of TP and KJEL_N on the exact dates. Most of the time, the TP and KJEL_N levels remain at their normal average concentrations and low values; The Chl‒a concentrations also remain at their low values, usually below the threshold level of the eutrophic conditions of 7 μg/L.

Time series analysis of spatial distributions of the chlorophyll-a using Landsat-8 data over the last decade revealed the spatial dynamics of summer blooms between 2013 and 2023. It illustrated the rising intensity of the chlorophyll-a concentration in the mapped images starting from the summer of 2015, which puts warnings to the necessity of controlling such increasing blooms year by year. Through the reservoir spatial geometry, relatively high concentrations were observed where the reservoir is shallow and meandering and where restricted water circulation occurs due to the low level of mixing and near inlets of the Vaal and Wilge rivers. The results of this time series analysis highlighted the effect of high levels of anthropogenic impacts on the dam water. The discharge of nutrient-rich aggravated the periodic deterioration of the dam water quality over some stochastic dates, as well as poor-quality effluents from the wastewater treatment works in the dam catchment.

6.4 Assessing the utility of using current generation high-resolution satellites (Sentinel-2 and Landsat-8) to monitor a large water supply dam in South Africa

The chapter aimed to examine the use of Landsat-8 and sentinel-2 satellite data in mapping and monitoring algal blooms in the Vaal Dam; it illustrated the potential of remote sensing algorithms to estimate chlorophyll-a concentration in biologically highly productive in-land open water bodies. Many band ratios from visible to near-infrared wavelengths were tested to find the best ratios that correlated with the measured chlorophyll-a from the satellite-based estimation. The results showed that the blue-green-based band ratio performed reasonably well with the Landsat-8 data, while the Sentinel-2 data gave poor results with blue-green-based band ratios; this is because the MSI sensor has a lower signal-to-noise ratio compared to the OLI sensor. The Sentinel-2 MSI performed well using the red/NIR-based model. This is because the MSI sensor has a red edge band centred at 705 nm, is more sensitive to chlorophyll–a signal, and is less sensitive to other non-algal organic matter.

6.5 Using Sentinel-2 satellite data and In-situ Hyperspectra for Cyanobacteria Assessment in the Vaal Dam, South Africa.

In this chapter, the study aimed to test some band ratios and a semi-analytical model for retrieval of the concentration of the cyanobacteria in the Vaal Dam. The algal community in the dam water is significantly varying, and different algal types were identified through the spatial extension of the dam. It also varies depending on the weather conditions through the year's seasons. Around six algal species were identified from the water samples collected between June 21st and June 27th, 2023. Some species were found in deficient concentrations and limited to one or two sampling sites. Noticeably, Microcystis sp. is found in the whole 17 analysed samples alongside other species. In-situ measurements taken by fluoroprobe also showed the composition of the algal community within the dam water. The cyanobacteria composed over 40% of the total algal community during the sampling period, and it will rise to nearly 50% when excluding the samples located in the most eastern part of the reservoir at the inlet of Vaal River, diatoms dominated such site. The Sentinel-2 satellite data performed very well in detecting chlorophyll–a (R^2 = 0.78). However, because Sentinel-2 has no band centred at 620 nm wavelength, it was not used to check the phycocyanin pigment concentration. In-situ hyperspectral data performed very well with both tested methods, band ratio indices and the semi-analytical algorithm (R^2 = 0.96 and 0.97, respectively). The semi-analytical algorithm estimating cyanobacteria by using a specific absorption coefficient $a^*_{PC}(620)$. It showed an overestimation of the cyanobacteria. The study suggests that the semi-analytical algorithm has a potential to be an effective method for assessing the cyanobacterial bloom in highly productive inland waters using hyperspectral remote sensing.

6.6 Conclusions

Based on the objectives of the research, the following findings are the foundation for the conclusions:

- Landsat series satellite data are valuable tools for land use, land cover mapping, and change detections using random forest classifiers. It’s been successfully used for accurate classification and change detection of the Vaal Dam catchment area and revealed the effect of some anthropogenic activities on the water quality, leading to more nutrient loads on the Vaal Dam Reservoir.

- Total phosphorus and organic nitrogen are the driving factors for algal blooms in the Vaal Dam water, and they enhance the growth of HABs in the reservoir. The temperature of the Vaal Dam water is ideal for algal growth, mainly in the summertime, when the cyanobacteria dominate over other species. At the same time different species, such as green algae and diatoms, occur alongside the cyanobacteria during the rest of the year.
- Both sensors, Landsat-8 OLI, and Sentinel-2 MSI are suitable for mapping chlorophyll-a, the proxy of algal blooms in the Vaal Dam, using band ratios in blue-green and red-NIR spectral regions, respectively. They performed well with reliable accuracy.
- The band ratios in red-NIR spectral regions of the sentinel-2 MSI sensor were not correlated well with cyanobacterial concentrations during June because the algal community was composed of multiple species, and the detected chlorophyll–a represents the whole phytoplankton community. Suppose cyanobacteria dominate over the other species during the summer months and control the optical properties of the water column. In that case, chlorophyll-a can be employed as a proxy to determine cyanobacteria biomass.
- The Semi-analytical algorithm is a promising tool for estimating cyanobacterial abundance and concentrations; it needs reflectance data with band centres at 620, 709, and 778 nm to be applied, which are not available in the Landsat-8 OLI and Sentinel-2 MSI sensors. The model successfully estimates the cyanobacteria in the Vaal Dam with robust correlation to the measured cyanobacteria but with a little bit of overestimation; however, it requires adjustments of the phycocyanin pigment-specific absorption coefficient for the Vaal Dam site to give more precise estimations.

The research has successfully illustrated the potential of different band ratio indices for detecting chlorophyll–a as a proxy of HABs in the study site using OLI and MSI sensors. However, the same band ratio indices can be used for estimating cyanobacteria in conditions of its domination over other species; in such cases, Chl–a will be the proxy of the cyanobacterial bloom.

6.7 Limitations

Nevertheless, in the conditions of non-domination of cyanobacteria over other species, the only key indicator of its abundance in the algal community is phycocyanin pigment, it absorbs spectrum at a wavelength of approximately 620 nm. Therefore, there are limitations to the multispectral satellite data in detecting phycocyanin pigment and even chlorophyll–a detection, including:

- One of the main challenges is the spectral resolution of the sensors used in multispectral satellite imagery, which may not be sufficient to accurately differentiate the specific absorption features of phycocyanin, they lack bands centred around 620 nm which is crucial to detecting the phycocyanin pigment.
- Another challenge is the presence of other substances in the water, such as dissolved organic matter or suspended sediments, tripton and CDOM can affect the accuracy of chlorophyll–a and cyanobacteria concentration estimates derived from the multispectral satellite data.
- Additionally, the spatial resolution of multispectral satellite imagery may not be high enough to accurately detect the spatially small-scale variations in the concentration of chlorophyll–a and cyanobacteria in the water body, leading to potential inaccuracies in the detection of pigments.
- Furthermore, atmospheric conditions such as cloud cover, water column effects, and the sun glint can distort the signals received by satellite sensors, making it difficult to isolate the signal related to phycocyanin and chlorophyll–a pigments in some conditions.
- The spatial variations of the algal community and turbidity within the Vaal Dam reservoir create more complications to the optically suitable effecting water column depth. This complicates the model build-up and composition from satellite remote sensing data compared to in-situ field measurements.

Despite these limitations, ongoing research and technological advancements continue to improve the capabilities of satellite remote sensing for detecting cyanobacteria in inland freshwater bodies.

6.8 Recommendations

Generally, this research highlighted the significant threats associated with harmful algal blooms in the Vaal Dam water to the aquatic environment and public health. To mitigate such threats, Policy-related recommendations should focus on prevention, monitoring, and response strategies to address this issue effectively through:

- Implementing regulations to reduce nutrient runoff from mining, agricultural, and urban sources. This will help to reduce the external nutrient input to the dam water, this requires to be implemented throughout the watershed area.
- Regular monitoring of the water quality parameters, such as nutrient levels and algal or cyanobacteria biomass, is essential for the early detection of potential bloom events.
- Detecting the temporal and spatial bloom changes in real-time using remote sensing methods, the study recommends testing hyperspectral satellite data to detect phycocyanin pigment in the dam water. This might help to identify the volume of cyanobacteria within the algal biomass.
- Additionally, establishing protocols for timely response actions can help mitigate the impacts of harmful algal blooms on ecosystems and human health. For example, eradicating blooms of cyanobacteria by adding rapidly degrading chemicals such as low concentrations of hydrogen peroxide. Hydrogen peroxide eliminates cyanobacterial blooms and degrades to water and oxygen within a few days. This can be applied to sites where excessive blooms occur if they are far from the water intake.
- Collaboration among government agencies, researchers, and stakeholders is crucial for successfully implementing these policy recommendations.

References

ABDULKAREEM, J. H., SULAIMAN, W. N. A., PRADHAN, B. & JAMIL, N. R. 2018. Long-term hydrologic impact assessment of non-point source pollution measured through Land Use/Land Cover (LULC) changes in a tropical complex catchment. *Earth systems and environment,* 2**,** 67-84.

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,* 35**,** 3440-3458.

ADU, J. T. & KUMARASAMY, M. V. 2018. Assessing Non-Point Source Pollution Models: a Review. *Polish Journal of Environmental Studies,* 27, 1913-1922.

AHMED, U., MUMTAZ, R., ANWAR, H., MUMTAZ, S. & QAMAR, A. M. 2020. Water quality monitoring: from conventional to emerging technologies. *Water Supply,* 20**,** 28-45.

AIKEN, J. & MOORE, G. 2000. ATBD case 2 s bright pixel atmospheric correction. *Center for Coastal & Marine Sciences, Plymouth Marine Laboratory, UK, Rep, PO-TN-MEL-GS,* 4.

AKAR, Ö. & GÜNGÖR, O. 2012. Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation,* 1**,** 105-112.

ALI, K., ABIYE, T. & ADAM, E. 2022a. Integrating In Situ and Current Generation Satellite Data for Temporal and Spatial Analysis of Harmful Algal Blooms in the Hartbeespoort Dam, Crocodile River Basin, South Africa. *Remote Sensing,* 14**,** 4277.

ALI, K., WITTER, D. & ORTIZ, J. 2014a. Application of empirical and semi-analytical algorithms to MERIS data for estimating chlorophyll a in Case 2 waters of Lake Erie. *Environmental Earth Sciences,* 71**,** 4209-4220.

ALI, K. A., ORTIZ, J., BONINI, N., SHUMAN, M. & SYDOW, C. 2016. Application of Aqua MODIS sensor data for estimating chlorophyll a in the turbid Case 2 waters of Lake Erie using bio-optical models. *GIScience & Remote Sensing,* 53**,** 483-505.

ALI, K. A., WITTER, D. L. & ORTIZ, J. D. 2014b. Multivariate approach to estimate colour producing agents in Case 2 waters using first-derivative spectrophotometer data. *Geocarto International,* 29**,** 102-127.

ALI, U., ESAU, T. J., FAROOQUE, A. A., ZAMAN, Q. U., ABBAS, F. & BILODEAU, M. F. 2022b. Limiting the Collection of Ground Truth Data for Land Use and Land Cover Maps with Machine Learning Algorithms. *ISPRS International Journal of Geo-Information,* 11**,** 333.

ANSPER, A. & ALIKAS, K. 2018. Retrieval of chlorophyll a from Sentinel-2 MSI data for the European Union water framework directive reporting purposes. *Remote Sensing,* 11**,** 64.

ARAR, E. J. & COLLINS, G. B. 1997. *Method 445.0: In vitro determination of chlorophyll a and pheophytin a in marine and freshwater algae by fluorescence*. (p. 22). Cincinnati: United States Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory.

ARCHIBALD, R. 1971. Diatoms from the Vaal dam catchment area Transvaal, South Africa.

ATTRI, P., CHAUDHRY, S. & SHARMA, S. 2015. Remote sensing & GIS based approaches for LULC change detection–a review. *Int. J. Curr. Eng. Technol,* 5**,** 3126-3137.

BABIN, M., MOREL, A. & GENTILI, B. 1996. Remote sensing of sea surface sun-induced chlorophyll fluorescence: consequences of natural variations in the optical characteristics of phytoplankton and the quantum yield of chlorophyll a fluorescence. *International Journal of Remote Sensing,* 17**,** 2417-2448.

BALASUBRAMANIAN, S. V., PAHLEVAN, N., SMITH, B., BINDING, C., SCHALLES, J., LOISEL, H., GURLIN, D., GREB, S., ALIKAS, K. & RANDLA, M. 2020. Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. *Remote Sensing of Environment,* 246**,** 111768.

BALLOT, A., SANDVIK, M., RUNDBERGET, T., BOTHA, C. J. & MILES, C. O. 2013. Diversity of cyanobacteria and cyanotoxins in Hartbeespoort Dam, South Africa. *Marine and Freshwater Research,* 65**,** 175-189.

BĂNĂDUC, D., SIMIĆ, V., CIANFAGLIONE, K., BARINOVA, S., AFANASYEV, S., ÖKTENER, A., MCCALL, G., SIMIĆ, S. & CURTEAN-BĂNĂDUC, A. 2022. Freshwater as a sustainable resource and generator of secondary resources in the 21st century: stressors, threats, risks, management and protection strategies, and conservation approaches. *International Journal of Environmental Research and Public Health,* 19**,** 16570.

BANDE, P., ADAM, E., ELBASIT, M. & ADELABU, S. Comparing Landsat 8 and Sentinel-2 in mapping water quality at Vaal dam. 2018 IEEE International geoscience and remote sensing symposium, 2018. 9280-9283.

BEAULAC, M. N. & RECKHOW, K. H. 1982. AN EXAMINATION OF LAND USE-NUTRIENT EXPORT RELATIONSHIPS 1. *JAWRA Journal of the American Water Resources Association,* 18**,** 1013-1024.

BELLACICCO, M., VOLPE, G., BRIGGS, N., BRANDO, V., PITARCH, J., LANDOLFI, A., COLELLA, S., MARULLO, S. & SANTOLERI, R. 2018. Global distribution of non-algal particles from ocean color data and implications for phytoplankton biomass detection. *Geophysical Research Letters,* 45**,** 7672-7682.

BIGGS, R. & SCHOLES, R. 2002. Land-cover changes in South Africa 1911-1993: research in action. *South African Journal of Science,* 98**,** 420-424.

BINDING, C. E., JEROME, J. H., BUKATA, R. P. & BOOTY, W. G. 2008. Spectral absorption properties of dissolved and particulate matter in Lake Erie. *Remote sensing of environment,* 112**,** 1702-1711.

BIRHANE, E., ASHFARE, H., FENTA, A. A., HISHE, H., GEBREMEDHIN, M. A. & SOLOMON, N. 2019. Land use land cover changes along topographic gradients in Hugumburda national forest priority area, Northern Ethiopia. *Remote Sensing Applications: Society and Environment,* 13**,** 61-68.

BRAUNE, E. & ROGER, K. 1987. *The Vaal River catchment: problems and research needs*.

BREIMAN, L. 2001. Random forests. *Machine learning,* 45**,** 5-32.

BROOKS, B. W., LAZORCHAK, J. M., HOWARD, M. D., JOHNSON, M. V. V., MORTON, S. L., PERKINS, D. A., REAVIE, E. D., SCOTT, G. I., SMITH, S. A. & STEEVENS, J. A. 2016. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? *Environmental toxicology and chemistry,* 35**,** 6-13.

BUKATA, R. P., JEROME, J. H., KONDRATYEV, A. S. & POZDNYAKOV, D. V. 1995. *Optical properties and remote sensing of inland and coastal waters*, CRC press.

BUMA, W. G. & LEE, S.-I. 2020. Evaluation of sentinel-2 and landsat 8 images for estimating chlorophyll-a concentrations in lake Chad, Africa. *Remote Sensing,* 12**,** 2437.

CAO, Z., MA, R., DUAN, H., PAHLEVAN, N., MELACK, J., SHEN, M. & XUE, K. 2020. A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. *Remote Sensing of Environment,* 248**,** 111974.

CAREY, C. C., IBELINGS, B. W., HOFFMANN, E. P., HAMILTON, D. P. & BROOKES, J. D. 2012. Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. *Water research,* 46**,** 1394-1407.

CAREY, R. O. & MIGLIACCIO, K. W. 2009. Contribution of wastewater treatment plant effluents to nutrient dynamics in aquatic systems: a review. *Environmental management,* 44**,** 205-217.

CHAFFIN, J. D., BRATTON, J. F., VERHAMME, E. M., BAIR, H. B., BEECHER, A. A., BINDING, C. E., BIRBECK, J. A., BRIDGEMAN, T. B., CHANG, X. & CROSSMAN, J. 2021. The Lake Erie HABs Grab: A binational collaboration to characterize the western basin cyanobacterial harmful algal blooms at an unprecedented high-resolution spatial scale. *Harmful Algae,* 108**,** 102080.

CHENG, C., WEI, Y., XU, J. & YUAN, Z. 2013. Remote sensing estimation of Chlorophyll a and suspended sediment concentration in turbid water based on spectral separation. *Optik,* 124**,** 6815-6819.

CHINYAMA, A., OCHIENG, G. M., SNYMAN, J. & NHAPI, I. 2016. Occurrence of cyanobacteria genera in the Vaal Dam: implications for potable water production. *Water SA,* 42**,** 415-420.

CHISLOCK, M. F., DOSTER, E., ZITOMER, R. A. & WILSON, A. 2013. Eutrophication: causes, consequences, and controls in aquatic ecosystems. *Nature Education Knowledge,* 4**,** 10.

CHUGHTAI, A. H., ABBASI, H. & KARAS, I. R. 2021. A review on change detection method and accuracy assessment for land use land cover. *Remote Sensing Applications: Society and Environment,* 22**,** 100482.

CHUTTER, F. M. 1970. Hydrobiological studies in the catchment of Vaal Dam, South Africa. Part 1. River zonation and the benthic fauna fauna, int. revue ges. hydrobiol. 55 3, 445-494.

CLARK, M. L., AIDE, T. M., GRAU, H. R. & RINER, G. 2010. A scalable approach to mapping annual land cover at 250 m using MODIS time series data: A case study in the Dry Chaco ecoregion of South America. *Remote Sensing of Environment,* 114**,** 2816-2832.

DAHINDEN, C. & ETHZ, M. 2011. An improved Random Forests approach with application to the performance prediction challenge datasets. *Hands-on Pattern Recognition, Challenges in Machine Learning,* 1**,** 223-30.

DAVIDSON, K., TETT, P. & GOWEN, R. 2011. Harmful algal blooms. *Marine pollution and human health***,** 95-127.

DAVIS, T. W., BULLERJAHN, G. S., TUTTLE, T., MCKAY, R. M. & WATSON, S. B. 2015. Effects of increasing nitrogen and phosphorus concentrations on phytoplankton community growth and toxicity during Planktothrix blooms in Sandusky Bay, Lake Erie. *Environmental science & technology,* 49**,** 7197-7207.

DEBÉN, S., ABOAL, J., CARBALLEIRA, A., CESA, M., REAL, C. & FERNÁNDEZ, J. 2015. Inland water quality monitoring with native bryophytes: A methodological review. *Ecological Indicators,* 53**,** 115-124.

DEKKER, A., MALTHUS, T., WIJNEN, M. & SEYHAN, E. 1992. Remote sensing as a tool for assessing water quality in Loosdrecht lakes. *Hydrobiologia,* 233**,** 137-159.

DEKKER, A. G. 1993. Detection of optical water quality parameters for eutrophic waters by high resolution remote sensing, PhD, Vrije Universiteit Amsterdam.

DENG, X., ZHAO, C. & YAN, H. 2013. Systematic modeling of impacts of land use and land cover changes on regional climate: a review. *Advances in Meteorology,* 2013(1), p.317678.

DIALLO, Y., HU, G. & WEN, X. 2009. Applications of remote sensing in land use/land cover change detection in Puer and Simao Counties, Yunnan Province. *Journal of American Science,* 5**,** 157-166.

DING, S., CHEN, M., GONG, M., FAN, X., QIN, B., XU, H., GAO, S., JIN, Z., TSANG, D. C. & ZHANG, C. 2018. Internal phosphorus loading from sediments causes seasonal nitrogen limitation for harmful algal blooms. *Science of the total environment,* 625**,** 872-884.

DIVIŠ, P., DOČEKALOVÁ, H., BRULÍK, L., PAVLIŠ, M. & HEKERA, P. 2007. Use of the diffusive gradients in thin films technique to evaluate (bio) available trace metal concentrations in river water. *Analytical and bioanalytical chemistry,* 387**,** 2239-2244.

DOERFFER, R. & FISCHER, J. 1994. Concentrations of chlorophyll, suspended matter, and gelbstoff in case II waters derived from satellite coastal zone color scanner data with inverse modeling methods. *Journal of Geophysical Research: Oceans,* 99**,** 7457-7466.

DOERFFER, R. & SCHILLER, H. 2007. The MERIS Case 2 water algorithm. *International Journal of Remote Sensing,* 28**,** 517-535.

DÖRNHÖFER, K. & OPPELT, N. 2016. Remote sensing for lake research and monitoring–Recent advances. *Ecological Indicators,* 64**,** 105-122.

DOXARAN, D., FROIDEFOND, J.-M., LAVENDER, S. & CASTAING, P. 2002. Spectral signature of highly turbid waters: Application with SPOT data to quantify suspended particulate matter concentrations. *Remote sensing of Environment,* 81**,** 149-161.

DU PLESSIS, A. 2017. *Freshwater Challenges of South Africa and Its Upper Vaal River*, Springer.

DU PLESSIS, A., HARMSE, T. & AHMED, F. 2015. Predicting water quality associated with land cover change in the Grootdraai Dam catchment, South Africa. *Water International,* 40**,** 647-663.

DU PLESSIS, A., HARMSE, T. & AHMED, F. 2018. Predicting water quality associated with land cover change in the Grootdraai Dam catchment, South Africa. *The Water Legacies of Conventional Mining.* Routledge, pp. 178-194.

DUAN, H., MA, R., ZHANG, Y., LOISELLE, S. A., XU, J., ZHAO, C., ZHOU, L. & SHANG, L. 2010. A new three-band algorithm for estimating chlorophyll concentrations in turbid inland lakes. *Environmental Research Letters,* 5**,** 044009.

DUTTA, D., GUPTA, S. & CHAKRABORTY, A. 2022. Effect of different land use land cover on surface heat budget–A case study from a tropical humid region of India. *Remote Sensing Applications: Society and Environment,* 25**,** 100675.

DVORNIKOV, Y., LEIBMAN, M., HEIM, B., BARTSCH, A., HERZSCHUH, U., SKOROSPEKHOVA, T., FEDOROVA, I., KHOMUTOV, A., WIDHALM, B. & GUBARKOV, A. 2018. Terrestrial CDOM in lakes of Yamal peninsula: connection to lake and lake catchment properties. *Remote Sensing,* 10**,** 167.

DWAF 2004. Internal Strategic Perspective: Upper Vaal Water Management Area. Compiled by PDNA, WRP Consulting Engineers (Pty) Ltd, WMB and Kwezi-V3 on behalf of the Directorate: National Water Resource Planning, 2004. Department of Water Affairs and Forestry South Africa.

EBRAHIMY, H., MIRBAGHERI, B., MATKAN, A. A. & AZADBAKHT, M. 2021. Per-pixel land cover accuracy prediction: A random forest-based method with limited reference sample data. *ISPRS Journal of Photogrammetry and Remote Sensing,* 172**,** 17-27.

EUGENIO, F., MARCELLO, J. & MARTÍN, J. 2020. Multiplatform Earth observation systems for monitoring water quality in vulnerable inland ecosystems: Maspalomas water lagoon. *Remote Sensing,* 12**,** 284.

EWERTS, H. 2010. *Effectiveness of purification processes in removing algae from Vaal Dam water at the Rand Water Zuikerbosch treatment plant in Vereeniging* (Doctoral dissertation, North-West University).

FENG, Q., LIU, J. & GONG, J. 2015. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote sensing,* 7**,** 1074-1094.

FRANZ, B. A., BAILEY, S. W., KURING, N. & WERDELL, P. J. 2014. Ocean color measurements from Landsat-8 OLI using SeaDAS (No. GSFC-E-DAA-TN21755).

GHOLIZADEH, M. H., MELESSE, A. M. & REDDI, L. 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors,* 16**,** 1298.

GILERSON, A. A., GITELSON, A. A., ZHOU, J., GURLIN, D., MOSES, W., IOANNOU, I. & AHMED, S. A. 2010. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. *Optics Express,* 18**,** 24109-24125.

GITELSON, A., GARBUZOV, G., SZILAGYI, F., MITTENZWEY, K., KARNIELI, A. & KAISER, A. 1993. Quantitative remote sensing methods for real-time monitoring of inland waters quality. *International Journal of Remote Sensing,* 14**,** 1269-1295.

GITELSON, A. & YACOBI, Y. 2004. Monitoring quality of productive aquatic ecosystem: requirements for satellite sensors. *BALWOIS: Ohrid, North Macedonia*.

GITELSON, A. A., DALL'OLMO, G., MOSES, W., RUNDQUIST, D. C., BARROW, T., FISHER, T. R., GURLIN, D. & HOLZ, J. 2008. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: Validation. *Remote Sensing of Environment,* 112**,** 3582-3593.

GITELSON, A. A., SCHALLES, J. F. & HLADIK, C. M. 2007. Remote chlorophyll-a retrieval in turbid, productive estuaries: Chesapeake Bay case study. *Remote Sensing of Environment,* 109**,** 464-472.

GLIBERT, P. M., MARANGER, R., SOBOTA, D. J. & BOUWMAN, L. 2014. The Haber Bosch–harmful algal bloom (HB–HAB) link. *Environmental Research Letters,* 9**,** 105001.

GONG, Y., CAI, M., YAO, L., CHENG, L., HAO, C. & ZHAO, Z. 2022. Assessing Changes in the Ecosystem Services Value in Response to Land-Use/Land-Cover Dynamics in Shanghai from 2000 to 2020. *International Journal of Environmental Research and Public Health,* 19**,** 12080.

GONS, H. J. 1999. Optical teledetection of chlorophyll a in turbid inland waters. *Environmental science & technology,* 33**,** 1127-1132.

GONS, H. J., RIJKEBOER, M. & RUDDICK, K. G. 2002. A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *Journal of Plankton Research,* 24**,** 947-951.

GORDON, H. & MOREL, A. 1983. Lecture notes on coastal and estuarine studies. *Remote assessment of ocean color for interpretation of satellite visible imagery. A review.* Springer-Verlag, Berlin, 1983. Pp. 114.

GORDON, H. R. & WANG, M. 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. *Applied optics,* 33**,** 443-452.

GOWER, J., KING, S., BORSTAD, G. & BROWN, L. 2005. Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing,* 26**,** 2005-2012.

GRIFFIN, N., PALMER, C. & SCHERMAN, P. 2014. Critical analysis of environmental water quality in South Africa: Historic and current trends. *Report submitted to the Water Research Commission on WRC Project, 1(5), Report No. 2184/1/14, ISBN 978-1-4312-0536-3*.

GROBLER, D., TOERIEN, D. & ROSSOUW, J. 1987. A review of sediment/water quality interaction with particular reference to the Vaal River system. *Water S. A.,* 13**,** 15-22.

GYEDU-ABABIO, T. & VAN WYK, F. 2004. Effects of human activities on the waterval river, Vaal river catchment, South Africa. *African Journal of Aquatic Science,* 29**,** 75-81.

HAARHOFF, J. & TEMPELHOFF, J. 2007. Water supply to the Witwatersrand (Gauteng) 1924-2003. *Journal for Contemporary History,* 32**,** 95-114.

HAJI GHOLIZADEH, M., MELESSE, A. M. & REDDI, L. 2016. Spaceborne and airborne sensors in water quality assessment. *International Journal of Remote Sensing,* 37**,** 3143-3180.

HARA, M. M. & BACKEBERG, G. R. 2014. An institutional approach for developing South African inland freshwater fisheries for improved food security and rural livelihoods. *Water SA,* 40**,** 277-286.

HARDING, W. R. 2008. *The determination of annual phosphorus loading limits for South African dams (p. 68)*, Water Research Commission Pretoria, South Africa.

HELLWEGER, F. L., MILLER, W. & OSHODI, K. S. 2007. Mapping turbidity in the Charles River, Boston using a high-resolution satellite. *Environmental monitoring and assessment,* 132**,** 311-320.

HIERONYMI, M., KRASEMANN, H., MÜLLER, D., BROCKMANN, C., RUESCAS, A., STELZER, K., NECHAD, B., RUDDICK, K., SIMIS, S. & TILSTONE, G. 2016. Ocean colour remote sensing of extreme case-2 waters. *spectrum,* 2**,** 56.

HO, L. T. & GOETHALS, P. L. 2019. Opportunities and challenges for the sustainability of lakes and reservoirs in relation to the Sustainable Development Goals (SDGs). *Water,* 11**,** 1462.

HUA, A. K. 2017. Land use land cover changes in detection of water quality: a study based on remote sensing and multivariate statistics. *Journal of environmental and public health,* 2017(1), p.7515130.

HUISMAN, J., CODD, G. A., PAERL, H. W., IBELINGS, B. W., VERSPAGEN, J. M. & VISSER, P. M. 2018. Cyanobacterial blooms. *Nature Reviews Microbiology,* 16**,** 471-483.

HÜTT, C., KOPPE, W., MIAO, Y. & BARETH, G. 2016. Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images. *Remote sensing,* 8**,** 684.

IBRAHIM, E. & MONBALIU, J. 2011. Suitability of spaceborne multispectral data for inter-tidal sediment characterization: A case study. *Estuarine, Coastal and Shelf Science,* 92**,** 437-445.

ILORI, C. O., PAHLEVAN, N. & KNUDBY, A. 2019. Analyzing Performances of Different Atmospheric Correction Techniques for Landsat 8: Application for Coastal Remote Sensing. *Remote Sensing,* 11**,** 469.

ILUNGA, M. 2017. Assessing catchment resilience using entropy associated with mean annual runoff for the Upper Vaal catchment in South Africa. *Entropy,* 19**,** 147.

JEFFREY, S. T. & HUMPHREY, G. 1975. New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. *Biochemie und physiologie der pflanzen,* 167**,** 191-194.

KENEA, U., ADEBA, D., REGASA, M. S. & NONES, M. 2021. Hydrological responses to land use land cover changes in the Fincha'a Watershed, Ethiopia. *Land,* 10**,** 916.

KERRIGAN, K. & ALI, K. A. 2020. Application of Landsat 8 OLI for monitoring the coastal waters of the US Virgin Islands. *International Journal of Remote Sensing,* 41**,** 5743-5769.

KHAN, R. M., SALEHI, B., MAHDIANPARI, M., MOHAMMADIMANESH, F., MOUNTRAKIS, G. & QUACKENBUSH, L. J. 2021. A meta-analysis on harmful algal bloom (HAB) detection and monitoring: a remote sensing perspective. *Remote Sensing,* 13**,** 4347.

KHARBUSH, J. J., ROBINSON, R. S. & CARTER, S. J. 2023. Patterns in sources and forms of nitrogen in a large eutrophic lake during a cyanobacterial harmful algal bloom. *Limnology and Oceanography,* 68**,** 803-815.

KHATRI, N. & TYAGI, S. 2015. Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas. *Frontiers in Life Science,* 8**,** 23-39.

KIM, H. H., KO, B. C. & NAM, J. Y. 2016. Predicting chlorophyll-a using Landsat 8 OLI sensor data and the non-linear RANSAC method–a case study of Nakdong River, South Korea. *International Journal of Remote Sensing,* 37**,** 3255-3271.

KLOIBER, S. M., BREZONIK, P. L. & BAUER, M. E. 2002. Application of Landsat imagery to regional-scale assessments of lake clarity. *Water Research,* 36**,** 4330-4340.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai, 1995. Montreal, Canada, 1137-1145.

KRUGER, A. & SEKELE, S. 2013. Trends in extreme temperature indices in South Africa: 1962–2009. *International Journal of Climatology,* 33**,** 661-676.

KUMMU, M., DE MOEL, H., WARD, P. J. & VARIS, O. 2011. How close do we live to water? A global analysis of population distance to freshwater bodies. *PloS one,* 6**,** e20578.

KUTSER, T., PAAVEL, B., VERPOORTER, C., LIGI, M., SOOMETS, T., TOMING, K. & CASAL, G. 2016. Remote sensing of black lakes and using 810 nm reflectance peak for retrieving water quality parameters of optically complex waters. *Remote Sensing,* 8**,** 497.

KUTSER, T., PIERSON, D. C., TRANVIK, L., REINART, A., SOBEK, S. & KALLIO, K. 2005. Using satellite remote sensing to estimate the colored dissolved organic matter absorption coefficient in lakes. *Ecosystems,* 8**,** 709-720.

KUTSER, T., VAHTMÄE, E., PAAVEL, B. & KAUER, T. 2013. Removing glint effects from field radiometry data measured in optically complex coastal and inland waters. *Remote Sensing of Environment,* 133**,** 85-89.

LANDMAN, W. A., MALHERBE, J. & ENGELBRECHT, F. 2017. South Africa's present-day climate. *Understanding the social and environmental implications of global change***,** 7-12.

LANEVE, G., BRUNO, M., MUKHERJEE, A., MESSINEO, V., GIUSEPPETTI, R., DE PACE, R., MAGURANO, F. & D'UGO, E. 2021. Remote sensing detection of algal blooms in a lake impacted by petroleum hydrocarbons. *Remote Sensing,* 14**,** 121.

LYU, H., WANG, Q., WU, C., ZHU, L., YIN, B., LI, Y. & HUANG, J. 2013. Retrieval of phycocyanin concentration from remote-sensing reflectance using a semi-analytic model in eutrophic lakes. *Ecological informatics,* 18**,** 178-187.

MAKARIGAKIS, A. K. & JIMENEZ-CISNEROS, B. E. 2019. UNESCO's contribution to face global water challenges. *Water,* 11**,** 388.

MALAHLELA, O. E., OLIPHANT, T., TSOELENG, L. T. & MHANGARA, P. 2018. Mapping chlorophyll-a concentrations in a cyanobacteria-and algae-impacted Vaal Dam using Landsat 8 OLI data. *South African Journal of Science,* 114**,** 1-9.

MAMBA, B., KRAUSE, R., MALEFETSE, T., MHLANGA, S., SITHOLE, S., SALIPIRA, K. & NXUMALO, E. 2007. Removal of geosmin and 2-methylisorboneol (2-MIB) in water from Zuikerbosch Treatment Plant (Rand Water) using â-cyclodextrin polyurethanes. *Water SA,* 33.

MANANDHAR, R., ODEH, I. O. & ANCEV, T. 2009. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. *Remote Sensing,* 1**,** 330-344.

MARKOGIANNI, V., KALIVAS, D., PETROPOULOS, G. P. & DIMITRIOU, E. 2020. Estimating chlorophyll-a of inland water bodies in Greece based on landsat data. *Remote Sensing,* 12**,** 2087.

MATETE, M. & HASSAN, R. 2006. Integrated ecological economics accounting approach to evaluation of inter-basin water transfers: An application to the Lesotho Highlands Water Project. *Ecological Economics,* 60**,** 246-259.

MATTHEWS, M. W. & BERNARD, S. 2015. Eutrophication and cyanobacteria in South Africa's standing water bodies: A view from space. *South African journal of science,* 111**,** 1-8.

MATTHEWS, M. W., BERNARD, S. & WINTER, K. 2010. Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. *Remote Sensing of Environment,* 114**,** 2070-2087.

MATTHEWS, M. W. & ODERMATT, D. 2015. Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters. *Remote Sensing of Environment,* 156**,** 374-382.

MENKEN, K. D., BREZONIK, P. L. & BAUER, M. E. 2006. Influence of chlorophyll and colored dissolved organic matter (CDOM) on lake reflectance spectra: Implications for measuring lake properties by remote sensing. *Lake and Reservoir Management,* 22**,** 179-190.

MISHRA, S. & MISHRA, D. R. 2012. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment,* 117**,** 394-406.

MNGUNI, E. S. 2022. Pollution of the Vaal River system in South Africa: A case study. *Waste Management and Environmental Impact XI,* 257**,** 17.

MOREL, A. & PRIEUR, L. 1977. Analysis of variations in ocean color 1. *Limnology and oceanography,* 22**,** 709-722.

MOSES, W. J., GITELSON, A. A., BERDNIKOV, S. & POVAZHNYY, V. 2009. Estimation of chlorophyll-a concentration in case II waters using MODIS and MERIS data—successes and challenges. *Environmental research letters,* 4**,** 045005.

MUDALY, L. & VAN DER LAAN, M. 2020. Interactions between irrigated agriculture and surface water quality with a focus on phosphate and nitrate in the middle olifants catchment, South Africa. *Sustainability,* 12**,** 4370.

NAIDOO, S. 2017. *Acid mine drainage in South Africa [electronic resource]: Development actors, policy impacts, and broader implications*, Springer.

NAZEER, M., BILAL, M., NICHOL, J. E., WU, W., ALSAHLI, M. M., SHAHZAD, M. I. & GAYEN, B. K. 2020. First experiences with the Landsat-8 aquatic reflectance product: evaluation of the regional and ocean color algorithms in a coastal environment. *Remote Sensing,* 12**,** 1938.

NCHABELENG, T., CHENG, P., OBERHOLSTER, P. J., BOTHA, A., SMIT, W. J. & LUUS-POWELL, W. 2014. Microcystin-LR equivalent concentrations in fish tissue during a postbloom Microcystis exposure in Loskop Dam, South Africa. *African journal of aquatic science,* 39**,** 459-466.

NDLELA, L. L., OBERHOLSTER, P. J., VAN WYK, J. & CHENG, P.-H. 2016. An overview of cyanobacterial bloom occurrences and research in Africa over the last decade. *Harmful Algae,* 60**,** 11-26.

NIROUMAND-JADIDI, M., BOVOLO, F. & BRUZZONE, L. 2019. Novel spectra-derived features for empirical retrieval of water quality parameters: Demonstrations for OLI, MSI, and OLCI Sensors. *IEEE Transactions on Geoscience and Remote Sensing,* 57**,** 10285-10300.

O'REILLY, J. E., MARITORENA, S., MITCHELL, B. G., SIEGEL, D. A., CARDER, K. L., GARVER, S. A., KAHRU, M. & MCCLAIN, C. 1998. Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research: Oceans,* 103**,** 24937-24953.

O'REILLY, J. E. & WERDELL, P. J. 2019. Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6. *Remote sensing of environment,* 229**,** 32-47.

OBAID, A., ADAM, E. & ALI, K. A. 2023. Land Use and Land Cover Change in the Vaal Dam Catchment, South Africa: A Study Based on Remote Sensing and Time Series Analysis. *Geomatics,* 3**,** 205-220.

OBAID, A., ALI, K., ABIYE, T. & ADAM, E. 2021. Assessing the utility of using current generation high-resolution satellites (Sentinel 2 and Landsat 8) to monitor large water supply dam in South Africa. *Remote Sensing Applications: Society and Environment,* 22**,** 100521.

OBERHOLSTER, P., BOTHA, A.-M. & CLOETE, T. 2008. Biological and chemical evaluation of sewage water pollution in the Rietvlei nature reserve wetland area, South Africa. *Environmental Pollution,* 156**,** 184-192.

OBERHOLSTER, P. J. 2009. Impact on ecotourism by water pollution in the Olifants River catchment, South Africa.

OBERHOLSTER, P. J. & ASHTON, P. J. 2008. State of the nation report: An overview of the current status of water quality and eutrophication in South African rivers and reservoirs. *Parliamentary Grant Deliverable. Pretoria: Council for Scientific and Industrial Research (CSIR)***,** 2006.

OBERHOLSTER, P. J. & BOTHA, A.-M. 2007. Use of PCR based technologies for risk assessment of a winter cyanobacterial bloom in Lake Midmar, South Africa. *African journal of biotechnology,* 6, 1684-5315.

OBERHOLSTER, P. J., BOTHA, A.-M. & ASHTON, P. J. 2009a. The influence of a toxic cyanobacterial bloom and water hydrology on algal populations and macroinvertebrate abundance in the upper littoral zone of Lake Krugersdrift, South Africa. *Ecotoxicology,* 18**,** 34-46.

OBERHOLSTER, P. J., BOTHA, A.-M. & CLOETE, T. E. 2005. An overview of toxic freshwater cyanobacteria in South Africa with special reference to risk, impact and detection by molecular marker tools. *Biokemistri,* 17**,** 57-71.

OBERHOLSTER, P. J., JAPPIE, S., CHENG, P.-H., BOTHA, A. & MATTHEWS, M. 2015. First report of an Anabaena Bory strain containing microcystin-LR in a freshwater body in Africa. *African journal of aquatic science,* 40**,** 21-36.

OBERHOLSTER, P. J., MYBURGH, J. G., GOVENDER, D., BENGIS, R. & BOTHA, A.-M. 2009b. Identification of toxigenic Microcystis strains after incidents of wild animal mortalities in the Kruger National Park, South Africa. *Ecotoxicology and Environmental Safety,* 72**,** 1177-1182.

OUMA, Y. O., NOOR, K. & HERBERT, K. 2020. Modelling Reservoir Chlorophyll-a, TSS, and Turbidity Using Sentinel-2A MSI and Landsat-8 OLI Satellite Sensors with Empirical Multivariate Regression. *Journal of Sensors,* 2020**,** 8858408.

OWOJORI, A. and XIE, H., 2005, March. Landsat image-based LULC changes of San Antonio, Texas using advanced atmospheric correction and object-oriented image analysis approaches. In *5th international symposium on remote sensing of urban areas, Tempe, AZ*, ES 6973.

PAERL, H. W. & PAUL, V. J. 2012. Climate change: links to global expansion of harmful cyanobacteria. *Water research,* 46**,** 1349-1363.

PAHLEVAN, N., LEE, Z., WEI, J., SCHAAF, C. B., SCHOTT, J. R. & BERK, A. 2014. On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing. *Remote Sensing of Environment,* 154**,** 272-284.

PAHLEVAN, N., ROGER, J.-C. & AHMAD, Z. 2017a. Revisiting short-wave-infrared (SWIR) bands for atmospheric correction in coastal waters. *Optics express,* 25**,** 6015-6035.

PAHLEVAN, N., SARKAR, S., FRANZ, B., BALASUBRAMANIAN, S. & HE, J. 2017b. Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. *Remote sensing of environment,* 201**,** 47-56.

PAHLEVAN, N., SCHOTT, J. R., FRANZ, B. A., ZIBORDI, G., MARKHAM, B., BAILEY, S., SCHAAF, C. B., ONDRUSEK, M., GREB, S. & STRAIT, C. M. 2017c. Landsat 8 remote sensing reflectance (Rrs) products: Evaluations, intercomparisons, and enhancements. *Remote sensing of environment,* 190**,** 289-301.

PAHLEVAN, N., SMITH, B., SCHALLES, J., BINDING, C., CAO, Z., MA, R., ALIKAS, K., KANGRO, K., GURLIN, D. & HÀ, N. 2020. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sensing of Environment,* 240**,** 111604.

PALMER, S. C., KUTSER, T. & HUNTER, P. D. 2015. Remote sensing of inland waters: Challenges, progress and future directions. *Remote sensing of Environment*, *157*, pp.1-8.

PIAO, Y., JEONG, S., PARK, S. & LEE, D. 2021. Analysis of land use and land cover change using time-series data and random forest in North Korea. *Remote Sensing,* 13**,** 3501.

PLESSL, C., GILBERT, B. M., SIGMUND, M. F., THEINER, S., AVENANT-OLDEWAGE, A., KEPPLER, B. K. & JIRSA, F. 2019. Mercury, silver, selenium and other trace elements in three cyprinid fish species from the Vaal Dam, South Africa, including implications for fish consumers. *Science of the Total Environment,* 659**,** 1158-1167.

PRIJITH, S. S., SRINIVASARAO, K., LIMA, C. B., GHARAI, B., RAO, P. V. N., SESHASAI, M. V. R. & RAMANA, M. V. 2021. Effects of land use/land cover alterations on regional meteorology over Northwest India. *Science of the Total Environment,* 765**,** 142678.

PROBST, P., WRIGHT, M. N. & BOULESTEIX, A. L. 2019. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery,* 9**,** e1301.

QIN, B., LI, W., ZHU, G., ZHANG, Y., WU, T. & GAO, G. 2015. Cyanobacterial bloom management through integrated monitoring and forecasting in large shallow eutrophic Lake Taihu (China). *Journal of hazardous materials,* 287**,** 356-363.

RANDOLPH, K., WILSON, J., TEDESCO, L., LI, L., PASCUAL, D. L. & SOYEUX, E. 2008. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll a and phycocyanin. *Remote Sensing of Environment,* 112**,** 4009-4019.

RASTOGI, R. P., SINHA, R. P. & INCHAROENSAKDI, A. 2014. The cyanotoxin-microcystins: current overview. *Reviews in Environmental Science and Bio/Technology,* 13**,** 215-249.

RESOLUTION, U. 1993. Analysis of suspended solids in water using remotely sensed high resolution derivative spectra [J]. *Photogrammetric Engineering & Remote Sensing,* 59**,** 505-510.

REVENGA, C. & KURA, Y., 2003. Status and trends of biodiversity of inland water ecosystems. Secretariat of the Convention on Biological Diversity Montreal.

RICHARDSON, L. L. 1996. Remote sensing of algal bloom dynamics. *BioScience,* 46**,** 492-501.

ROBARTS, R. D. & ZOHARY, T. 2018. Limnology and the future of African inland waters. *Inland Waters,* 8**,** 399-412.

RODRIGUEZ-GALIANO, V. F., GHIMIRE, B., ROGAN, J., CHICA-OLMO, M. & RIGOL-SANCHEZ, J. P. 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing,* 67**,** 93-104.

ROUAULT, M., ROY, S. S. & BALLING JR, R. C. 2013. The diurnal cycle of rainfall in South Africa in the austral summer. *International journal of climatology,* 33**,** 770-777.

SAKUNO, Y., YAJIMA, H., YOSHIOKA, Y., SUGAHARA, S., ABD ELBASIT, M., ADAM, E. & CHIRIMA, J. 2018a. Evaluation of Unified Algorithms for Remote Sensing of Chlorophyll-a and Turbidity in Lake Shinji and Lake Nakaumi of Japan and the Vaal Dam Reservoir of South Africa under Eutrophic and Ultra-Turbid Conditions. *Water,* 10**,** 618.

SAKUNO, Y., YAJIMA, H., YOSHIOKA, Y., SUGAHARA, S., ABD ELBASIT, M. A., ADAM, E. & CHIRIMA, J. G. 2018b. Evaluation of unified algorithms for remote sensing of chlorophyll-a and turbidity in lake Shinji and lake Nakaumi of Japan and the Vaal Dam reservoir of South Africa under eutrophic and ultra-turbid conditions. *Water,* 10**,** 618.

SANDERSON, E. W., JAITEH, M., LEVY, M. A., REDFORD, K. H., WANNEBO, A. V. & WOOLMER, G. 2002. The human footprint and the last of the wild: the human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. *BioScience,* 52**,** 891-904.

SCHALLES, J. F. 2006. Optical remote sensing techniques to estimate phytoplankton chlorophyll a concentrations in coastal waters with varying suspended matter and CDOM concentrations. *Remote sensing of aquatic coastal ecosystem processes. Science and management applications*, pp.27-79.

SCOTT, W. 1991. Occurrence and significance of toxic cyanobacteria in Southern Africa. *Water Science and Technology,* 23**,** 175-180.

SHANG, R. & ZHU, Z. 2019. Harmonizing Landsat 8 and Sentinel-2: A time-series-based reflectance adjustment approach. *Remote Sensing of Environment,* 235**,** 111439.

SHEN, M., DUAN, H., CAO, Z., XUE, K., LOISELLE, S. & YESOU, H. 2017. Determination of the downwelling diffuse attenuation coefficient of lake water with the sentinel-3A OLCI. *Remote Sensing,* 9**,** 1246.

SHETTY, S., GUPTA, P. K., BELGIU, M. & SRIVASTAV, S. 2021. Assessing the effect of training sampling design on the performance of machine learning classifiers for land cover mapping using multi-temporal remote sensing data and google earth engine. *Remote sensing,* 13**,** 1433.

SHI, K., ZHANG, Y., QIN, B. & ZHOU, B. 2019. Remote sensing of cyanobacterial blooms in inland waters: present knowledge and future challenges. *Science Bulletin,* 64**,** 1540-1556.

SHI, W. & WANG, M. 2009. An assessment of the black ocean pixel assumption for MODIS SWIR bands. *Remote Sensing of Environment,* 113**,** 1587-1597.

SIMIS, S. G., PETERS, S. W. & GONS, H. J. 2005. Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water. *Limnology and oceanography,* 50**,** 237-245.

SMITH, V. H. 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research,* 10**,** 126-139.

SNYDER, P., DELIRE, C. & FOLEY, J. 2004. Evaluating the influence of different vegetation biomes on the global climate. *Climate Dynamics,* 23**,** 279-302.

SOKOLETSKY, L. G., LUNETTA, R. S., WETZ, M. S. & PAERL, H. W. 2011. MERIS retrieval of water quality components in the turbid Albemarle-Pamlico Sound Estuary, USA. *Remote Sensing,* 3**,** 684-707.

SONZOGNI, W. C., CHESTERS, G., COOTE, D., JEFFS, D., KONRAD, J., OSTRY, R. & ROBINSON, J. 1980. Pollution from land runoff. *Environmental Science & Technology,* 14**,** 148-153.

SÒRIA-PERPINYÀ, X., VICENTE, E., URREGO, P., PEREIRA-SANDOVAL, M., TENJO, C., RUÍZ-VERDÚ, A., DELEGIDO, J., SORIA, J. M., PEÑA, R. & MORENO, J. 2021. Validation of water quality monitoring algorithms for sentinel-2 and sentinel-3 in mediterranean inland waters with in situ reflectance data. *Water,* 13**,** 686.

STEVENSON, J. 2014. Ecological assessments with algae: a review and synthesis. *Journal of Phycology,* 50**,** 437-461.

SUN, D., LI, Y., WANG, Q., GAO, J., LE, C., HUANG, C. & GONG, S. 2013. Hyperspectral remote sensing of the pigment c-phycocyanin in turbid inland waters, based on optical classification. *IEEE Transactions on Geoscience and Remote Sensing,* 51**,** 3871-3884.

SWANEPOEL, A., BARNARD, S., RECKNAGEL, F. & CAO, H. 2016. Evaluation of models generated via hybrid evolutionary algorithms for the prediction of Microcystis concentrations in the Vaal Dam, South Africa. *Water SA,* 42**,** 243-252.

TAHERI TIZRO, A., GHASHGHAIE, M., GEORGIOU, P. & VOUDOURIS, K. 2014. Time series analysis of water quality parameters. *Journal of Applied Research in Water and Wastewater,* 1**,** 40-50.

TALUKDAR, S., SINGHA, P., MAHATO, S., PAL, S., LIOU, Y.-A. & RAHMAN, A. 2020. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sensing,* 12**,** 1135.

TANG, X., WOODCOCK, C. E., OLOFSSON, P. & HUTYRA, L. R. 2021. Spatiotemporal assessment of land use/land cover change and associated carbon emissions and uptake in the Mekong River Basin. *Remote Sensing of Environment,* 256**,** 112336.

THENKABAIL, P. S., SCHULL, M. & TURRAL, H. 2005. Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data. *Remote Sensing of Environment,* 95**,** 317-341.

TOMING, K., KUTSER, T., LAAS, A., SEPP, M., PAAVEL, B. & NÕGES, T. 2016. First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sensing,* 8**,** 640.

TOPP, S. N., PAVELSKY, T. M., JENSEN, D., SIMARD, M. & ROSS, M. R. 2020. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. *Water,* 12**,** 169.

TRAN, M. D., VANTREPOTTE, V., LOISEL, H., OLIVEIRA, E. N., TRAN, K. T., JORGE, D., MÉRIAUX, X. & PARANHOS, R. 2023. Band ratios combination for estimating chlorophyll-a from sentinel-2 and sentinel-3 in coastal waters. *Remote Sensing,* 15**,** 1653.

TWISA, S. & BUCHROITHNER, M. F. 2019. Land-use and land-cover (LULC) change detection in Wami River Basin, Tanzania. *Land,* 8**,** 136.

TYLER, A., SVAB, E., PRESTON, T., PRÉSING, M. & KOVÁCS, W. 2006. Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. *International Journal of Remote Sensing,* 27**,** 1521-1537.

VAN GINKEL, C. 2004. A national survey of the incidence of cyanobacterial blooms and toxin production in major impoundments. *Rapport South Africa*.

VAN GINKEL, C. 2011. Eutrophication: Present reality and future challenges for South Africa. *Water SA,* 37**,** 693-702.

VERHEUL, J. K. 2012. *Assessing the effects of different land uses on water quality in the Upper Wilge River Catchment.* University of Johannesburg, South Africa.

VISHNU PRASANTH, B., SIVAKUMAR, R. & RAMARAJ, M. 2022. A study on algae bloom pigment in the eutrophic lake using bio-optical modelling: hyperspectral remote sensing approach. *Bulletin of Environmental Contamination and Toxicology,* 109**,** 962-968.

WALMSLEY, R. 2002. A review and discussion document: Perspectives on eutrophication of surface waters: Policy or Research Needs in South Africa. WRC Report No. KV.

WANG, J. & CHEN, X. 2024. A new approach to quantify chlorophyll-a over inland water targets based on multi-source remote sensing data. *Science of The Total Environment,* 906**,** 167631.

WANG, M. & SHI, W. 2005. Estimation of ocean contribution at the MODIS near-infrared wavelengths along the east coast of the US: Two case studies. *Geophysical research letters,* 32.

WANG, M. & SHI, W. 2007. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing. *Optics Express,* 15**,** 15722-15733.

WANG, M., SON, S. & HARDING JR, L. W. 2009. Retrieval of diffuse attenuation coefficient in the Chesapeake Bay and turbid ocean regions for satellite ocean color applications. *Journal of Geophysical Research: Oceans,* 114, 2156-2202.

WANG, X., DONG, X., LIU, H., WEI, H., FAN, W., LU, N., XU, Z., REN, J. & XING, K. 2017. Linking land use change, ecosystem services and human well-being: A case study of the Manas River Basin of Xinjiang, China. *Ecosystem services,* 27**,** 113-123.

WANG, Y., XIA, H., FU, J. & SHENG, G. 2004. Water quality change in reservoirs of Shenzhen, China: detection using LANDSAT/TM data. *Science of the Total Environment,* 328**,** 195-206.

WARDLOW, B. D. & EGBERT, S. L. 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains. *Remote sensing of environment,* 112**,** 1096-1116.

WATANABE, F., ALCANTARA, E., RODRIGUES, T., ROTTA, L., BERNARDO, N. & IMAI, N. 2017. Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil). *Anais da Academia Brasileira de Ciências,* 90**,** 1987-2000.

WATANABE, F. S. Y., ALCÂNTARA, E., RODRIGUES, T. W. P., IMAI, N. N., BARBOSA, C. C. F. & ROTTA, L. H. D. S. 2015. Estimation of chlorophyll-a concentration and the trophic state of the Barra Bonita hydroelectric reservoir using OLI/Landsat-8 images. *International journal of environmental research and public health,* 12**,** 10391-10417.

WELCH, E. B. 1969. Factors initiating phytoplankton blooms and resulting effects on dissolved oxygen in Duwamish River estuary, Seattle, Washington. US Govt. Print. Off.

WELCH, E. B. & COOKE, G. D. 2005. Internal phosphorus loading in shallow lakes: importance and control. *Lake and reservoir management,* 21**,** 209-217.

WINSLOW, L. A., LEACH, T. H. & ROSE, K. C. 2018. Global lake response to the recent warming hiatus. *Environmental Research Letters,* 13**,** 054005.

WU, D., LI, R., ZHANG, F. & LIU, J. 2019. A review on drone-based harmful algae blooms monitoring. *Environmental monitoring and assessment,* 191**,** 1-11.

XU, Y., YU, L., ZHAO, F. R., CAI, X., ZHAO, J., LU, H. & GONG, P. 2018. Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: Experiments from three sites in Africa. *Remote Sensing of Environment,* 218**,** 13-31.

YUAN, F. 2008. Land-cover change and environmental impact analysis in the Greater Mankato area of Minnesota using remote sensing and GIS modelling. *International Journal of Remote Sensing,* 29**,** 1169-1184.

ZHANG, T., HUANG, M. & WANG, Z. 2020. Estimation of chlorophyll-a Concentration of lakes based on SVM algorithm and Landsat 8 OLI images. *Environmental Science and Pollution Research,* 27**,** 14977-14990.