



UNIVERSITY *of the*
WESTERN CAPE

**Developing an integrated remotely sensed framework for the detection and
monitoring of seasonally-flooded wetlands in semi-arid environments of
southern Africa**

by

Siyamthanda Gxokwe

A thesis submitted to the Faculty of Natural Science, University of the Western Cape, in
academic fulfilment of the requirements for the degree

Doctor of Philosophy

in

Environmental and Water Science

Supervisor: Prof Timothy Dube

Co-Supervisor: Prof Dominic Mazvimavi

October 2022

ABSTRACT

Developing an integrated remotely-sensed framework for the detection and monitoring of seasonally-flooded wetlands in semi-arid environments of southern Africa

Siyamthanda Gxokwe

PhD thesis, Department of Earth Science, University of the Western Cape

Wetlands are among the most important ecosystems on earth; they cover approximately 4-6% of the earth's surface and offer critical eco-hydrological services. However, these ecosystems are under threat from anthropogenic activities, droughts and climate variability, as well as from global environmental change. It is estimated that over 60% of the world's wetlands have been lost due to climate change and variability, as well as other anthropogenic influences. There is, therefore, a need for their routine monitoring and assessment to ensure the sustainable use and management of these systems on a national, regional and local scale, and prevent their further degradation and loss. This study aimed at developing an integrated cloud-computing-based, remotely-sensed framework for the detection and monitoring of small and seasonally-flooded wetlands along the semi-arid Limpopo Transboundary River Basin of southern Africa, which was previously a challenging task when using the traditional assessment and monitoring methods. The availability of cloud-computing platforms, like the Google Earth Engine (GEE) and the presence of advanced image-processing algorithms is perceived to offer unique opportunities for improving the detection and monitoring of small and seasonally-flooded wetlands in semi-arid environments. More specifically, four key objectives were set: firstly, the capability of the GEE and advanced machine-learning algorithms were explored for characterising and mapping small and seasonally-flooded wetlands in semi-arid environments. Secondly, an evaluation was conducted on the volume of freely-available archival remote sensing data in the GEE catalogue that is capable of monitoring small and seasonally-flooded wetlands in semi-arid southern Africa. Thirdly, the potential of using the GEE cloud-computing platform was assessed for the monitoring of the long-term variations in the selected eco-hydrological attributes of seasonally-flooded wetlands in semi-arid southern Africa. Finally, the GEE techniques and machine-learning algorithms for the large-scale monitoring of the impacts of adjacent land use and land cover changes on seasonally-flooded wetlands in the area were upscaled. The findings of the study underscore the relevance of the GEE and its advanced image-processing techniques for detecting and monitoring the eco-hydrological dynamics of small and seasonally-flooded

wetlands in semi-arid areas. The integration of multi-temporal spatial data, and the synergic use of spectral data and remotely-sensed indices, such as the Normalised Difference Vegetation Index, the Normalised Difference Water Index and the Normalised Difference Phenology Index (NDVI, NDWI and NDPI) allowed for the increased enhancement of wetland features. The study also noted that the Random Forest, the Support Vector Machine and the Classification and Regression Tree are the most appropriate GEE machine-learning algorithms for detecting and monitoring these systems. Furthermore, a significant volume of remotely-sensed data was retrieved, capable of monitoring small and seasonally-flooded wetlands in semi-arid areas. In addition, the integration of multi-source spatial data, such as the Synthetic Aperture Radar and Landsat series datasets, improved the detection and monitoring of these wetlands. Overall, this work provides the necessary baseline framework for the eco-hydrological monitoring and assessment of wetlands in semi-arid data-scarce environments. However, it was noted that the large-scale monitoring of these systems on the GEE present challenges associated with computational time costs, which results in computational time-out errors. The study therefore recommends split and parallel processing for the larger-scale mapping and monitoring of these systems.



October 2022

PREFACE 1

The goal of the study was to develop an integrated cloud-based, remotely-sensed framework for the detection and monitoring of small and seasonally-flooded wetlands in the semi-arid environments of southern Africa. The approach of the study included a succession of independent, but linked, papers which form the different chapters of this thesis. This thesis comprises of seven chapters, with four chapters that are conceptualised as stand-alone research papers that address the objectives of the study, which are listed in Chapter one, and one review paper, which is presented as chapter two.

Each chapter reads independently from the rest of the thesis, but draws conclusions that are linked and relevant to the scope of the whole thesis. Although the structure of this thesis conforms to the standards of the University of the Western Cape, a degree of repetition is certain, given the shared thread of the papers.

- **Chapter 1:** Gives a general introduction, background and contextualisation of the study.
- **Chapter 2:** This chapter is based on publication 1 and provides a detailed review of the progress, challenges and future research directions regarding the use of freely-available multispectral datasets for understanding the dynamics of wetlands within semi-arid environments.
- **Chapter 3:** This chapter addresses objective 1 of the study, which presents the findings on the characterisation and mapping of two small seasonally-flooded wetlands in the Limpopo Transboundary Basin in South Africa, using new-generational Sentinel-2 data, coupled with advanced machine-learning algorithms on the Google Earth Engine cloud-computing platform are presented. This chapter is based on publication 2.
- **Chapter 4:** This chapter addresses objective 2 of the study, which presents the findings on an evaluation of the available and useful remotely-sensed data in the Google Earth Engine catalogue for monitoring the long-term eco-hydrological dynamics of the small, seasonally-flooded wetlands in semi-arid South Africa. This chapter is based on manuscript 1.
- **Chapter 5:** This chapter addresses objective 3 which presents the findings on the use of cloud-computing artificial intelligence techniques for monitoring the long-term variations in the eco-hydrological dynamics of small seasonally-flooded wetlands in semi-arid South Africa. This chapter is based on publication 3.

- **Chapter 6:** This chapter addresses objective 4 of the study, which presents findings on the use of multi-source data, coupled with cloud-computing techniques, to assess the large-scale impacts of land use and land cover changes on seasonally-flooded wetlands in semi-arid southern Africa. This chapter is based on manuscript 2
- **Chapter 7:** Provides a synthesis of the study, as well as the conclusions and recommendations.



PREFACE 2

This research study was conducted in the Department of Earth Sciences (ESD) at the University of the Western Cape, South Africa, from August 2018 to July 2022, under the supervision of Prof. Timothy Dube and Prof. Dominic Mazvimavi (Earth Sciences, University of the Western Cape). I declare that the work presented in this thesis has never been submitted, in any form, to any other institution. This work represents my original work, except where due acknowledgements are made.

Candidate:

Siyamthanda Gxokwe

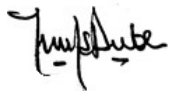


Date: October 2022

We certify that the above statement is correct and that, as the candidate's Supervisors, we have approved this thesis for submission.

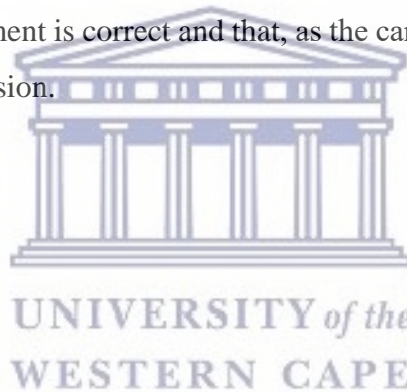
Supervisor:

Prof Timothy Dube



Co-supervisor:

Prof Dominic Mazvimavi



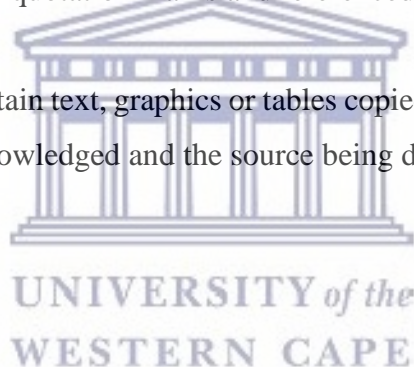
Date: 07 October 2022

Date:

DECLARATION 1: PLAGIARISM

I, Siyamthanda Gxokwe, declare that:

1. the research reported in this thesis, except where otherwise indicated, is my original research;
2. this thesis has not been submitted for any degree or examination at any other university;
3. this thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;
4. this thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - their words have been re-written, but the general information attributed to them has been referenced; and
 - where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged and the source being detailed in the thesis and in the References sections.



Signed 

Siyamthanda Gxokwe

October 2022

DECLARATION 2: PUBLICATIONS AND MANUSCRIPTS

The details of the publications and manuscripts forming this thesis are as follows:

Publication 1:

Gxokwe, S., Dube, T., Mazvimavi, D. 2020. Multispectral remote sensing of wetlands in semi-arid and arid areas: A review on applications, challenges and possible future research directions. *Remote Sensing* **12** (4190): [DOI:rs12244190](https://doi.org/10.3390/rs12244190) pp. 1-17 (**IF = 5.349**)

Publication 2:

Gxokwe, S., Dube, T., Mazvimavi, D. 2021. Leveraging the Google Earth Engine platform to characterise and map small seasonally-flooded wetlands in the semi-arid environments of South Africa. *Science of the Total Environment*. **803** (2022): <http://doi.org/10.1016/j.scitotenv.2021.150139> (**IF = 10.753**)

Manuscript 1: *under review*

Gxokwe, S., Dube, T., Mazvimavi, D. 2022. Available satellite data capable of monitoring small and seasonally-flooded wetlands in semi-arid southern Africa. *Ecohydrology*, **Manuscript no. ECO-22-0119** (**IF = 3.166**)

Publication 3:

Gxokwe, S., Dube T., Mazvimavi, D., Grenfell, M.C. 2022. Using cloud-computing techniques to monitor the long-term variations in eco-hydrological dynamics of the small seasonally-flooded wetlands in semi-arid South Africa. *Journal of Hydrology*. **612** (2022): <https://doi.org/10.1016/j.jhydrol.2022.128080> (**IF = 6.708**)

Manuscript 2: *Under review*

Gxokwe, S., Dube, T., Mazvimavi, D. 2022. Large-scale analysis of Land use Land cover changes impacts of seasonally-flooded wetlands extent in the semi-arid southern Africa. *Hydrology and Earth Science Systems*. **Manuscript no. hess-2022-272** (**IF = 6.617**)

All the data collection and analysis, experiments and preparation of the above publications and manuscripts were completed by the candidate, Siyamthanda Gxokwe, under the supervision of Prof Timothy Dube and Prof Dominic Mazvimavi.

ACKNOWLEDGEMENTS

In a Doctoral study, one is indebted to so many people and sources; therefore, it is practically impossible to list every individual and source. At the risk of omission, I would like to extend my profound gratitude to the following people and organisations:

- My supervisory team, Prof Timothy Dube and Prof Dominic Mazvimavi, from the Department of Earth Science, University of the Western Cape, for their unwavering support, guidance, constructive criticism and suggestions throughout the period of this PhD. I am grateful to them for providing me with the opportunity to work under them, and for believing in me, when I doubted my capabilities. I have learnt a lot from them; not only things relating to academics, but also how to tackle life's challenges.
- My co-authors in some of the papers published, particularly Prof Michael Grenfell from the Department of Earth Science at the University of the Western Cape, and Dr Jane Olwoch from the Southern African Science Services Centre for Climate Change and Adaptive Land Management (SASSCAL), for their valuable input, comments and suggestions.
- GMES-AFRICA and SASSCAL for funding this work under the WeMAST project, as well as the National Research Foundation (NRF) of South Africa. My gratitude also goes to the Department of Economic Development, Environment and Tourism in Limpopo and the owners of the Lindani Private Game Lodge for granting me permission to access the studied wetlands, as well as the to the Agricultural Research Council (ARC) for providing the long-term weather data.
- My friends, Dr Bester Mudereri, Dr Humphrey Thamaga, Dr Cynthia Laar, Dr Mandla Dlamini, Dr Zanele Ntshidi, Dr Faith Jumbi, Mr Kudzai Mpakairi, Mr Eugene Maswanganye, Ms Tatenda Dzurume, Ms Yonela Mkunyana, Ms Retang Mokuua, Mr Bongolethu Mtengwana, Mr James Machingura and Mr Abongile Xaza, to name a few, for your support and encouragement, as well as the night-shifts we spent together at the Earth Science post-graduate laboratory.
- The SASSCAL-Namibia team, particularly Ms Bianca Mutale, Ms Nancy Makwatikizo and Mr Panduleni Hamukwaya, as well as Mrs Mandy Naidoo ("*Prof Naidoo*") from the Department of Earth Science, University of the Western Cape, for ensuring that my finances reached me on time.

- My family, particularly my mother, for their love, committed support and prayers during the course of this PhD; and for understanding why I had to be away from them for a while.
- Above all, I would like to thank the Lord Almighty for the love and strength that He has given me to successfully complete this PhD.



DEDICATION

This thesis is dedicated to the family of the “*Amancotshe*” clan, and serves as a motivation to my brothers and sisters that ‘it does not matter where you come from, but what matters is where you are going.’

Jeremiah 29:11 - For I know the plans I have for you, declares the LORD, plans to prosper you and not to harm you, plans to give you hope and a future.



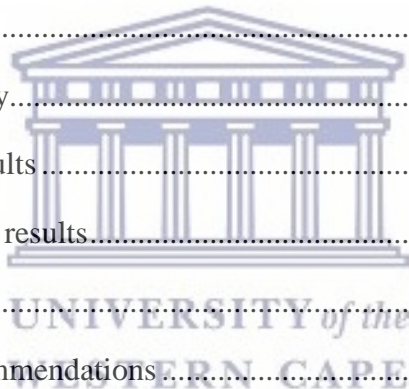
TABLE OF CONTENTS

ABSTRACT	i
PREFACE 1	iii
PREFACE 2	v
DECLARATION 1: PLAGIARISM	vi
DECLARATION 2: PUBLICATIONS AND MANUSCRIPTS	vii
AKNOWLEDGEMENTS	viii
DEDICATION	X
LIST OF FIGURES	xvi
LIST OF TABLES	xviii
CHAPTER 1	1
GENERAL INTRODUCTION	1
1.1 Background and Problem Statement.....	2
1.2 Research Question.....	4
1.3 Main Hypothesis	4
1.4 Aim and Objectives	5
1.4.1 Aim.....	5
1.4.2 Objectives	5
1.5 Description of the Study Area.....	5
CHAPTER 2	8
MULTISPECTRAL REMOTE SENSING OF WETLANDS IN SEMI-ARID AND ARID AREAS: A REVIEW ON THE APPLICATIONS, CHALLENGES AND FUTURE RESEARCH DIRECTIONS	8
Abstract	9
2.1 Introduction	10

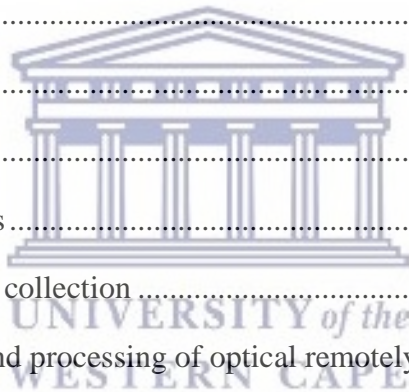


2.2 Semi-arid and Arid Wetlands Characteristics and Key Monitoring and Management Challenges	13
2.3 Commonly-used, Freely-available Multispectral data for Semi-arid and Arid Wetland Inventories	15
2.4 Mapping Semi-arid and Arid Wetland Vegetation, using Freely-available Multispectral Images	19
2.5 Mapping Wetland Inundation by using Freely-available Multispectral Images	22
2.6 Mapping Land-use and Land-cover Changes Impacts of Semi-arid and using Freely-available Multispectral Images.....	23
2.7 Low to Medium vs High Resolution Remote Sensing for Wetland Monitoring and Assessment	24
2.8 Available Satellite Image Processing Techniques for Accurate Wetland Monitoring ..	27
2.9 Summary of Key Challenges and Future Research Directions	28
2.10 Conclusion.....	30
CHAPTER 3	31
LEVERAGING GOOGLE EARTH ENGINE PLATFORM TO CHARACTERISE AND MAP SMALL SEASONALLY-FLOODED WETLANDS IN THE SEMI-ARID ENVIRONMENTS OF SOUTH AFRICA	31
Abstract	32
3.1 Introduction	33
3.2 Materials and Methods	35
3.2.1 Field data.....	35
3.2.2 Remote sensing data acquisition and processing	37
3.2.3 Adopted wetland classification scheme	40
3.2.4 Accuracy assessment	41
3.3 Results	42
3.3.1 Class separability results from the spectral data used.....	42
3.3.2 Classification results and accuracies.....	44

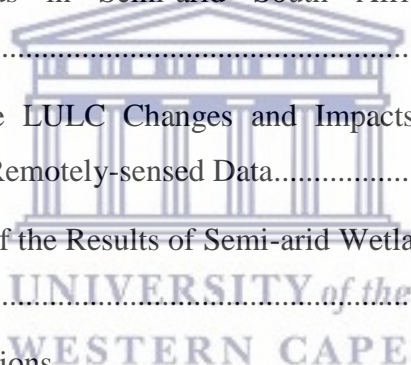
3.4	Discussion	49
3.5	Conclusion and Recommendations	52
CHAPTER 4	54
AVAILABLE SATELLITE DATA CAPABLE OF MONITORING SMALL AND SEASONALLY-FLOODED WETLANDS IN SEMI-ARID SOUTHERN AFRICA		
	Abstract	55
4.1	Introduction	56
4.2	Materials and Methods	58
4.2.1	Ground truth data collection	58
4.2.2	Remotely-sensed data collection and analysis	59
4.2.3	Image classification	63
4.2.4	Accuracy assessment	63
4.3	Results	64
4.3.1	Image availability.....	64
4.3.2	Classification results.....	65
4.3.3	Accuracy assessment results.....	68
4.4	Discussion	73
4.5	Conclusion and Recommendations	77
CHAPTER 5	78
USING CLOUD-COMPUTING TECHNIQUES TO MONITOR LONG-TERM VARIATIONS IN ECO-HYDROLOGICAL DYNAMICS OF THE SMALL SEASONALLY-FLOODED WETLANDS IN SEMI-ARID SOUTH AFRICA.....		
	Abstract	79
5.1	Introduction	80
5.2	Materials and Methods	82
5.2.1	Field data collection.....	82
5.2.2	Ancillary data.....	83
5.2.3	Remote sensing data acquisition and processing.....	83



5.2.4	Adopted wetland classification approach	87
5.2.5	Accuracy assessment	87
5.2.6	Change analysis, variations and major drivers	88
5.3	Results	88
5.3.1	Variations in the eco-hydrological dynamics	90
5.3.2	Accuracy assessment	92
5.3.3	Drivers of eco-hydrological variations	93
5.4	Discussion	95
5.5	Conclusion and Recommendation.....	98
CHAPTER 6	100
LARGE-SCALE ANALYSIS OF LAND USE LAND COVER CHANGES ON SEASONALLY-FLOODED WETLANDS EXTENT IN SEMI-ARID SOUTHERN AFRICA	100
Abstract	101
6.1	Introduction	103
6.2	Materials and Methods	104
6.2.1	Ground truth data collection	104
6.3.2	Pre-preparation and processing of optical remotely-sensed data.....	106
6.2.3	Pre-preparation and processing of Synthetic Aperture Radar data	108
6.2.4	Pixel-based image analyses.....	110
6.2.5	Accuracy assessment	111
6.2.5	Land use and land cover change analysis	111
6.3	Results	112
6.3.1	Accuracy assessment results	112
6.3.2	Land cover change analysis	114
6.3.3	LULC transitions and impacts on the extent of wetlands	117
6.4	Discussion	119
6.5	Conclusions	121



CHAPTER 7	123
AN INTEGRATED REMOTE-SENSING FRAMEWORK FOR THE DETECTION AND MONITORING OF SMALL SEASONALLY-FLOODED WETLANDS IN SEMI-ARID SOUTHERN AFRICA: A SYNTHESIS	123
7.1 Introduction.....	123
7.2 A Review of the Progress, Challenges and Future Research Directions of the Multispectral Remote Sensing of Wetlands in Semi-arid and Arid Areas.....	124
7.3 The Characterisation and Mapping of Small Seasonally-flooded Wetlands in the Semi-arid Environments of South Africa, using the Google Earth Engine	125
7.4 Available Remotely-sensed Data in the GEE Catalogue, which are capable of Monitoring the Long-term Variations in the Eco-hydrological Dynamics of Small and Seasonally-flooded Wetlands in Semi-arid South Africa	126
7.5 Monitoring the Long-term Variations in the Eco-hydrological Dynamics of Small and Seasonally-flooded Wetlands in Semi-arid South Africa, using Cloud-computing Techniques.	127
7.6 Assessing Large-scale LULC Changes and Impacts on Wetlands using Cloud-computing and Multisource Remotely-sensed Data.....	129
7.7 Overall Implications of the Results of Semi-arid Wetland Detection and Monitoring, using Remotely-sensed Data	130
7.8 Overall Recommendations	131
REFERENCES	133
APPENDICES	156



LIST OF FIGURES

Figure 1.1 Locality of the Limpopo Transboundary River Basin as the selected case study wetlands	7
Figure 2.1 Global extent of semi-arid and arid areas, with a number of studies in each region (modified from Scanlon et al., 2010).....	15
Figure 2.2 The number of remote sensing publications on semi-arid and arid wetlands (a) and the number of publications per area of focus (b). Characterisation includes studies on classification and wetland mapping, an impact analysis includes studies focusing of both climatic and anthropogenic impacts on wetlands, and an LC analysis refers to all studies on the wetland cover change analysis (n=196).....	16
Figure 3.1 Steps taken to characterise and map the two wetlands, using GEE.....	37
Figure 3.2 Wetland cover class spectral reflectance values for a) the Lindani valley bottom and b) the Nylsvley floodplain, using pixel values extracted from the training data	43
Figure 3.3 Overall accuracy comparison between the algorithms used for the two studied wetlands.....	45
Figure 3.4 Wetland cover class areas for a) the Lindani valley bottom and b) the Nylsvley floodplain.....	46
Figure 3.5 Producer’s and user’s accuracies for the two wetlands classes	47
Figure 3.7 Nylsvley floodplain Sentinel 2-derived wetland cover classes, based on the four algorithms used	49
Figure 4.1 Steps taken to execute the classifications using GEE (Red box indicates the image stack not included in the classification)	62
Figure 4.2 Landcover classes for the Nylsvley floodplain in different seasons and cloud cover between 2016 and 2017, based on Sentinel-2 Level 1C, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover	66
Figure 4.3 Landcover classes for the Nylsvley floodplain at different seasons and cloud cover between 2016 and 2017, based on Landsat 7, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover	67
Figure 4.4 Overall accuracies based on a) Sentinel-2 level 1C and b) Landsat 7 data	68
Figure 5.1 Image classification process using the GEE cloud-computing platform.....	86

Figure 5.2 Spatial distribution of wetland classes for the studied periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020), produced by using Object-based Random Forest classification of Landsat-7 data.....90

Figure 5.3 Nylsvley floodplain land cover surface areas per time periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020).....91

Figure 5.4 The relationship between biomass (indices) and annual ET rates for a) Period 1: 2000-2005, b) Period 2: 2006- 2010, c) Period 3: 2011-2015 and d) Period 4: 2016-202094

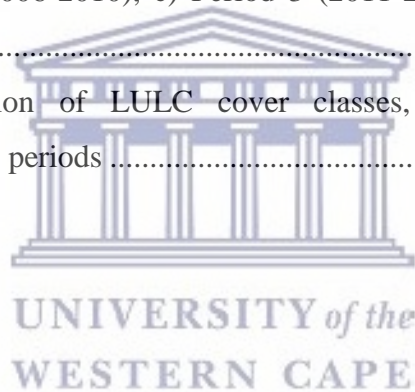
Figure 6.1 Steps followed on the GEE to process the remotely-sensed data used in this study..... 106

Figure 6.2 Overall accuracy for each time period (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020)..... 112

Figure 6.3 User’s and producer’s accuracy per studied time period. a) Period 1 (2000-2005), b) Period 2 (2006-2010), c) Period 3 (2011-2015) and d) Period 4 (2016-2020)..... 113

Figure 6.4 Pontius commission and omission errors for each class per time period. a) Period 1 (2000-2005), b) Period 2 (2006-2010), c) Period 3 (2011-2015) and d) Period 4 (2016-2020)..... 114

Figure 6.5 Spatial distribution of LULC cover classes, based on the pixel-based classification for the study time periods 116



LIST OF TABLES

Table 2.1 Commonly-used, freely-available sensors specifications for wetland inventories (modified from Ozesmi and Bauer, 2002)	17
Table 2.2 Selected low to high spatial resolution sensors for wetland monitoring and assessment (freely-available data sensors are highlighted in bold)	26
Table 2.3 Available algorithms for wetlands remote sensing	28
Table 3.1 Optical features extracted from the remotely-sensed data used	40
Table 3.2 JM distances between wetland-cover classes in the Lindani valley bottom.....	44
Table 3.3 JM distances between wetland-cover classes in the Nylsvley floodplain	44
Table 4.1 Indices extracted from the time series Landsat-7 and Sentinel-2 data	61
Table 4.2 Satellite images available for the Nylsvley floodplain between 2000 and 2020 per cloud-cover percentage	64
Table 4.3 Producer's and user's accuracies per cloud cover range for dry season scenes	68
Table 4.4 Producer's and user's accuracies per cloud cover range for wet season scenes....	69
Table 5.1 Number of remotely-sensed images with no cloud cover obtained per time intervals with their details.....	85
Table 5.2 Rate of change of the Nylsvley floodplain between the periods	91
Table 5.3 Error metrics results for classified image composites representing Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020.....	93
Table 6.1 Number of training and test points for LULC analysis	105
Table 6.2 Spectral indices extracted from the SAR and optical remotely-sensed data	109
Table 6.3 Land cover class areas (ha) per time period	115

CHAPTER 1

GENERAL INTRODUCTION



Photo: Courtesy of Siyamthanda Gxokwe, 2020

UNIVERSITY OF
WESTERN CAPE

1.1 Background and Problem Statement

Wetlands are some of the most critical ecosystems on earth. They cover about 4-6% of the earth's surface and occur at the transition zones between terrestrial and aquatic ecosystems (Tooth et al., 2002; Li et al., 2015). These ecosystems offer critical eco-hydrological services, which are categorised as provisioning, regulating and cultural services. The provisioning services include the provision of water for domestic use and livestock, raw materials and the production of wild food and medicine (Day et al., 2010; McCartney and Houghton-Carr, 2009; Thamaga et al., 2022). Regulation services include, amongst other things, flood attenuation, groundwater recharge, carbon sequestration, sediment retention, waste treatment and pest and pathogen regulations (Cape et al., 2015; Day et al., 2010). The cultural provisions include providing opportunities for cultural activities and heritage, as well as for recreational use and social amenities (Ewart-Smith et al., 2006; Ghobadi et al., 2012; Turpie and Malan, 2010). Despite offering such important services, over 60% of wetlands globally have been lost due to anthropogenic activities and climate variability and change (Millennium Ecosystem Assessment Programme, 2005).

Wetlands in semi-arid areas are largely small and intermittently flooded (mostly <10 ha -2500 ha) (Liu et al., 2015); they are therefore more susceptible to climate variability and anthropogenic influences. It has been reported that a significant proportion of the wetland systems in arid environments have been lost, due to anthropogenic influences, climate change and variability (Gebreslassie et al., 2014; Guo et al., 2017; Uluocha and Okeke, 2004), and that over 50% of the wetlands in semi-arid South Africa have been lost to date (Adeeyo et al. 2022; Barbier, 1993; Ewart-smith et al. 2006; Turpie and Malan, 2010). It has also been reported that over 30% of the wetlands in semi-arid China have been lost over the past 50 years, mostly due to anthropogenic influences (Chen and Liu, 2015). The same has been observed in semi-arid Ethiopia, where wetland degradation has been observed due to the unregulated use and over-utilisation of these systems by the surrounding communities (Bahilu and Tadesse, 2017). The sustainable use and management of wetlands on a national, regional and local scale is therefore necessary, in order to prevent further degradation and loss.

The basis for the sustainable use and management of wetlands hinges on the frequent monitoring and assessment of their extent, their status and the environmental impacts over time and space. This provides the baseline information that will inform the preventative measures and protection policies. Conventional traditional field mapping and assessment methods for

wetlands have previously provided baseline information on these systems; however, using these approaches for monitoring presents many difficulties, including their high cost and lack of accessibility of these wetlands for in situ monitoring (de Roeck et al., 2008). In addition, the traditional field assessments used relatively complex methods to compile wetland inventories, thus giving incomparable and inconsistent results (Rabelo et al., 2007).

Advances in remote sensing techniques have resulted in an increase in the use of robust geospatial techniques, with data inputs from different satellite sensors like MODIS, Landsat, IKONO, AVHRR, etc., for mapping and understanding the eco-hydrological dynamics of different wetland systems (Ozesmi and Bauer, 2002). The advantages of remote sensing techniques for monitoring wetland eco-hydrological dynamics include their cost-effectiveness, the fact that wetlands located in inaccessible sites can be accessed and that changes in wetland dynamics can be monitored at no cost, over a longer period of time and over a larger spatial extent (Mishra, 2014). However, the isolated nature of small and seasonally-flooded wetlands located in drylands presents challenges that are associated with the spatial resolution of the type of remotely-sensed data that are used, where coarse- to medium-resolution data, like MODIS, result in these wetlands being either missed or confused with other landcover classes during the classification process (Gómez-Rodríguez et al., 2010). This results in inaccurate information being derived regarding their extent and eco-hydrological dynamics. Moreover, the unavailability of cloud-free images for certain parts of the year further complicates the detection and monitoring of these wetlands from the coarse- to medium-resolution remote sensing data across the seasons (Mwita, 2013; Sakané et al., 2011).

Various studies have demonstrated how the sensor characteristics of different remote sensing products and cloud coverage influence the detection and monitoring of small and seasonally-flooded wetlands, particularly when using freely-available data. For example, Powell et al. (2019) used Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper plus (ETM+), coupled with LiDAR data, to map the wetland types in the semi-arid Barwon-Darling River floodplain in New South Wales. Although the study achieved acceptable overall accuracies, it failed to distinguish certain wetland types that were located along the boundaries of the other landcover classes. This was attributed mainly to the sensing characteristics of the remotely-sensed data that were used. A study by Mwita et al. (2012) examined the potential of the optical Landsat-7, coupled with ALOS-PALSAR data and Shuttle Radar Topographical Mission Digital Elevation (SRTM -DEM), to map semi-arid seasonally-flooded wetlands in Laikipia and Pangani in the humid Mt Kenya and Usambara Highlands in Tanzania. The results

also showed acceptable overall accuracies; however, they failed to delineate the semi-arid inland valley systems that were long, narrow and fragmented. This was also attributed to the sensing characteristics of the remotely-sensed data that were used, wetland size, as well as the use of dry season images because of unavailable cloud-free wet season scenes. Dry season images influenced the detection of wetland features, due to the absence of surface inundation and healthy wetland vegetation at the time when the images were acquired. These studies by Mwita et al., (2012) and Powell et al., (2019) demonstrate some challenges that are associated with the remote sensing of small and seasonally-flooded wetlands in semi-arid environments. They also demonstrate that mapping these systems at an improved precision, presents challenges that are associated with the sensing characteristics of the commonly-used coarse- to mid-resolution products, as well as the lack of seamless cloud-free remotely-sensed data that can enable the better detection and monitoring of these systems. There is therefore a need to develop robust methodologies and frameworks for improving the detection and monitoring of small and seasonally-flooded wetlands. These will integrate the strengths of remotely-sensed data from different sources and time periods, with different spectral characteristics, water and vegetation indices, as well as contextual parameters, which will further enhance the detection and monitoring of these systems. Currently, remotely-sensed frameworks for detecting wetlands rely either on a single date image analysis, or they are based on the spectral characteristics or indices, as stand-alone model input parameters (Amani et al., 2021; Mahdianpari et al., 2020). There is therefore a need to integrate these variables for the improved detection and monitoring of small and seasonally-flooded wetlands in semi-arid areas. Generating precise routine spatial and explicit wetland scale information will accurately inform the conservation, planning and management of these systems, which has been overlooked by the existing wetland management and conservation strategies in these regions, due to the lack of appropriate data.

1.2 Research Question

To what extent can integrated multisource data, coupled with a synergy of spectral characteristics, vegetation and water indices, enhance the detection and monitoring of small and seasonally-flooded wetlands that are located in semi-arid environments?

1.3 Main Hypothesis

In this study, it is hypothesised that the development of an integrated remotely-sensed framework, recently introduced cloud-computing platforms, such as Google Earth Engine and

its advanced processing algorithms, with higher-resolution and freely-available optical data and spectral indices synergies, has the potential to improve the remote detection and monitoring of small and seasonally-flooded wetlands in semi-arid environments.

1.4 Aim and Objectives

1.4.1 Aim

The main aim of this study was to develop an integrated cloud-computing-based, remote sensing framework for the detection and monitoring of small seasonally-flooded wetland along the semi-arid basin of southern Africa, using the Limpopo Transboundary River Basin as the study case.

1.4.2 Objectives

The objectives of this thesis are as follows:

1. to assess the capabilities of the Google Earth Engine cloud-computing platform and its advanced processing algorithms in characterising and mapping the wetlands types in semi-arid southern Africa;
2. to evaluate the amount of freely-available archival remote sensing data that are capable of monitoring small and seasonally-flooded wetlands in semi-arid southern Africa;
3. to assess the potential of using the Google Earth Engine cloud-computing platform to monitor the long-term variations in the selected eco-hydrological attributes of seasonally-flooded wetlands in semi-arid southern Africa; and
4. to evaluate the use of the Google Earth Engine platform and its advanced machine-learning algorithm in the large-scale monitoring of the impacts of adjacent land-use and land-cover changes in seasonally-flooded wetlands in semi-arid southern Africa.

1.5 Description of the Study Area

The study was conducted in the Limpopo Transboundary River Basin (LTRB) (Figure 1.1), which is shared by four countries in southern Africa, namely, Botswana, Mozambique, Zimbabwe and South Africa. The basin covers a total surface area of 412 938 km² and drains to northern South Africa, eastern Botswana, southern Zimbabwe and a bisect of southern Mozambique (Chapman, 2017; Dzurume et al., 2021). The LTRB is characterised by semi-arid climate conditions with the annual precipitation varying between 200 mm and 1 500 mm, and averaging about 500 mm. Ninety-five percent (95%) of this precipitation is received during the period from October to April. Evaporation in the LTRB varies between 800 mm/year to 2 000

mm/year, with an average of 1 970 mm/year (Mosase et al., 2019; Mosase and Ahiablame, 2018). The hydrology of the LTRB is represented by a number of major rivers, including the Crocodile and Marico Rivers in South Africa, the Notwane River in Botswana and the Sheshe River between Zimbabwe and Botswana, which are all tributaries of the Limpopo River that runs between these four countries. Several inter-basin transfer schemes also contribute water to the LTRB (Mosase et al., 2019; Chapman, 2017). The basin is mostly dominated by mixed land cover types, which include forested lands, built-up areas, savannah, mixed grasslands and croplands. Although the study focuses on the LTRB, in some sections of this thesis, individual wetlands, such as the Lindani valley bottom (Figure 1.1c) and the Nylsvley floodplain (Figure 1.1d), were used as case studies.

The Lindani valley bottom wetland is located between 24°03'01.36"E and 28°41'43.37"S within the boundaries of the Lindani Private Game Reserve at Vaalwater, in the Limpopo Province of South Africa. The wetland covers an area of about 28 ha and receives water mostly from the local rainfall, and groundwater seeps from several springs in the area. The dominant vegetation species include the *Oryza longistaminata* (rice grass), *Phragmites australis* (common reeds), *Scirpoides dioecus* (Kunth), as well as *Cynodon dactylon* (Bermuda grass). The Nylsvley floodplain is a Ramsar-protected system that is located at 24°39'17"S and 28°41'28"E near the towns of Mookgopong and Modemolle in the Limpopo Province of South Africa. The wetland forms a 70 km-long floodplain along the Mogalakwena River, which is a tributary of the Limpopo River (Dzurume, 2021). The dominant vegetation species in the Nylsvley floodplain include common grass species, such as *Oryza longistaminata* (rice grass) and *Phragmites australis* (common reeds), and tree species, such as *Acacia tortilis*, *Acacia nilotica* and *Acacia karoo*. The Nylsvley floodplain receives most of its inflow from seasonal rivers, such as the Olifantsspruit and the Groot and Klein Nyl (Dzurume, 2021). Although the Nylsvley is 70 km long, the study focused in the upper middle reach of the system located within the boundaries of the Nylsvley nature reserve.

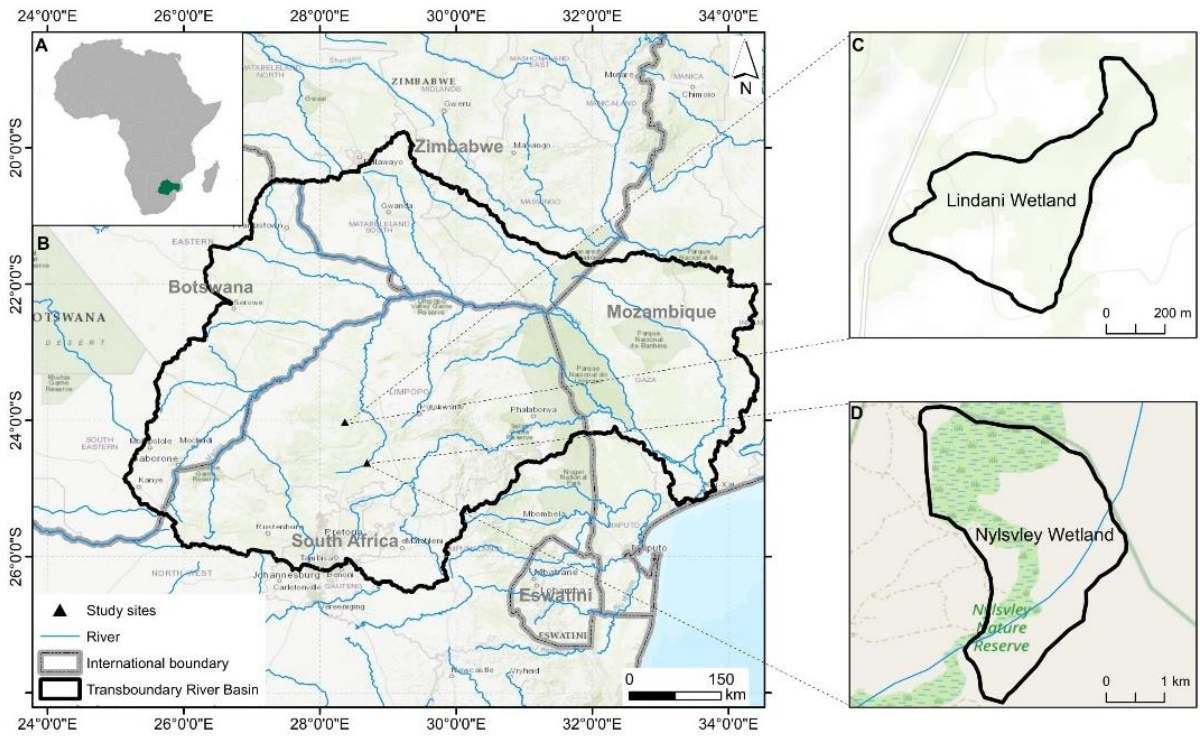


Figure 1.1 Locality of the Limpopo Transboundary River Basin as the selected case study wetlands



CHAPTER 2

MULTISPECTRAL REMOTE SENSING OF WETLANDS IN SEMI-ARID AND ARID AREAS: A REVIEW ON THE APPLICATIONS, CHALLENGES AND FUTURE RESEARCH DIRECTIONS



Photo: Courtesy of Siyamthanda Gxokwe, 2020

This chapter is based on the following publication:

Gxokwe, S., Dube, T., Mazvimavi, D. 2020. Multispectral Remote Sensing of Wetlands in Semi-arid and Arid Areas: A Review on Applications, Challenges and Possible Future Research Directions. *Remote Sensing* **12** (4190): [DOI:rs12244190](https://doi.org/10.3390/rs12244190) pp. 1-17 (*IF* = 5.349)

Abstract

Wetlands are ranked as very diverse ecosystems, and they cover about 4-6% of the global land surface. They occupy the transition zones between the aquatic and terrestrial environments, and share the characteristics of both zones. Wetlands play a critical role in the hydrological cycle, sustaining peoples' livelihoods, the aquatic life and biodiversity. The poor management of wetlands results in the loss of critical ecosystem goods and services. Globally, wetlands are degrading at a rapid rate, due to environmental change and anthropogenic activities. This requires holistic monitoring, assessment and management of wetlands to prevent their further degradation and loss. Remote sensing data offer an opportunity to assess changes in the status of these wetlands, including their spatial coverage. Thus far, a number of studies have used remotely-sensed data to assess and monitor the wetland status in semi-arid and arid regions. A literature search shows a significant increase in the number of papers published during the 2000-2020 period, with most of them being in the semi-arid regions of Australia and China, and a few in sub-Saharan Africa. This chapter reviews the progress that has been made in the use of remote sensing for detecting and monitoring the semi-arid and arid wetlands, and it focuses particularly on gaining new insights by using freely-available multispectral sensors. It begins by describing the important characteristics of wetlands in semi-arid and arid regions that require monitoring, in order to improve their management. Secondly, the use of freely-available multispectral imagery for compiling the wetland inventories is reviewed. Thirdly, the challenges of using freely-available multispectral imagery in the mapping and monitoring of wetland dynamics, like inundation, vegetation cover and extent, are examined. Lastly, algorithms for image classification as well as challenges associated with their uses and possible future research are summarised. There are, however, concerns regarding whether the spatial and temporal resolution of some of the remote sensing data enable the accurate monitoring of wetlands of varying sizes. Furthermore, it was noted that there were challenges associated with the both the spatial and spectral resolutions of data, used when mapping and monitoring wetlands. However, advancements in remote sensing and data analytics provide new opportunities for further research into the monitoring and assessment of wetlands across various scales.

Keywords: Data integration, inundation; multispectral imagery; semi-arid; seasonal wetland; vegetation dynamics

2.1 Introduction

There are several definitions of wetlands and most of them include abiotic and biotic factors, hydrological regimes, geomorphology and vegetation factors that control their existence. The Ramsar Convention definition is widely-used and includes these factors. However, this definition excludes areas with marine water greater than six meters at low tide. Wetlands exist where soils are saturated or inundated with water for varying durations and at different frequencies (Day et al., 2010). The Ramsar definition of wetlands includes not only those systems falling within the traditional concept of wetlands, such as mangrove swamps, peat bogs, tidal flats and water meadows, but also many other natural and man-made features, like flooded gravel peats, reservoirs, rice paddies and coastal beaches (Bowman, 2002). While it does not refer to the hydrological system, the definition includes components of the natural inland systems, and it pre-dates the recent conceptual developments and management of coastal and water systems (Shine and Klemm, 1999).

In North America, wetlands are defined as “lands that are either inundated by shallow water less than 2 meters deep during low water events, or that have soils that are saturated long enough during the growing season to become anoxic and that support specialised wetland plants (hydrophytes)” (Rochefort et al., 2012). Unlike the Ramsar definition, the North American definition takes into consideration the fact that wetland water can be at the soil surface, or below it, in some seasons. Several definitions of wetlands are used in South Africa, such as those used in the National Water Act (36 of 1998) and by SANBI (the South African National Biodiversity Institute). The South African National Water Act (SANWA) (36 of 1998) defines wetlands as “areas which are transitional between terrestrial and aquatic systems, where the water table is usually at, or near, the surface, or the land is periodically covered with shallow water and which in normal circumstances supports, or would support, vegetation typically adapted to life in saturated soils” (Ewart-Smith et al., 2006). For the purposes of this review, the Ramsar definition of wetlands will be used, as it is accepted globally. Since different countries have different wetland definitions, some ecosystems that are not considered as wetlands, based on the Ramsar definition, will be included.

Wetlands are ranked among the very diverse ecosystems that cover a proportion of about 4-6% of the land surface (Tooth et al., 2002; Li et al., 2015), and they provide an array of ecosystem services, which are categorised as provisioning, regulating and cultural services. The provisioning services include the provision of water for livestock and domestic use, raw materials, as well as genetic resources, and the production of wild foods and medicine,

(McCartney and Houghton-Carr, 2009). The regulating services include carbon sequestration, flood attenuation, groundwater replenishment, sediment retention, waste treatment and the regulation of pests and pathogens (Day et al., 2010). Cultural services include providing opportunities for cultural activities and heritage, recreational use and social interaction (Ghobadi et al., 2012; Turpie and Malan, 2010)

Globally, wetlands have been undergoing changes that result from natural and human anthropogenic causes (Fang et al., 2019). The natural causes include the severe droughts experienced in certain parts of the world, which have led to the drying and degradation of wetlands (Fang et al., 2019). The anthropogenic causes include the conversion of wetlands to agricultural land and pollution. There are, however, uncertainties regarding the global extent of wetland losses (Hu et al., 2017). These uncertainties are caused by inconsistencies in the data on the spatial changes in wetlands and the various sizes of these systems (Davidson, 2014). Liu et al. (2017) reported that natural wetlands in semi-arid China have been lost over the past 50 years and that about 30% of these systems disappeared between 1990 and 2000, mostly due to anthropogenic causes. In Africa, wetlands are considered to be among the threatened and degraded ecosystems (Landmann et al., 2010), while in the semi-arid parts of South Africa, at least 50% of all wetlands have been eradicated in some catchments (Day et al., 2010). Riddell et al., (2012) highlighted the fact that about 30-60% of the wetland losses in South Africa have been experienced in several major catchments, due to poor land-use management practices, which is also the case across sub-Saharan Africa. Uluocha and Okeke (2004) reported that wetlands in semi-arid Nigeria, which were previously recharging the groundwater systems, are undergoing degradation at an alarming rate, without any measures being taken to protect them.

Given the significance of the ecological services provided by wetlands, it is imperative that they are sustainably managed. One of the key elements for the sustainable management of wetlands is the continuous monitoring of the changes in their eco-hydrological dynamics (Slagter et al., 2020). This is a challenge for wetlands in semi-arid and arid areas, as most wetlands are seasonal or temporary, and they are inaccessible because of their remoteness (Schmid et al., 2004; Li et al., 2015). Remote sensing offers a unique opportunity for providing information about wetlands in a spatially-explicit manner, where other monitoring programs are not available (Chen and Liu, 2015), and it offers input data from various satellite sensors, ranging from multispectral to hyperspectral sensors. There are, however, concerns about whether the spatial and temporal resolution of some of the remote sensing data enables the

accurate detection and monitoring of wetlands of varying sizes, especially in semi-arid and arid areas (Li et al., 2019).

A significant number of reviews have been published on the remote sensing of wetlands. These reviews highlight the major progress that has been made on the use of remote sensing data, which ranges from low spatial resolution to hyperspectral imagery for inventorying wetlands in different climatic zones (Guo et al., 2017; Klemas, 2014; Dronova, 2015; Amler et al., 2015; Adam et al., 2010; Adeli et al., 2020; Wohlfart et al., 2018; Klemas, 2013; Kuenzer et al., 2011). Although these reviews have shown the progress in the use of various remotely-sensed data for inventorying different wetland types, geographical locations and climatic zones, they did not only focus on the application of freely-available multispectral data in the remote sensing of wetlands in semi-arid and arid areas. The reviews by Guo et al. (2017) and Dronova (2015) incorporated semi-arid and arid studies in their reviews; however, they also included wetlands in humid areas. In addition, they included studies on the hyperspectral remote sensing of wetlands and either focused on only one aspect of wetland ecosystem type or the application of only one type of multispectral data. Reviews by Adeli et al. (2020) and Wohlfart et al. (2018) focused only on the application of Synthetic Aperture Radar (SAR) on the remote sensing of different wetland types, while Kuenzer et al., (2011) and Klemas (2013) focused on the remote sensing of one specific wetland type, namely, the coastal marsh. Adam et al., (2010) provided a comprehensive review on the status of remote sensing applications for differentiating and mapping the biochemical and biophysical parameters of wetland vegetation. With all of the above as a background, this chapter seeks to provide a comprehensive review of the progress and development of remote sensing in the detection and monitoring of semi-arid and arid wetlands. Attention is drawn to new insights in the detection and monitoring of all wetland types located within the semi-arid and arid regions, using freely-available multispectral data.

In order to achieve the objective of this chapter, a literature search was conducted by using search engines, such as Google Scholar, Scopus and the Web of Sciences, to get an overview of the remote sensing applications on wetlands. The targeted journals were internationally-recognised, peer-reviewed journals that covered Geographical Information System (GIS), Remote Sensing and Water Resource Science. The journal information was supplemented by books and reports from the European Union (EU), the South African Water Research Commission (WRC), the International Union for Conservation of Nature (IUCN), the African Union (AU) and the South African National Biodiversity Institute (SANBI), amongst others. The search criteria were used to find studies that were published between the years 2000 and

2020. For the Level 1 search, the key words ‘Remote sensing’ and ‘wetlands’ were used to find publications from within the specified time-frame. A total of 32 500 publications were retrieved. These included 17 500 from Google Scholar, 8 500 from Scopus and 6 500 from the Web of Science. The above articles were further subjected to a Level 2 search, or screening, using the key words ‘Multispectral sensors’, ‘Semi-arid wetlands’ and ‘Arid’, and the year range of 2000-2020. A total of 6 380 were retrieved from the Google Scholar articles, 3 870 from Scopus and 3 200 from Web of Science. Further screening was conducted on these articles using the keywords ‘Wetlands inundation and extent’, ‘Wetlands vegetation cover’, ‘Wetlands degradation extent’, ‘Land use Land cover changes’, ‘Wetlands monitoring challenges’ and ‘Wetland classification’, ‘Sentinel’, ‘SAR’, ‘Landsat’, ‘MODIS’ and ‘RADAR’ on Level 3, and a total of 196 articles within the scope of this review were retrieved.

2.2 Semi-arid and Arid Wetlands Characteristics and Key Monitoring and Management Challenges

Semi-arid and arid areas (Figure 2.1) host a diverse range of perennial to non-perennial wetlands, with most of them being visible during the wet season (Fang et al., 2019). These wetlands include swamps, peatlands, marshes and floodplains (Day et al., 2010; Powell et al., 2019; Tooth et al., 2002). The existence of wetlands in semi-arid and arid areas is controlled by a positive surface water balance for the whole, or part, of the year, and inundation is mostly due to the frequent rainfall from the upper humid basins. Groundwater also contributes to the inundation of wetlands in semi-arid areas (Hollis, 1990; Tooth et al., 2002). Outflows from wetlands are usually higher, due to the high evaporation rates experienced over prolonged dry periods in semi-arid and arid regions (Hollis, 1990). The dominating vegetation species that characterise each of these wetlands vary according to their locality. In semi-arid South Africa, the common wetland vegetation species include the short grass family species, such as *Cynodon dactylon*, and common reeds, such as *Phragmites australis*. These species are able to adapt to inundation, drying and sediment deposition. In some seasons, the semi-arid and arid wetlands soils tend to be oxygenated, due to the episodic nature of the inundation. For this reason, these wetlands tend to host more animal species than other wetlands that are permanently inundated (Sieben et al., 2016; Li et al., 2015; Jenkins et al., 2005).

Despite the ecological significance of semi-arid and arid wetlands, their conservation has not been prioritised (Minckley et al., 2013), due to their ephemeral nature and small size (Li et al., 2015), and this results in their poor management, their degradation and a loss of the species and ecosystem services that they provide (Chen and Liu, 2015). Gebreslassie et al. (2014)

reported that the lack of policies to protect wetlands in semi-arid Ethiopia has resulted in the loss of socio-economic services provided by these systems. The regular monitoring of their eco-hydrological dynamics (Liu et al., 2017) is critical for formulating appropriate management measures. The frequent monitoring of semi-arid and arid wetland systems presents some challenges, which are associated with the methods that are used and the wetland type. Traditional field monitoring methods provided baseline information about semi-arid and arid wetlands; however, due to the cost of these methods, regular monitoring has not been possible, which has created problems in the tracking of the changes that occur within these wetlands over time (Powell et al., 2019).

Remote sensing images from satellite sensors, such as Landsat and MODIS, provide cost-effective datasets for wetland monitoring, in both space and time. However, these sensors have a limitation, based on their spatial resolution, since most semi-arid and arid wetlands are fairly small, namely <10 and 2 500 ha (Li et al., 2015), they are confined to small depressions and they have no definite boundaries. These wetlands merge with the surrounding terrestrial ecosystems (Day et al., 2010), which poses a challenge when mapping their spatial extent by using optical sensors, especially during the dry period when the surrounding and wetland vegetation are not very healthy; this results in a similar spectral reflectance of the soils and other land-cover classes. Rapinel et al. (2019) reported that the inventorying and characterisation of wetlands in semi-arid and arid areas is limited to mostly small basins. Furthermore, Cape et al. (2015) reported that the semi-arid wetlands are lost over a short period of time because of anthropogenic activities, including the over-exploitation of their water for irrigation. This necessitates the development of an integrated and cost-effective monitoring and management approach, which will enable the generation of information for wetlands of different sizes, thus informing the management strategies for wetlands in semi-arid and arid areas.

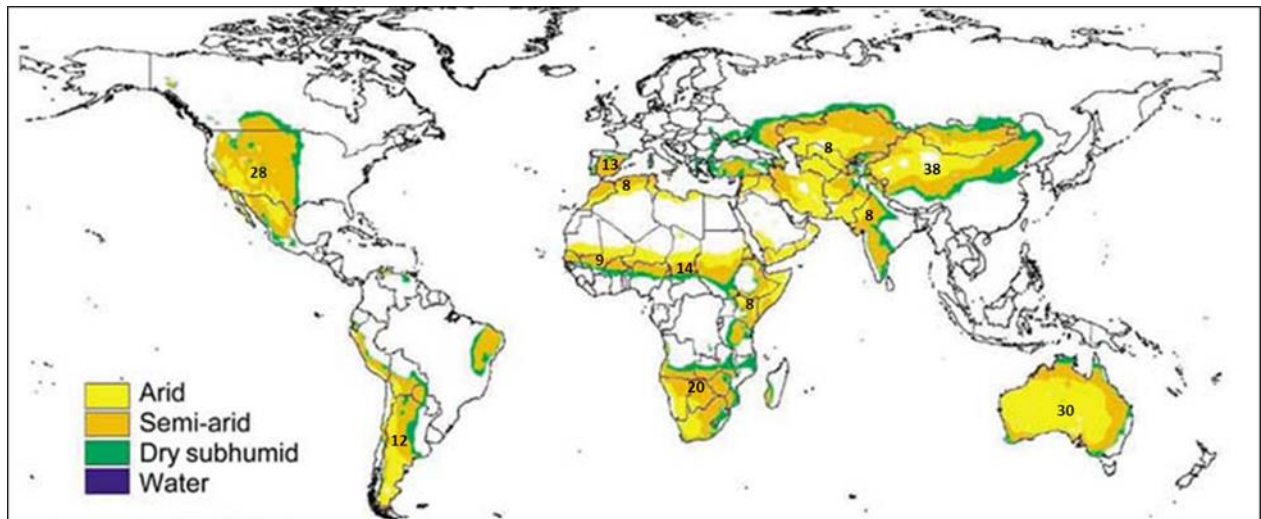


Figure 2.1 Global extent of semi-arid and arid areas, with a number of studies in each region (modified from Scanlon et al., 2010)

2.3 Commonly-used, Freely-available Multispectral data for Semi-arid and Arid Wetland Inventories

Over the past few decades, different types of wetlands, ranging from inland freshwater marshes, coastal tidal marshes, mangrove ecosystems and forested wetlands or swamps, have been studied by using remotely-sensed datasets with different spatial, spectral and temporal resolutions (Ozesmi and Bauer, 2002). The number of studies conducted in semi-arid and arid areas on the application of freely-available multispectral datasets for inventorying wetlands has increased exponentially, which is evident by the number of publications between 2000 and 2020 ($R^2 = 0.76$) (Figure 2.2a). A significant increase was noted between 2008 and 2020, with the highest number of publications in 2020. An analysis of the number of publications per region reveals that most of them were from the semi-arid Australia and China (Figure 2.1), with a total of 30 and 38 publications, respectively, while the semi-arid India and North Africa had the lowest number of publications (Figure 2.1). In all these studies, the commonly-used data sources were the Landsat Thematic Mapper (TM), the Landsat Enhanced Thematic Mapper Plus (ETM+), the Landsat Operational Land Imager (OLI), the Moderate Resolution Imaging Spectroradiometer (MODIS), the Landsat Multispectral Scanner System (MSS) and the Synthetic Aperture Radar in the form of Sentinel-1 and ALOS PALSAR. There has been an increase in the use of Landsat OLI for mapping different aspects of semi-arid and arid wetlands. The justification is that Landsat 8 OLI uses the push-broom feature, which has improved the noise-to-signal ratio; this is an advantage, when compared to Landsat TM and

ETM+. The most studied wetland aspect was their characterisation, which includes wetland classification and mapping, as well as inundation (Figure 2.2b).

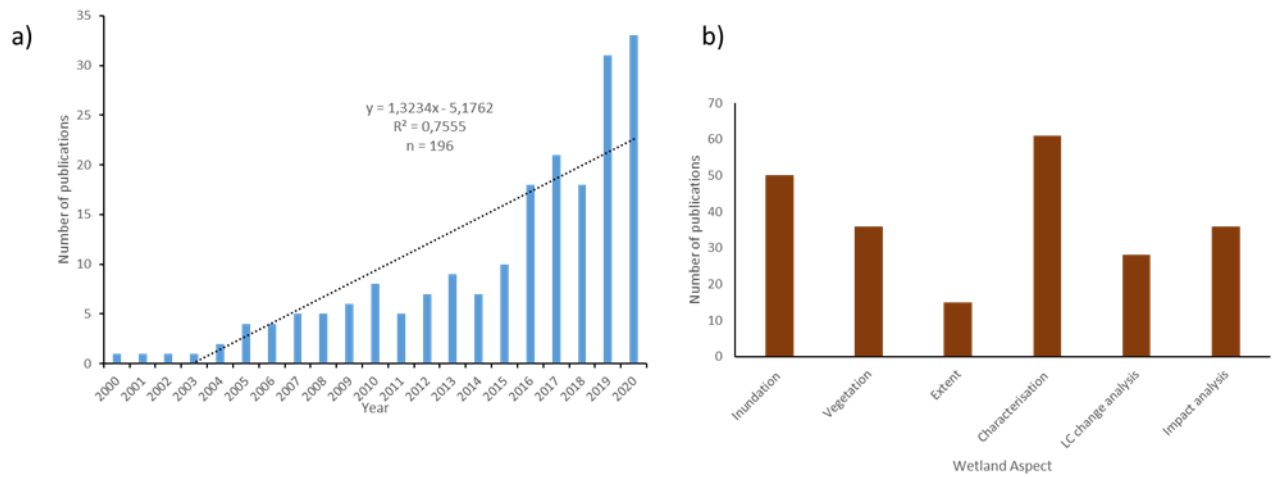


Figure 2.1 The number of remote sensing publications on semi-arid and arid wetlands (a) and the number of publications per area of focus (b). Characterisation includes studies on classification and wetland mapping, an impact analysis includes studies focusing of both climatic and anthropogenic impacts on wetlands, and an LC analysis refers to all studies on the wetland cover change analysis (n=196).

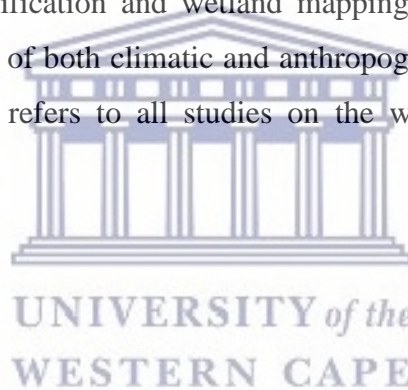


Table 2.1 Commonly-used, freely-available sensors specifications for wetland inventories (modified from Ozesmi and Bauer, 2002)

Resolution	Landsat MSS	Landsat TM	Landsat ETM+	Landsat OLI	Sentinel-1	Sentinel-2	MODIS
Spectral bands (μm)					C-band (3.75-75 cm)		
Band 1		0.45-0.52	0.45-0.515	0.43-0.45		0.443	0.62-0.67
Band 2		0.52-0.62	0.525-0.605	0.45-0.51		0.490	0.841-0.876
Band 3		0.63-0.69	0.63-0.69	0.53-0.59		0.560	0.459-0.479
Band 4	0.5-0.6	0.76-0.90	0.775-0.90	0.64-0.67		0.665	0.545-0.565
Band 5	0.6-0.7	1.55-1.75	1.55-1.75	0.85-0.88		0.705	1.23-1.25
Band 6	0.7-0.8	10.4-12.5	10.4-12.5	1.57-1.65		0.740	1.628-1.652
Band 7	0.8-1.1	2.08-2.35	2.08-2.35	2.11-2.29		0.783	2.105-2.155
Band 8			0.52-0.9	0.50-0.68		0.842	
Band 8A						0.865	
Band 9				1.36-1.38		0.945	0.438-0.448
Band 10				10.6-11.19		1.375	
Band 11				11.5-12.5		1.610	
Band 12						2.190	
Band 19							0.915-0.965
Band 31							10.78-11.28
Band 32							11.77-12.27
Temporal	180 days	16 days	16 days	16 days	12 days	5 days	1-2 days
Spatial (pixel-sizes)	80 m	30 m & 120 for Band 6	30 m, 15 m B8 & 60 m B6	30m B1-7 & 9 15m B8 100m B10-11	5m X 5m	60m B1,9,10 10m B2,3,4,8 20m- B5,6,7,11,12	250m B1-2, 500m B8-36 1000m B8-36
Period	1972-1992	1982-Present	2003-Present	2013- Present	2014-present for 1A 2016-present For 1B	2015-Present For 2A and 2017 for 2B	2000-present for Terra 2002- present for Aqua

Although the multispectral sensors in Table 2.1 have been providing crucial information on wetlands, at no cost, with the repeated coverage over the large spatial areas, fine-detailed wetland detection is still a major challenge (Schmid et al., 2005; Gallant 2015), especially for the fairly small semi-arid and arid wetlands (<10 ha) with varied vegetation and other land-cover characteristics. This results in some of the wetlands being missed, or confused with other land-cover classes, during classification. Various studies have been undertaken to establish the use of these sensors in order to understand the different dynamics of wetlands in semi-arid and arid areas. These studies include Li et al., (2015), Landmann et al., (2010), Powell et al., (2019) and Chen et al., (2013), amongst others. Powell et al., (2019) demonstrated the use of Landsat TM and ETM+ datasets, coupled with digital elevation and LIDAR data, which were set to classify and map the land-cover classes of the semi-arid wetland in Barwon-Darling River system, using the stochastic gradient-boasting algorithm and the fractional cover model. The study identified five land-cover classes, which included tree-dominated forest and woodlands, shrub lands, vegetated swamps and non-flood-dependent terrestrial communities, with an overall accuracy of 88%. However, the study failed to distinguish between the certain types of wetland located at the boundaries of the drier wetlands, from the Landsat TM and TM+ images that were used. Li et al., (2015) evaluated the use of MODIS derived spectral indices for monitoring the hydrological dynamics of a small seasonally-flooded wetland (1 364 ha) in semi-arid southern Spain. An analysis of the relationship between the MODIS inundation area and the field-measured water levels showed a positive linear relationship between the two variables, with a R^2 determinant of 0.96, which suggests the success of the MODIS dataset for monitoring the hydrological dynamics of seasonal wetlands. However, the study focused on a single seasonal wetland, with only varying soil characteristics. The other semi-arid and arid seasonal wetlands have other diverging characteristics e.g. marshes with dense emergent vegetation, and even smaller in size and they are cover only a few MODIS pixels.

Chen et al., (2013) used a MODIS dataset with a 250 m resolution, coupled with the daily field water levels, to investigate the applicability of the MODIS time series dataset for monitoring the wetland cover dynamics over time. Four land-cover classes were identified, namely, water, mudflats, submerged and emergent vegetation, with an overall accuracy of 80.18% and a Kappa coefficient of 0.734. There were, however, omission errors of about 30%, where water was confused with other classes, such as mudflats and the emergent vegetation. Much of this water was located at the interface of the mudflats and other classes. Landmann et al., (2010) also used MODIS, coupled with the topographical landform dataset, to map the basic wetland

classes in semi-arid Burkina Faso and Mali. The results showed a total of five wetland classes that covered a total area of 9 350 km², but they demonstrated low accuracies for the mapped land-cover classes. Although the studies demonstrated the success of these freely-available datasets in the detection and mapping of different wetland-cover classes, the finely-detailed differentiations between the wetland classes were a major challenge. The use of pixel- and sub-pixel-based approaches offer a great opportunity to improve the accuracy of wetland cover detection and monitoring from freely-available multispectral data. These approaches provide an analysis of the spectral characteristics of each class within a pixel, as such pixels with the same spectral characteristics are grouped together as one object. This has the potential to minimise the spectral confusion between classes. In the case of wetlands, however, the challenge is that wetlands vary spatially and temporally; the inner wetland may be permanently inundated, when compared to the seasonally-inundated edges of the same wetland, but the temporal pattern of the entire system is what distinguishes it from other landscapes. Using the pixel-based approach will, however, not permit the capture of that temporal pattern (Halabisky et al., 2018). In addition, wetlands have similar spectral characteristics to other landscapes i.e. a flooded wetland may resemble the shadows of trees, hills and other features, since these have a low surface reflectance.

The use of an Object-Based Image analysis (OBIA) offers the opportunity to improve wetland detection and classification, despite the inherent computational costs, which are higher than those of a pixel-based approach. The introduction of cloud-computing systems, such as Google Earth Engine (GEE), NASA Earth Exchange and Amazon web services, amongst others, presents an opportunity to simplify the use of the OBIA approach in wetland mapping, thus improving the accuracy of the classification. In addition, the use of fine spatial resolution images from commercial sensors may also assist in improving the classification accuracy.

2.4 Mapping Semi-arid and Arid Wetland Vegetation, using Freely-available Multispectral Images

Wetland vegetation provides a habitat for a variety of aquatic animal species (Adam et al., 2010). Changes in the condition of the wetland vegetation can be used as a proxy for the early signs of any chemical and physical wetland degradation (Mishra, 2014). The assessment of wetland vegetation is considered to be an important aspect for evaluating the ecological status of a particular wetland (Bhatnagar et al., 2020; Martínez-López et al., 2014; Thakur et al., 2012), and the management of wetland biodiversity relies heavily on the accurate assessment

of its vegetation (Sieben et al., 2016). The assessment of wetland vegetation includes an understanding of its components, such as its structure, species types and composition. The use of optical freely-available multispectral remote sensing data for understanding the wetland vegetation components is a common practice for semi-arid and arid wetland inventories, and most studies have successfully mapped wetland vegetation by using this type of dataset e.g. Gong et al. (2014), Landmann et al. (2013) and Thakur et al. (2012). The challenge, however, lies with the spatial resolution of these datasets, which is often too coarse to accurately map the mixed vegetation in small semi-arid and arid wetlands.

The spectral reflectance of vegetation types in a mixed-vegetated wetland are similar, and they are usually combined with the spectral reflectance from the underlying soils, water and top of the atmospheric effects, which results in complications during the classification process (Gallant, 2015). In addition, steep environmental gradients cause short ecotones and sharp demarcations between vegetation species in the wetlands and result in a high spectral and spatial variability, thus presenting difficulties in the identification of the boundaries between the vegetation communities and types during optical mapping (Adam et al., 2010; Gallant, 2015). Therefore, the use of optical freely-available multispectral imagery may present challenges with regard to separating and understanding the wetland vegetation components, such as the different species types, their composition and structure, due to their low to medium spatial and spectral resolution. McCarthy et al., (2005) mapped the eco-regions of the Okavango Delta in Botswana from the Landsat TM imagery, using Maximum Likelihood Classification (MLC) and Rule Based Classification (RBC) with 6 and 10 classes. The results showed an overall accuracy of 46% and a Kappa co-efficient of 0.37 for all 10 classes, based on the MLC, an overall accuracy of 63% and Kappa co-efficient of 0.59, based on the RBC for all 10 classes, and an overall accuracy of 74% and Kappa co-efficient of 0.67, based on RBC 6 class map. Based on these findings, the study deemed Landsat TM to be unsatisfactory for the classification of land-cover classes, including the wetland vegetation, in the Okavango Delta.

The spectral reflectance of vegetation types in a mixed-vegetated wetland are similar, and they are usually combined with the spectral reflectance from the underlying soils, water and top of the atmospheric effects, which results in complications during the classification process (Gallant, 2015). In addition, steep environmental gradients cause short ecotones and sharp demarcations between the vegetation species in the wetlands, which results in a high spectral and spatial variability, thus making it difficult to identify the boundaries between the vegetation

communities and types during the optical mapping (Adam et al., 2010; Gallant, 2015). Therefore, the use of optical freely-available multispectral imagery may present challenges for separating and understanding the components of wetland vegetation, such as the different species types and their composition and structure, due to their low to medium spatial and spectral resolution. McCarthy et al., (2005) mapped the eco-regions of the Okavango Delta in Botswana from the Landsat TM imagery, using the Maximum Likelihood Classification (MLC) and Rule-Based Classification (RBC) with 6 and 10 classes. The results showed an overall accuracy of 46% and a Kappa co-efficient of 0.37 for all 10 classes, based on the MLC, an overall accuracy of 63% and a Kappa co-efficient of 0.59, based on the RBC for the 10 classes, and an overall accuracy of 74% and a Kappa co-efficient of 0.67, based on the RBC 6 class map. Based on these findings, the study deemed Landsat TM to be unsatisfactory for the classification of the land-cover classes, including the wetland vegetation in the Okavango Delta.

Carreño et al. (2008) utilised Landsat TM and ETM+ to assess the spatiotemporal changes in the area and the internal components of the Mar Menor coastal wetland in semi-arid Spain, from 1984 to 2001. The results for the classification of land-cover classes showed three natural vegetation sub-classes, namely, the salt steppe, salt marsh and reed bed. A user accuracy of 90% was achieved for the salt marshes and 80% for both the reed beds and salt marshes, based on the 1984 image, while a user accuracy of 62% was reported for the salt steppe, 96.5% for the salt marshes and 76.92% for the reed beds, respectively, based on the 1997 image. The study reported high commission errors for salt steppe (37.5%) based on the 1997 image. This was attributed to the spatial resolution of the Landsat images that were used. Mazzarino (2014) used Landsat 5 TM-derived NDVI to investigate the multi-decadal (1985-2010) vegetation dynamics of the Andean wetland system in the Nuñoa watershed. The classification results showed that two classes which were named wetland (characterised by wetland vegetation) and a non-wetland, with an accuracy of 93% for the wetland system and 87% for non-wetland system. Although the Landsat 5 TM that was used in the study proved to be successful in the separation of the non-wetland and wetland areas, the images that were used represented the dry season, when the wetland was not inundated and when the vegetation classes can be easily identified.

The literature (Carreño et al., 2008; Mazzarino, 2014; McCarthy et al., 2005) shows the success of the application of multispectral images for mapping the vegetation of semi-arid wetlands; however, mapping vegetation communities and specific species in the finest detail, is still a

major gap that requires the use of high spatial resolution satellite images. In an attempt to resolve this issue, the fusion of different datasets, which combine the strengths of the different sensors, has the potential to improve the mapping of wetland vegetation. Data fusion may, however, cause a distortion of the information, as a result of the mismatching of different-sized pixels in the fused data, thus lowering the quality of the produced image.

2.5 Mapping Wetland Inundation by using Freely-available Multispectral Images

Inundation plays a critical role in expressing the hydrological dynamics of wetlands (Li et al., 2015). Having an understanding of this process plays a key role in water management, in ecosystems assessment and in biodiversity conservation (Klein et al., 2014). The use of in-situ gauge datasets has been the backbone for the current understanding of surface water dynamics, including wetland inundation (Huang et al., 2011). However, it presents challenges because most wetlands in semi-arid and arid areas are not gauged, partly due to their episodic nature (Li et al., 2015; Jenkins et al., 2005). Advancements in the remote sensing approaches and products improves the mapping of surface water features, including the changes in inundation. The most utilised products include Landsat TM, Landsat ETM+ and MODIS. The use of freely-available, medium-spatial resolution products, like MODIS, presents difficulties when mapping the inundation of heterogeneous seasonally-flooded wetlands, when the water is beneath the vegetation, and when they are small in size and have very dynamic eco-hydrological changes, especially during the dry season (Chen et al., 2013). This is evident in studies by Klein et al., (2014), Xie et al. (2016) and Moser et al., (2014).

Moser et al., (2014) used the MODIS time series dataset to establish the spatio-temporal variability of the water coverage of a semi-arid wetland in sub-Saharan West Africa, and the coverage of surface water was slightly over-estimated. Klein et al. (2014) evaluated the spatial extent of the seasonal water bodies in semi-arid central Asia from 1968-2001, by using the coupled AVHRR and MODIS multispectral data. The accuracy assessment showed an overall classification accuracy of 0.83, based on the AVHRR data, and an accuracy of 0.91, based on the MODIS data. Lower accuracies were observed for the month of April in the northern region of the basin, including the Tengiz-Kolgalzhyn lake system, and they were attributed to the presence of ice and snow. In addition, water masks were over-estimated, due to the coarse spatial resolution of the datasets, particularly at the interface of the land and water surfaces.

Xie et al. (2016) used Landsat (TM, ETM+ and OLI), coupled with the Gravity Recovery and Climate Experiment datasets, to investigate the hydrological dynamics and ecosystem

functioning of the Coongie Lake in arid central Australia, over a 24-year period. The analysis of the flooding extent indicated a variable water regime, with episodes of long-term droughts and short periods of flooding over the Coongie Lake. Although the study successfully mapped the inundation dynamics of the Lake, there were uncertainties regarding the magnitude of the monthly inundation derived from the Landsat images. Klein et al., (2014), Moser et al., (2014) and Xie et al., (2016) demonstrated the capabilities of the freely-available multispectral datasets for mapping and understanding the inundation of semi-arid and arid wetlands; however, it was noted that the presence of other wetland features influenced the level of accuracy when mapping inundation by using coarse spatial resolution images, such as MODIS. In addition, the issue of spectral confusion between the water and soils at their boundaries remained unresolved, as such water pixels were over-estimated. When attempting to resolve the issue of spectral mixing between wetland water and soils at the boundaries of these two classes, the use of high spatial resolution images is likely to improve the latter. Moreover, the use of the Synthetic Aperture Radar (SAR) dataset has the potential to improve the detection of inundation patterns, since the sensor has the ability to penetrate the wetland vegetation canopy.

2.6 Mapping Land-use and Land-cover Changes Impacts of Semi-arid and using Freely-available Multispectral Images

An understanding of Land use and Land cover changes (LULC) assists with the development of effective environmental management strategies to counteract the degradation and loss of wetlands (Valdez and Ruiz-Luna 2016; Zhang et al., 2011), and an LULC change analysis is pertinent for the better understanding of land dynamics (Mwita et al., 2012). Studies have revealed that changes in the LULC are significant drivers for wetland degradation. Alam et al., (2011) reported that the continuous inflow of sediment loads and nutrients in the Hokar Sar wetlands in India has led to their degradation. The inflow of sediments was attributed to changes in the LULC, due to the anthropogenic activities in the upper basin from 1986-2005. Martínez-López et al., (2014) also noted that the expansion of irrigated lands in the semi-arid Mediterranean catchments has altered the input of water and nutrients to lowland wetlands, which has resulted in their degradation. The regular monitoring of the LULC is necessary for developing measures to manage the degradation of wetlands (Butt et al., 2015). Data from different satellite sensors, such as the multispectral MODIS and Landsat, have been widely used and are recognised as powerful tools for studying and monitoring the dynamic impacts of

LULC changes on the wetlands in semi-arid and arid environments. The challenge, however, is to detect these impacts with adequate precision (Schmid et al., 2005).

Peter et al., (2020) used four decades of Landsat data to assess the impact of anthropogenic activities and climate variability on the spatiotemporal patterns of Lake Babati in Tanzania. The study achieved an overall classification accuracy of 87%. The extent of the water surface area was not accurately captured, due to the unavailability of continuous usable Landsat data which was caused by a significant cloud cover for most of the year. Wang et al., (2011) used Landsat data, coupled with topographical land cover maps, to investigate the shrinkage and fragmentation of marshes in West Songnen in China for the period between 1954 and 2008. The study reported an overall classification accuracy of 90%. Mwita (2013) used Landsat MSS, TM, ETM and ETM+ over a 30-year period with a sequence of 10 years (1976-2003) to assess the land use and land cover dynamics of the Rumuruti and Malindi wetlands in Kenya and Tanzania. The classification achieved an overall accuracy of between 88.28% and 95.17%, respectively, for both wetlands. Although the overall accuracy results were higher, the producer's accuracy for the open water class for the 1976 scene was low (33.3%) for both wetlands. This was attributed to the spatial resolution of the Landsat MSS that was used in 1976. Although other studies by Mwita (2013), Peter et al., (2020) and Wang et al., (2011) reported high classification accuracies, there were challenges associated with the type of data used in these studies. In an attempt to avoid the issue of cloud cover, as reported by Peter et al., (2020), the use of the SAR dataset proved to be the solution, as the sensor can penetrate through the cloud cover.

2.7 Low to Medium vs High Resolution Remote Sensing for Wetland Monitoring and Assessment

Low to medium spatial resolution remote sensing datasets are characterised by their pixel size, which ranges between 30 and > 200 meters. These datasets have been successfully used in many wetland inventories in different climatic zones globally. This is because they are readily available, come at no cost and provide mostly timeous datasets, thus providing the opportunity to monitor changes in wetlands over a longer period of time. Despite the highlighted advantages of low to medium spatial resolution data, their application is somehow challenging, especially on wetlands with an aerial extent of less than 1-ha (Adeli et al., 2020; Dronova, 2015). Advancements in the development of satellite technology have led to the introduction of new-generation multispectral sensors, which are both spaceborne and airborne, with high spatial resolution datasets (<10 meters pixel size) (see Table 2.2). The sensors are unique, with

improved sensing characteristics, including the presence of strategically-positioned spectral bands e.g. red-edge and near-infra red II, as well as an improved signal-to-noise ratio, amongst others (Shoko et al., 2016). They are freely available and come with a fine spatial resolution with strategically-organized bands that can detect the finer wetland features. These sensors have been explored in vegetation monitoring, as well as biomass and surface water mapping studies, and the findings have been commendable (Shoko et al., 2016). It is upon this premise that these datasets are likely to improve the monitoring and understanding of wetlands and wetland dynamics in semi-arid and arid areas, which was previously a challenging task when using broadband sensors, as the majority have an aerial extent of below 1-ha (Adeli et al., 2020). However, the challenge is the cost implications of some of these datasets, because most of them expensive. For example, high spatial resolution datasets from sensors, such as Worldview-2, QuickBird and RapidEye, amongst others, are very costly to acquire, thus making it difficult to map semi-arid and arid wetlands that are distributed in resource-limited environments.



Table 2.2 Selected low to high spatial resolution sensors for wetland monitoring and assessment (freely-available data sensors are highlighted in bold)

Sensor	Pixel size (m)	Bands	Revisit time	Acquisition cost	Scale of Application	Spatial Resolution
AVHRR	1100	5	1	Readily available	Regional to global	Low
Hyperspectral	<1	>100	-	Very expensive	Plot	High
IKONOS	4	5	1-2	Expensive	Local	High
Landsat TM	30	7	16	Readily available	Local to regional	Medium
Landsat ETM+	30	8,11	16	Readily available	Local to regional	Medium
Landsat MSS	80	4	180	Readily available	Local to regional	Low
Landsat OLI	30	11	16	Readily available	Local to regional	Medium
MERIS	300	15	3	Readily available	Regional	Low
MODIS	500,1000	7	1	Readily available	Regional to global	Low
QuickBird	2.4	5	1-3.5	Expensive	Local	High
RapidEye	5	5	5.5	Expensive	Local	High
Sentinel-2	10,20,60	13	5	Readily available	Local to regional	High/medium
SPOT	10,20	4	26	Readily available	Local to regional	High
Worldview-2	<1	8	1	Very expensive	Local	High
Sentinel-1	5m	1	12	Readily available	Local to regional	high

2.8 Available Satellite Image Processing Techniques for Accurate Wetland Monitoring

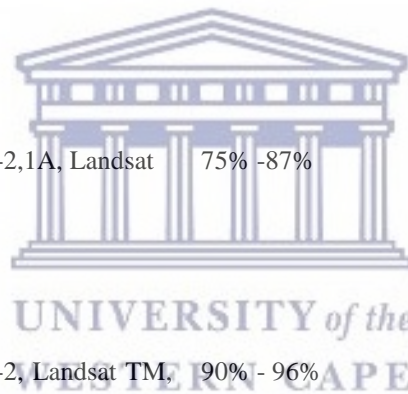
Mapping wetlands with optical sensors is challenging, since wetlands have a heterogeneous mixture of land-cover classes, which may produce a similar spectral reflectance thus resulting in complications during the classification process (Amani et al., 2017; Gallant, 2015). In addition, wetlands are highly dynamic with regards to the presence of water, plants and land surface, which alters their reflectance and energy back-scattering properties. The classification of wetlands can be achieved through a pixel-based analysis or an Object-Based Image Analysis (OBIA). In a pixel-based classification, pixels are analysed by their spectral information and they require imagery that extend beyond the visible spectrum (Halabisky, 2011). Although the pixel-based classification has been used for a long time (Amani et al., 2017), this approach does not fully utilise the spatial information of the multispectral imagery. Unlike a pixel-based classification, OBIA aggregates the pixels with similar characteristics into objects or segments, which are then classified by using analyst rules, machine-learning algorithms and statistical approaches (Halabisky et al., 2018). One of the advantages of using OBIA over the pixel-based approach is that additional features, such as the shape, size and texture, are considered during the classification process, and it reduces the within class spectral variation, thus improving the accuracy of the classification outputs (Blaschke, 2010). Although OBIA is the most preferred approach, one of the limitations in wetland mapping is that there is no clear standard for the pre-classification assessment of the segmentation effects on the final outcomes, in either the wetlands or other heterogeneous landscapes (Dronova, 2015).

Different machine-learning algorithms are available for classifying wetlands using remotely-sensed data. These include supervised machine-learning algorithms, such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Maximum Likelihood Classification (MLC), Random Forest (RF), Artificial Neural Network and Classification and Regression Tree (CART), as well as the Unsupervised K-means and ISODATA. The selection of an appropriate algorithm depends on the objective of the classification. A number of studies have used these algorithms to study wetland systems in semi-arid and arid areas. The supervised machine-learning algorithms have been proved to perform better than the unsupervised classification algorithms (Table 2.3). However, despite their better performance, these algorithms have some limitations. For example, ANN and SVM were reported to be too difficult to automate and require the adjustment of a large number of parameters (Shoko et al., 2016), while RF sometimes tends to overfit and for every data set the size of a tree can take up memory (Na et al., 2010). The availability of image processing techniques, such as the Google

Earth Engine and image processing on cloud, simplifies the use of supervised machine-learning algorithms. However, according to the literature, these platforms have been under-utilized in the remote sensing of semi-arid and arid wetlands, using the freely-available multispectral sensors.

Table 2.3 Available algorithms for wetlands remote sensing

Algorithm	Remote sensing data	Performance range	Reference
RF	Landsat TM, Sentinel 1A, 2A, MODIS, LiDAR, SAR, ALOS-PALSAR, RADARSAT	80% - 98%	(Dubeau et al., 2017; Fu et al., 2017; Judah and Hu, 2019; Millard and Richardson, 2013; Shoko et al., 2016; Tian et al., 2016; Westra and de Wulf, 2007)
CART	Landsat TM, Sentinel-1A, 2A, PALSAR, Landsat ETM+	89.2 % -92%	(Na et al., 2010; Simioni et al., 2020; Corcoran et al., 2013; Otukei and Blaschke 2010)
MLC	Landsat TM, MODIS, Landsat MSS, Landsat ETM+	83.6% - 94%	(Frohn et al., 2011; Na et al., 2010; Otukei and Blaschke, 2010; Ramachandra and Kumar, 2008; Westra and de Wulf, 2007)
SVM	Sentinel-2,1A, Landsat OLI	75% -87%	(Abdi, 2020; Berberoglu et al., 2004; Cao et al., 2018; Hakdaoui et al., 2019; Otukei and Blaschke, 2010; Simioni et al., 2020)
ANN	Sentinel-2, Landsat TM, ETM+, OLI	90% - 96%	(Berberoglu et al., 2004; Simioni et al., 2020; Soltani et al., 2020; Westra and de Wulf, 2007)
KNN	Sentinel-2, Landsat TM, ETM+, OLI, RADARSAT-2, Sentinel-1	83% - 97%	(Cao et al., 2018; Judah and Hu, 2019; Simioni et al., 2020; Soltani et al., 2020)
Unsupervised classification	Landsat TM, ETM+, MSS, Sentinel-2	82% - 96%	(Na et al., 2010; Ramachandra and Kumar, 2008; Simioni et al. 2020)



2.9 Summary of Key Challenges and Future Research Directions

Although the semi-arid and arid wetland systems tend to host most invertebrate and vertebrate species that would not survive in the surrounding landscapes, their conservation is still

overlooked. Remote sensing approaches offer an opportunity to understand these wetlands from various aspects, ranging from their characterisation, inundation, vegetation, extent and land cover changes. The majority of these studies used the freely-available multispectral sensors, such as Sentinel-2 MODIS and Landsat (MSS, TM, ETM+, OLI). Although progress has been made regarding the use of freely-available multispectral sensors for understanding the dynamics of wetlands in semi-arid and arid areas, it is still challenging to map these wetlands with the finest precision, because of their complex edaphic nature and small hydrological gradients. Moreover, the spatial resolution of these freely-available multispectral sensors limits the detection and monitoring of these wetlands.

Wetlands are complex systems, and understanding their eco-hydrological dynamics cannot be based solely on a single data source or validated by using the in-situ measured data. In most arid and semi-arid environments, in-situ datasets are limited. In addition, the sharing of such data among different institutions in most of these regions is poor, particularly in the sub-Saharan Africa. The availability and free opening of numerous spatial data sources with varying sensing characteristics e.g. the global mapper Landsat series, Sentinel Copernicus and freely-available Radar and weather dataset products, offer opportunities to improve the mapping of wetlands, where in-situ data are limited. This provides new opportunities for the monitoring and assessment of fairly small semi-arid and arid wetlands, which were previously ignored, due to the lack of the required spatial data. This challenge, or knowledge gap, can easily be addressed by exploring different spatial dataset integration techniques, which was previously a challenging task with the broadband and coarse spatial resolution multispectral datasets. Improvements in data analytic techniques, such as the introduction of advanced computer processing methods, also provide new opportunities for the detection and monitoring of wetlands.

The literature shows that the introduction of advanced machine-learning algorithms and cloud-computing, such as the Google Earth Engine (GEE) and Petascale image-processing techniques amongst other, provide new avenues for multisource data integration and fusion (Mahdianpari et al., 2019). Although few studies have explored the applicability of these techniques in vegetation monitoring and other related fields of study, there is need for future studies to shift towards embracing these methods to enhance the detection and monitoring of wetlands, particularly in data-poor regions. One advantage of these techniques is the quick processing of large datasets. However, challenges, such as inadequate network and internet connectivity, as well as the lack of high-performance computing systems for cloud-computing and the lack of

skilled personnel, limit the application of such techniques, especially in developing countries, which are mostly in sub-Saharan Africa and other parts of the world. Despite some of the highlighted challenges, this review advocates a paradigm shift in satellite data applications for wetland monitoring, by embracing multi-data and advanced data processing techniques to improve our understanding of these systems.

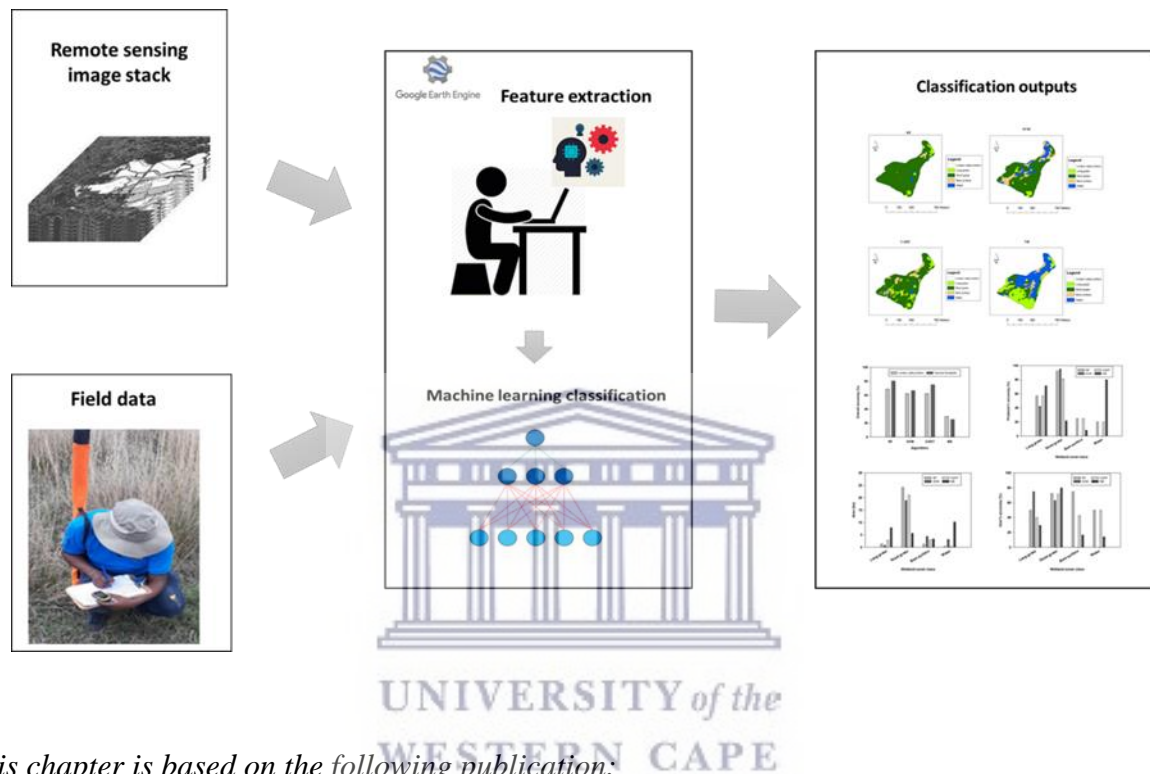
2.10 Conclusion

The current review was aimed at providing a comprehensive overview of the progress and development of multispectral remote sensing for the detection and monitoring of semi-arid and arid wetlands. The literature search showed that there has been great improvement in the use of freely-available multispectral datasets for monitoring semi-arid and arid wetlands, but more is required for the monitoring and assessment of smaller wetlands in these regions. This is evident by the number of studies that were published between the period of 2000-2020, and the significant increase between 2008 and 2020. Although there has been a significant increase in the number of published papers that focus on the use of freely-available multispectral datasets for the remote sensing of semi-arid and arid wetlands, it was noted that the monitoring of key wetland aspects presented some challenges, mainly due to spectral mixing and the poor data quality for determining the inherent wetland characteristics. These challenges include mapping inaccuracies, which were either attributed to the poor spatial resolution of the wetlands, the inadequate validation data, or the incorrect classification method that was used. Therefore, the introduction of advanced machine-learning algorithms and cloud-computing systems, such as the Google Earth Engine and Petascale, provide a great opportunity for improving the monitoring and assessment of wetlands, particularly in data-poor regions and in semi-arid or arid environments. Thus far, the use of these machine-learning algorithms and cloud-computing techniques, as well as data integration methods for semi-arid and arid wetlands, is still in its infancy, but the increased application of these methods provides a new window of hope. Further investigations are therefore required to assess the use of these programs and platforms for understanding the distribution, dynamics and the status of wetlands in semi-arid and arid regions, to enhance their management and conservation, as well as to safe-guard their ecosystem goods and services and, above all, their livelihoods.

CHAPTER 3

LEVERAGING GOOGLE EARTH ENGINE PLATFORM TO CHARACTERISE AND MAP SMALL SEASONALLY-FLOODED WETLANDS IN THE SEMI-ARID ENVIRONMENTS OF SOUTH AFRICA

Graphical abstract:



This chapter is based on the following publication:

Gxokwe, S., Dube, T., Mazvimavi, D. 2021. Leveraging Google Earth Engine platform to characterise and map small seasonally-flooded wetlands in the semi-arid environments of South Africa. *Science of the Total Environment*. **803** (2022) <http://doi.org/10.1016/j.scitotenv.2021.150139> (IF =10.753)

Abstract

Although significant scientific research strides have been made in mapping the spatial extent and eco-hydrological dynamics of wetlands in semi-arid environments, the focus on small wetlands remains a challenge. This is due to the characteristics of the remote sensing platforms and the lack of robust data-processing techniques. Advancements in data analytic tools, such as the introduction of the Google Earth Engine (GEE) platform, provide unique opportunities for the improved assessment of small and scattered wetlands. This study thus assessed the capability of the GEE cloud-computing platform for characterising small and seasonally-flooded wetlands, using the new generation Sentinel-2 data from 2016 to 2020. More specifically, the study assessed the spectral separability of different land-cover classes for the two different wetlands that were detected, by using Sentinel-2 multi-year composite water and vegetation indices, and it identified the most suitable GEE machine-learning algorithm for accurately detecting and mapping semi-arid seasonally wetlands. This was achieved by using the object-based Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART) and Naïve Bayes (NB) advanced algorithms in GEE. The results demonstrated the capability of the GEE platform to characterise wetlands with an acceptable accuracy. All algorithms showed superiority in the mapping of the two wetlands, except for the NB method, which had lowest overall classification accuracy. These findings underscore the relevance of the GEE platform, Sentinel-2 data and the advanced algorithms for characterising small and seasonally-flooded semi-arid wetlands.

Keywords: Limpopo River Basin; object-based classification; machine-learning algorithm; wetland mapping; wetland condition

3.1 Introduction

Wetlands play a critical role in the hydrological cycle and they sustain peoples' livelihoods, the aquatic life and the biodiversity. They occupy the transition zones between the aquatic and terrestrial environments and share the characteristics of both zones (Gxokwe et al., 2020). Wetlands cover about 4-6% of the global land surface and are ranked amongst the most diverse ecosystems on earth (Mahdianpari et al., 2019). Semi-arid areas are dominated by small seasonally- or intermittently-flooded wetlands, with the flooded area depending on the balance between the precipitation and evapotranspiration (Ruiz, 2008). They often host more invertebrates than permanently-inundated systems because of their oxygenated period, as a result of the episodic inundation. However, their conservation is not prioritised (Chen and Liu, 2015), mainly because of their small size and ephemeral nature, which causes them to be neglected by the monitoring and management programmes, and has led to loss of inherent ecosystems goods and services provision (Li et al., 2015).

The abundance and quality of wetlands in semi-arid environments are reported to be declining globally, due to climate change and variability, as well as poor land management practices (Mahdianpari et al., 2019). Gebreslassie et al. (2014) reported significant wetland losses in semi-arid Ethiopia, due to lack of policies to safeguard these systems, which has resulted in the loss of socio-economic services. It has been reported that about 30% of the wetlands in the semi-arid parts of China have been lost over the past 50 years, due to anthropogenic activities, with most of them disappearing between 1990 and 2000 (Liu et al., 2017), while over 50% of wetlands in South Africa have been eradicated in some catchments, due to climate change and anthropogenic activities (Day et al., 2010). Given the significance of the ecological services provided by wetlands, it is imperative that these systems are sustainably managed.

The basis for the sustainable management of wetlands hinges on the frequent monitoring of their eco-hydrological dynamics, in order to derive consistent and comparable information, which is lacking in most semi-arid regions, especially in sub-Saharan Africa (Mahdianpari et al., 2019). The availability of earth observation data offers an opportunity to map and monitor wetlands in a spatially-explicit manner in different climatic regions that lack monitoring systems (Gxokwe et al., 2020). The challenge, however, is that these data come in a range of spatial, spectral and temporal resolutions, which presents difficulties when coarse- to medium-resolution data, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), are used to map semi-arid wetlands.

Wetlands in semi-arid regions are mostly heterogeneous, with no definitive boundaries and they are spectrally similar to the surrounding landscapes. This results in difficulties when coarse- to medium-spatial resolution data are used to separate these systems from the surrounding landscapes (Mahdianpari et al., 2020). High-resolution data, such as Worldview-2 and Satellite Pour l'Observation de la Terre (SPOT 6-7), are commercially-available; they require complex processing algorithms and are therefore not feasible for monitoring the spatial characteristics of semi-arid wetlands over large areas and over time (Gxokwe et al., 2020).

Advancements in data analytic tools and platforms and the development of cloud-computing platforms, such as Microsoft Azure (MA), Amazon Web Services (AWS) and Google Earth Engine (GEE), provide opportunities for wetland monitoring and assessments across various scales. The AWS was launched in 2006 and contains several remote sensing data, ranging from Sentinel-1, Sentinel-2, Landsat 8, National Oceanographic and Atmospheric Administration (NOAA) and the Advanced High-Resolution Rapid Refresh Model (AHRRRM) (Tamiminia et al., 2020). Although AWS provides unique benefits by providing access to a large suite of machine-learning algorithms and artificial intelligence, the platform offers pay-as-you-go services. Microsoft Azure was launched in 2010 for building, deploying and managing applications, as well as services, through Microsoft-managed data centres. The platform consists of advanced machine-learning algorithms, like Landsat and Sentinel-2 data, for only North America, from 2013 to the present, as well as MODIS data, from 2000 to present.

The recently-introduced GEE platform offers parallelised processing on Google Cloud, which enables the processing of a stack of images at once, rather than relying on a single image. The platform hosts a 40-year petabyte scale of pre-processed remotely-sensed data, which includes Landsat, MODIS, National Oceanographic and Atmospheric Administration Advanced Very High-Resolution Radiometer (NOAA AVHRR), Sentinel 1, 2, 3 and 5-P, and Advanced Land Observing Satellite (ALOS) data, as well as advanced machine-learning algorithms (Amani et al., 2020). The other data-types available on the GEE cloud-computing platform include climate and geophysical data, as well as ready-to-use products, such as the Normalised Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) (Gorelick et al., 2017).

Although the GEE was launched over a decade ago, its application in remote sensing of wetlands, including the small seasonally-flooded systems in semi-arid regions, is still limited. A review by Tamiminia et al., (2020) reported 13 wetlands and mangroves studies that utilised

the GEE platform between the year 2010 and 2019 in all climatic zones, with most of them using Landsat 8 data. A review by Kumar and Mutanga (2018) also reported that 8% of the studies (out of the 300 identified articles) utilised the GEE platform for wetlands and hydrologically-related research globally. The most recent studies by Mahdianpari et al., (2020; 2019) used the GEE to monitor semi-arid wetlands in Canada, with reasonable accuracies (70%-90%); however, the studies focused on large-scale mapping. Due to the strengths of the GEE cloud-computing platform, there is a need to fully explore its capabilities in mapping and for determining the characteristics of semi-arid wetlands on a site-specific scale. This is particularly relevant in sub-Saharan Africa, where small wetlands are poorly documented; however, the application of the GEE in small and seasonally-flooded wetlands is still lacking, and the data on these systems are inconsistent and incomparable, due to the limited research.

Owing to this background, the overarching goal of this chapter was to characterise and map two small and seasonally-flooded wetlands (Nylsvley floodplain and Lindani valley bottom) in the semi-arid Limpopo Transboundary River Basin in South Africa, using the GEE cloud-computing platform and the multi-year Sentinel-2 composite data. More specifically, the objectives were: (1) to assess the spectral separability of the different wetland-cover classes detected, using the GEE and multi-year Sentinel-2 composite derivatives; (2) to evaluate the capability of the GEE cloud-computing platform to produce customised wetland cover maps with a reasonable accuracy, using the high-resolution Sentinel-2 data, and (3) to identify a suitable GEE machine-learning algorithm for accurately detecting and mapping semi-arid seasonally-flooded wetlands characteristics, using the multi-year Sentinel 2 composite data.

3.2 Materials and Methods

3.2.1 Field data

Land cover field data were collected towards the end of the dry season and at the beginning of the wet season between the 28th of September 2020 and the 1st of October 2020. The collected data included 600 ground-truth points that were collected on both wetlands, with their locations being determined by using a hand-held Geographical Positioning System (GPS) with an error margin of less than 3.5 m. The points were collected by using a stratified random sampling approach. This approach involved the division of the population into smaller sub-groups, based on certain attributes named strata, and random sampling was implemented on the strata (Ding et al., 1998). Stratified random sampling was selected because of the possibility that each sample is equally likely to occur. During the implementation of the stratified random sampling

in this study, the wetlands were subdivided into 30 m x 30 m quadrants, based on the size of the studied wetlands. The quadrants were 12 m apart, in order to minimise the overlapping of samples on the satellite image. In each quadrant, a minimum of 20 ground-truth points was collected, depending on the dominating landcover classes. Moreover, the wetland vegetation communities were visually identified on site and then grouped as either the short or long grass species, based on their structural features (canopy height and cover). The collected data were used in the GEE wetland model training and validating satellite derived wetland-cover classes.



3.2.2 Remote sensing data acquisition and processing

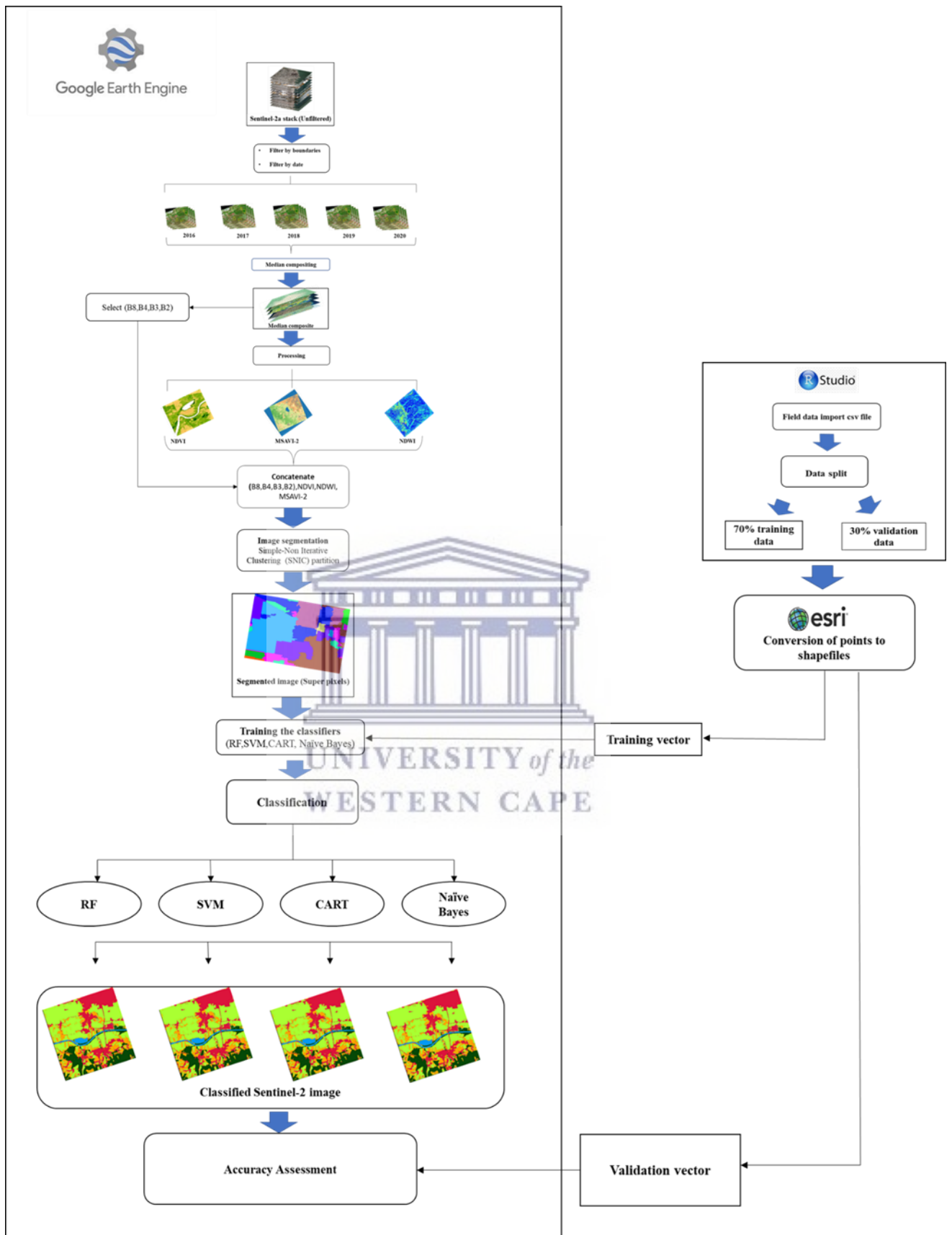


Figure 3.1 Steps taken to characterise and map the two wetlands, using GEE

The acquisition and processing of remotely-sensed data in this study were executed by following the steps shown in Figure 3.1. The Sentinel-2 MultiSpectral Instrument (MSI), level 2A (COPERNICUS/S2_SR/20151128T002653_20151128T102149_T56MNN) Surface reflectance images were acquired from the GEE database and used in this study. These products are already atmospherically corrected by using the Sen2cor toolbox, and they contain twelve UTM16 spectral bands that are scaled by 10000, as well as three QA bands, where one (QA60) is a bitmask band with cloud mask information. The image stack was filtered to represent the period of 2016-01-01 to 2020-12-31, and by the boundaries of the selected wetland using the codes `ee.Filter.Date ()` and `Image.filterBounds ()`. This period was chosen because a drought was reported in the Limpopo Basin in 2016, which caused significant changes in most surface water systems, including the wetlands in the area. This study aimed to establish whether the drought-induced changes within the wetlands could be determined. Two-hundred-and-ninety-six (296) images were obtained after the filtering process. The acquired image stack was then normalised for illumination effects (i.e. shade) and the minimisation of clouds, using median compositing. The median composite works by reducing a stack of images through the calculation of the median of all the values at each pixel across the stack of all matching bands, thus minimising the effects of shades and clouds (Mahdianpari et al., 2019). The median composite in this study was executed by using the code "Median ()" on the GEE. This was then used to calculate the Normalised Difference Vegetation Index (NDVI), the Normalised Difference Water Index (NDWI) and the Modified Soil Adjusted Vegetation Index 2 (MSAVI-2), using the equations given in Table 3.1. The NDVI is one of the most widely-used vegetation indices in wetland studies, due to its sensitivity to photosynthetically-active biomass; it can also discriminate between vegetation and non-vegetation, as well as wetlands from non-wetland features (Liu and Huete, 1995). The NDWI was selected due to its sensitivity to open water, which enables discrimination between the water and land surfaces (McFeeters, 1996). MSAVI-2 was chosen to improve the limitations of the NDVI. In addition to the extracted indices, the Near Infrared (NIR), Red, Green and Blue bands were selected and concatenated to the NDVI, NDWI and MSAVI-2 outputs to produce an image with only the bands for wetland classification. The NIR band was chosen because of its usefulness for distinguishing water from the land surface, as well as its ability to discern the biomass content of vegetation and its health, because water absorbs strong NIR light, while vegetation strongly reflects NIR light. The Red band was chosen because of its ability to delineate the wetland classes and to detect chlorophyll absorption in vegetation. The Blue band has the ability to differentiate between the soils and vegetation. The produced composite with integrated indices was then

subjected to an Object-Based Image Analysis (OBIA), which was selected because of its superiority over a pixel-based classification, as shown in various wetlands mapping studies (e.g. Berhane et al., 2018; Kamal and Phinn, 2011; Kaplan and Avdan, 2017). Moreover, the approach not only relies on the spectral characteristics of each pixel, but it also considers other pixel characteristics, such as the size, shape and contextual information, thus improving the spectral separability within the classes of the heterogeneous wetlands (Halabisky, 2011). The first step in OBIA is image segmentation. The process involves partitioning the image into multiple discrete and non-overlapping segments, based on a specific criterion (Dlamini et al., 2021). In this process, individual pixels are merged to produce larger objects. This increases the discrimination of spectrally-similar objects by using the texture, shape and contextual features and prevents the ‘salt and pepper’ noise in the final classification map (Dlamini et al., 2021a; Mahdianpari et al., 2020). In this study, a Simple Non-Iterative Clustering (SNIC) algorithm was used to segment the composite. The SNIC algorithm was chosen because of its simplicity, its memory efficiency, its processing speed, as well as its ability to incorporate the connectivity between pixels after the algorithm has been initiated (Achanta and Süsstrunk, 2017). The SNIC algorithm starts the process of image segmentation by initialising the centroid pixels on a regular grid image; the dependency of each pixel is then established relative to the centroids, by using the distance in a five-dimensional space of colour and spatial coordinates. In particular, the distance integrates the normalised spatial and colour distances to produce uniform super pixels (Achanta and Süsstrunk, 2017). The candidate pixel is selected, based on the shortest distance from the centroid (Achanta and Süsstrunk, 2017). The SNIC algorithm was executed by using the code ‘`ee.Algorithms.Image.Segmantation.SNIC ()`’ on the GEE and the output was an image with super pixels, calculated textures, areas, sizes and perimeters of all the super pixels.

Table 3.1 Optical features extracted from the remotely-sensed data used

Data	Data	Formula	Band	Reference
	extracted		Width	
			(nm)	
	B8 - Near-infrared		842	
	(NIR)		665	
	B4 - Red		560	
	B3 - Green		490	
	B2 - Blue			
	Normalized Difference Vegetation Index (NDVI)	$\frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}$	-	(Liu and Huete, 1995)
Sentinel-2	Normalized Difference Water Index (NDWI)	$\frac{\text{NIR} - \text{Green}}{\text{NIR} + \text{Green}}$	-	(McFeeters, 1996)
	Modified Soil Adjusted Vegetation index (MSAVI-2)	$\frac{2 \times \text{NIR} + 1 - \sqrt{(2 \times \text{NIR} + 1)^2 - 8 \times (\text{NIR} - \text{Red})}}{2}$		(Qi et al. 1994)



3.2.3 Adopted wetland classification scheme

Satellite image classification was executed by using the Random Forest (RF), Support Vector Machine (SVM), Classification and Regression Tree (CART) and Naïve Bayes (NB) algorithms. The RF is an ensemble classifier, which consists of a combination of tree classifiers, and each classifier is generated by using the random factor, which is sampled

independently from the input vector data. Each tree casts a unit vote for a popular class to classify an input vector (Ao et al., 2019; Pal, 2005). An advantage of the RF algorithm is its ability to handle large differentiations within the land-cover classes, and the noise in data can be neutralised (Slagter et al., 2020). Moreover, the algorithm does not require an understanding of the data distribution, unlike the parametric classification algorithms, such as the maximum likelihood, where the data distribution needs to be known. The SVM applies a sophisticated kernel function to classify the datasets with a complex decision surface. One of its strengths is that the uncertainty in the model structure is decreased and, similar to RF, the algorithm does not rely on the data distribution (Oommen et al., 2008). CART is a tree-based classification algorithm that measures the dependent relationship of one variable on other variables (Simioni et al., 2020). An advantage of CART is its ability to naturally model the non-linear boundaries because of its hierarchical structure. NB forms part of simple probabilistic classifiers, which is based on applying the Bayes theorem with a strong independence assumption between the features (Shelestov et al., 2017a; Simioni et al., 2020). An advantage of the NB algorithm includes its ability to solve multi-class prediction problems, it is time-saving and it has less training data requirements. The four algorithms were chosen because of their above advantages, in addition to their acceptable performance that was demonstrated in several land cover studies (e.g. Dlamini et al., 2021; Rana and Venkata Suryanarayana, 2020; Rodriguez-Galiano et al., 2012; Simioni et al., 2020; Slagter et al., 2020). Moreover, studies such as Hayri Kesikoglu et al., (2019) and Tian et al., (2016) reported that the application of the above algorithms on mapping semi-arid wetlands is limited, hence they were selected. Prior to classification, the algorithms were trained by using 70% of the field data. The data were randomly split in R-studio to produce a 70% training set, as well as a 30% validation set. The code that was used to split the data in R-studio is `'wasdt = sort(sample(nrow(data), nrow(data)*.7))'`. After splitting, the data were converted to GIS files in Esri ArcGIS 10.2, and they were then imported to GEE for model training and validation. The training of the classifiers was executed by using the code `'ee.Classifier.train ()'`. Classification using the latter algorithms was then implemented on the segmented image by using the code `'Image.classify()'` in GEE.

3.2.4 Accuracy assessment

Three evaluation indices were used to evaluate the performance of the classification algorithms. These were the Overall Accuracy (OA), the producer's accuracy and the user's accuracy. In addition, line plots, as well as Jeffries-Matusita (JM) distances, were used to establish the

separability of different wetland classes, by using the selected spectral bands. JM is a parametric criterion that ranges between 0 and 2. This criterion uses distances (Equations 3.1 and 3.2) between the class means, and the distribution of values from the means, to assess the separability of one class from the other (Dabboor et al., 2014; Wang et al., 2018). Distances approaching 2 indicate a greater average distance between two classes and that they are therefore separable, using the data type. The OA was used to measure the efficiency of the used algorithms and it was quantified as the ratio of the total correctly-labelled samples and the total number of testing samples. The producer's accuracy was used to measure the probability of the reference sample being correctly classified on the map. The user's accuracy was used as an indicator of the probability that a classified pixel in the wetland cover classification map accurately represents that category on the ground. The accuracy assessments with the latter indices were computed and executed in GEE. The JM distance presented in Dabboor *et al.* (2014) is given as:

$$JM = 2 (1 - e^{-B}) \quad (3.1)$$

where B is the Bhattacharyya distance and is quantified as:

$$B = \frac{1}{8} (\mu_i - \mu_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{|\Sigma_i + \Sigma_j|/2}{\sqrt{|\Sigma_i| |\Sigma_j|}} \right) \quad (3.2)$$

where μ_i and Σ_i are the mean and covariance matrix of class i and μ_j and Σ_j are the mean and covariance of class j . In evaluating the classification accuracy of the output image, using the latter indices, the codes 'Image.accuracy()' for OA, 'Image.producersAccuracy()' for the producer's accuracy and 'Image.consumersAccuracy()' for the user's accuracy were implemented in the GEE. During the implementation of the latter codes, the validation vector data were first imported to the GEE and used to sample out the corresponding regions on the classified image by using the code 'Image.sampleRegions()'. The sampled regions were then used in the latter codes as input image to test the accuracy of the classification outputs. The line plots and JM distances were implemented by using a series of interconnected codes presented in the supplementary material.

3.3 Results

3.3.1 Class separability results from the spectral data used

Line plots (Figure 3.2) showing the spectral reflectance values per class of the studied wetlands indicate that most wetland classes are not distinguishable when using Band 2, Band 3 and Band

4 for both wetlands, except the long grass in the Nylsvley flood plain, which was discernible from the rest of the classes by using B4 and B8. The results also indicated that the water class in the Lindani valley bottom has a high spectral reflectance in the NIR region, although a low reflectance was anticipated in this spectral region.

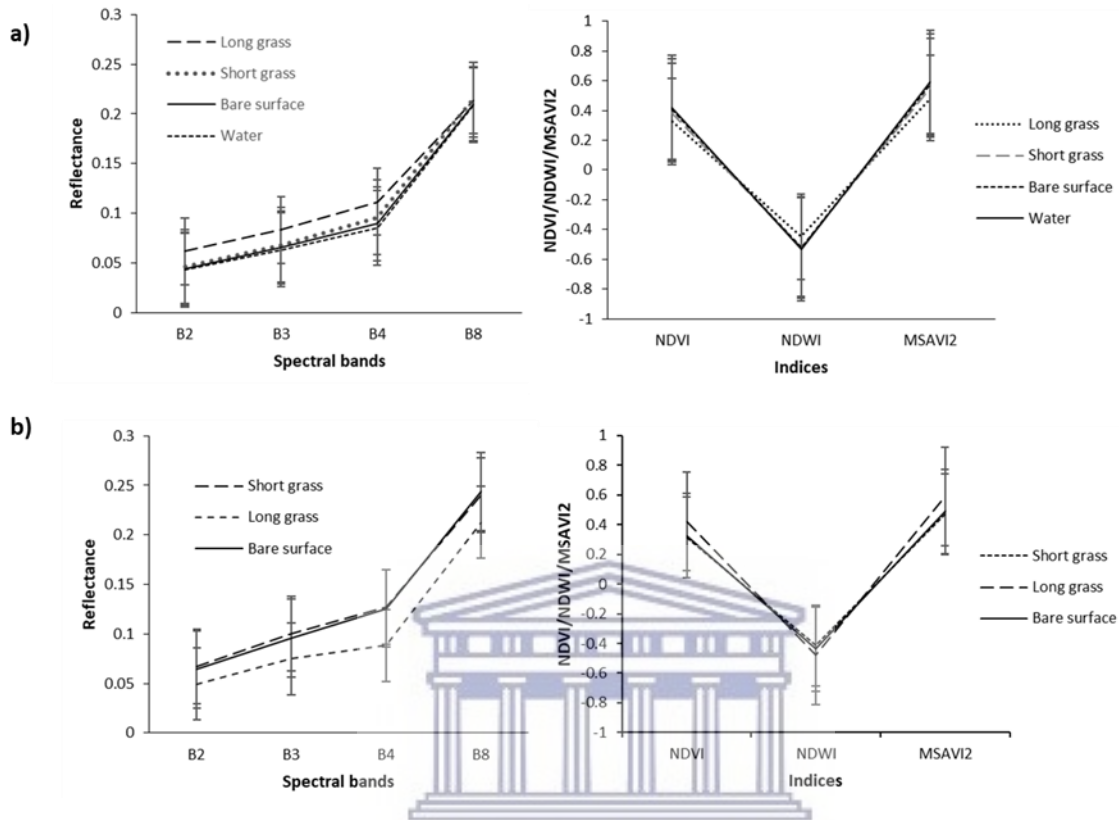


Figure 3.2 Wetland cover class spectral reflectance values for a) the Lindani valley bottom and b) the Nylsvley floodplain, using pixel values extracted from the training data

The JM distances obtained from the multi-year median composite (Tables 3.2 and 3.3) indicate that the wetland features were hardly distinguishable from the single optical bands for both wetlands. The least distinguishable classes were the bare surface and water from the Lindani valley bottom, as well as the bare surface and short grass from the Nylsvley floodplain. The JM distances for all the least distinguishable classes did not exceed 1.4 for both wetlands. The results also showed that the synergic use of all optical features significantly increased the separability between the classes, with the JM distances exceeding 1.8 for both wetland classes.

Table 3.2 JM distances between wetland-cover classes in the Lindani valley bottom

Optical features	D1	D2	D3	D4	D5	D6
NIR	0.0016	0.0131	0.1036	0.0135	1.272	0.1398
Red	0.3146	0.4311	0.5635	0.0502	0.1576	0.0314
Green	0.4331	0.6508	0.8978	0.0600	0.2714	0.0966
Blue	0.4675	0.6381	0.8861	0.0418	0.2193	0.0858
NDVI	0.2728	0.3044	0.6300	0.0385	0.1823	0.1127
NDWI	0.3478	0.4028	0.4346	0.0036	0.0168	0.0089
MSAVI2	0.3124	0.3509	0.6849	0.0041	0.1948	0.1187
ALL	2	2	2	2	2	2

*** D1: Long grass and short grass, D2: Long grass and bare surface, D3: Long grass and water, D4: Short grass and water, D5: Short grass and water, D6: Bare surface and water*

Table 3.3 JM distances between wetland-cover classes in the Nylsvley floodplain

Optical features	D1	D2	D3
NIR	0.2586	0.3839	0.0342
Red	0.6208	0.5998	0.0016
Green	0.6913	0.5289	0.0226
Blue	0.7211	0.5828	0.0251
NDVI	0.5663	0.5534	0.0358
NDWI	0.6675	0.3876	0.1351
MSAVI2	0.5986	0.0038	0.0524
ALL	2	2	1.990

***D1: Short grass and long grass, D2: Long grass and bare surface, D3: Bare surface and short grass*

3.3.2 Classification results and accuracies

Four GEE algorithms were applied to a median composite of Sentinel-2 images to produce custom maps for two seasonally-flooded wetlands of variable sizes in the LTRB. Figures 3.6 and 3.7 show the custom maps for the studied wetlands. The Overall Accuracies, based on the algorithms used, ranged between 20% and 80% for both wetlands, with Random Forest (RF)

having a high OA for both the Lindani valley bottom and the Nylsvley floodplain (68.80% and 80.55%), and Naïve Bayes (NB) having low OA values for both wetlands (29.50% and 25.00%) (Figure 3.3). The other algorithms had a reasonable accuracy, with the Support Vector Machine (SVM) attaining 66.60% for Lindani and 62.29% for Nylsvley. The Classification and Regression Tree (CART) achieved an OA of 62.30% for Lindani and 75.00% for Nylsvley, thus proving the superiority of RF among the other algorithms used in this study.

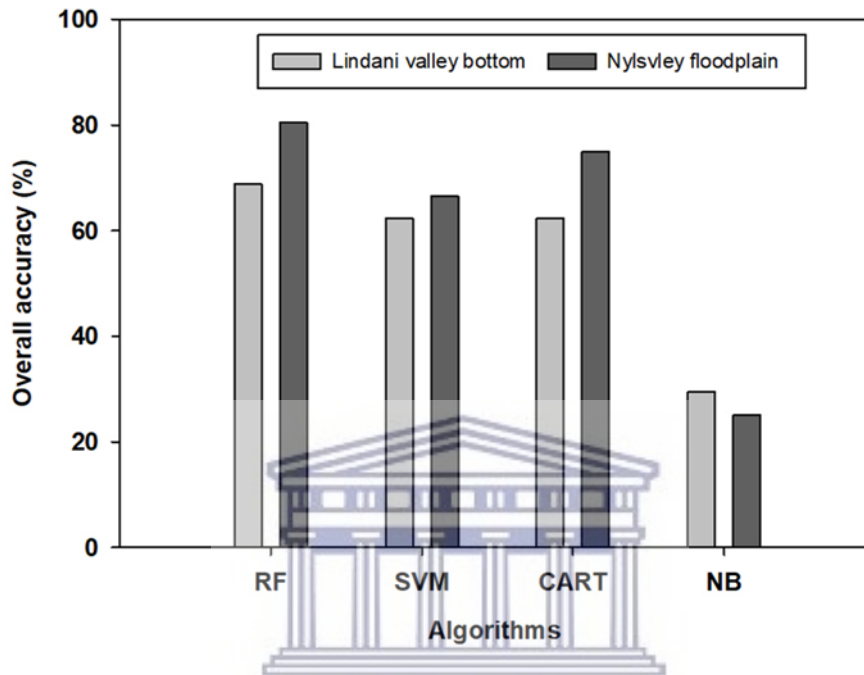


Figure 3.3 Overall accuracy comparison between the algorithms used for the two studied wetlands

Figures 3.6 and 3.7 show the distribution of wetland-cover classes at a 10 m spatial resolution. The maps illustrate the fine separation between the wetland classes for all the algorithms. However, the water class could not be detected for the Nylsvley floodplain, due to the unavailability of training data representing this class. The areas of the various land cover types in Figure 3.4 show that, based on all the algorithms, except NB, the most dominating class in the Lindani valley bottom is short grass, consisting of *Cynodon dactylon* and *Oryza longistaminata*. In contrast, the NB identified water as being the most dominant class. The area of the short grass ranged between 5 ha and 25 ha, with the NB model identifying the smallest area for this cover type. The producer's and user's accuracies for the dominating short grass ranged between 20% and 91%, with NB model having the lowest producer's accuracy (Figure 3.5), and the user's accuracy was from 60% to 80%, with SVM having the lowest accuracy. The least dominating classes in the Lindani valley bottom wetland were the water and bare

surface with areas, with a range of 0.3-5.0 ha for the bare surface and 0.2-10.0 ha for water. The producer's accuracies for these two classes ranged between 0% and 75% for water and 0% and 30% for the bare surface. The SVM algorithm achieved a 0% producer's accuracy for both classes. The user's accuracy for the two classes was from 0% to 75% for the bare surface and 0% to 50% for water. The SVM algorithm had a 0% user's accuracy for both classes (Figure 3.5). The dominating landcover class in the Nylsvley floodplain was bare surface, with an estimated area ranging from 362 ha to 495 ha. The CART algorithm estimated the lowest area. The least-dominating class was long grass, comprising mainly the common reed species, such as *Phragmites australis*. The area for the least dominating wetland cover class ranged between 130 ha and 352 ha. The highest area was estimated by CART, and the lowest by NB. The producer's accuracy for the dominating class ranged between 62% and 88%, with the highest being attained by the NB and the lowest by the SVM model. The producer's accuracy for the least-dominating class ranged between 33% and 66.6%, with the highest attained by NB model and the lowest by the CART and RF models. The user's accuracy for the most dominating class ranged between 22% and 71%, with the highest recorded by SVM and CART, and the lowest by the NB model. The user's accuracy for the least dominating class ranged between 20% and 66.6%, with the highest recorded by RF and the lowest by CART.

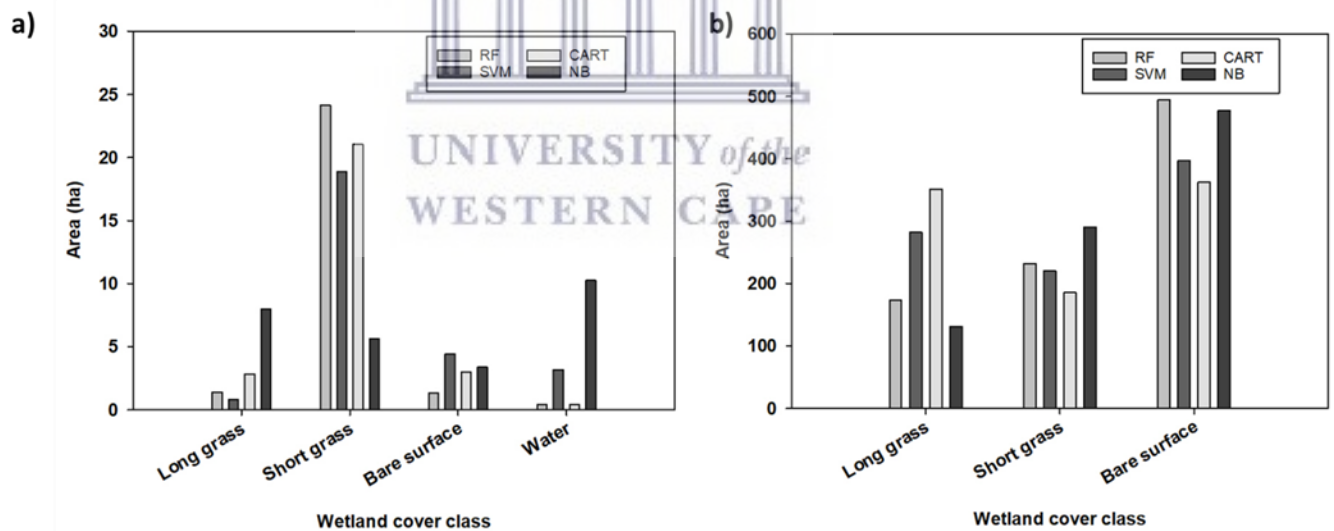


Figure 3.4 Wetland cover class areas for a) the Lindani valley bottom and b) the Nylsvley floodplain

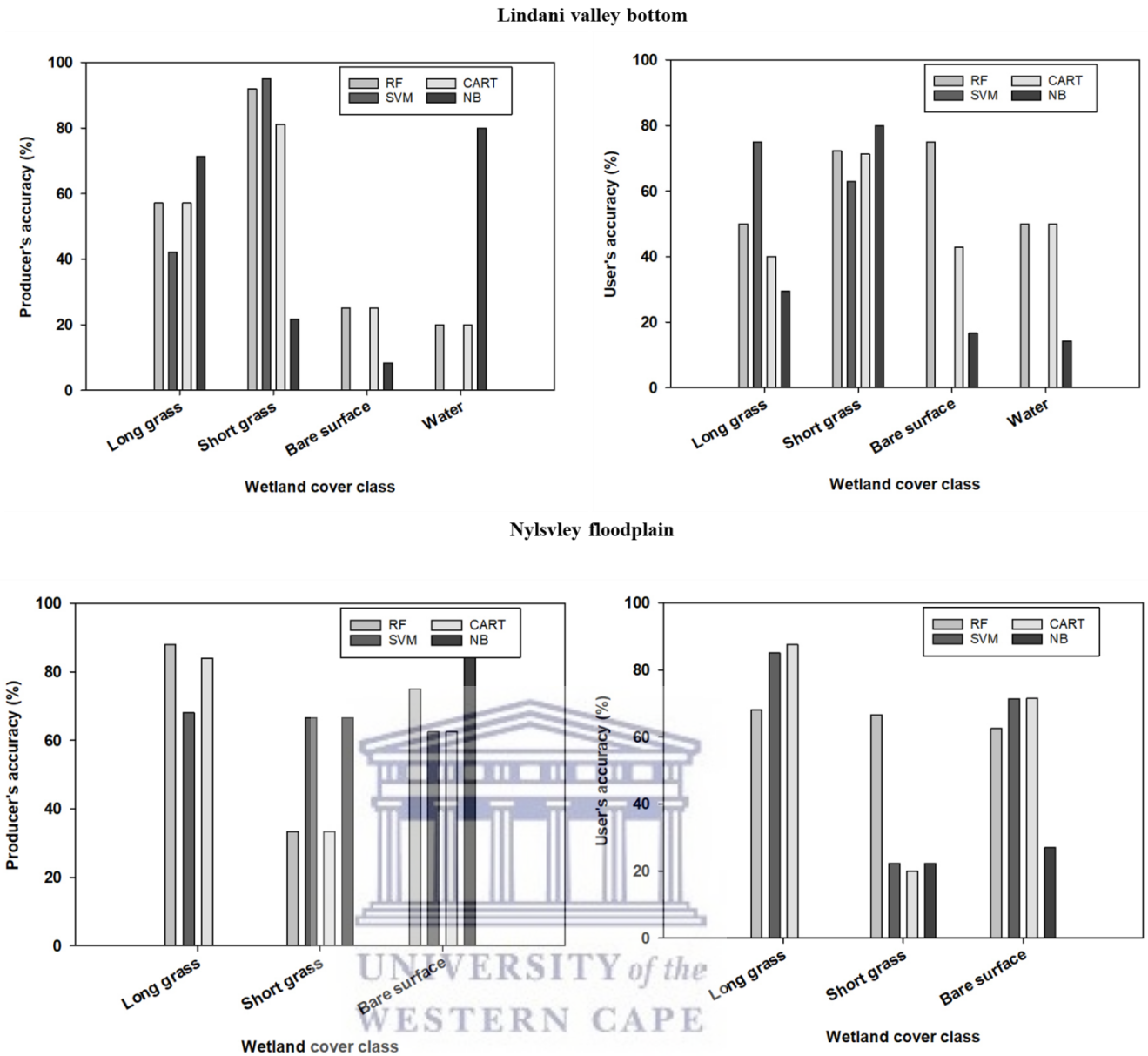


Figure 3.5 Producer's and user's accuracies for the two wetlands classes

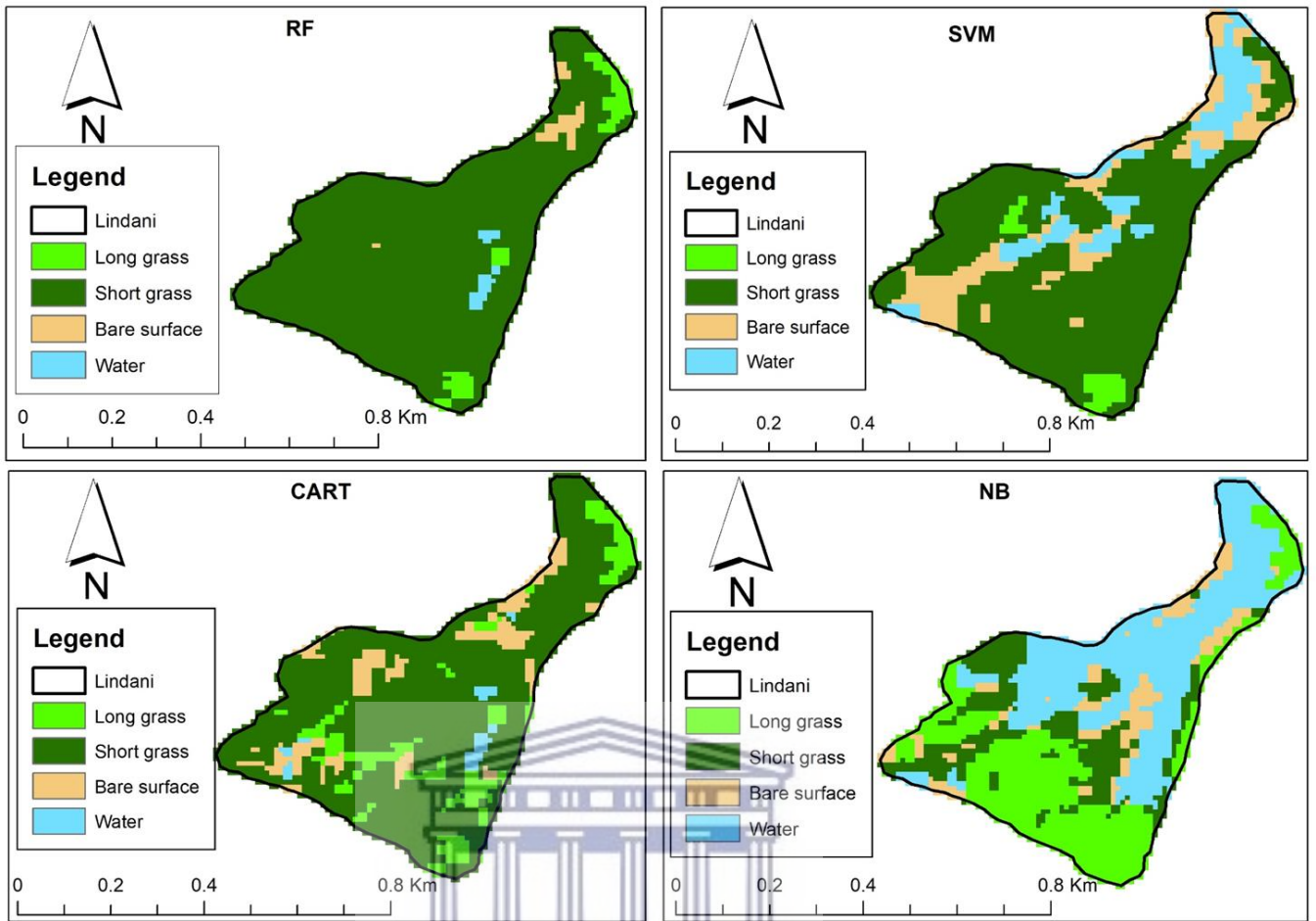


Figure 3.6 Lindani valley bottom Sentinel-2 derived wetland-cover classes, based on the four used algorithms



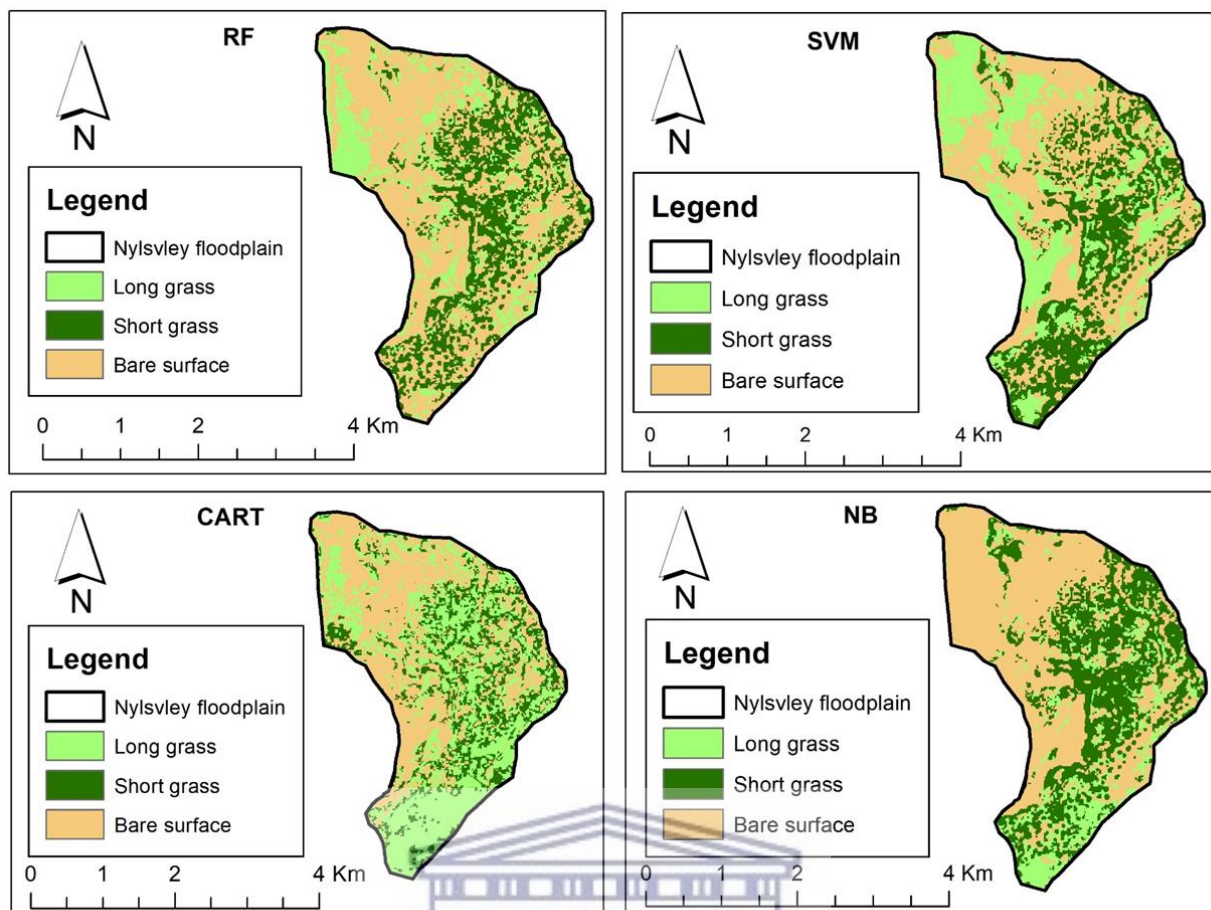


Figure 3.6 Nylsvley floodplain Sentinel 2-derived wetland cover classes, based on the four algorithms used

3.4 Discussion

The accurate detection and monitoring of small seasonal and heterogeneous wetlands in semi-arid regions is important for understanding the eco-hydrological dynamics of these systems, as most of these wetlands, particularly in sub-Saharan Africa, offer socio-economic benefits to the surrounding communities (Gardner et al., 2009; Kabii and Kabii, 2005; Thamaga et al., 2021). Advancements in data analytic tools provide unique opportunities to improve the detection and monitoring of semi-arid wetlands of variable sizes, which were not feasible when using the traditional remote sensing techniques. The introduction of cloud-computing platforms, such as the Google Earth Engine (GEE), offers advantages, such as advanced machine-learning algorithms and parallel processing, memory efficiency and a fast image processing power. This study sought to assess the capabilities of the GEE cloud-computing platform for characterising and mapping semi-arid seasonally-flooded wetlands on a site-

specific scale, as well as suggesting a suitable GEE machine-learning algorithm for characterising and mapping such systems.

In general, the results demonstrate the capabilities of the Google Earth Engine cloud-computing platform for characterising and mapping the semi-arid wetland systems of variable sizes with acceptable overall accuracies. In addition, the RF, CART and SVM algorithms proved to be superior to the NB model, with low OA values for both the studied wetlands. Although higher OA values were obtained using the latter algorithms, low producer's and user's accuracies were reported in the Lindani valley bottom for the bare and water classes, especially when using SVM model. The contributing factors of the low producer's and user's accuracies for the two classes were the few training and validation sample points representing these classes, due to less than a pixel spatial coverage for each class within the wetland boundary. Fewer training and validation points tend to reduce the level of accuracy during classification (Zhen et al., 2013; Corcoran et al., 2015; Mahdianpari et al., 2020). In addition to the few training and validation data points, the multi-year images used during the computation of the median composite did not consider seasonality and the yearly differences, which may have reduced the producer's and user's accuracies. The study by Noi Phan et al., (2020) analysed the impact of different composition methods, as well as the input images, on the classification accuracy of the different landcover classes using GEE. The results showed that the temporal aggregation that was considered during median compositing produced high accuracy values for the classification outputs. This demonstrates the significance of temporal aggregations during the median compositing stage. Although the results underscore the relevance of the GEE cloud-computing platform for characterising and mapping small seasonally-flooded wetlands in the semi-arid regions, there are some limitations associated with its use. These include the computational restrictions, where the large trainings required by complex machine-learning algorithms cannot be performed, due to space limitations, the unavailability of complex and accurate image segmentation algorithms, as well as restricted image-processing tools. Moreover, the algorithms that were used have limitations, such as the slow training and biasness when dealing with categorical data, in the case of RF, under-performance when dealing with large data, in the case of SVM, more time required to train the model, in the case of CART, and the dependencies amongst the classes that cannot be modelled, when using NB.

The class spectral separability results show that the use of single optical bands and indices for differentiating between the wetland classes for all bands, was not feasible in this case, although spectral bands, such as NIR and Red, are known to be useful in wetland delineation, vegetation,

soil and geology studies (Amani et al., 2018). Water was expected to be easily-discernible from other classes, as it is known to have a strong NIR adsorption and a Low NIR reflectance, which makes it easy to discriminate it from other classes, such as the vegetation and soils, which reflect more light in the NIR region. A study by Amani et al., (2018) reported the great separability of wetland-cover classes, when using the NIR band, particularly in the shallow waters of Newfoundland in semi-arid Canada. The Red band is known to be useful for detecting chlorophyll absorption in vegetation, as well as the composition of soils, where soils with rich iron oxide will have a stronger reflectance and healthy vegetation will absorb red light, making it easy to separate these classes from the water. The results of Amani et al., (2019) showed that the shallow water of the Newfoundland wetlands in Canada were discernible from the other classes, such as the soils and vegetation, by using the Red Band; however, it was highlighted that it may not be easy to discriminate these classes in some instances, due to similarities in their spectral reflectance.

The synergic use of all spectral features significantly increased the separability of different classes, with JM distances of above 1.9, for both wetlands. However, the logical expectation was that vegetation and water in the NIR region would have a strong reflectance and adsorption, but in the case of the Lindani valley bottom, water had a stronger reflectance in this spectral region. This could be due to the submerged and floating wetland vegetation that interfered with the water signals, thus causing the stronger reflectance of water in the NIR region (de Vries et al., 2017). In addition, materials at the bottom of the shallow waters are known to affect the absorption and reflectance of light by shallow waters (Vinciková et al., 2015). This could also be the case for this wetland with shallow water. Jones (2015) reported increased errors in mapping the spatial extent of water in areas within the greater Everglades, where the vegetation is floating. This showed the problem of discriminating between water and vegetation in such wetlands. In this study, both wetlands had short and long grass classes, with higher reflectance values in the visible red-light region, which is indicative of water-stressed vegetation. Water-stressed vegetation is known to have stronger reflectance signals in the visible red-light region (Adam et al., 2010; MacArthur, 1975). Caturegli et al. (2020) assessed the effects of water stress on the spectral reflectance of Bermuda grass (*Cynodon dactylon*) under controlled laboratory conditions. The results showed an increase in red light reflectance with the increasing water stress, thus proving that water-stressed vegetation has a stronger reflectance in the visible red-light region. The indices were found to be the least useful for separating the classes for both wetlands, partly because the seasonality and yearly differences were not

considered during the median compositing of the selected images. In most cases, the maximum NDVI values tend to correspond to the growing season. A study by Wang et al., (2020) examined the response of the maximum NDVI values to precipitation during the period of active growth, from 2000 to 2013, in the Alpine grassland site of the Tibetan Plateau. The results showed a positive linear relationship between the precipitation and the maximum NDVI, thus proving that NDVI is most useful during the peak growing season.

The findings of this chapter prove that the GEE platform and advanced machine-learning algorithm have the potential to improve the detection and monitoring of small seasonally-flooded wetlands in semi-arid regions, using the Sentinel-2 multi-year composite image. This has previously been a daunting task, when using the conventional mapping methodologies and optical data. In addition, the results demonstrated that the most detectable wetland features were mostly wetland vegetation communities, although there were some challenges relating to accuracy, particularly for the Lindani valley bottom system. The study provides baseline information and new insights into enhancing small and seasonally-flooded wetlands from optical data at a reasonable accuracy and moderately high-resolution, thus underscoring the significance of freely-available optical data for monitoring semi-arid, seasonally-flooded systems. This is important for semi-arid regions with limited data access, particularly in sub-Saharan Africa, where less attention is given to these systems due to the limited information regarding their status, even though they serve as an important source of water for most communities. The findings also contribute towards the ongoing global wetland monitoring programmes, such as Wetland Monitoring and Assessment Services for Transboundary Basins in Southern Africa (WeMAST), which is funded by European Union-Africa Global Monitoring for Environmental Security (EU Africa GMES). This programme aims to develop an integrated platform for wetland assessment and monitoring that will support the sustainable management of transboundary basins. Furthermore, the study contributes to Goal 6.6 of the Sustainable Development, which seeks to halt the degradation and destruction of ecosystems, including wetlands, and to assist in the recovery of the already-degraded systems.

3.5 Conclusion and Recommendations

The current study was aimed at characterising and mapping two seasonally-flooded wetlands in the Limpopo Transboundary River Basin, with the objective of assessing the usefulness of the GEE cloud-computing platform in producing maps of the studied wetlands, as well as suggesting possible GEE algorithms for the detection and mapping of such systems. The main

findings indicate the capabilities of GEE in mapping seasonal semi-arid wetland systems of variable size, with a reasonable overall accuracy, and that the RF, CART and SVM algorithms are superior to the NB model. Although reasonable overall accuracies were attained, there were poor producer's and user's accuracies for some classes, such as the water and bare surfaces, especially for the Lindani valley bottom wetland. This can be attributed to less than a pixel of spatial coverage of these classes within the perimeter of the wetlands, which results in them being difficult to detect with the highest precision, using Sentinel-2 composite data. In addition, the seasonality and yearly difference were not considered, which could have significantly affected the results, because some features, such as water, tend to correspond to the seasonal changes, especially for semi-arid areas. It is therefore recommended that temporal variability be considered, in order to capture the peak growing season of the systems and to better enhance the wetland features. Spectral confusion was also observed between water and some wetland vegetation, which resulted in a higher reflectance of water in the NIR region. In order to avoid this, the study recommends the integration of Synthetic Aperture Radar (SAR) data with the optical data, since these data can penetrate through forested vegetation, thus minimising the effect of floating vegetation in the detection of the water class. Moreover, the testing of other machine-learning algorithms, such as the Artificial Neural Network (ANN), as well as the inclusion of Short-wave infrared and thermal bands, is recommended. In addition, there is a need to evaluate the available remotely-sensed data and its capability in long-term monitoring of small and seasonally-flooded wetlands located in semi-arid environments.

CHAPTER 4

AVAILABLE SATELLITE DATA CAPABLE OF MONITORING SMALL AND SEASONALLY-FLOODED WETLANDS IN SEMI-ARID SOUTHERN AFRICA

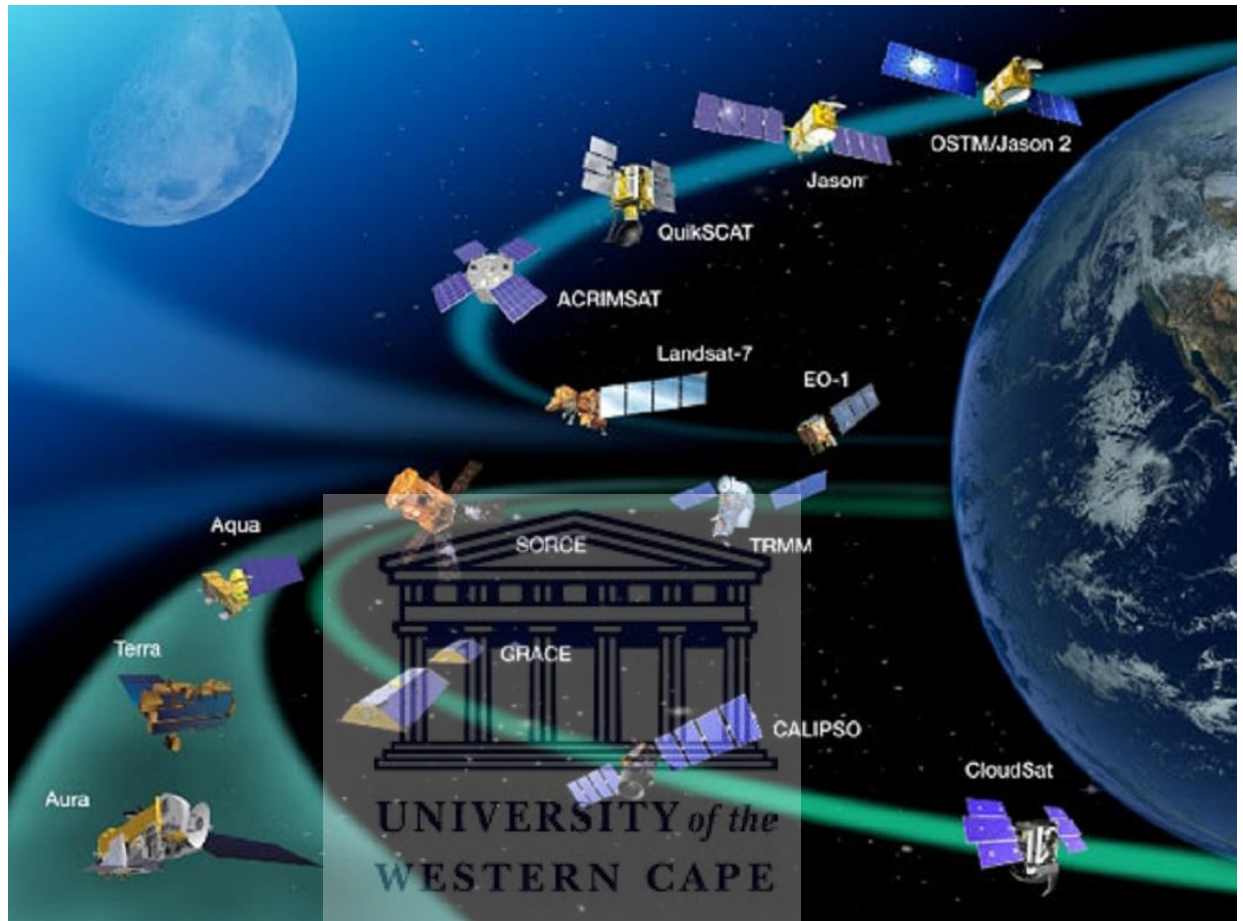


Photo: Courtesy of NASA 2020

This chapter is based on the following manuscript: (Under review)

Gxokwe, S., Dube, T., Mazvimavi, D. 2022. Available satellite data capable of monitoring small and seasonally-flooded wetlands in semi-arid southern Africa. *Ecohydrology*, Manuscript no. **ECO-22-0119** (IF =3.166)

Abstract

The use of remote sensing for the time-series monitoring of wetlands remains an attractive and practical tool for monitoring the eco-hydrological dynamics of wetlands, given its ability to overcome the challenges associated with in-situ data availability. Nonetheless, obtaining seamless and cloud-free, time-sensitive data for the accurate routine monitoring of wetlands condition remains a challenge. This study thus evaluated the amount of satellite scenes available on in the Google Earth Engine (GEE) catalogue that are capable of monitoring the seasonal eco-hydrological dynamics of small and seasonally-flooded wetlands in the semi-arid environments of southern Africa, over a 20-year period (2000-2020). The Nylsvley floodplain was thus used as a case study. More specifically, the study assessed all the available products on the GEE platform (Landsat TM, ETM+, OLI, Sentinel- 1 and 2), and the identified images were then filtered and screened, based on varying cloud cover percentages (i.e. 0%, 1-10%, 11-25%, 26-50%) on the GEE platform. The screening and filtering results showed a good number of satellite products available on this platform. Approximately 1 376 images were found for the period under study. The Landsat series data had 492 images and Sentinel-1 had 394 images, whereas Sentinel-2 had 490 images. Sentinel-2 and Landsat-7 had the highest number (69% and 76%) of images, with a cloud cover percentage range of 0-20%. However, images with a cloud cover percentage of above 26% were not included in the analysis. In addition, the use of satellite images with 0% cloud cover yielded an Overall Accuracy (OA) ranging between 69-72%, a 1-10% cloud cover had an OA ranging between 68-70% and an 11-25% cloud cover had an OA ranging between 69-80.55% for both the dry and wet seasons. Generally, the classification results indicated satisfactory overall accuracies (68-80%) for all the scenes, although there were some inaccuracies for some classes (bare surface and long grass), particularly when using Landsat-7 scenes. The results indicated that there is a reasonable amount of archival satellite data available that is capable of monitoring small and seasonally-flooded wetlands, which provide useful insights into the eco-hydrological dynamics of these ecosystems. In addition, the use of cloud-computing platforms, such as the Google Earth Engine (GEE), provide a unique opportunity to address the problems associated with big data filtering, processing and analytics, and which will enhance environmental monitoring and assessments.

Keywords: Archival satellite data; Cloud cover; Cloud-computing; Satellite data availability; Spatial data filtering; Wetland condition

4.1 Introduction

The significance of wetland functions has always been recognised by the scientific community. Despite this, over 50% of wetlands globally have been degraded and lost during the 20th and 21st century, due to environmental change and anthropogenic factors (Davidson, 2014). There is, therefore, an urgent need for holistic intervention, in order to prevent their further degradation and loss, especially in the semi-arid environments where most of these systems serve as important sources of water and other ecological services to the surrounding communities (Collins et al., 2014). The basis for proper interventions is centred on an understanding of the dynamic changes that occur in wetlands over time, which requires the frequent monitoring of their eco-hydrological dynamics.

Currently, wetland inventory data (geospatially referenced information on their status, extent, characteristics and functions) is inconsistent, scanty, incomparable and incomplete; it is usually outdated and non-existent for most regions, which makes it difficult to track wetland changes over time, especially in semi-arid sub-Saharan African regions where wetland data inventories are limited (Halabisky, 2011; Hird et al., 2017). The lack of frequent monitoring and assessment results in there being limited explicit, reliable and current spatial and temporal information about the wetland characteristics (Thamaga et al., 2021). The use of remote sensing in the time-series monitoring of wetlands remains an attractive and feasible tool for understanding these systems, since it can overcome the main issues associated with coverage and accessibility, when compared to the traditional onsite monitoring approaches. The challenge, however, is obtaining time-sensitive data and the generation of precise and accurate information about wetlands (Wang and Yésou 2018). Significant efforts have been devoted to developing time-sensitive remote sensing databases and approaches for the long-term monitoring of wetlands (Hird et al., 2017); for example, in Wang and Yésou (2018), Huang et al., (2018) and Gallant (2015).

The availability of freely-accessible, high-resolution archival optical data, such as the Sentinel-2 and Landsat series data, offer an unprecedented opportunity for the time-sensitive and multi-decadal analysis and monitoring of all wetland systems at improved spatial, spectral and temporal resolutions (Kandus et al., 2018). Moreover, the introduction of cloud-based platforms, like the Google Earth Engine (GEE), provides an opportunity to address the problems associated with huge satellite data filtering, processing and analysis by a great number of users (Hird et al., 2017; Inman and Lyons, 2020; Kandus et al., 2018; Thamaga et

al., 2021). Currently, the GEE hosts 40 years of peta-bytes of remotely-sensed data, like Sentinel-1,2,3 and 5P, Landsat 4-8, Moderate Resolution Image Spectroradiometer (MODIS), Advanced Land Observing Satellite (ALOS), as well as the National Oceanographic and Atmospheric Administration Advanced Very High-Resolution Radiometer (NOAA AVHRR), that is freely available, as well as advanced machine-learning algorithms. This makes the GEE cloud-computing platform more relevant in the study of natural resources, such as wetlands in most regions and over longer periods of time. However, challenges and uncertainties relating to the continuity of data still exist (Wang and Yésou, 2018). Furthermore, the capability (in terms of the image quality) of the remotely-sensed data that are accessible on these platforms, for monitoring the long-term wetland eco-hydrological dynamics, remains unknown, particularly in cloud-prone areas (Wua et al., 2019).

Several studies have utilised these platforms in several long-term environmental monitoring programmes, such as global forest mapping and earth surface water monitoring, amongst others (Hansen, 2013; Pekel et al., 2016; Yamazaki and Trigg, 2016); nonetheless, wetlands are among the most under-represented systems using these platforms (Wua et al., 2019). The long-term comprehensive monitoring of wetlands is critical for their sustainable management, and thus for achieving the Sustainable Development Goals (SDG), such as SDG 6.6, which are aimed at ensuring the conservation and restoration of freshwater ecosystems, including wetlands, by monitoring their status and change over time. The long-term monitoring of wetlands also contributes to SDG 6.4, which emphasises the monitoring of freshwater availability, environmental requirements, as well as water use. This is done in order to analyse, in detail, the water scarcity issue and its impact on both the economy and the population, and to encourage people to consider the health of ecosystems (Kates et al., 2005). Programmes such as the African Space Policy, Space, Science and Technology, National and Regional Space Strategies and Science, Technology, and Innovation Strategy for Africa (STISA) advocate the use of space science technology tools for monitoring the sustainable use of natural resources, including wetlands. The main challenge with the space monitoring of environmental change are the limited consistent earth observation data that are available for some regions, particularly in semi-arid sub-Saharan Africa, which makes it difficult to track environmental changes in these regions over time. This study is largely motivated by the lack of data consensus on the usefulness of remotely-sensed satellite data in environmental and ecological monitoring, particularly in data-scarce environments like sub-Saharan Africa. Thus far, the capability of remotely-sensed data for resolving pertinent environmental and ecological questions is

contested by diverse non-remote sensing communities and policy-makers, mainly due to the incompleteness, or the lack of, seamless cloud-free days for the most part of the year. This has thus restricted the use of multispectral remotely-sensed data in the dry season, which hampers the full-cycle understanding of wetlands throughout the year (in both the wet and dry seasons).

Owing to this background, this chapter is aimed at evaluating the number of satellite images available in the GEE catalogue that are capable of monitoring small and seasonally-flooded wetlands located in semi-arid environments of southern Africa over a 20-year period (2000-2020), using the upper middle reach of Nylsvley floodplain as a case study. The specific objectives are: 1) to establish the number of optical scenes available with different cloud cover ranges for various remotely-sensed products, and 2) to determine how slight variations in the cloud cover ranges (0%, 1-10%, 11-25%) affect the monitoring of wetland eco-hydrological dynamics across different seasons.

4.2 Materials and Methods

4.2.1 Ground truth data collection

The collection of ground truth data was conducted from the 28th of September 2020 to the 01st of October 2020 in the upper middle reach of the Nylsvley floodplain located within the boundaries of the Nylsvley nature reserve. The collated data included 300 ground control points, which represented the location of different wetland cover data (vegetation types and bare surface (*open space*)), by using a hand-held Geographical Positioning System (GPS), with an error margin of ± 3.65 m. The field data that were collected coincided with the dry season images used in this study. The wet season data were from the high-resolution Google Earth images (November-April). Three-hundred-and-twenty wet season points, representing the location of the same wetland-cover classes, were also collected. These were guided by the dry season field sampled points. The observed wetland cover data included the short wetland grass vegetation, comprised of *Oryza longistaminata* (rice grass) and *Cynodon dactylon* (Bermuda grass), the long wetland grass vegetation, comprised of mainly *Phragmites australis* (common reeds), and the bare surface. Water could not be identified, as there was no surface inundation visible during that period and it also not identifiable from the wet season Google Earth images, due to the presence of dense wetland vegetation. The field points representing the locations of the above-mentioned landcover classes were collected by using the stratified random-sampling approach. During the collection of these points, the floodplain was sub-divided into 900 m² quadrants, based on the size of the wetland (13.69 km²), as well as the heterogeneity in the

landcover observed in the wetland. A minimum of 20 random points in each quadrant were gathered, depending on the mixing in the identified landcover types. The quadrants were separated 12 m apart because of the size of the section of the Nylsvley under study, in order to reduce the overlaps in the collected ground truth points. The collected field points were then used in the grid-point intercept approach to establish the dominating landcover classes within each pixel, and they were used to label the area covered by the particular pixel. The data collected from field work and the Google Earth images representing both seasons were split into 70% training and 30% validation reference data and they were used to train (70%) and validate (30%) the GEE model.

4.2.2 Remotely-sensed data collection and analysis

The collection and processing of remotely-sensed data was executed on the GEE platform by using the steps presented in Figure 4.1. The initial step involved the manual identification of various remotely-sensed optical products that are available on the GEE platform for mapping small and seasonally-flooded wetlands in the semi-arid environments of southern Africa. The identified products were then filtered by date (01-01-2000 to 31-12-2020) to obtain the images available within the period of interest (2000-2020). Further, the images of all the identified products were clipped, using the Nylsvley floodplain boundary. This was done to obtain images that covered the studied wetland. The images obtained, as per the products identified in the GEE, are presented in Table 4.2. Due to the spatial resolution of Moderate Resolution Imaging Spectroradiometer (MODIS) (250 m for Bands 1-2, 500 m for Bands 3-7 and 1 km for Bands 8-36), these products were excluded from this study, as the study monitors small and seasonally-flooded wetlands, which are poorly detected or completely masked-out by this data type, because of its coarse resolution vis-à-vis wetland size. The images obtained were further filtered by the cloud coverage, based on the following ranges: 0% representing no cloud coverage, 1-10% representing a low cloudy coverage, 11-25% representing a moderate cloud coverage and 26-50% representing a high cloud coverage. This gave an indication of the number of archival images available within each cloud-cover range. The next step was to process the available images, in order to illustrate the capability of detecting the studied wetland under different cloud-cover conditions and in different seasons. In this stage, image classification was only restricted to composites with limited cloud coverage (0%, 1-10% and 11-25%), as the use of images with cloud cover above 25% is not recommended. Prior to the processing stage, dry (May-October) and wet (November-April) season image stacks of Sentinel-2 level 1C and Landsat-7, which represent each cloud cover range, were chosen. The

above-mentioned products were selected because of the large number of scenes available for these products after filtering. Although the study focused on the available images over a 20-year period (2000-2020), the chosen images for the classification, to test the impact of different cloud-cover ranges on the classification outputs, were collected over a 5-year period (2016-2020) for both the dry and wet seasons, because Sentinel-2 products were unavailable before 2016. A total of 71 and 86 dry and wet season scenes were obtained for Landsat-7 and Sentinel-2 Level 1C, representing the 5-year period. During the processing stage of the Landsat-7 and Sentinel-2 scenes, the acquired images were first reduced and normalised for the illumination and clouds effects, using the median composite algorithm, which reduces a stack of images by the computation of median values across the matching bands of a pixel in an image stack, and consequently reduces the cloud cover and illumination effects (Gorelick et al., 2017). Traditional classification approaches require cloud-masking before image-processing, to improve the accuracy of the classification. In this study, it was noted that a significant proportion of the studied wetland area was lost when the clouds were masked out; hence, the selection of median compositing approach to minimise the clouds and to consequently reduce the stack of images. This approach has been successfully adopted in literature (Mahdianpari et al., 2020; Mudereri et al., 2021). The produced median composites were then used to compute indices, such as the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Water Index (NDWI) and the Modified Soil Adjusted Vegetation Index 2 (MSAVI2) (see Table 4.1). NDVI was selected because of its sensitivity to photosynthetically-active biomass (Mahdianpari et al., 2020), and it can therefore differentiate between vegetated and non-vegetated areas, as well as wetland and non-wetland areas (White et al., 2016; Zaitunah et al., 2018). mNDWI was chosen because of its sensitivity to open water and it can differentiate between open water and land surface (Chen et al., 2014). MSAVI-2 was selected to improve the NDVI limitations. The indices were then added as extra bands to the selected near-infrared and visible bands (red, green and blue), to enhance the wetland features.

The images with extra NDVI, mNDWI and MSAVI2 bands were then used in an Object-Based Image Analysis (OBIA). OBIA was selected because of its superiority over pixel-based image classification in several wetland studies, such as those of Mahdianpari et al., (2019) and Pande-Chhetri et al., (2017), amongst others. The initial step to object-based classification is image segmentation, which partitions the image into multiple objects that are discreet and non-overlapping, by using a specified criterion, therefore increasing the separability of spectrally-similar classes and minimising the ‘salt and pepper’ effects on the output image (Mahdianpari

et al., 2019). Image segmentation was implemented in this study by using a Simple Non-Iterative Clustering (SNIC) algorithm. The SNIC algorithm was selected because it is simple, memory-efficient and it can integrate the connection between pixels after it has been implemented. The SNIC initiates the image segmentation by the initialisation of centroids in a regular gridded image, and then the dependency of each pixel relative to the centroids is determined by using the calculated distance in five dimensions of colour and the spatial coordination of the pentagon, by forming super pixels. The calculated distance joins the normalised spatial and colour distances to compute the unvarying super-pixels. The next candidate pixel to be added to a cluster is determined by its shortest distance to the centroids (Gorelick et al., 2017). The output from the SNIC were super-pixels, which were used in computing their contextual parameters, such as the area, standard deviation (texture), height and perimeters, amongst others. The contextual parameters, super pixels, optical features and indices were then concatenated to produce an image that was later used as an input during the classification process.

Table 4.1 Indices extracted from the time series Landsat-7 and Sentinel-2 data

Data	Data extracted	Formula	Reference
	NDVI	$\frac{NIR - Red}{NIR + Red}$	Tucker, (1979)
Landsat-7, Sentinel-2	NDWI	$\frac{NIR - Green}{NIR + Green}$	McFeeters, (1996)
	NDPI	$\frac{NIR - (0.74 \times Red + 0.26 \times SWIR1)}{NIR + (0.74 \times Red + 0.26 \times SWIR1)}$	Wang et al. (2017)

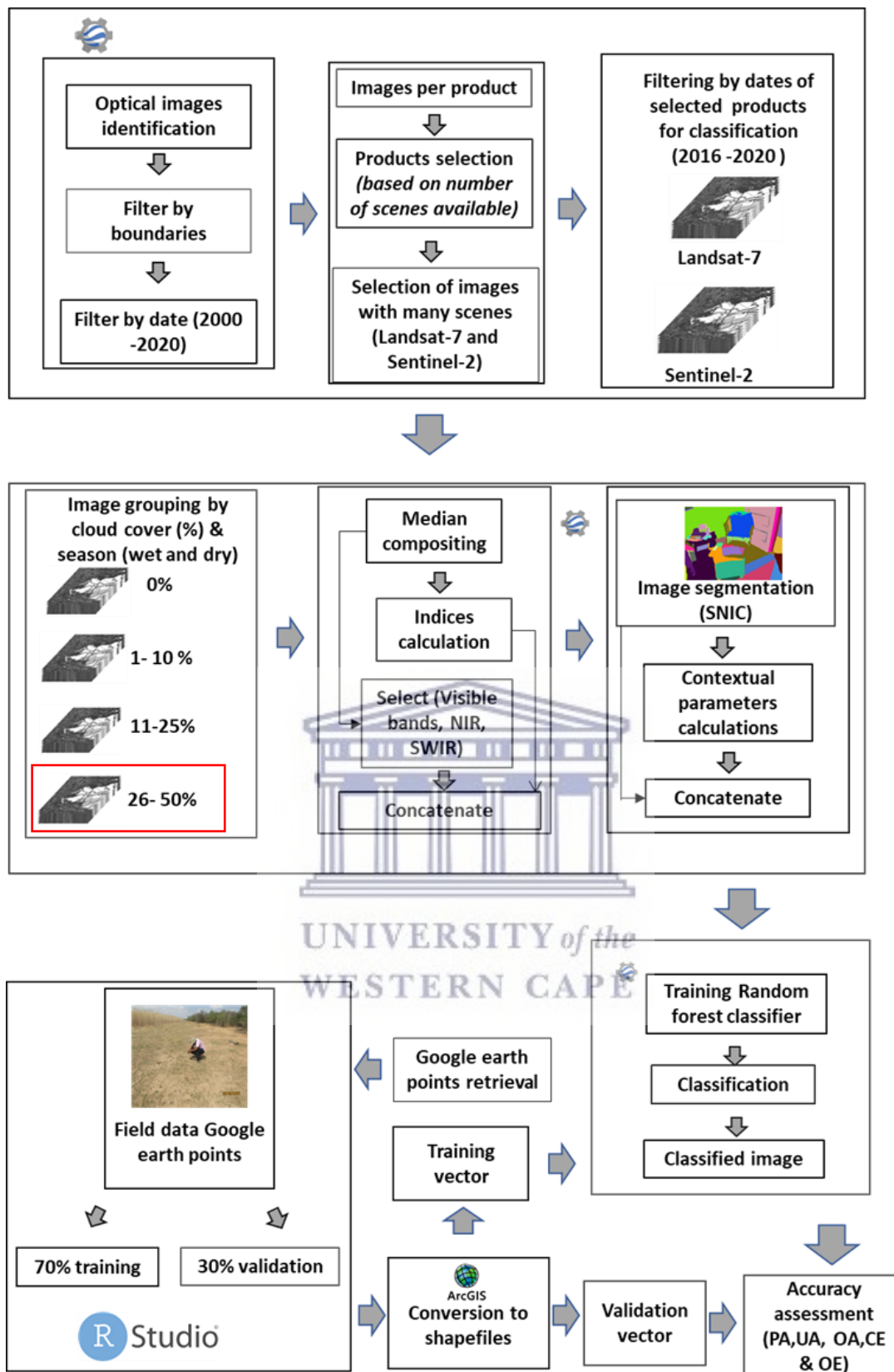


Figure 4.1 Steps taken to execute the classifications using GEE (Red box indicates the image stack not included in the classification)

4.2.3 Image classification

The classification of the images in this study was implemented by using the Random Forest algorithm on the GEE platform. The algorithm is an ensemble classifier consisting of many trees. Each tree casts a unit vote to split the samples. The algorithm was chosen, firstly, because it is capable of handling large differentiations between the landcover classes, thus neutralising the data noise, and secondly, because of its superiority to other GEE algorithms in studies e.g. Simioni et al., (2020). Moreover, the selection of the RF was informed by a study by Gxokwe et al., (2021), which assessed the performance of different GEE machine-learning algorithms in the mapping of semi-arid wetlands including the studied. In that research, RF was among the best-performing algorithms, with a high overall and class accuracies; therefore, it was selected for this study. Prior to classification, the collected field data were randomly split into 70% training and 30% validation data on the R-Studio. The principle behind the splitting of data to 70/30 was to ensure that they represented a large training dataset, while the remaining data were preserved, in order to compute accurate statistics. After splitting, the training and validation data points were then imported and converted to shapefiles on Esri ArcGIS, and then imported to the GEE platform, in order to train and validate the Random Forest model.

4.2.4 Accuracy assessment

The derived classification maps were assessed for their accuracy. This was achieved by using 30% of the data (300 points for the dry season and 350 for the wet season) for each season, which was randomly split in R-studio. The 30% of the 300 field-collected points were used to validate the dry season classification outputs, and the 30% of the 350 points derived from the high-resolution Google Earth images were used to validate the wet season classification outputs. Prior to the accuracy assessment, the 30% points were first imported to ESRI 10.2 ArcGIS, and then converted to shapefiles (validation vectors). The validation vector files were exported from ArcGIS and imported to the GEE platform and then used to sample out the corresponding regions to the location of those points on the classified images. The sampled regions, coupled with the validation vectors, were then used to compute the producer's and user's accuracies. The error matrix was also extracted by using the sampled regions and validation vectors, and then used to provide a comprehensive evaluation of the commission and omission errors amongst the classification results and training data, including evidence of how the classification errors had occurred.

4.3 Results

4.3.1 Image availability

Table 4.2 shows the total number of remotely-sensed images filtered in the GEE catalogue for the Nylsvley floodplain for various optical and active sensors, between the period of 2000 and 2020. The results indicated that the Sentinel-1 GRD C-Band had the most (394) images available, when compared to the other optical data. For the Landsat 5 thematic mapper, only a few images (44) were recorded. Moreover, the results also reveal that there were no Landsat 4 images available for the specified time-frame. Filtering by the cloud-cover percentage indicated that most optical products had a higher total number of images with a clear cloud coverage (0%), except for Sentinel-2 level 2A, which had no cloud-free images available for the specified time-frame. Landsat-7 and Sentinel-2 level 1C had the most images, with no cloud coverage. Moreover, the study noted that as the cloud cover percentage increased during the filtering process, the total number of images decreased significantly, especially for the 26-50% range products, such as Landsat-8 and Landsat-5. However, satellite image scenes with a cloud-cover percentage above 25% were excluded from the analysis.

Table 4.2 Satellite images available for the Nylsvley floodplain between 2000 and 2020 per cloud-cover percentage

Remote sensing product	# of images	0%	1- 10%	11-25%	26- 50%	>50%
Landsat 4	0	0	0	0	0	0
Landsat 5	44	13	13	4	3	11
Landsat 7	284	109*	77*	31*	37	30
Landsat 8	164	11	88	18	4	43
Sentinel-2 level 2A	148	0	82	14	11	41
Sentinel-2 MSI level 1C	342	109*	103*	29*	27	26
Sentinel-1 GRD C-Band (SAR)	394	-	-	-	-	-

* *this denotes images that were used for further processing*

4.3.2 Classification results

The classification results, based on dry and wet season Sentinel-2 Level 1C and Landsat-7 composites with variable cloud coverage, are presented in Figures 4.2 and 4.3. The results show the spatial distribution of three wetland-cover classes, namely, short grass, long grass and the bare surface. Short grass consists of mainly *Oryza longistaminata* (rice grass) and *Cynodon dactylon* (Bermuda grass), and long-grass consists of mainly *Phragmites australis* (common reeds). The water class could not be detected because of the unavailable training and validation data resulting from the absence of surface inundation during the field visits. Moreover, the dense wetland vegetation resulting in difficulties in detecting water from the high-resolution Google Earth image. Although water could not be detected, the presence of vegetation, such as *Phragmites australis*, indicated that the soils were semi-saturated. Based on the Sentinel-2 dry season composites, the most dominating class was short grass, except for the composite with 0% cloud cover, where long grass was the most dominating and short grass the least dominating. The least dominating class for the other composites was long grass. Based on the wet season Sentinel-2 composites, the most dominating class was short grass for all the composites, except the composite with 0% cloud coverage, where the dominating class was long grass. The least dominating class was the bare surface for all the wet season composites. The results based on the Landsat-7 composites showed that the most dominating class is short grass for all the dry season composites, except for the composite with 11-25% cloud coverage, where long grass dominates the most. The least dominating class varied, based on the dry season composites. The results based on the dry season composites showed that for composites with a cloud cover of 1-10%, long grass was the least dominating class. In addition, composites with a cloud coverage of 0% and 11-25% showed that short grass was the least dominating class. The wet season Landsat-7 composites showed that the most dominating class was long grass for data with a 0% cloud cover and 1-10% cloud coverage. Satellite image composites with 11-25% cloud cover indicate that the bare surface is the most dominating class. Furthermore, the short grass was identified as the dominating class in the composite with a cloud coverage 0% and 11-25%. The least dominating class for the wet season varied, depending on the composite image that was used. The bare surface was reported to be the least dominating class for a 0% cloud-cover composite. For the composite satellite data with a cloud-cover range of 1-10%, short grass was the least dominating class reported. The least dominating class for satellite composite data with a cloud coverage of 11-25% was long grass and the bare surface.

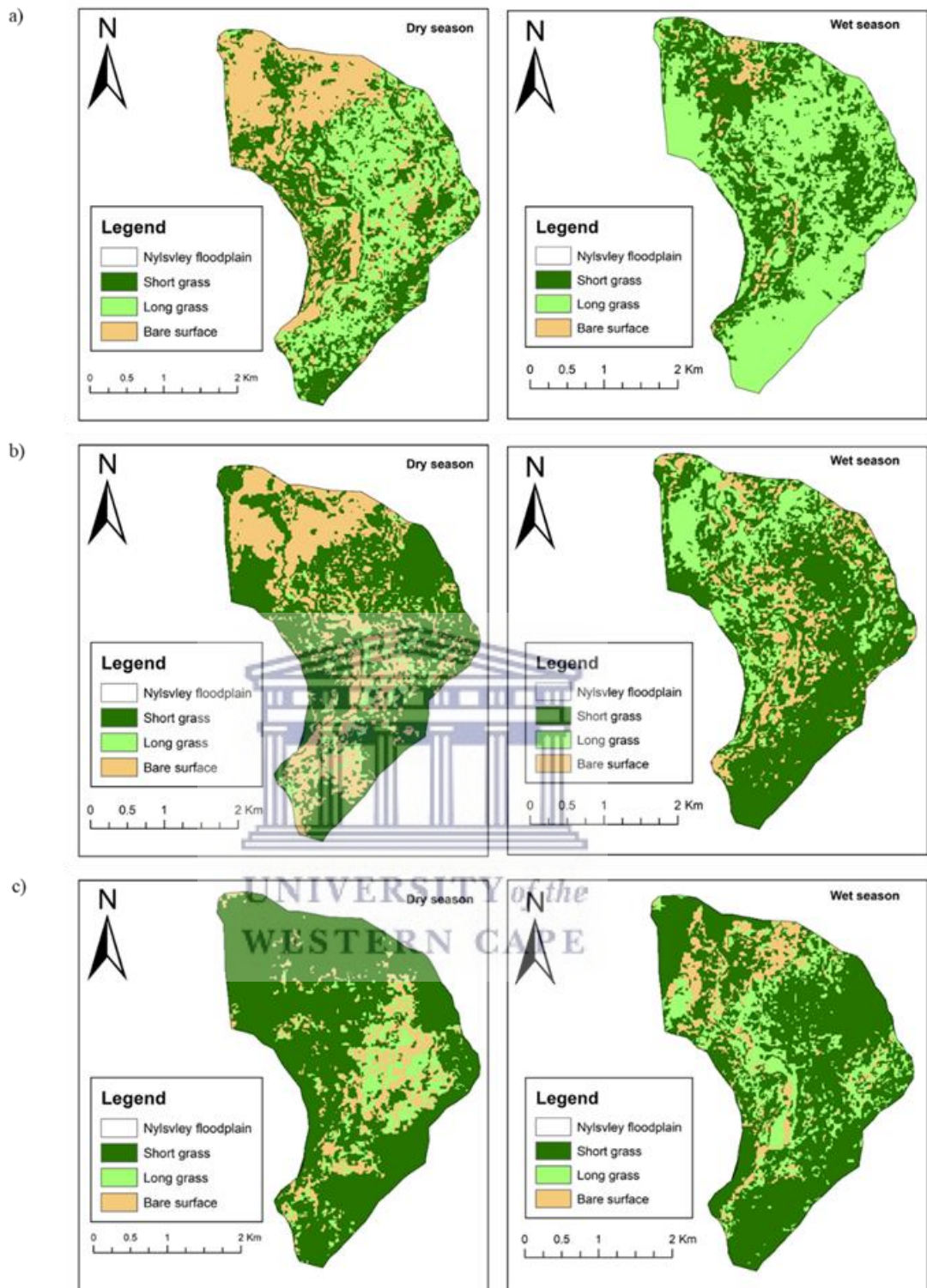


Figure 4.2 Landcover classes for the Nylsvley floodplain in different seasons and cloud cover between 2016 and 2017, based on Sentinel-2 Level 1C, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover

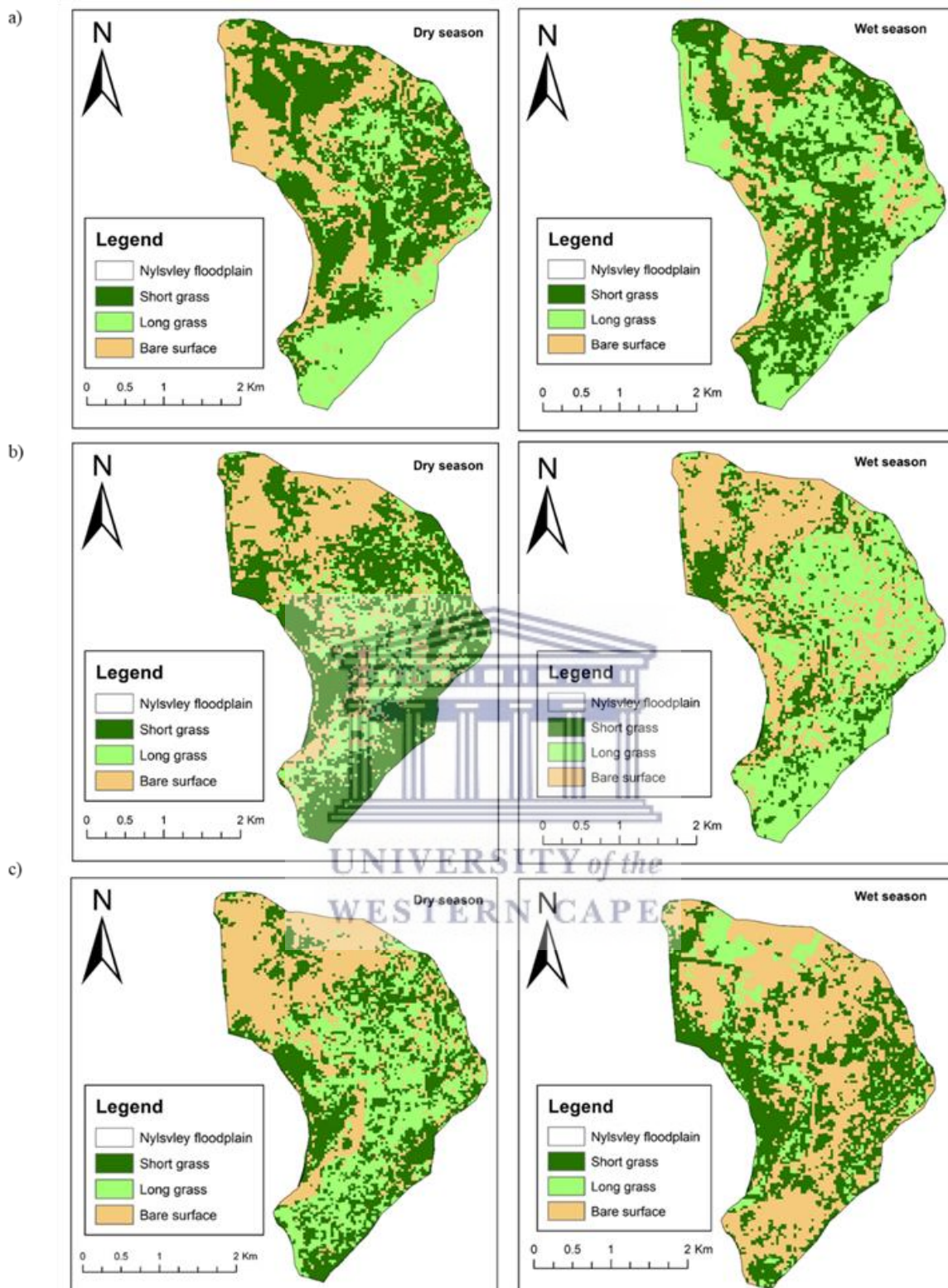


Figure 4.3 Landcover classes for the Nylsvley floodplain at different seasons and cloud cover between 2016 and 2017, based on Landsat 7, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover

4.3.3 Accuracy assessment results

Figure 4.4 shows the overall accuracy results for the classifications based on a) the Sentinel-2 level 1C data and b) the Landsat-7 data, for both the wet and dry seasons between 2016 and 2017. The results show overall accuracies that are within an acceptable range. Based on the Sentinel-2 data, the results show that the overall accuracies ranged between 72% and 80.55% for the wet season and between 77% and 80.55% for the dry season. The highest overall accuracy, based on Sentinel-2 optical data type, was observed for the dry season composite, with a cloud-cover range of 11-25%, and the lowest was observed for the wet season composite, with a cloud cover range of 11-25%. The Landsat-7 composite data produced overall accuracies ranging between 72% and 77% for both the wet and dry seasons, respectively. The highest overall accuracies were observed for both the dry and wet season composites, with a cloud-cover range of 11-25%, and the lowest was observed for the wet season composite, with a 1-10% cloud cover.

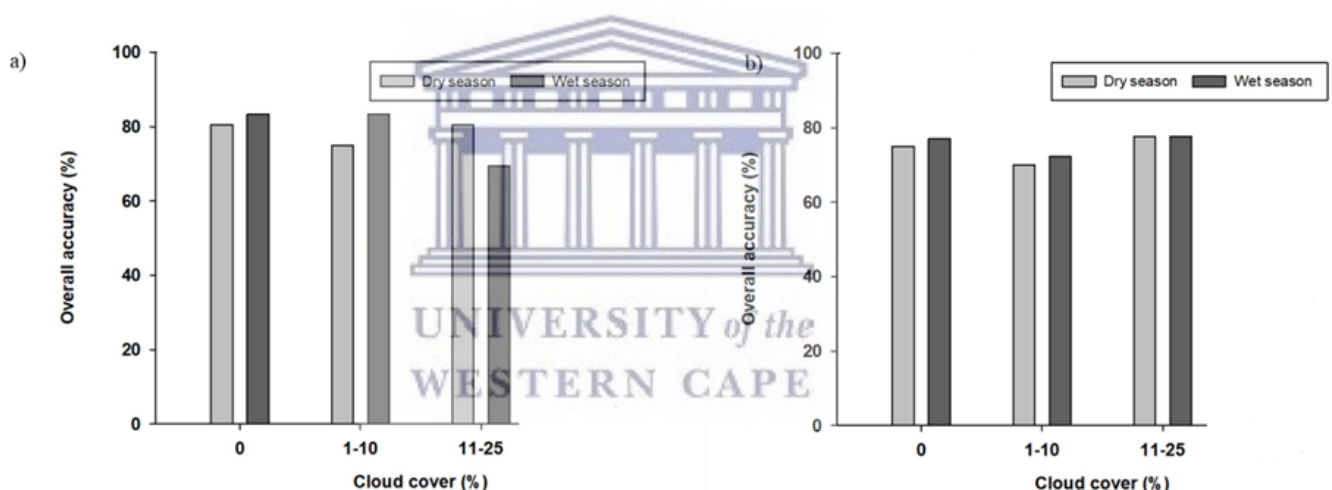


Figure 4.4 Overall accuracies based on a) Sentinel-2 level 1C and b) Landsat 7 data

The PA and UA for the three mapped wetland-cover classes ranged between 33.3% to 92% and 33.3% to 100% for the wet season, and between 33.3% to 100% and 25% to 100% for dry season, based on the Sentinel-2 composites (Tables 4.3 and 4.4). Moreover, the results showed that the long grass landcover class had the lowest PA and UA for both the dry and wet season composites, and that the short grass class had the highest PA and UA, in most cases. A similar case was observed for the PA of this class, for the dry season composite with 1-10% cloud cover range, where this accuracy improved significantly to 100%. Based on the Landsat composites, the PA and UA ranged between 33.3% and 100%, and between 40% and 84.61%

for the dry season composites. Long grass had the lowest PA and UA in most cases, while short grass had the highest PA and UA. The wet season composites reported a PA and UA ranging between 37.3% and 92%, and 33.3% and 100%. In most cases, the long grass had the lowest PA and UA, except for the composite data with an 11-25% cloud cover, which reported a PA of 100%. Moreover, the composite with a 1-10% cloud coverage reported a UA of 100% for long-grass, while short grass had the highest PA and UA, in most cases.

Table 4.3 Producer’s and user’s accuracies per cloud cover range for dry season scenes

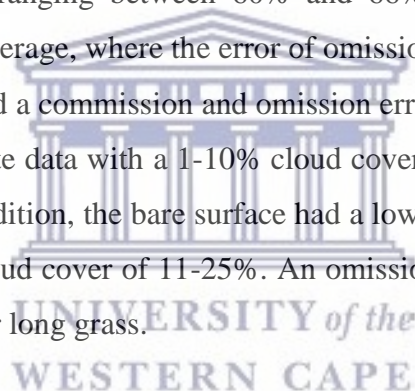
Remote sensing product	Landcover type	PA (0%)	PA (1-10%)	PA (11-25%)	UA (0%)	UA (1-10%)	UA (11-25%)
Sentinel-2 level 1C	Short grass	95	96	84	88.46	88.8	75
	Long grass	66.6	42.3	33.3	66.6	33.3	33.3
	Bare surface	62.5	62.5	37.5	71.43	62.5	60
Landsat-7	Short grass	84	92	84	84	82.14	91
	Long grass	33.3	33.3	100	33.3	100	37.5
	Bare surface	50	50	50	50	57.14	80

Table 4.4 Producer’s and user’s accuracies per cloud cover range for wet season scenes

Remote sensing product	Landcover type	PA (0%)	PA (1-10%)	PA (11-25%)	UA (0%)	UA (1-10%)	UA (11-25%)
Sentinel-2 level 1C	Short grass	84	84	92	91.3	87.5	85.19
	Long grass	66.6	100	33.3	66.6	42.86	100
	Bare surface	75	35.7	62.5	60	60	62.5
Landsat-7	Short grass	88	92	92	84.61	76	85.18
	Long grass	66.6	33.3	66.6	40	50	50
	Bare surface	37.5	37.5	37.5	60	75	60

Pontius’ commission and omission error results, based on Sentinel-2 level 1C composite data for both dry and wet seasons, are presented in Figure 4.5. The results indicate that, for both seasons and for the composites with different cloud cover ranges, short grass had the lowest commission and omission errors, when compared to the bare surface and long grass, which had

high commission and omission errors, in most cases. More specifically, the long grass had higher commission and omission errors. Moreover, a higher error of omission was observed for this class when the dry season composite with an 11-25% cloud coverage was used, where the reported omission error was 66.6%. The bare surface class had a high error of omission, when the dry season composite, with a cloud coverage ranging between 1-10%, was used. The wet season Sentinel-2 composite data also showed that the long grass had high commission and omission errors in all cases, with the recorded errors being above 50%. The results also showed that the bare surface reported a high omission error for the wet season composite data with a cloud coverage of 11-25%. The error of commission and omission, based on Landsat-7 dry and wet season composites, are presented in Figure 4.6. The results show that short grass had low commission and omission errors, for all cases, with reported errors below 20% for both dry and wet season composites. The results also showed that the bare surface and long grass had high commission and omission errors, in most cases, with errors exceeding 50%. More specifically, long grass had the highest commission and omission errors for the dry season Landsat-7 composite, ranging between 60% and 66%, except for the dry season composite with a 0% cloud coverage, where the error of omission was 33.3%. Moreover, bare surface and long grass class had a commission and omission errors above 50%, in most cases, except for the case of composite data with a 1-10% cloud cover, where the commission error for both was below 45%. In addition, the bare surface had a low commission error for the wet season composite data with cloud cover of 11-25%. An omission error of 0% was also noted for the same composite data for long grass.



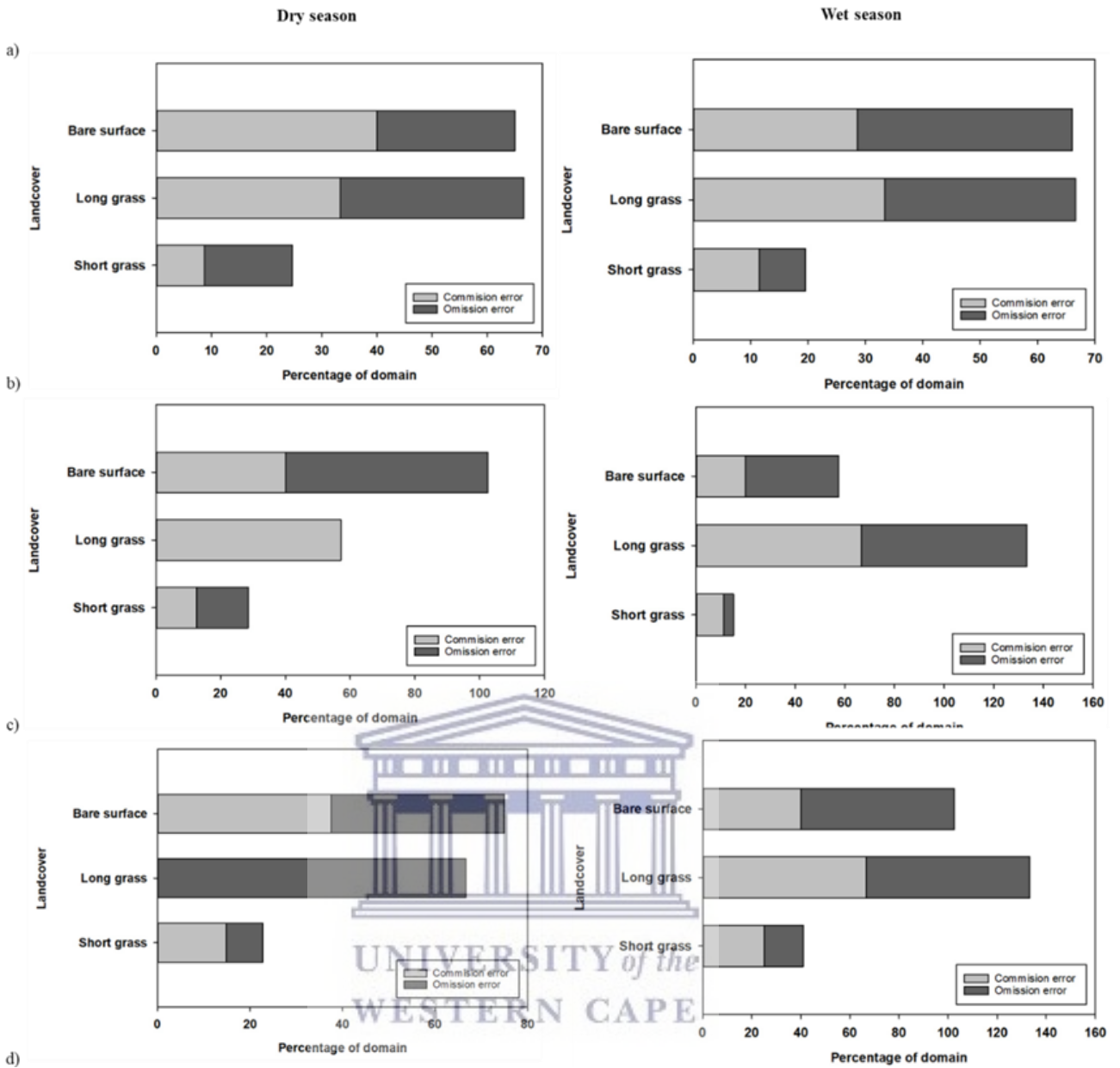


Figure 4.5 Pontius commission and omission errors for different seasons and cloud cover between 2016 and 2017, based on Sentinel-2 level 1C, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover

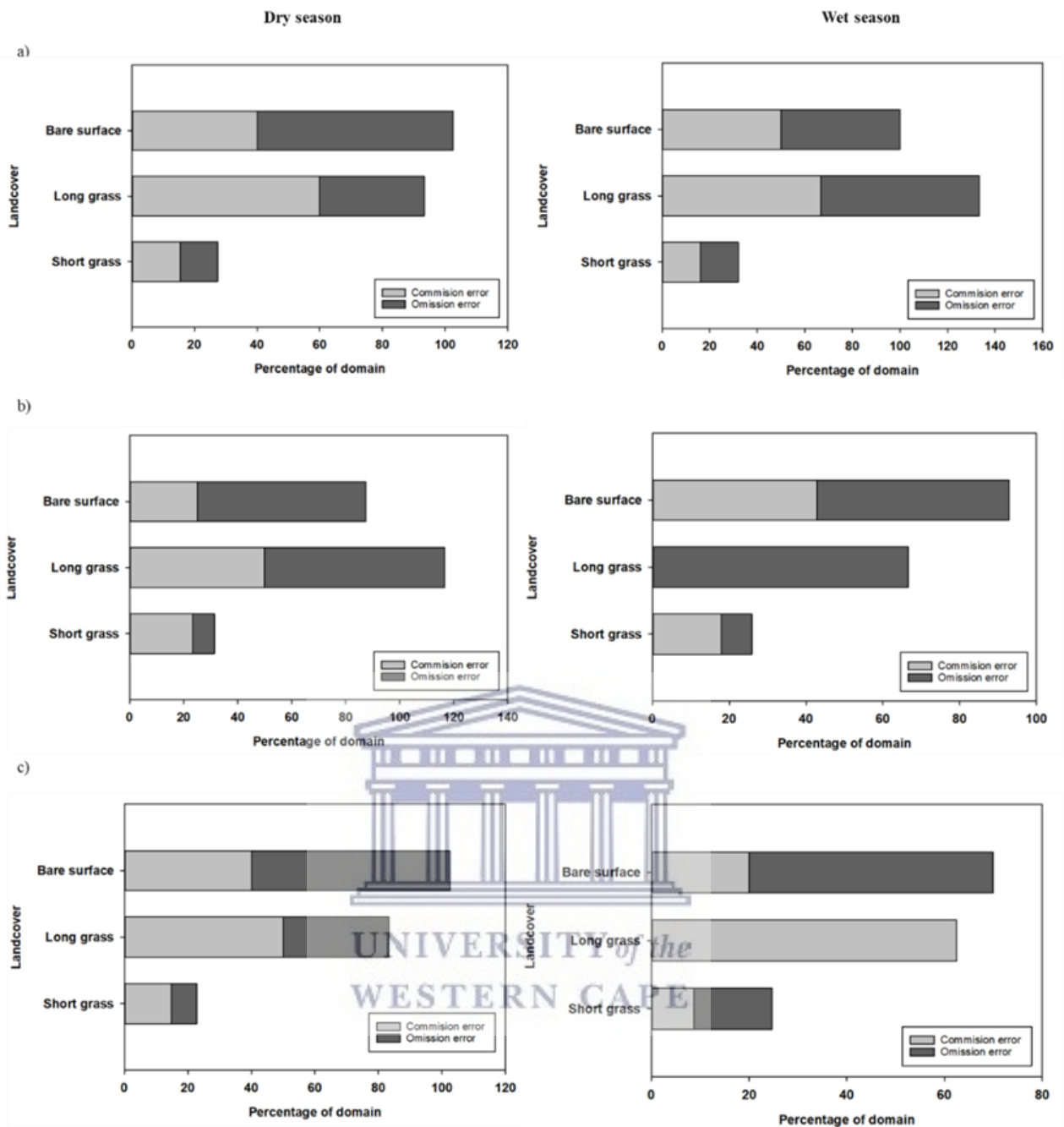


Figure 4.6 Pontius commission and omission errors for different seasons and cloud cover between 2016 and 2017, based on Landsat 7, a) 0% cloud cover, b) 1-10% cloud cover, and c) 11-25% cloud cover

4.4 Discussion

This study sought to determine the number of archival remotely-sensed optical images available on the GEE cloud-computing platform that are capable of monitoring small and seasonally-flooded wetlands in the semi-arid environments of southern Africa over a 20-year period (2000–2020). The results showed that various optical remote-sensing products are available on the GEE platform. These images range from Landsat 4-5-7-8, Sentinel-1-2 and MODIS. The results have shown that Landsat-7 data has the highest number of images available, followed by Sentinel-2 level 1C. This can be attributed to the fact that Landsat-7 was launched in 1999 and has been acquiring spatial data since then. Although data from this sensor became freely-available after 2008, the GEE catalogue hosts a range of Landsat-7 data, dating from 1999-01-01 to the present. Moreover, Sentinel-2 level 1C and Sentinel-1 C-Band Synthetic Aperture Radar (SAR) Ground Range Detected (GRD) were also reported to have a high number of images available in the GEE catalogue. This is mainly due to the high temporal resolution of both Sentinel-1 and -2. The temporal resolution for Sentinel-1 and -2 is 6 and 10 days, with Sentinel-2 having a constellation revisit time of five days, thus resulting in more images being collected by these optical sensors (Thamaga and Dube, 2018). The collection of these products in the GEE catalogue date back to the year 2013. Although, there are large quantities of Sentinel-1 and Sentinel-2 scenes in the GEE catalogue, these data do not provide continuous and timely long-term monitoring, since it has only been available from 2014 to the current date (Schott et al., 2012). This is a problem, because the long-term back-tracking of the wetland changes over longer terms (20 years) is not possible with such data, as it only became available after 2013. The filtering by date results also showed that there were no images available for Landsat 4 between the year 2000 and 2020. This can be attributed to the life-span of this sensor. Landsat 4 was launched in 1982 and has been in operation since then, until 1993; therefore, the GEE catalogue hosts Landsat-4 images corresponding to the life-span of the sensor (Schott et al., 2012; Thamaga and Dube, 2018). The results also suggest that the Landsat 5 product had a low image count, which can be attributed to the life-span of the sensor, as well as the temporal resolution of the data. Landsat has been in operation since it was launched in 1984, until 2012, and it has been collecting data on a repeated 16-day cycle (Markham and Helder, 2012). These results reveal that the long-term monitoring of wetlands is feasible when using Landsat-7 products, as it offers continuous data, from before 2000 to the present. However, the issues relating to the spatial resolution of Landsat-7 data persist, especially when monitoring the small and seasonally-flooded wetlands of the semi-arid environments, as the

strips in some of the images, which were caused by the failing of Scan Line Corrector, resulted in the loss of about 22% of the Landsat-7 data. There are, therefore, trade-offs between the long-term monitoring of wetlands and the accurate analysis of their characteristics.

Among the other factors influencing the long-term monitoring of wetlands by using remote sensing products, is the availability of good remotely-sensed data for minimum cloud coverage. Satellite data filtering by the cloud-cover percentage showed that there are a high number of remotely-sensed images with a clear cloud coverage (0%) for most products. This was expected for semi-arid regions because these regions are characterised mostly by cirrus clouds, which are scattered in different zones and which result in most regions being clear most of the time (Schott et al., 2012). Moreover, the times at which most images are collected may result in the collection of more cloud-free scenes.

The classification results indicated a high overall accuracy for all remotely-sensed products, seasons and cloud-cover percentages. Although this is the case, a low PA and UA were observed for some classes, such as long grass and bare surface, particularly with the Landsat 7 composite data. This may be due to the limited number of training and validation points for these classes, resulting from low spatial coverage of the classes in the floodplain. Various studies, e.g. Corcoran et al., (2013), Jing et al., (2010), Mahdianpari et al., (2019) and Schott et al., (2012), reported that limited training and validation data may influence the level of the class accuracy during classification and validation. Moreover, the spatial resolution of the data that are used may have contributed to the low PA and UA for the classes, particularly for Landsat 7, which has a spatial resolution of 30 m. A report by the FGDC (Federal Geographic Data Committee (FGDC) (1992) indicated that the identification of an object on a Landsat image requires at least 9 pixels, which is a total area of about 0.9 ha and, at most, 25 pixels. This is also evident in a study by Dvoretz et al., (2016), who reported that 30 m resolution Landsat 7 images were unsuccessful in detecting many small wetlands in the sand-dune ecoregion of Oklahoma, which were previously mapped manually on the National Wetland Inventory map for the region. However, Thakur et al., (2012) showed the effectiveness of Landsat 7 data in mapping the semi-arid Macquarie Marshes in Australia; it had a higher producer's and user's accuracies, which revealed that additional factors, such as the wetland size, need to be carefully considered before selecting Landsat data to study such systems. The high commission and omission errors reported in this study further confirm the underestimation of the bare and long grass classes for both products, especially when using wet season composites with a cloud coverage of 11-25%. Clouds are known to reduce the quality

of a remotely-sensed image, therefore various authors, such as Li et al., (2007; 2020) suggested that an image should not exceed 10% of the cloud coverage and that a single data image will certainly be affected by a high proportion of haze and clouds, thus degrading the quality of the classification outputs. The study further assessed the usability of various images with varying cloud coverage. The results showed that there were quite a lot of classifications errors for some classes, such as for a bare surface and long grass, when composite satellite data with a cloud cover of 11-25% for both the Sentinel-2 and Landsat-7 scenes. This was expected, although median compositing was implemented to minimise the illumination and cloud effects, as it does not mask out the clouds (Gorelick et al., 2017), which therefore became a limitation in this study. These results corroborate the findings of Shen et al., (2018), who reported the misclassifications of some wetlands classes due to the use of poor-quality images (>20% cloud cover).

The findings of the study prove that there are a large number of good quality (0-20% cloud cover) optical remote sensing scenes available in the GEE catalogue for monitoring small and seasonally-flooded wetlands in semi-arid environments. Moreover, the findings underscore the relevance of Landsat-7 data in the long-term monitoring of small and seasonally-flooded wetlands in semi-arid environments. However, some class inaccuracies were associated with the limited training and validation data, as well as the spatial resolution of the Landsat-7 data, thus necessitating the careful consideration of the training data and the classification algorithm, to achieve the optimum results. The study provides new insights into the available and useable data accessed via the GEE catalogue for monitoring small and seasonally-flooded wetlands in semi-arid areas, which was previously unknown for such regions. This is important for semi-arid regions with limited data, such as sub-Saharan Africa, where limited attention is given to these wetlands, due to the lack of available information, even though these systems serve as important water sources and provide other eco-hydrological benefits for the surrounding communities. In particular, the results indicate the possibilities of long-term monitoring and assessment of small and seasonally-flooded wetlands in semi-arid regions by using optical data, specifically Landsat on the GEE cloud-computing platform, as it offers consistent and continuous data, which makes it easy to track changes over the longer term. Moreover, the availability of more images with a cloud cover of less than (25%) indicate that the available data are of good quality, and that the monitoring and assessment of these systems will therefore have an acceptable accuracy, even though certain issues associated with spatial resolution may influence their accuracy. The results also show that careful attention needs to be given to the

seasonal differences, when selecting images, as some wetland features tend to be more separable in some seasons than in others (i.e. water is more visible during the wet season than when there is a bare surface).

Overall, the results show that the remote sensing of small and seasonally-flooded wetlands, using time-series optical data that are available on GEE platform, is the most feasible approach for generating information on the spatio-temporal changes of these systems over time, in semi-arid regions. Thus, the issues associated with limited data can be addressed, as well as the knowledge gaps about their eco-hydrological dynamics. The results also show that images with a cloud cover of less than 25% can assist in monitoring the status of small and seasonally-flooded wetlands, in the absence of the high-resolution data. Furthermore, the results highlight the relevance of remote sensing data in the monitoring of wetlands; there are currently conflicting arguments and debates with conflicting argument currently because of the spatial resolution of these data, as well as the availability of good quality images, in terms of their cloud coverage. This study therefore improves the confidence and trust of non-remote communities and policy makers, regarding the usefulness of remote sensing in natural resources management. Moreover, the study contributes to the current global wetland research programme, such as the Wetland Monitoring and Assessment Services for Transboundary Basins in southern Africa (WeMAST), which is funded by European Union-Africa Global Monitoring for Environmental Security (EU Africa GMES), and which focuses on establishing integrated wetland inventorying platforms that will inform the sustainable management and utilisation of wetlands in transboundary basins in southern Africa. Moreover, the study informs the Sustainable Development Goal (SDG) 6.6, the objective of which is to halt the degradation and destruction of all ecosystems, including wetlands, and to rehabilitate those that are already degraded. In addition, SDG 6.4 emphasises the need to monitor the availability of freshwater, the environmental requirements, as well as water use, with the aim of properly analysing the water-scarcity issue and the impact that it has on both the economy and the population. It also stresses the need to consider the health of ecosystems. This study achieves the contributions of the latter SDGs by providing knowledge about freely-available and usable data types that are accessible for monitoring the long-term changes of small seasonally-flooded wetlands in semi-arid regions. It will therefore assist in the consistent tracking of any negative changes in these systems over time, and it will address the issues that are associated with their degradation and their protection.

4.5 Conclusion and Recommendations

This chapter was aimed at evaluating the number of archival optical images available in the GEE catalogue that are capable of monitoring small and seasonally-flooded wetlands in the semi-arid environments of South Africa over a 20-year period (2000–2020), using the Nylsvley floodplain as a case study. The findings showed that a significant number of optical products are available on the platform and that the majority of these are of good quality (0-20% cloud cover). Moreover, the findings underscore the relevance of Landsat-7 data scenes with a cloud cover of 0%, 1-10% and 11-25% for the long-term monitoring of small and seasonally-flooded wetlands in semi-arid environments, although there were some class inaccuracies, particularly in long grass and bare surfaces and with high cloud coverage (11-25%). These can be attributed to limited training and validation data, the spatial resolution of the data used particularly Landsat, as well as the effect of cloud cover from some of the retrieved scenes. Although the Landsat series data provide a timely monitoring framework for studying the long-term changes in such wetlands, the study recommends the use of multisource data for monitoring these changes in semi-arid wetlands, in order to address the inherent spatial gaps. Moreover, the study recommends that the careful selection of images with a cloud cover of less than 25% should be considered, as it was noted that cloud cover contributed to the misclassification of some image scenes.



CHAPTER 5

USING CLOUD-COMPUTING TECHNIQUES TO MONITOR LONG-TERM VARIATIONS IN ECO-HYDROLOGICAL DYNAMICS OF THE SMALL SEASONALLY-FLOODED WETLANDS IN SEMI-ARID SOUTH AFRICA



Photo: Courtesy of Tatenda Dzurume 2020

This chapter is based on the following publication:

Gxokwe, S., Dube T., Mazvimavi, D., Grenfell, M.C. 2022. Using cloud-computing techniques to monitor long-term variations in eco-hydrological dynamics of the small seasonally-flooded wetlands in semi-arid South Africa. *Journal of Hydrology*. **612** (2022): DOI: <https://doi.org/10.1016/j.jhydrol.2022.128080> (IF = 6.708)

Abstract

Wetlands in drylands have high inter- and intra-annual eco-hydrological variations that are driven, to a great extent, by climate variability and anthropogenic influences. The Ramsar Convention on Wetlands encourages the development of frameworks for national action and international cooperation with regard to the conservation and wise use of wetlands and their resources, on a local, national and regional scale. However, the implementation of these frameworks remains a challenge. This is mainly due to the limited high-resolution data and suitable big data-processing techniques for assessing and monitoring the eco-hydrological dynamics of wetlands on a large spatial scale, particularly in the sub-Saharan African region. The availability of cloud-computing platforms, such as Google Earth Engine (GEE), offers unique big data handling and processing opportunities for these challenges to be addressed. In this study, the GEE cloud-computing platform was applied to monitor the long-term eco-hydrological dynamics of a seasonally-flooded part of the Nylsvley floodplain wetland complex in north-eastern South Africa, over a 20-year period (2000–2020). The specific objectives of the study were: 1) to evaluate the wetland eco-hydrological dynamics by using 20-year multi-date Landsat composite data, coupled with the Random Forest machine-learning algorithm; and 2) to establish the major drivers of wetland eco-hydrological changes by using selected spectral indices (i.e. the Normalised Difference Vegetation Index: NDVI, the Normalised Difference Water Index: NDWI and the Normalised Difference Phenology Index: NDPI), coupled with the climate data. The ecohydrology of the wetland changed over time, with some classes increasing twice as much as the previous measurements, while others decreased significantly during the study period. Notably, the bare surface class increased at a rate of 230% and 350% between 2006-2010 and 2016-2020, respectively. Moreover, the indices showed a similar trend throughout the 20-year period, with NDWI having the lowest values, namely, less than zero in all cases. This implied that there was no surface inundation, although the presence of some wetland vegetation indicates that there were seasonal to semi-permanent soil-saturation conditions. A comparative analysis of the climate data and remotely-sensed indices showed that the annual changes of precipitation and evapotranspiration were the main drivers of the eco-hydrological variations in wetlands. The findings of the study underscore the relevance of cloud-computing artificial intelligence techniques, and particularly the GEE platform, for evaluating the eco-hydrological dynamics of wetland systems in semi-arid southern Africa, which are deteriorating due to their unsustainable use and poor

management, as a result of the limited amount of available knowledge on these changes over time.

Keywords: Artificial intelligence; Dryland wetland; Ephemeral wetland; Machine-learning algorithm; Wetland condition; Wetland management

5.1 Introduction

Semi-arid areas commonly host small intermittently-flooded wetlands, which are highly sensitive to climate variability and anthropogenic influences (The Wetlands in Dry Lands Research Network, 2014). Anthropogenic influences and climate variability are the major causes of the high inter- and intra-annual eco-hydrological variability in these systems. In many cases, the disappearance of surface inundation during the dry season makes such wetlands susceptible to habitat destruction and loss (Blanckenberg et al., 2020). Semi-arid wetlands have been reported to be declining at an alarming rate globally, mostly due to the erratic rainfall, climate change and their high utilisation by the surrounding communities (Thamaga et al., 2021). Reports in Ethiopia, which is predominantly arid, indicate that several seasonal wetlands have been lost (Bahilu and Tadesse, 2017). This has been attributed to the unregulated and over-utilization of these systems, which have had a negative impact on the condition of the wetlands. It has been reported that over 30% of wetlands in China have been lost over the past 50 years, due to anthropogenic activities (Liu et al., 2017). The estimates of wetland loss in South Africa range from 50 to 58%, but the full extent of the transformation is difficult to quantify over large spatial scales necessary to inform national policy responses (van Deventer et al., 2019). There is therefore an urgent need for interventions, in order to halt the further degradation of these important ecosystems.

In order to sustain these systems, programs such as the South Africa's National Biodiversity Assessment (van Deventer et al., 2019) and Working for Wetlands (Dini and Bahadur, 2016) have been devoted to promoting the conservation, wise use and restoration of all wetlands on a local, national and regional scale. There are also several regional and international efforts that are aimed at reversing the degradation of wetlands, through the dissemination of information and the involvement of local communities in establishing proper management plans (Rebelo, 2010). These efforts are, however, hindered by a lack of data on the eco-hydrological dynamics of wetlands, on an appropriate spatial and temporal scale, particularly in the sub-Saharan African region (Gxokwe et al., 2020; Thamaga et al., 2021).

Projects such as the current Wetland Monitoring and Assessment Services for Transboundary Basins in Southern Africa (WeMAST), which is funded by the European Union–Africa Global Monitoring for Environmental Security (EU Africa GMES), have been initiated to ensure the effectiveness of monitoring and assessment of wetlands by using earth observation data. Such data provide opportunities for producing detailed wetland inventories and understanding wetland dynamics on a scale that is relevant for the management of wetlands of varying sizes (van Deventer et al., 2019). Moreover, the availability of cloud-computing platforms, such as the Google Earth Engine (GEE), supports the efforts of regional wetland analyses by using earth observation data. The availability of over 40 years of petabytes of remotely-sensed and freely-accessible data, coupled with the advanced machine-learning algorithms on the GEE platform, makes it possible to consistently monitor the small to large wetlands and track their changes over time (Gorelick et al., 2017). This therefore provides much-needed information about the eco-hydrological dynamics of wetlands, which can be used to develop management actions.

Although the GEE cloud-computing platform offers numerous advantages for addressing the issue of limited and inconsistent data for wetland monitoring, its use for monitoring the long-term changes in the eco-hydrological dynamics of semi-arid wetlands requires further evaluation (Wua et al., 2019). Studies that have utilised this platform for the long-term monitoring of environmental change have included the mapping of forests, crops and open water (Kumar and Mutanga, 2018; Tamiminia et al., 2020), and the most recent studies that have used the GEE in wetland mapping have demonstrated the potential of this platform for wetland studies (Gxokwe et al., 2021; Zhou et al., 2019). Although these studies have underscored the relevance of utilising GEE for investigating small and seasonally-flooded wetlands, they have not considered the longer-term changes that can be assessed by using the time-series analysis in GEE.

There is therefore a need to assess the applicability of the GEE and multi-source spatial data for monitoring the long-term changes in the eco-hydrological dynamics of wetlands in semi-arid environments, particularly in sub-Saharan Africa. This will help to develop consistent and comparable wetland inventories across the region, and it will advance the understanding of the drivers of wetland dynamics (including the anthropogenic stressors that contribute to the degradation and loss of wetlands); it will also enhance the local, regional and international wetland conservation and management programme. Moreover, the long-term monitoring of eco-hydrological dynamics is critical for preventing and managing the loss of wetland

ecosystem services (Millennium Ecosystem Assessment, 2005). This chapter therefore aims to assess the use of the GEE cloud-computing platform for monitoring the long-term variations in the eco-hydrological dynamics of small and seasonally-flooded wetlands in a semi-arid region of South Africa over a 20-year period (2000-2020), by using the upper middle reach of the Nylsvley floodplain as a case study. The specific objectives are: 1) to evaluate the spatio-temporal variations in the eco-hydrological dynamics of wetlands, using time-series Landsat composite data, coupled with the GEE machine-learning algorithms; and 2) to establish the major drivers of the eco-hydrological changes in the studied systems, by using remotely-sensed metrics, coupled with climate data.

5.2 Materials and Methods

5.2.1 Field data collection

Field data were collected between 28-09-2020 and 1-10-2020 in the upper middle reach of the Nylsvley floodplain located within the boundaries of the Nylsvley nature reserve. A total of 300 ground truth points, which represented the different landcover classes, were collected during field surveys, by using a hand-held Geographical Positioning System (GPS) with an error margin of less than 3.65 m. The land cover that was observed included classes, such as the bare surfaces, the short wetland grass species, comprised of *Cynodon dactylon* (Bermuda grass) and *Oryza longistaminata* (Rice grass), as well as the long wetland grass species, comprised of only *Phragmites australis* (Common reeds). These vegetation communities were visually identified in the field. The floodplain was completely dry during the field visits; therefore, no water was detected. The collected field data coincided with the months of some images that were acquired during the study time periods. The ground truth points were collected following a stratified random sampling approach. Consequently, the floodplain was then subdivided into 30 m x 30 m quadrants, which were spaced about 10 m apart. These quadrants were chosen because of the pixel size of the remotely-sensed data that were used, as well as the size of the studied floodplain. This is supported by a number of remote sensing studies, such as Mudereri et al., (2021) and Thamaga and Dube (2018). A minimum of 20 points were randomly collected from each quadrant, depending on the heterogeneity of the landcover identified in each quadrant and the accessibility of some sections within the wetland. The field data were supplemented by 350 more points that were randomly extracted from higher-resolution Google Earth images and which coincided with the dates of some of the remotely-sensed images that were used. The data gathered from field visits and the higher-resolution Google Earth images were comprised of 552 points, which represented the short grass species,

68 points representing the bare surface and 30 points representing the long grass species. The collected points were then randomly split on the R-studio into 70% for training (386 points for short grass, 48 points for bare surface and 21 points for long grass) and 30 % for validation (136 points for short grass, 20 points for bare surface and 9 points for long grass) for the classification of images by using the Random Forest model.

5.2.2 Ancillary data

The 2000-2020 monthly rainfall and temperature data for the Mookgopong Weather Station, located at -24.42745 and 28.59419 north-west of the Nylsvley floodplain, were collected from the South African Agricultural Research Council. The station is located at an altitude of 820 m towards the higher elevation areas, and 26 km away from the studied floodplain. Due to the unavailability of evapotranspiration data, the evapotranspiration rates were estimated by using the following Equation 5.1 (Hargreaves and Samani, 1985):

$$E = 1.25 \times 0.0023 \times R_a T_r^{0.5} (T_a + 17.8) \quad (5.1)$$

where E is the potential evapotranspiration rate (mm/day), R_a is extra-terrestrial radiation (mm equivalent per day), T_r is monthly temperature range ($^{\circ}\text{C}$) and T_a is average daily temperature for the month ($^{\circ}\text{C}$). The evapotranspiration rates were used as the daily rates for a month. Other methods, such as the Penman-Monteith equation, could not be used because the solar radiation, wind speed and relative humidity were not available for the study area.

5.2.3 Remote sensing data acquisition and processing

The steps followed during the acquisition and processing of the remote sensing data are presented in Figure 5.1. Prior to the image classification, remotely-sensed products were extracted from the GEE catalogue, and Landsat-7 ETM+ was chosen because it has been available the longest on the platform. Although it is reported in literature that Landsat-7 ETM+ was adversely affected by the failure of the Scan Line Corrector (SLC) (Dube and Mutanga, 2015) after the 31st of May 2003, the stripes caused by the failed SLC did not affect the area covered by the wetland that was studied. The data collected were then filtered by the wetland boundaries and the date, using the codes 'Image.filterBounds ()' and 'ee.Filter.Date ()' on the GEE cloud-computing platform. After filtering the images by their dates and boundaries, they were then filtered by cloud cover. Images with a clear cloud coverage (0%) were selected because of the size (13.69 km²) of the studied wetland, while images with cloud cover were excluded, since cloud masking would result in severe error propagation. The 'ee.Filter.eq

('CLOUD_COVER', 0)' code was used to filter the images with cloud cover. The number of images obtained from the filtering process are presented in Table 5.1. The cloud-free images were then grouped in 5-year intervals, with the first interval being from 2000 to 2005, the second and third intervals from 2006 -2010 and 2011-2015, respectively, and the last interval from 2016 and 2020. After the grouping of the images, wetland indices, such as the Normalised Difference Vegetation Index (NDVI), the Normalised Difference Water Index (NDWI) and the Normalised Difference Phenology Index (NDPI), were computed for all the images obtained in each interval, using the equations presented in the supplementary material (Table A). They were then extracted to assess their variations over time, since they are used as a proxy for the eco-hydrological dynamics. NDVI was chosen because of its sensitivity to photosynthetically-active biomass, which enables the discrimination between wetland and non-wetland areas, as well as vegetated and non-vegetated areas (Liu and Huete, 1995). The NDWI was chosen because of its sensitivity to open water, as it can discriminate between water and non-water areas. Although it was noted that the system under study had very limited open water, the NDWI was used to capture the pixels that may have open water during any time period. The NDPI, which was recently developed by Wang et al., (2017), and it uses the weighted combination of Red and Short-Wave Infrared (SWIR) bands instead of using the red band, as in the NDVI. The weighted combination is almost equal to the value of the Near Infrared band (NIR) for different soils and is sensitive to vegetation. Moreover, the NDPI considers the absorption of the leaf water content in the SWIR band, which can thus account for variations in the canopy water content. This helps to improve the separability of soil and vegetation with the varying leaf water content, which was previously a daunting task, when using NDVI, and which is why the NDPI was chosen. Furthermore, the abovementioned indices were computed from the time series data and then used to assess the eco-hydrological dynamics of the wetland over time. In addition, visible (Red, Green and Blue) NIR and SWIR bands were extracted from the grouped images representing each interval, and concatenated to the extracted indices to form images with only the appropriate band combinations, which enhances the wetland features more. The concatenated images were then composited per interval to form single images that represented each interval, by using the median composite algorithm in GEE, which reduces a stack of images by calculating a median across the matching bands, thus minimising the illumination effects, such as shade, as well as the effects of cloud cover (Gorelick et al., 2017; Mahdianpari et al., 2020). The derived composite images were then subjected to the Object-Based Image Analysis (OBIA), which was chosen because of its superiority over a pixel-based analysis. Various studies (e.g. Mahdianpari et al., 2019; Pande-Chhetri et al., 2017)

have shown that OBIA relies not only on the spectral characteristics, but it also considers the contextual information (the height, texture, area, perimeter, etc.) of a pixel, which helps to improve the discrimination of wetland classes. The initial step in OBIA is image segmentation, which involves the partitioning of the image into multiple discrete and non-overlapping objects, based on a specific criterion. This maximises the separability of different landcover classes and prevents the ‘salt and pepper’ effect on the final classified image (Dlamini et al., 2021a). A Simple Non-Iterative Clustering (SNIC) Algorithm was then used to partition the composite images. The SNIC undertakes image segmentation by initialising the centroids in a regular gridded image, and then the dependency of each pixel is established, relative to the centroids, by using the distance in the five-dimensional space of colour and spatial coordinates. The distance joins the normalised spatial and colour distances to produce a uniform super pixel. The candidate pixel is added to the cluster, based on its shortest distance to the centroid (Achanta and Süssstrunk, 2017). This algorithm was selected because of its simplicity and memory efficiency, as well as its ability to incorporate connectivity between the pixels after the initiation of the algorithm. The SNIC algorithm was implemented using the code ‘ee.Algorithms.Image.Segmantation.SNIC()’ and the outputs were images with super pixels, as well as the contextual information of those pixels, such as the area, texture (standard deviation), perimeter and height. These were then integrated into the selected spectral bands and used as an input concatenate images in the selected machine-learning algorithm.

Table 5.1 Number of remotely-sensed images with no cloud cover obtained per time intervals with their details

Period	No. of images obtained	Remote sensing product
2000 - 2005	26	Landsat (ETM+)
2006 - 2010	20	Landsat (ETM+)
2011 - 2015	25	Landsat (ETM+)
2016 - 2020	37	Landsat (ETM+)

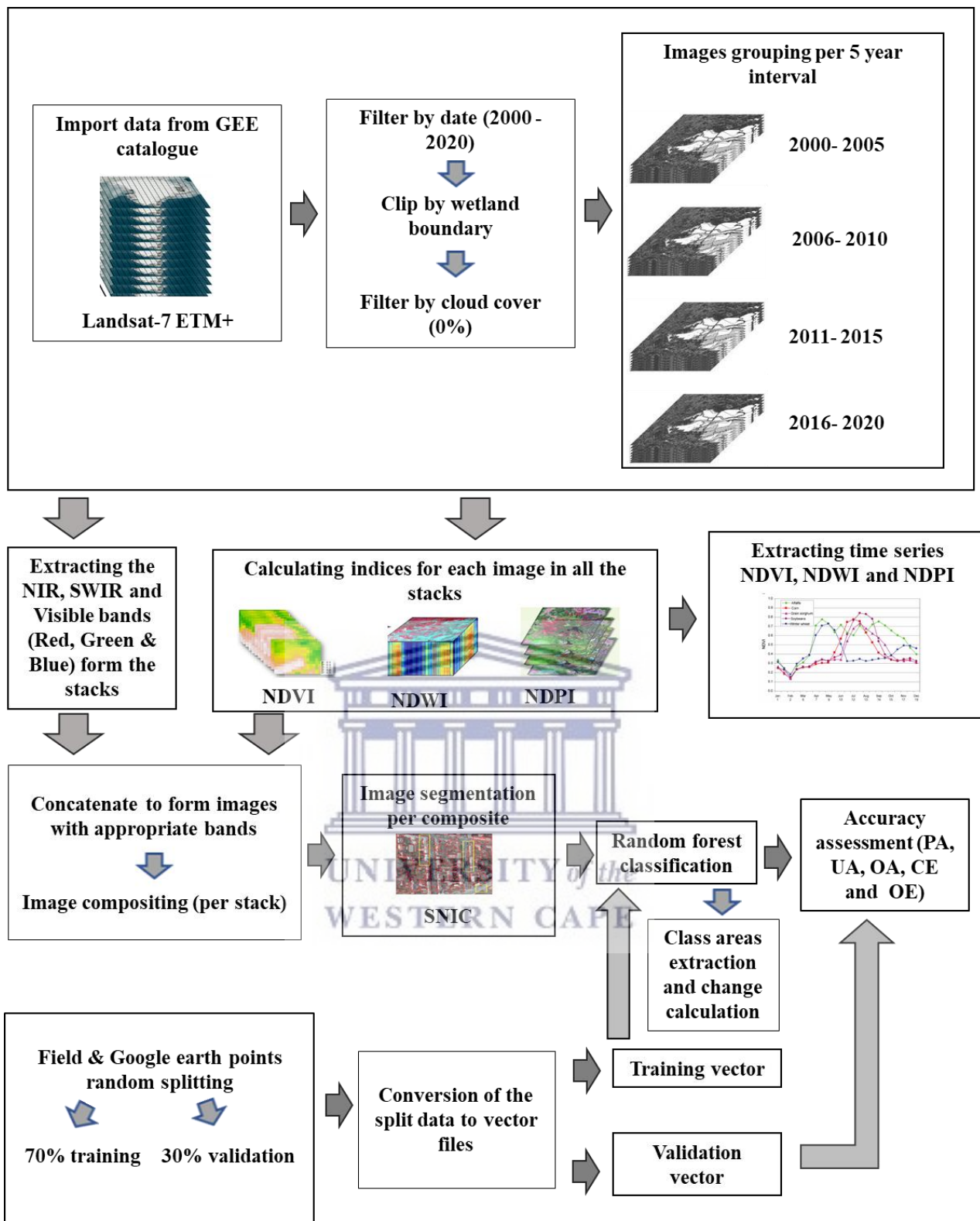
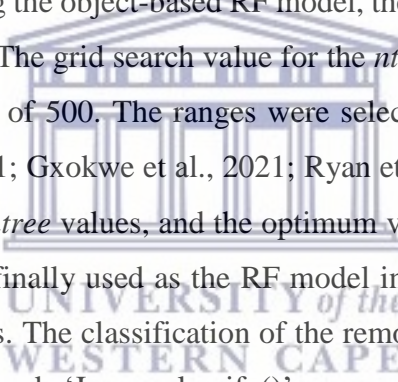


Figure 5.1 Image classification process using the GEE cloud-computing platform

5.2.4 Adopted wetland classification approach

The classification of the acquired images was executed by using the Random Forest (RF) machine-learning algorithm on the GEE cloud-computing platform. The RF algorithm is an ensemble non-parametric classifier that consists of many different trees from subsets of randomly-selected training data (Dlamini et al., 2021). Each tree casts a unit vote to decide on the final class of the object. The algorithm was selected because of its ability to handle the large differences between landcover classes, thereby neutralising the noise in the dataset (Simioni et al., 2020). Moreover, the selection of the RF algorithm in this study was informed by the study of Gxokwe et al., (2021), which identified RF as an appropriate GEE algorithm for classifying small and seasonally-flooded wetlands. Comparatively, the RF algorithms produced highest overall producer's and user's accuracies. Prior to the classification, the ground truth points were split on R-studio to 70% training and 30% validation samples. The split training data were then imported and converted to vector files in a GIS environment and imported to the GEE to train the classifier. The training of the RF classifier was executed by using the code 'ee.Classifier.train()'. In training the object-based RF model, the grid search value for the *mtry* parameter varied from 1 and 5. The grid search value for the *ntree* parameter was varied from 500 to 10 000, with an interval of 500. The ranges were selected, based on those in similar studies (e.g. Dlamini et al., 2021; Gxokwe et al., 2021; Ryan et al., 2014). The search yielded 100 combinations of *mtry* and *ntree* values, and the optimum values (1 500 *ntree* and 5 *mtry*) from these combinations were finally used as the RF model input parameters for classifying the Landsat-7 composite images. The classification of the remotely-sensed image composites was then executed by using the code 'Image.classify()'.


5.2.5 Accuracy assessment

The accuracy assessment was executed by using three metrics, namely, the Overall Accuracy (OA), the User's Accuracy (UA) and the Producer's Accuracy (PA). The OA measures the overall correctly-classified pixels (Story and Congalton, 1986) and UA measures the probability that a classified pixel of a particular category belongs to that category on the ground, while PA measures how well the reference groundcover type pixels were classified (Story and Congalton, 1986). The OA was implemented by using the code 'Image.accuracy()' in GEE, while PA and UA were implemented by using the codes 'Image.producersAccuracy' and 'Image.consumersAccuracy()'. During the implementation of the above-mentioned algorithms in GEE, the validation vector produced from the 30% field data was first imported to the

platform, and the zones on the classified image that corresponded to the vector points were sampled by using the code 'Image.sampleRegions()'. The sampled zones, as well as the validation vector points, were then used as input files for the latter algorithms to compute each accuracy metric (OA, UA and PA).

5.2.6 Change analysis, variations and major drivers

In order to assess the eco-hydrological changes in the wetlands, the area covered by each land cover class was estimated from the classified image per interval. Moreover, the rate of change for each class (RAC) was established by using the method of Shen et al. (2019) (see Equation 5.2).

$$RAC = \left(\frac{EA - IA}{IA} \right) \times 100\% \quad (5.2)$$

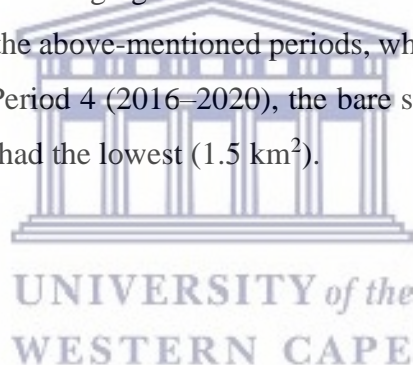
where *RAC* is given as the rate of change of the wetland area; *EA* is the area of the period considered (it refers to the areas in years 2000-2005, 2006-2010, 2011-2015 and 2016-2020); and *IA* is the initial wetland area (it refers to the total area for the period of 2000-2005 calculated as the sum of the surface areas covered by the identified wetland classes). The variations in eco-hydrological dynamics were established by analysing the NDVI, NDWI and NDPI time series, which were used as a proxy for the eco-hydrological dynamics of the wetland.

5.3 Results

The changes in spatial distribution of the wetland-cover classes for the time periods are presented in Figure 5.2. The water class could not be identified by using the object-based RF model because the training and validation data for this class were unavailable. This was because there was no visible spatial coverage of the class during the ground-truthing period, which may have been caused by short-term surface inundation. Moreover, the indices that were used could detect no water classes, mainly because of the forested vegetation within the wetland, as well as the spatial resolution of the data used. Although water could not be detected, most of the identified vegetation species, such as *Phragmites australis*, are associated with the presence of semi-permanent saturated wetland soils, which provided eco-hydrological information about the Nylsvley floodplain. This is supported by the findings of Kotze and O'Connor (2000), who reported that species such as *Phragmites australis*, were associated with the wettest zones within the studied wetland, therefore implying that ecohydrology information could be deduced from this type of species.

The class distribution shows that short grass was the most dominant class, except for the period 2016 to 2020, when the bare surface was the most dominant class. Long grass was the least dominant class for all the cases, except for the 4th period (2016-2020), where short grass was the least dominant. The visual assessment of the changes from the derived spatial distribution maps of the wetland classes shows variations in these classes over time. These results show that classes, such as short grass and long grass, declined between Periods 1 and 2 (2000-2005 and 2006-2010). However, a resurgence was observed in these classes during Period 3 (2011-2015). For the bare surface, an increase was noted for the entire monitoring period, except for Period 3 (2011-2015), where a decline of about 1.5 km² was noted. A sharp decline in the spatial distribution of short and long grass were observed for the Period 4 (2016–2020), and that for the bare surface sharply increased for the same period.

The class area results, which are shown in Figure 5.3, corroborate the visual observations of the spatial distribution of the identified wetland classes during fieldwork. The results show that the short grass class had the highest surface area for Periods 1, 2 and 3 (2000-2005, 2006-2010 and 2011-2015), with surface areas ranging between 5 km² and 8 km², respectively. Long grass had the lowest surface area for the above-mentioned periods, which ranged between 0.1 km² to 0.5 km², respectively. During Period 4 (2016–2020), the bare surface had the highest surface area of 4.8 km² and short grass had the lowest (1.5 km²).



5.3.1 Variations in the eco-hydrological dynamics

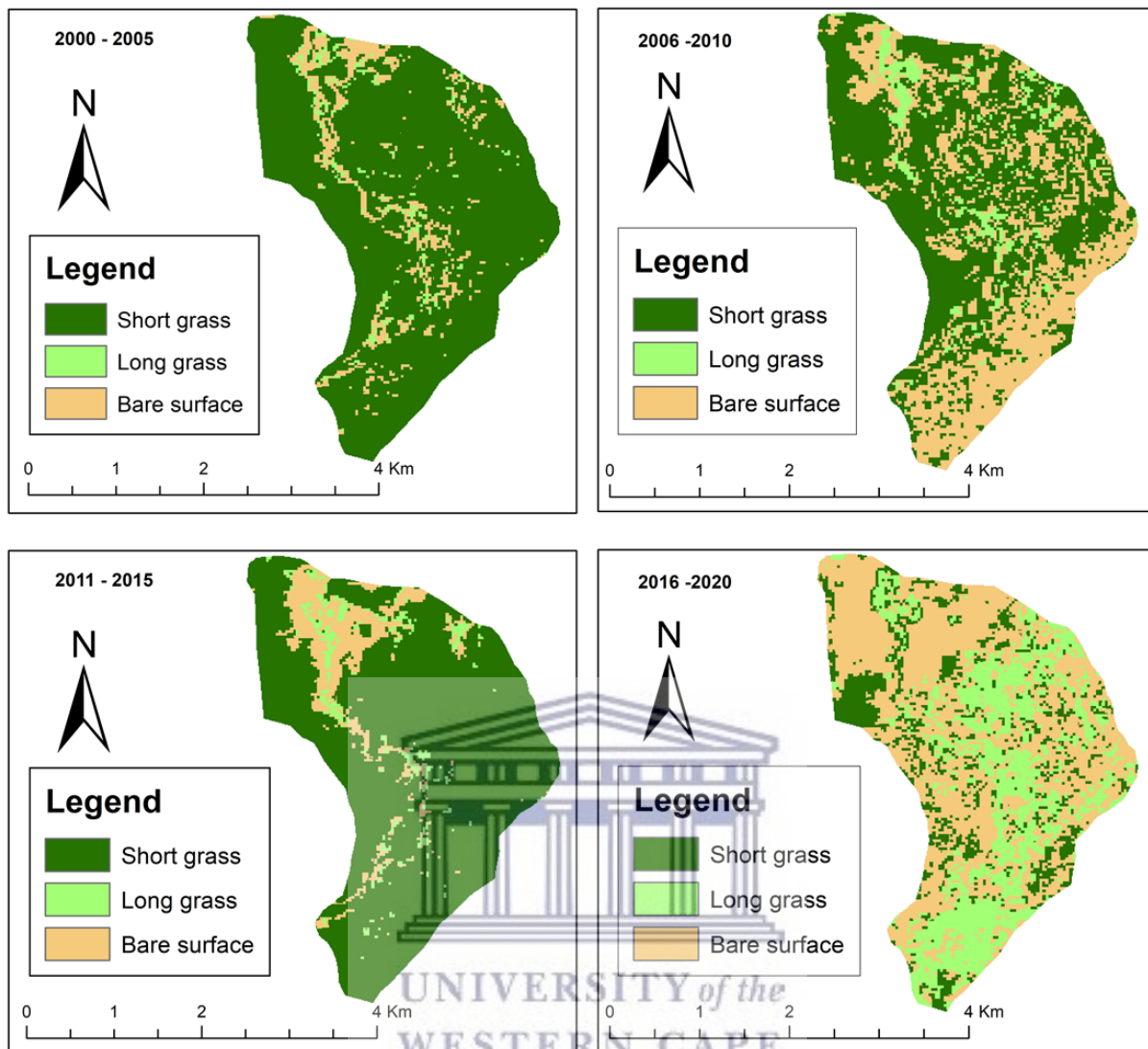


Figure 5.2 Spatial distribution of wetland classes for the studied periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020), produced by using Object-based Random Forest classification of Landsat-7 data

The results of the rate of change per class are shown in Table 5.2. Drastic changes occurred at a high rate in the wetland-cover classes between the periods. The results show that the short grass declined at a rate of about 38% between Periods 1 and 2 (2000-2005 and 2006-2010), whereas it increased for the other classes at a rate of 228% and 230%, particularly for the bare surface, which doubled, compared to Period 1 (2000–2005). Although a decline in short grass was observed for the above-mentioned periods, there was a 51% increase in this class during Period 3 (2011-2015). The bare surface and long grass decreased at a rate of 69% and 46%,

respectively, for Period 3 (2011-2015). Furthermore, the results show that the short grass declined by 81%, while it doubled for the other classes such as long grass.

Table 5.2 Rate of change of the Nylsvley floodplain between the periods

Class	Rate of change from P1 to P2 (%)	Rate of Change from P2 to P3 (%)	Rate of Change from P3 to P4 (%)
Short grass	-38	51	-81
Long grass	228	-46	896
Bare surface	230	-69	350

*P1: 2000 -2005, P2: 2006-2010, P3: 2011-2015 & P4: 2016-2020

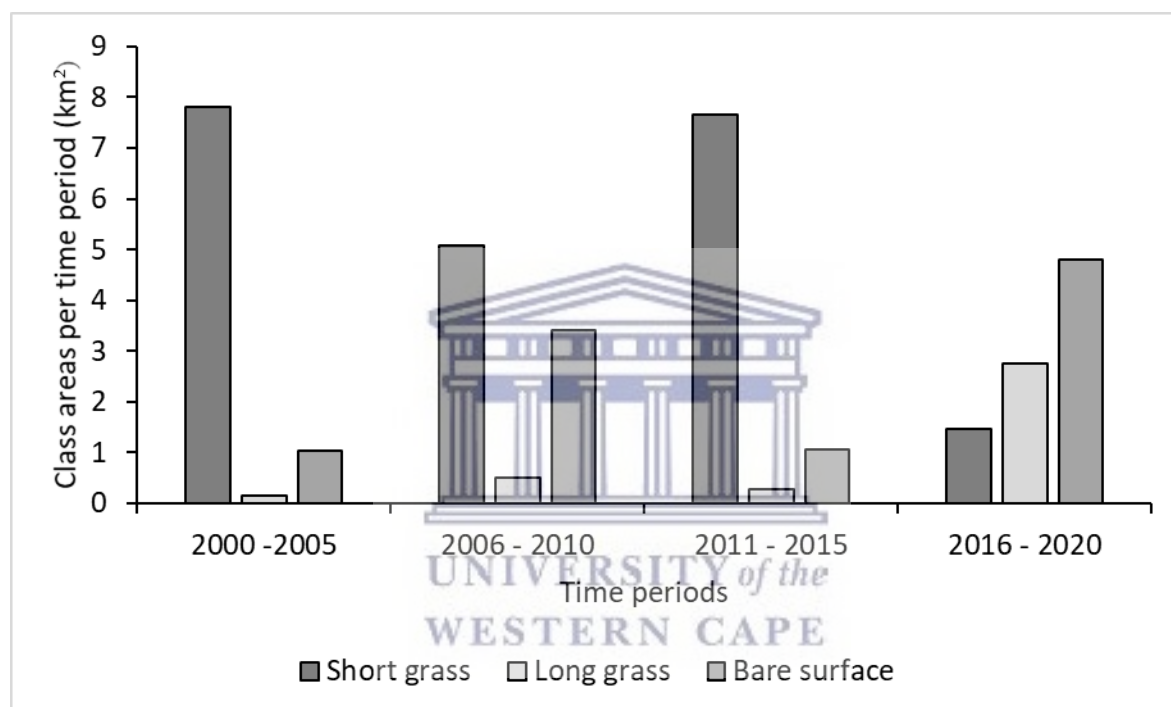


Figure 5.3 Nylsvley floodplain land cover surface areas per time periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020)

There were no major changes in the temporal variations of NDVI, NDPI and NDWI for the Nylsvley floodplain (Supplementary material: Figure A). However, the NDWI had low values of less than 0 throughout, which implies that no surface inundation that was detected. In most cases, NDPI and NDVI had values greater than 0, particularly for the 2000-2005, 2006-2010 and 2016-2020 time periods (Supplementary material: Figure A), which implies that vegetation was detected, although it was not healthy.

5.3.2 Accuracy assessment

The accuracy assessment results showed that the OA (Figure 5.3) for all the periods were within an acceptable range of between 72% and 78%, with 2011-2015 and 2016-2020 having the highest values, and 2006-2010 having the lowest.

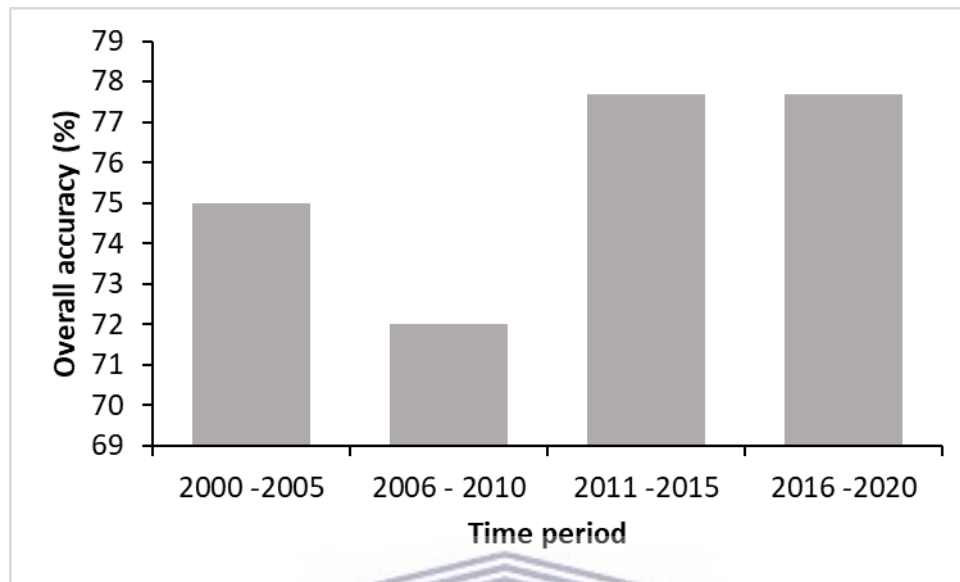


Figure 5.3 Overall accuracy for the classified image composites for each period (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020)

The class accuracy results are presented in Table 5.3. The results show that the PA ranged between 33% and 88% for all the periods, and that short grass had the highest PA in all the cases and long grass had the lowest PA, particularly for Periods 1 and 2 (2000-2005 and 2006-2010), during which the PA was 33%. For Periods 3 and 4 (2011-2015 and 2016-2020), the bare surface had the lowest PA, with 37.5% for both periods. The UA results ranged from 40% to 100%, with long grass having the highest UA for Periods 1 and 2 (2000-2005 and 2006-2010). The lowest UA was for the bare surface during Period 1 (2006-2010), with 40% (Table 5.3).

Table 5.3 Error metrics results for classified image composites representing Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020

Period	Class	PA (%)	UA (%)
2000 - 2005	Short grass	88	78.57
	Long grass	33.3	100
	Bare surface	50	57.14
2006 - 2010	Short grass	84	84
	Long grass	33.3	100
	Bare surface	50	40
2011 -2015	Short grass	92	85.18
	Long grass	66.6	66.6
	Bare surface	37.5	50
2016 - 2020	Short grass	92	88.46
	Long grass	66.6	50
	Bare surface	37.5	50

5.3.3 Drivers of eco-hydrological variations

Statistical summary of the NDVI, NDWI and NDPI results for all the periods are presented in the supplementary material (see Figure B). The results indicate that the indices data are mostly asymmetric (mostly positively skewed) for most of the time periods, except for some NDWI data, for example, Period 2 in 2006 and Period 3 in 2014, where the data show a symmetric distribution (bell-curved). In most cases, the data also show larger variations with large ranges (stretched whiskers), except for some years, such as 2003 in Period 1 (2000 -2005), where the whiskers for NDVI, NDWI and NDPI are squeezed, as well as 2008 and 2009 in Period 2 (2006 -2010), where the whiskers are short. The outliers were observed for the NDVI, NDWI and NDPI data for the year 2004 in Periods 1 and 4 for the years 2017, 2018 and 2019. For Period 3, outliers were observed for NDVI during the year 2015. In addition, an inverse relationship between the biomass (indices) and the annual ET rate (Figure 5.5) was also noted. The NDPI decreased considerably during the years when the ET increased e.g. 2002 and 2004 (Figure 5.5a).

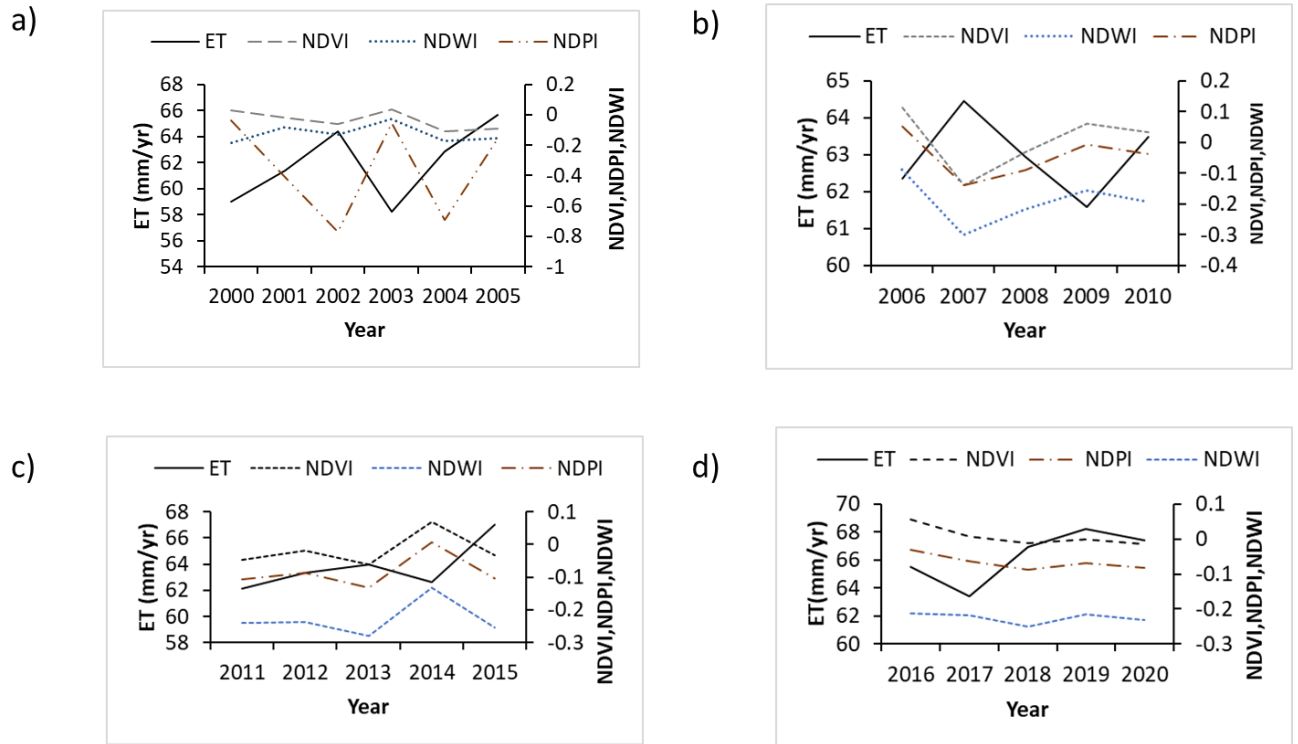


Figure 5.4 The relationship between biomass (indices) and annual ET rates for a) Period 1: 2000-2005, b) Period 2: 2006- 2010, c) Period 3: 2011-2015 and d) Period 4: 2016-2020

The biomass indices, except for NDPI, tended to increase, as expected, with the annual amount of precipitation e.g. from 2007 to 2009 (Figure 5.6). However, this relationship is not evident during some years, such as Period 1 from 2002 to 2004. A wetland typically stores both surface and subsurface water and, therefore, a decrease in rainfall in a single year, such as in 2002 and 2003, may not have a significant effect on the growth of plants, as they may be utilising water that is already in storage.

5.4 Discussion

This study assessed the use of the GEE cloud-computing platform for monitoring the long-term (2000-2020) variations in the eco-hydrological dynamics of the small and seasonally-flooded wetlands in a semi-arid environment of South Africa, using the Nylsvley floodplain as a case study. The main classification results showed that the following three landcover classes could be identified: a) short grass vegetation, characterised by species such as *Cynodon dactylon* and *Oryza longistaminata*; b) long grass species, such as *Phragmites australis*; and c) bare surfaces. These land-cover classes represent the zones with a semi-permanent soil saturation, as shown by the presence of *Phragmites australis*, and temporary to seasonal saturation, as represented by short grass (*Oryza longistaminata*) (Kotze and O'Connor, 2000). While the Nylsvley floodplain may generally be considered as a seasonally-flooded system, the results of this study show that the system contains both seasonal and semi-permanent hydro-periods because of the presence of the abovementioned wetland vegetation species, which are viewed as indicators of an environment with semi-permanent saturated soils and seasonal inundation. The water class could not be detected, as the training and validation data for this class were unavailable, because of the absence of surface inundation during the period of ground truthing. Moreover, the remotely-sensed spectral indices that were used failed to detect the presence of water, mostly due to the dense vegetation in parts of the studied system, as well as the spatial resolution of the Landsat-7 ETM+, which was selected because of the length of its image records and the number of good-quality images (Table B in supplementary material) that are available for this product in the GEE catalogue.

The accuracy assessment results showed that the overall accuracy was within the acceptable range of 69%-79%. However, there were some inaccuracies for the bare surface and long grass classes for some periods. These can be mostly attributed to the spatial resolution of the remotely-sensed product used, which resulted in significant spectral mixing between the classes. Several studies (e.g. Dvoretz et al., 2016; Gxokwe et al., 2021), have demonstrated that the spatial resolution of the remotely-sensed data does impact the accuracy of classification results significantly, particularly when mapping small areas, such as the wetlands in this study. Moreover, the Federal Geographic Data Committee report (1992) highlighted that, with the Landsat data that are used in this study, the accurate identification of an object requires at least 9 pixels, which covers a total surface area of about 0.9 ha and, at most, 25 pixels. The limited training data of the bare surface and long grass species, relative to the short grass species, due to the spatial coverage of these classes within the wetland, led to an unequal distribution of

training points, which affected the ability of the training data to adequately characterise the classes of interest.

Due to the small size of the bare surface and long grass within the wetland, fewer points were collected for these classes, than for short grass, thus resulting in an imbalance in the training samples. These imbalances are likely to have caused inaccuracies in the classification of the bare surface and long grass, as opposed to that of short grass, which had more training data. Several studies (e.g. Millard and Richardson, 2013; Ustuner et al., 2016; Xia et al., 2019) have demonstrated that an unequal distribution, or imbalance, in the training data is likely to influence the classification accuracy of the classes in such a way that the over-represented classes may dominate the under-represented classes, which may result in classification inaccuracies for the under-represented classes. However, the imbalances in training samples in this case did not result from a sampling error, but rather, from the lower number of training points available for the bare surface and long grass, which was caused by the small areal coverage of land-cover classes in the area under study.

The spatiotemporal variation results showed that eco-hydrological conditions varied from one period to another; some of the classes changed to more than double the previous area, and some dropped significantly to half of the previous period. The results indicated that, in most cases, the bare surface increased significantly for periods, such as 2006-2010 and 2016-2020. This can be attributed mostly to the seasonal changes, as most images used for these periods were collected during the dry season (Addendum A), which is the leaf-off season. Few suitable images were available for the wet (leaf-on) season, due to the frequent cloud cover. Moreover, below average precipitation was received in the LTRB during 2016-2017, which significantly impacted most ecosystems, including the wetlands in the area (Gxokwe et al., 2021) and which therefore explains the domination of the bare surface. This is also evident in Figure A in the supplementary material. The temporal variations in the indices results showed a similar trend between NDVI, NDWI and NDPI, and the floodplain was dry most of the time. These findings corroborate those of Nhamo et al., (2017), who reported high seasonal variabilities in the wetland of the Witbank Catchment in South Africa, although they focused on the inundation dynamics on a larger scale than in this study.

As expected, there was an inverse relationship between the ET and wetland biomass for all the periods. This is because ET tends to contribute to the water losses in wetlands, particularly in semi-arid regions, which therefore results in a loss of the available moisture in wetlands and

therefore contributes to the decline in the NDWI values for the wetlands. Due to the decline in moisture content available in the wetland, some vegetation become less healthy, because water stress results in a decline in the NDVI and NDPI values. Al-Shehhi et al., (2011) monitored the effects of a variable soil moisture content using the MODIS-derived NDVI for different vegetation species in Abu Dhabi. The findings showed that, with the increasing evaporation rates, less healthy vegetation dominated the studied area, due to the plants wilting, which resulted in decreased NDVI values. These results corroborate the findings of the current study, although the focus was not only on wetland vegetation. West et al., (2018) also studied the vegetation response to the varying soil moisture contents, using Sentinel-2 data, and the results showed a strong correlation between a Sentinel-2-derived NDVI and the soil moisture content, which therefore supports the findings of the current study.

A comparative analysis of the annual precipitation with average NDVI, NDWI and NDPI values showed that with an increase in the annual precipitation, the indices also increased for all the time periods, except for 2000-2001, where the indices showed a delayed response to the varying annual precipitation. The delayed response could have resulted from the use of dry season images, which corresponds to the leaf-off season for vegetation and the low-flow period for inundation. The study by Chen et al., (2020) evaluated the vegetation response to precipitation anomalies in the semi-arid Inner Mongolia Plateau of China over a 34-year period, using the Advanced Very High-Resolution Radiometer (AVHRR)-derived NDVI and multi-source precipitation data. The findings of the study showed that there is a time-lag in the vegetation's response to precipitation, although the duration is not visible.

This study has shown that the annual variations in the rainfall and evapotranspiration rates are the major drivers for the eco-hydrological dynamics of the Nylsvley floodplain. The annual rainfall will also influence the availability of both soil water and groundwater, which influence the vegetation growth in a wetland. These findings corroborate those of Birkhead et al., (2007), who modelled the hydraulic behaviour of the Nylsvley floodplain by using field-measured climate data, coupled with LIDAR data. The study identified the ET and annual rainfall as the driving factors of the eco-hydrological changes in the floodplain.

Overall, the findings underscore the relevance of cloud-computing artificial intelligence platforms, such as GEE, for monitoring the long-term variations in the eco-hydrological dynamics of semi-arid wetlands in the sub-Saharan region, by providing information on their degradation rates and losses, which are currently unknown for these regions. These findings

are critical for the development of sustainable use and management frameworks of small and seasonally-flooded wetlands in semi-arid regions. They also provide useful alternative ways and robust methodologies to generate routine and continuous wetland information about these changes, which was previously a daunting task when using traditional mapping and monitoring techniques. This will therefore help to develop the local- and global-scale frameworks and assist in setting up conservation policies and the use of effective management practices in wetlands, particularly in data-scarce regions with semi-arid climates and semi-permanent wetlands with a varying ecohydrology, which may typically be over-exploited during the dry periods.

The study also informs the ongoing efforts of various projects, for example, those of the Global Monitoring Environment and Security (GMES)-Africa, which is hosted by southern African Science Services Centre for Climate Change and Adaptive Land Management, and which promotes the use of earth observation data for assessing, monitoring and managing wetlands in transboundary basins, to prevent the loss of ecosystems services, which are beneficial to the communities living in the vicinity of these systems. This study also contributes towards achieving the key priority areas of Agenda 2063 of the African Union, which seeks to preserve at least 17% of the terrestrial and inland waters (including wetlands) and 10% of the marine areas by 2023. It also contributes to achieving Sustainable Development Goal (SDG 6.6), which advocates the prevention of the destruction and degradation of all ecosystems, and the rehabilitation of those that have already been destroyed (Sudmanns et al., 2020).

5.5 Conclusion and Recommendation

This study utilised remotely-sensed data, which is accessible on the GEE platform, coupled with advanced machine-learning algorithms, to monitor the long-term eco-hydrological dynamics of semi-permanent wetlands (over a 20-year period), using the Nylsvley floodplain in Limpopo, South Africa, as a case study. The findings of the study showed that the ecohydrology of the wetland varied over time and space. For example, the bare surface increased twice as much as the previous measurement period, whereas the short grass decreased significantly during Period 2. Moreover, the remotely-sensed-derived spectral indices showed a similar trend throughout the 20-year period, with NDWI having minimum values of less than zero, in all cases, which implies that there was no surface inundation. However, the presence of some wetland vegetation indicated that there were seasonal to semi-permanent soil-saturation conditions. The results also showed that the annual rainfall and ET were the major

drivers of the eco-hydrological dynamics of the Nylsvley floodplain. These findings underscore the relevance of using cloud-computing artificial intelligence techniques, coupled with optical remotely-sensed data, for monitoring the changes in the ecohydrology of small and seasonally-flooded wetlands, which is critical for the sustainable use and management of the systems in these regions. They provide useful alternative ways and robust methodologies to generate long-term baseline information about these changes over time, which was previously a daunting task when using traditional mapping and monitoring techniques. This study will therefore help to develop local- and global-scale frameworks that will assist in setting up conservation policies, and in the application of effective management practices in wetlands, particularly in data-scarce regions with semi-arid climates and semi-permanent wetlands with a varying ecohydrology that may typically be over-exploited during the dry periods. While the study offers new opportunities to generate continuous high-resolution, wetland-scale information, it is recommended that the potential of this platform for the large-scale monitoring of the impacts of adjacent land use and land cover changes to these systems should be investigated, as these systems are most likely to be influenced by the land use changes that occur on a regional and local level.



CHAPTER 6

LARGE-SCALE ANALYSIS OF LAND USE LAND COVER CHANGES ON SEASONALLY-FLOODED WETLANDS EXTENT IN SEMI-ARID SOUTHERN AFRICA



Photo: Courtesy of Siyamthanda Gxokwe 2020

This chapter is based on the following manuscript: (Under review)

Gxokwe, S., Dube T., Mazvimavi D. 2022. Large-scale analysis of Land use Land cover changes impacts of seasonally-flooded wetlands extent in the semi-arid southern Africa. *Hydrology and Earth Science Systems*. **Manuscript no. hess-2022-272** (IF = 6.617)

Abstract

Although significant research strides have been made in the analysis and modelling of Land Use Land Cover (LULC) changes, at various scales, the attention has been focused on the impact of individual aspects of LULC changes on wetlands or larger protected systems, rather than on the over-utilised small and seasonally-flooded systems. It is therefore imperative that LULC changes, and their impact on wetlands, be monitored at all levels, and in wetlands of varying size and nature, in order to properly inform the future wetland conservation strategies. Although Earth Observations (EO) data offer useful opportunities for addressing the issues associated with the monitoring of LULC changes and their impact on wetlands, they are still contested by the non-remote sensing community because of unavailability of seamless and cloud-free data throughout the year, which creates difficulties in continuous monitoring of the environment. Improvements in remote sensing data analytic tools, such as the introduction of various cloud-computing platforms, offer the opportunity to address the limitations of EO data with their advanced processing power and algorithms, which allow for the integration of multi-source and multitemporal data and which may improve environmental monitoring. Therefore, this study analysed the impact of large-scale LULC changes on the extent of semi-arid wetlands over a 20-year period, using the Limpopo Transboundary Basin in southern Africa as a case study. The specific objectives were: 1) to examine the LULC changes at the studied basin over a 20-year period, using time series multisource remotely-sensed data (Landsat-5 TM, Landsat-8 Operational Land Imager (OLI) and Sentinel-1), coupled with the Random Forest machine-learning algorithm; and 2) to assess the relationship between the LULC transitions and the extent of the wetlands, by using land-cover class area changes. Overall, 9 land-cover classes were characterised; these included shrublands, croplands, bare surface, wetlands, sparse vegetation, grasslands, built-up area and savanna grasslands. The shrubland was the most dominating class throughout the study periods, covering between 76% and 82% of the total surface area of the basin, while wetlands and sparse vegetation were the least dominating classes, with the proportions ranging between 0.9% and 2%, and 0.3% and 0.04%, respectively. The results showed an overall accuracy within an acceptable range (77% -78%), although some classes, such as wetlands and sparse vegetation, presented low class accuracies for some of the periods. This was attributed to the unbalanced training and validation data caused by the low spatial coverage of these classes. It was also observed that wetlands and sparse vegetation continued to decline at an average rate of 19% and 44%, while shrublands, croplands and savanna grasslands continued to increase at an average rate of 0.4%, 12.4% and 4.25%,

respectively. Most of the wetland area (40%) was replaced by built-up areas, indicating that urbanization is a major driving factor that is leading to the loss of wetlands in the area. The study provides new insights into the diminishing state of the semi-arid wetlands in southern Africa, particularly in the LTRB. These findings are critical for environmental planners and conservation specialists, as they provide the baseline information that is required for developing proper strategies and that will assist in curbing the negative impacts that these LULC cover changes have on wetlands. Moreover, the study provides robust, cost-effective and efficient ways to assist the generation of acceptable and precise knowledge, to inform the LULC management policies, as well as wetlands conservation and management strategies, on a larger scale.

Keywords: Big data analytic; Dryland wetland; Google Earth Engine; Wetland condition; Wetland loss; Wetland management



6.1 Introduction

Over the years, Land Use and Land Cover (LULC) changes have become a cause for concern to nature conservationists and decision planners, due to their disproportionate growth and impact on natural ecosystems, such as wetlands (Kulithalai et al., 2021). These changes are mostly the result of anthropogenic activities, such as urbanisation, artisanal mining and agricultural land expansion, which may lead to severe environmental concerns, such as ecological change, climate variability, biodiversity losses, as well as water and air contamination. These LULC changes result in a decline in the productivity and quality of wetlands and an increase in the wetland stress, such as a deterioration in the water quality and a shrinkage in the extent of the wetland, as well as the deterioration of its vegetation health. It has been reported that over 60% of the global wetlands are degraded, due to human-induced pressure associated with LULC changes (Millennium Ecosystem Assessment, 2005). This has resulted in the loss of the critical ecosystem goods and services provided by wetlands to the surrounding communities (Alam et al., 2011; Martínez-López et al., 2014). Various studies have demonstrated the impacts of adjacent LULC changes on the condition and loss of wetlands (e.g. Singh and Lin, 2015; Mohammadimanesh et al., 2018; Chen et al., 2020; Sakané et al., 2011; Thamaga et al., 2022; Dzurume, 2021); however, these studies have focused mostly on the impacts, in relation to the wetland water quality and vegetation dynamics, or the impacts on a smaller scale, such as at a sub-basin level. The attention has also mostly been drawn to the impacts on large protected systems, rather than on the over-utilised, small and seasonally-flooded wetlands (Thamaga et al., 2022), which has, in turn, resulted in these wetlands being replaced by other land cover types, thus causing a loss of the ecosystem goods and services that they provide. It is therefore imperative that a large-scale analysis of the LULC changes and their impacts on all wetlands of varying sizes and of a different nature, be monitored, in order to properly inform the conservation policies on the sustainable use of wetland systems.

The large-scale analysis and monitoring of the impacts of LULC changes on wetlands presents challenges that are largely associated with finances. Moreover, in transboundary basins, challenges like the differences in LULC management policies and the reconciliation of basin boundary politics, present difficulties in symmetrically LULC changes monitoring and impacts on wetlands. The availability of Earth Observation (EO) data and remote sensing approaches offers the opportunities to address these challenges; however, the use of EO data in environmental monitoring has been contested by the non-remote sensing community simply because of the lack of seamless and cloud-free data for some parts of the year, which therefore

restricts the use of EO data in environmental monitoring. Improvements in remote sensing data analytic tools, such as the introduction of cloud-computing platforms like Google Earth Engine, offer unprecedented opportunities to address these issues, through their specialised multi-source and multi-temporal, data-filtering, integration and -processing algorithms (Wang and Yésou 2018), which enhances the monitoring of environmental changes when using EO data.

Significant research strides have been made in the analysis of LULC changes and wetlands by using the newly-available cloud-computing platforms, such as the Google Earth Engine. For example, studies like those of Mahdianpari et al., (2020; 2019), Shelestov et al., (2017b; 2017a), Shafizadeh-Moghadam et al., (2021) and Ji et al.,(2020) have used these platforms to successfully analyse the LULC changes and impacts with reasonable accuracy in different parts of the world, although their focus has been either on one land-cover aspect, on the impacts on large protected systems, or the impacts at a sub-basin level. It is therefore imperative that the LULC changes, and their impacts on all wetlands of varying sizes and of a different nature, are monitored on a larger scale. This will help to properly inform wetland management programmes on the trends of wetlands ecosystems, and it will enable the setting up of proper conservation policies that will safe guide the sustainable use of all wetlands of different nature and of varying sizes.

Owing to this background, this chapter is aimed at analysing the impacts of LULC changes on the extent of semi-arid wetlands systems at regional scale over a 20-year period, using the Limpopo Transboundary Basin in southern Africa as a case study. Specifically, the objectives are: 1) to examine the LULC changes at the transboundary basin over a 20-year period, using time series multisource remotely-sensed data, coupled with GEE advanced machine-learning algorithms; and 2) to assess the LULC transitions and their impacts on the extent of the wetlands between the studied time periods, using a land cover class area change analysis.

6.2 Materials and Methods

6.2.1 Ground truth data collection

The ground truth data for this study were collected from field-based surveys and the review of the records of different land-cover products, from sources such as the South African Biodiversity Institute (SANBI), the European Space Agency Global land cover data, as well as the Copernicus global land cover data. The field points were collected during the period between 28-09-2020 and 1-10-2020, which coincided with the dates of some images that were used in this study. The data were collected by using a hand-held Geographical Positioning

System (GPS) with an error margin of less than 3.65 m. The data consisted of 428 ground truth points that represent the 9 land-cover classes presented in Table 6.1. The data were collected by using the stratified random sampling approach where the region was subdivided into 900 m² plot sizes, and a minimum of 20 points representing the different land-cover classes (see Table 6.1) were gathered in each plot. This was informed by a number of related studies (e.g. Mtengwana et al., 2020; Thamaga and Dube, 2018; Mudereri et al., 2021). In addition, the 900m² plots were spaced 1 km apart, to avoid the overlap of the collected samples. The field-collected ground truth points were further supplemented by points extracted from the land-cover products by SANBI (with a 30 m resolution), Copernicus (with a 100 m resolution) and the European Space Agency (with a 300 m resolution), as well as the previous land-use data collected by the South African Water Research Commission. The extracted data also coincided with the dates of some of the images used in this study. A total of 1 495 points representing the same land-cover classes as the field collected. The 1 495 points collected from the field surveys, as well as the records and reviews that were collected, were later randomly split into 70% training and 30% testing data, to train the pixel-based RF model on the GEE platform.

Table 6.1 Number of training and test points for LULC analysis

Land-cover classes identified	field samples	Total training points	Total validation points
Forest	133	93	40
Shrublands	582	408	174
Grasslands	157	110	47
Croplands	134	94	40
Wetlands	50	35	15
Sparse Vegetation	20	15	5
Bare surface	176	123	53
Built-up area	95	66	29
Open water	137	96	41

6.3.2 Pre-preparation and processing of optical remotely-sensed data

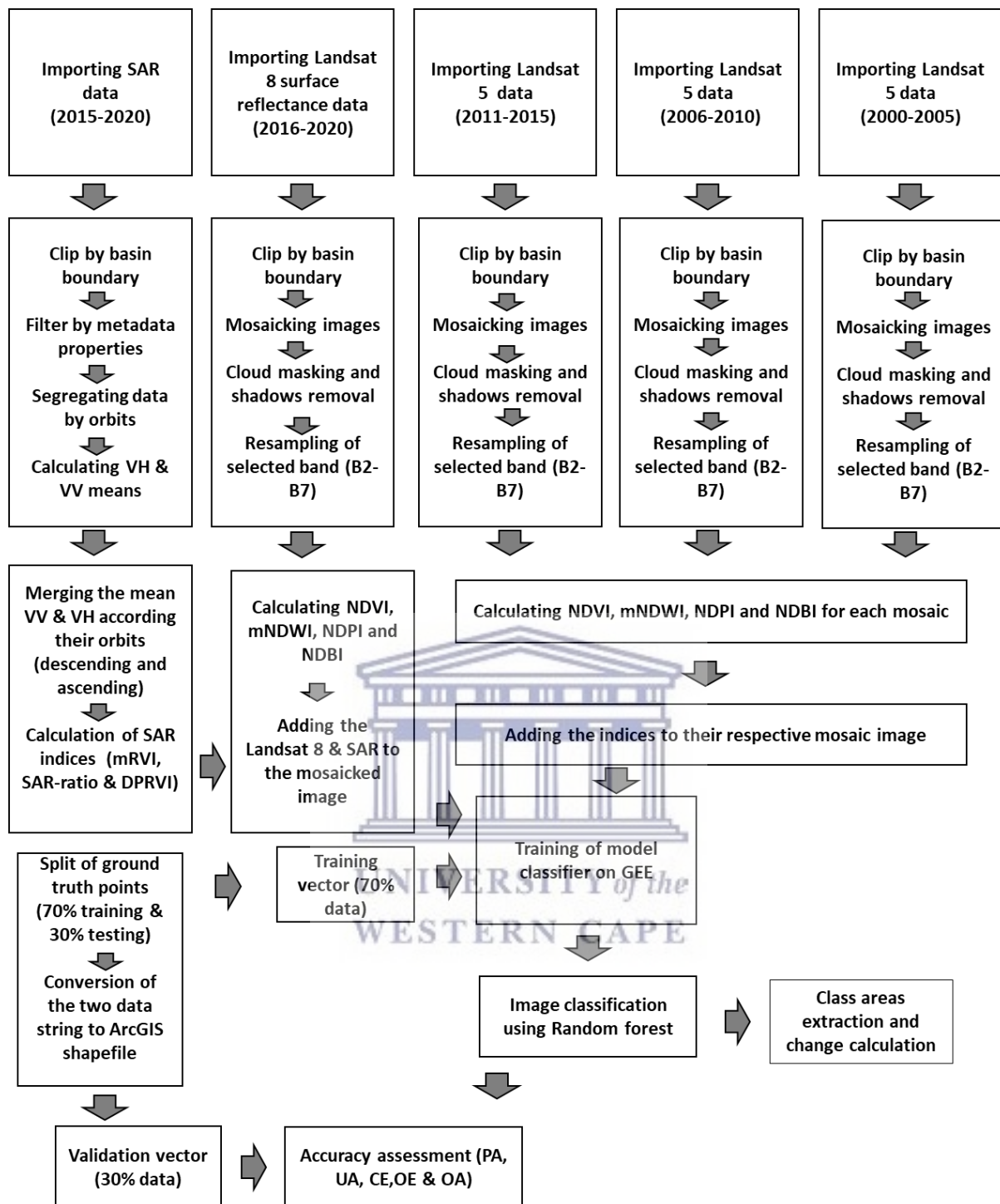


Figure 6.1 Steps followed on the GEE to process the remotely-sensed data used in this study

The steps that were followed to process the acquired data are presented in Figure 6.1. During this process, multisource data that were acquired in GEE catalogue were used. The remotely-sensed data included Landsat 8 data, which represented the 2016-2020 period, Sentinel-1 data, which represented the same period as Landsat 8, and Landsat 5 data, which represented the 2000-2005, 2006-2010 and 2011-2015 periods. Landsat products were chosen because of their continued availability on the GEE platform, which allowed for the continuous monitoring of the impacts of LULC change on the wetlands over a longer term and at a larger spatial scale. Sentinel-1 data were chosen to better enhance some LULC classes, such as forested wetlands, which may not be enhanced by the Landsat data only. Moreover, the study sought to demonstrate whether multisource data integration improves the LULC classification. Prior to the classification, the remotely-sensed Landsat data were pre-processed. During the pre-processing stage, the images were firstly clipped to the extent of the transboundary basin. After clipping the images, stacks representing each time period (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020) were mosaicked to produce an image scene that represented each time period. After the mosaicking of the images, cloud-masking and shadows removal were initiated on each composite. This was executed by using the Quality Assurance (QA) bands that are included in these products. These bands assist with indicating those pixels that are affected by instrument artefacts or by cloud contamination. During the process of cloud- and shadow-removal in this study, the QA bits were firstly computed, in order to get a single band image with the cloud and shadow scores. The cloud and shadow masks were then created by using a single band image with the cloud and shadow scores. The masks were then applied on the mosaicked images, in order to remove the cloudy and shadow pixels. After the removal of the clouds and shadows, the Landsat 5 and 8 bands were resampled to a 10 m spatial resolution, since the study wanted to also integrate Sentinel-1 data, with a spatial resolution ranging from 10 m, 20 m and 25 m, to better enhance the LULC features. The resampling of the Landsat bands was executed by using the nearest-neighbour resampling method, which was informed by a number of related studies (e.g. Mtengwana et al., 2020; Thamaga and Dube, 2018, amongst others). After resampling, the images were further used to compute the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Phenology Index (NDPI), the modified Normalized Difference Water Index (mNDWI) and the Normalized Difference Built-up Index (NDBI) by using the equations presented in Table 6.2. The NDVI was selected because it is sensitive to photosynthetically-active biomass, thus making it easy to discriminate between the wetland and non-wetland areas, as well as the vegetated and non-vegetated areas (Liu and Huete, 1995). The NDPI was selected because of

its ability to provide information on the vegetation water content, since it uses a weighted combination of Red and Short-Wave Infrared (SWIR), thus making it easy to differentiate between vegetation that is healthy and non-healthy, as well as the bare soils (Wang et al., 2017). The mNDWI was selected because of its ability to accurately extract open water features better than the standard Normalized Difference Water Index (NDWI), as reported in the study by Maswanganye et al., (2022). The NDBI was selected because of its success in mapping urban areas in various studies (e.g. Karanam and BabuNeela, 2018, amongst others). All the computed indices were added to their respective mosaicked images as extra bands and these images were then used in a pixel-based classification.

6.2.3 Pre-preparation and processing of Synthetic Aperture Radar data

This study used the already pre-processed Sentinel-1 GRD data, which are available in the GEE catalogue. The steps that were used to pre-process the data are made available in the GEE User's Guide and are similar to those of the ESA SNAP Sentinel-1 toolbox. These steps include updating the orbit metadata, GRD border noise removal, thermal noise removal, radiometric calibration, terrain correction, the conversion of the back-scattering intensity to normalised back-scattering coefficients, incidence angle correction and speckle reduction. In this study, the pre-processed Sentinel-1 data were firstly clipped to the boundary of the transboundary basin and filtered by their metadata properties, such as transmitter-receiver polarization, the instrument mode and orbit properties. The transmitter-receiver polarization used in this study were the vertically-transmitted, vertically-received SAR back-scattering coefficients (σ_{VV}^0) and the vertically-transmitted, horizontally-received SAR back-scattering coefficients (σ_{VH}^0). σ_{VV}^0 was used because of its sensitivity to soil moisture and its ability to discriminate between flooded and non-flooded vegetation, thus enhancing the detection of the wetland areas (Adeli et al., 2020). σ_{VH}^0 was used because of its cross-polarization which is known to be produced by volume scattering within the vegetation canopy and has a higher sensitivity to vegetation structures (Adeli et al., 2020). The instrument mode that was chosen was wide swath (IW) because of the size of the transboundary basin studies, and for both ascending and descending orbits. After filtering by metadata properties, the σ_{VH}^0 and σ_{VV}^0 were segregated using the orbit properties (ascending and descending), and the mean values for each polarisation and orbit property were computed. The computed mean σ_{VH}^0 and σ_{VV}^0 values were then merged for each orbital property and used to compute the SAR ratio, the Dual Polarisation Radar Vegetation Index (dual-pol RV), as well as Dual Polarimetric Synthetic Aperture Radar Vegetation Index (DPSVI), by using the equations presented in Table 6.2. The SAR ratio was computed because

was proven to be useful in the mapping of pasture lands, as demonstrated by Nicolau et al., (2021); thus, it will assist in discriminating such land cover from the other types. The dual-pol RV was chosen because the Sentinel-1 data used in this study has dual polarisation (VV, VH). Moreover, the index has been proved to be successful in retrieving soil moisture information by Li and Wang (2018) and Trudel et al., (2012). DPSVI was chosen because it calculates the rate of depolarisation in terms of the vertical dual depolarisation index, thus discriminating between the bare surface and vegetated areas (Mandal et al., 2020). The merged mean values for σ_{VH}^0 and σ_{VV}^0 , and the above-mentioned SAR indices were later added to the resampled Landsat-8 data, in order to test the potential of the integrated SAR and optical data for improving the classification accuracy.

Table 6.2 Spectral indices extracted from the SAR and optical remotely-sensed data

Remote sensing data	Indices extracted	Reference
Dual Polarisation Synthetic Aperture Radar	$dual_pol\ RV = \frac{4\sigma_{VV}^0}{\sigma_{VV}^0 + \sigma_{VH}^0}$	(Mandal et al., 2020)
Synthetic Aperture Radar Ratio (SAR_ratio)	$SAR_{ratio} = \frac{\sigma_{VV}^0}{\sigma_{VH}^0}$	(Mahdianpari et al., 2020)
Dual Polarimetric Synthetic Aperture Radar Vegetation Index (DPSVI)	$DPSVI = \frac{\sigma_{VV}^0 + \sigma_{VH}^0}{\sigma_{VV}^0}$	(Periasamy, 2018)
Normalised Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR - RED}{NIR + RED}$	(Tucker, 1979)
Modified Normalised Difference Water Index (mNDWI)	$mNDWI = \frac{Green - SWIR1}{Green + SWIR1}$	(Xu, 2006)
Normalised Difference	$NDPI = \frac{NIR - (0.74 \times Red + 0.26 \times SWIR1)}{NIR + (0.74 \times Red + 0.26 \times SWIR1)}$	(Wang et al., 2017)

Phenology Index**(NDPI)****Normalised****Difference Built-up****Index**

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$$

(Zha et al.,
2003)**6.2.4 Pixel-based image analyses**

In this chapter, pixel-based image analyses were implemented on the mosaicked images, although various studies (e.g. Mahdianpari et al., 2020; Amani et al., 2017; Dlamini et al., 2021) proved that the use of pixel-based images may result in significant spectral mixing between the various classes during the classification process, which may result in inaccuracies. In this study, a pixel-based analysis was chosen simply because an object-based analysis in large-scale mapping is associated with high computational costs, which results in computational time-out errors on the GEE platform (Gorelick et al., 2017; Shafizadeh-Moghadam et al., 2021); therefore, the implementation of pixel-based may reduce the high computational time costs. The pixel-based analysis was implemented on the Random Forest (RF) machine-learning algorithm, which is an ensemble classifier consisting of many different trees that are formulated from the subsets of randomly-selected training data (Dlamini et al., 2021b). In separating the land-cover classes, each tree casts a unit vote and the object with the highest number of votes will be classified as a particular class. The selection of RF was based on its performance in a number of related studies (e.g. Simioni et al., 2020; Gxokwe et al., 2021; Dlamini et al., 2021). During the implementation of the pixel-based RF model, the 1 428 ground truth points representing different landcover classes in Table 6.1 were randomly split into 70% training and 30% validation data in the GEE. The 70/30 split was informed by a number of related studies (e.g. Dzurume et al., 2021; Thamaga and Dube, 2018; Gxokwe et al., 2021). The 70% training points were then used to sampled out the regions corresponding to the location of these points. The sampled regions and the points were then used to train the RF classifier, and the classifier was used to classify the mosaicked images. During this process, the search for the input *mtry* and *ntree* parameters for the RF model varied between 1 to 5 for *mtry*, and 15 000 for the *ntree* parameter. The *ntree* search interval was 500. The selection of these ranges for the *mtry* and *ntree* parameters was informed by other similar studies (e.g. Adam et al., 2014; Simioni et al., 2020). The search yielded 300 combinations of *mtry* and *ntree* values

from several iterations, and the optimum values that yielded the higher overall accuracies were used.

6.2.5 Accuracy assessment

The accuracy assessment in this chapter was implemented by using five accuracy metrics, namely, the Overall Accuracy (OA), the Producer's Accuracy (PA), the User's Accuracy (UA), the Commission Errors (CE) and the Omission Errors (OE). The OA gave an indication of the correctly-classified pixels, while PA and UA gave an indication of how well the ground cover-type pixels were classified, as well as the probability that the classified pixel represents the actual feature on the ground (Story and Congalton, 1986). CE gave a review of the incorrectly-classified sites, while OE gave an indication of the number of sites that were omitted from the correctly-classified classes on a map (Mtengwana et al., 2020). During the implementation of the above-mentioned metrics, the remaining 30% of the data were used to sample regions on a classified image. The regions were then used with validation points as input GEE files to calculate the OA, UA and PA algorithms on this platform. Moreover, the files were used to extract the error matrix, which was then used to calculate the CE and OE.

6.2.5 Land use and land cover change analysis

To assess the changes in the LULC for the transboundary basin, the land cover maps were exported from the GEE and then imported into the GIS environment, and the class areas were extracted for all the time periods. After importing, the raster maps, showing the spatial distribution of each class, were then converted to vector files on the platform, and the class areas were extracted and utilised to assess the landcover transitions between the studied time periods, by using the Sankey diagram. Sankey diagrams visualise the flow from various nodes in a network. This diagram is mostly applied in the energy flow and materials. Although this is the case, various studies (e.g. Cuba, 2015; de Alban et al., 2018; Spruce et al., 2020) have successfully used Sankey diagrams to visualise the LULC transitions between the different time periods, which is why they were used in this study. The Sankey diagram was plotted using the e! Sankey software. In addition to the Sankey diagram, the percentage gains and losses were also calculated for each class, in relation to the total surface area of the transboundary basin.

6.3 Results

6.3.1 Accuracy assessment results

Figure 6.2 shows the Overall Accuracy (OA) results for the classification of the mosaicked image representing the studied time periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015, Period 4: 2016-2020). The results show that the OA values were within an acceptable range, namely, between 79% and 87%. The highest OA value was observed for Period 1 (2000-2005), which had an OA value of 86.59%, and the lowest OA was observed for Period 2 (2006-2010), which had an OA value of 79.64%.

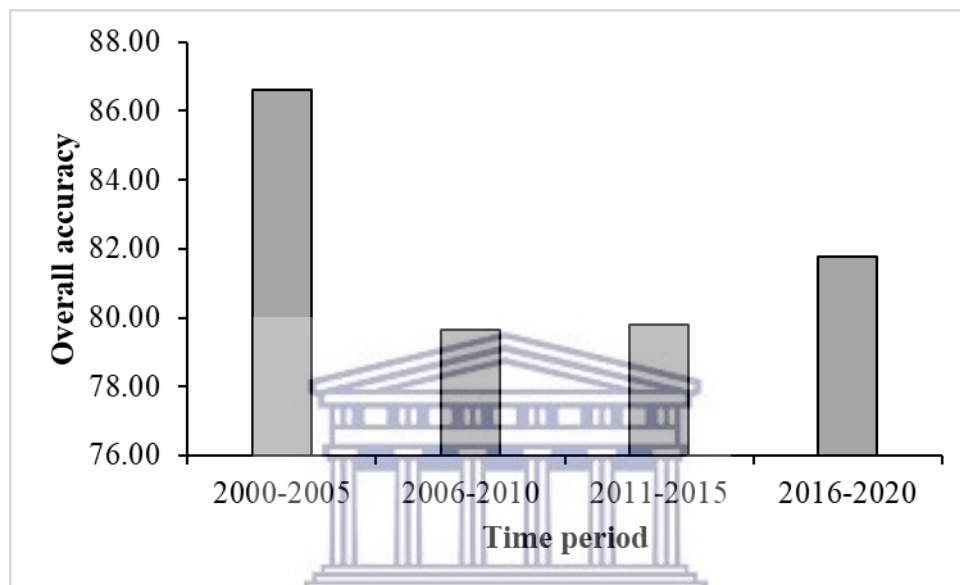


Figure 6.2 Overall accuracy for each time period (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020)

The producer's and user's accuracy results per time period are presented in Figure 6.3. The results show that the bare surface, shrublands and open water classes had a higher producer's and user's accuracy than all the other classes, for all the time periods. The producer's and user's accuracy for these classes were above 70% for all the time periods. Although other classes, such as the built-up and forest presented a higher producer's and user's accuracy (>60%) for some periods, such as 2000-2005, these classes also reported a low producer's and user's accuracy for the 2006-2010 and 2011-2015 time periods. The bare surface also reported a low producer's accuracy (<20%) for the 2016-2020 time period. Croplands presented an acceptable producer's and user's accuracy ranging between 50% and 60% for all time periods, except for the 2006-2010 time period, where both of these metrics were below 40%. The wetlands class had the lowest producer's and user's accuracy, which ranged between 30% to 45% for the all

periods. Further, sparse vegetation had a producer's and user's accuracy of less than 20% for all the cases.

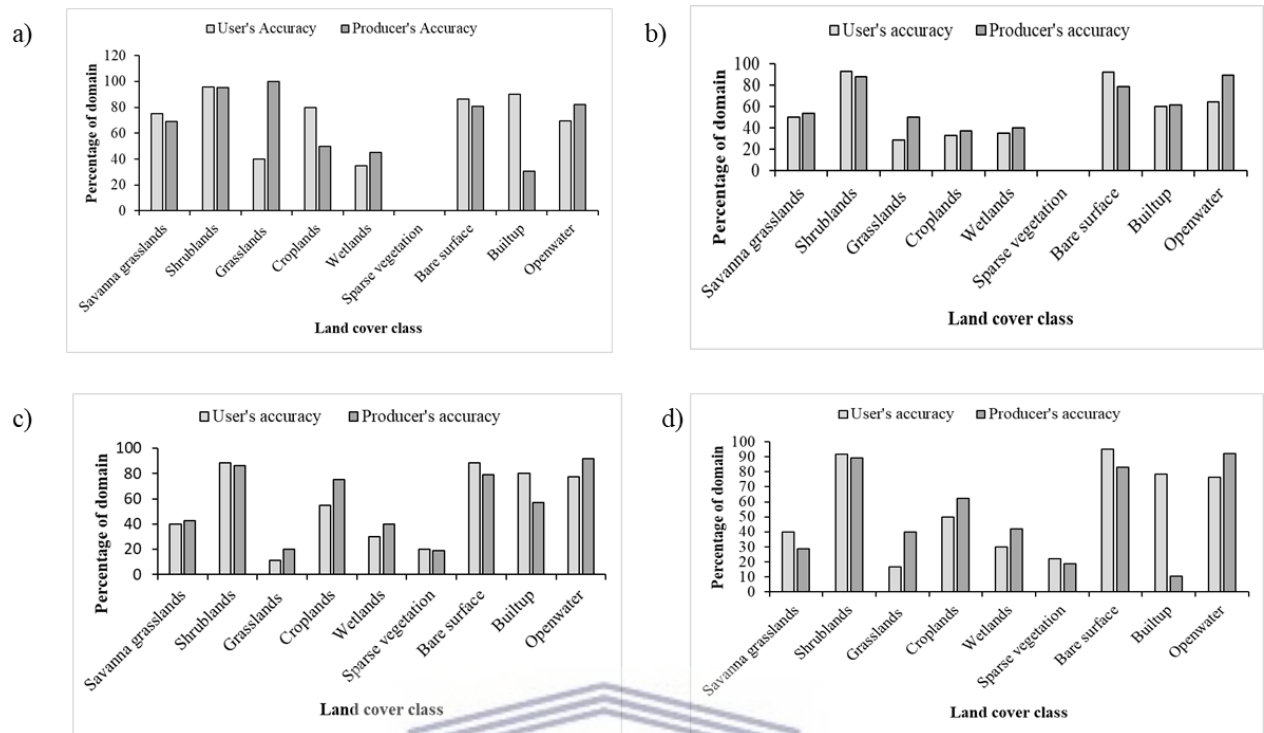


Figure 6.3 User's and producer's accuracy per studied time period. a) Period 1 (2000-2005), b) Period 2 (2006-2010), c) Period 3 (2011-2015) and d) Period 4 (2016-2020)

The Pontius commission and omission error results are presented in Figure 6.4; they show that classes, such as sparse vegetation and wetlands, had high commission and omission errors of above 50% for most of the time periods, which is indicative of the possible misclassification of these classes. Moreover, classes such as grasslands, presented high commission and omission errors of above 60% for the 2011-2015 and 2016-2020 time periods, while the forest class presented omission errors that were above 50% for the 2011-2015 and 2016-2020 time periods. Classes, such as shrublands, croplands, open water, built up and bare surfaces, reported low commission errors (<30%) for all the time periods, except for the 2006-2010 time period, where built-up areas reported a commission error of above 50%.

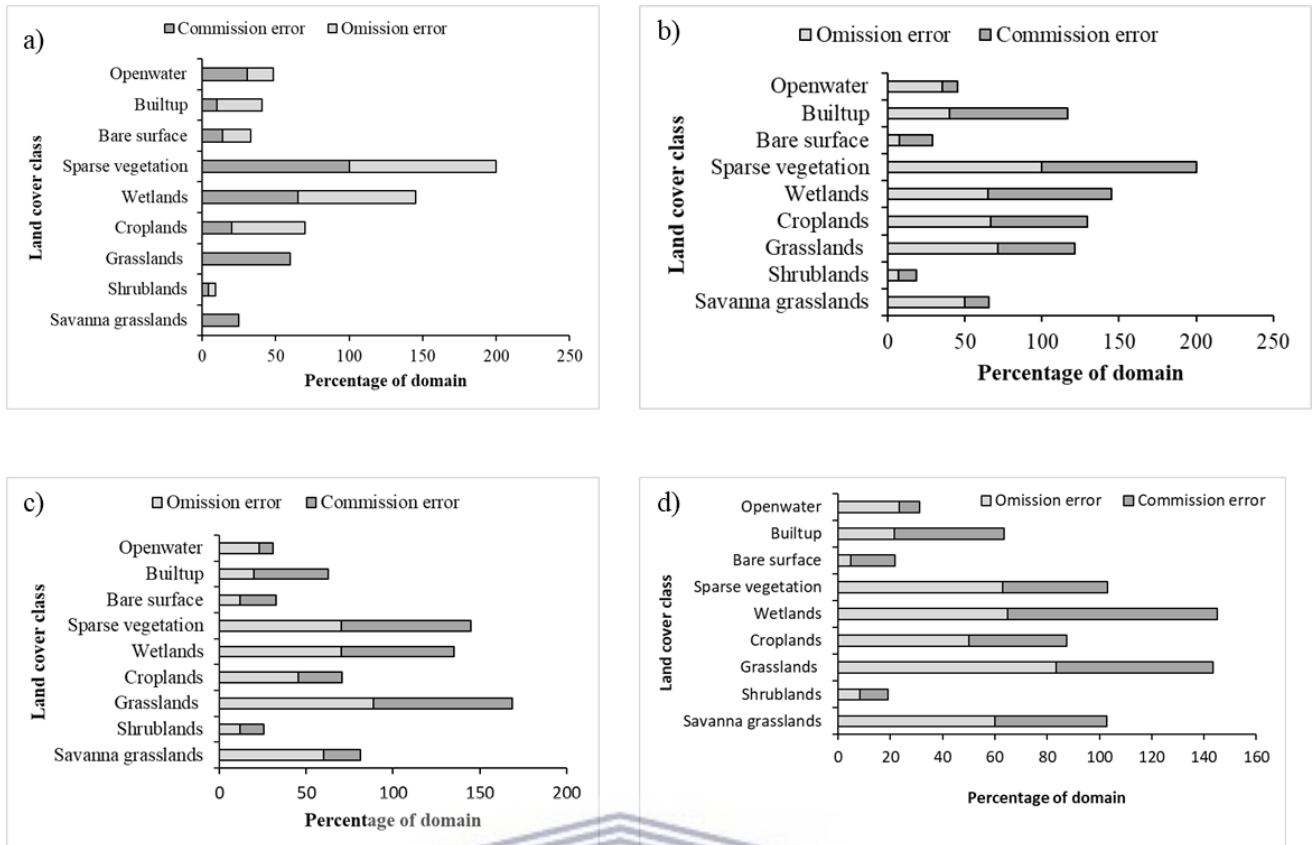


Figure 6.4 Pontius commission and omission errors for each class per time period. a) Period 1 (2000-2005), b) Period 2 (2006-2010), c) Period 3 (2011-2015) and d) Period 4 (2016-2020)

6.3.2 Land cover change analysis

The spatial distribution of different wetland-cover classes, based on the pixel-based classification for each time period, is shown in Figure 6.5. The results show that shrubland was the most dominating land cover class for all the study periods, with the class areas ranging between 36 548 811.54 ha and 3 491 650.28 ha. Wetlands and sparse vegetation were found to be the least dominating classes for all the time periods. The class areas ranged between 14 511.2 ha and 32 241 ha for wetlands, while the sparse vegetation ranged between 21.33 ha and 4 937.4 ha, with the lowest being recorded for Period 4 (2016-2020) for both classes (Table 6.3). The results also indicated a continued pattern of growth in croplands, savanna grasslands and shrublands from Period 1 (2000-2005) to Period 4 (2016-2020). Built-up areas and bare surfaces exhibited a growth pattern between Period 1 (2000-2005) and Period 2 (2006-2010). Although built-up areas exhibited a pattern of growth between Periods 1 and 2, an anomaly was observed for this class between Periods 2 and 4, where the area continued to decline. This anomaly was probably due to classification errors.

Table 6.3 Land cover class areas (ha) per time period

Landcover	2000-2005	2006-2010	2011-2015	2016-2020
Savanna grasslands	3095709.93	3481055.01	3487177.13	3491650.28
Shrublands	36548811.54	36779118.3	36842378.4	37224435.17
Grasslands	1193094.27	1432914.3	1454202.18	1539978.93
Croplands	1254567.78	1623800.97	1681331.32	1754400.24
Wetlands	32241	31401	14901.85	14511.2
Sparse vegetation	4937.4	4066.11	3440.14	21.33
Bare surface	1795877.37	2108955.33	1379135.16	366324.66
Built-up	682773.66	1020318.03	523968.66	480733.02
Open water	470450.25	395702.55	372540.96	231446.07

**hectares (ha)*



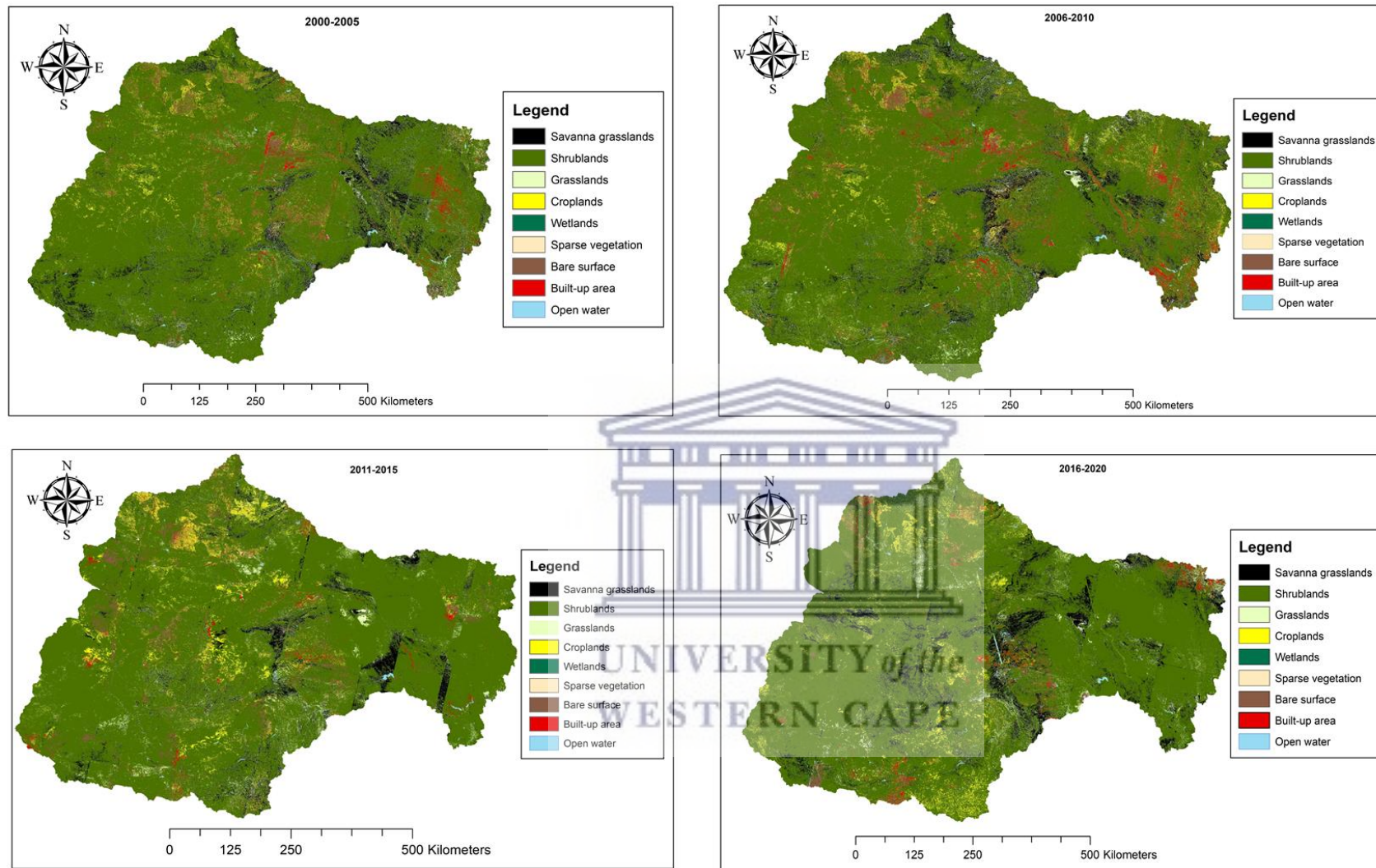


Figure 6.5 Spatial distribution of LULC cover classes, based on the pixel-based classification for the study time periods

6.3.3 LULC transitions and impacts on the extent of wetlands

Figure 6.6 shows the LULC transitions between the time periods, and the major transitions (>20% loss or gain) are labelled in their respective proportions. The results showed that major transitions were observed for classes such as wetlands, sparse vegetation grasslands, as well as the bare surface between Period 1 (2000-2005) and Period 2 (2006-2010). During the abovementioned time periods, about 40% of the wetland areas was replaced by built-up areas, 61% of sparse vegetation was replaced by savanna grasslands, 40% of the grasslands was also replaced by savanna grasslands, and 20% of the bare surface was replaced by croplands. Between Period 2 (2006-2010) and Period 3 (2011-2015), major transitions occurred between the sparse vegetation and savanna grasslands, the grasslands and savanna grasslands, the bare surface and croplands, as well as the built-up and bare surface. In between these time periods, 61% of the sparse vegetation was replaced by savanna grasslands, 37% of the grasslands was replaced by savanna grasslands, 25% of the bare surface was replaced by croplands and 20% of built-up areas was replaced by bare surface. The results also show that between Period 3 (2011-2015) and Period 4 (2016-2020), major LULC transitions were observed between the bare surface and croplands, the built-up area and croplands, the wetlands and shrublands, the grasslands and savanna grassland, as well as the sparse vegetation and savanna grasslands. The results indicate that between the abovementioned time periods, about 36% of bare surface was replaced by croplands and 32% of the built-up area was replaced by croplands. In addition, 46% of grasslands was replaced by savanna grasslands, 63% of sparse vegetation was replaced by savanna grasslands and about 30% of the wetland area was replaced by shrublands. Figure 6.6 also shows a continuous decline in the wetland area and sparse vegetation, while other landcover classes, such as croplands, savanna grasslands and shrublands, increased continuously between the Period 1 and Period 4 time periods, which are indicated by the size of the combined arrows.

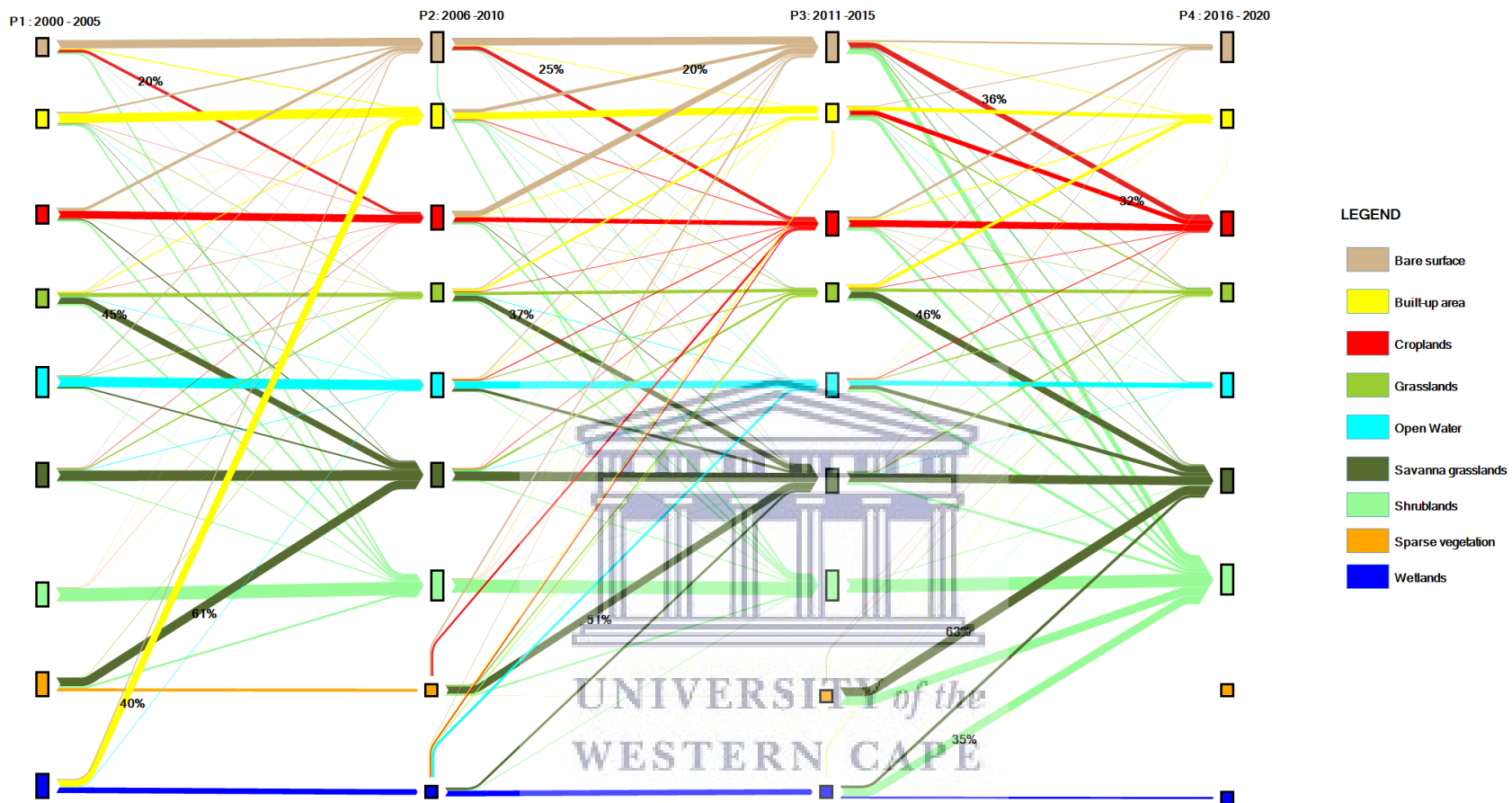


Figure 6.6 LULC transitions between the time periods, with the major transition labelled with the respective proportions

6.4 Discussion

This study analysed the impacts of the large-scale LULC changes on the extent of semi-arid wetlands in southern Africa over a 20-year period (2000-2020), using the Limpopo Transboundary River Basin as a case study. The classification results yielded a total of 9 land-cover classes, which included savanna grasslands, shrublands, grasslands, croplands, wetlands, sparse vegetation, bare surface and built-up areas, with shrublands being the dominating class for all the time periods, while wetlands and sparse vegetation were the least dominating. The wetlands were expected to be the least dominating because the study used dry season images, because of the unavailability of images with a minimum cloud coverage. Since semi-arid areas are dominated by small and intermittently-flooded systems, which tend to merge with the surrounding terrestrial ecosystem (Day et al., 2010; Fang et al., 2019), they may not have been picked from the images that were used because they were not visible at the time when these images were captured. A sub-basin study that was conducted within the LTRB in Mozambique, using the dry season Landsat 5 and 7 and ASTER images for the years 1999-2003, also reported a low spatial coverage of water bodies, including wetlands, for the regions, in response to the seasonal variability (Pereira, 2004).

The results also showed that the shrublands, savanna grasslands and croplands continued to increase, while the wetlands and sparse vegetation continued to decline. Moreover, they indicate that about 40% of the wetlands were converted to built-up areas, which shows that the continued decline was mainly driven by built-up expansion (anthropogenic activities) caused by the increasing population, which increases the demand for land, in order to further develop the residential areas. This has also been reported in a number of LULC analyses in the LTRB, at a sub-basin level. For example, a study by Thamaga et al., (2022) analysed the impacts of the LULC changes on unprotected systems in one of the sub-basins of the LTRB located on the South African side. The results showed that urbanisation was one of the major drivers for wetlands loss, thus corroborating the findings of the current study. Another study by Sibanda and Ahmed (2021) modelled the impacts of the LULC changes on the extent of wetlands in the Shashe Catchment, a sub-basin of the LTRB that is located on the Zimbabwean side. The results showed that some of the wetland areas were replaced by built-up areas, which they also corroborate the findings of this study. Other studies from different parts of the semi-arid Africa and beyond (e.g. Marambanyika et al., 2017; Mwita, 2013; Chikodzi and Mufori, 2018) also reported the impact of LULC changes on wetlands, particularly urbanisation, although these studies were not conducted on a larger scale. The other major LULC transition observed in this

study was between the bare surface and cropland areas. This can be attributed to the planting cycles of different crops and their growth stages. For some periods, the plant species were not fully-grown, which resulting in a low canopy cover that exposed the bare surface (Ziter et al., 2019). In this study, the images were collected on different dates as such, for some dates the crops were more mature than others. This finding was also reported in a study by Sibanda and Ahmed (2021). The explanation for the transition between wetlands, grasslands and shrublands is the prolonged moisture availability in the wetland soils, which favours the growth of different types of vegetation, including grasslands and shrublands, particularly because the study area is semi-arid (Day et al., 2010).

The Overall Accuracies (OA) for all the time periods were within an acceptable range, and were between 78% and 87%. However, some classes, such as wetlands and sparse vegetation, reported high commission and omission errors, as well as low producer's and user's accuracies for some time periods. These errors could have resulted from the dates on which the images were used and the spatial resolution of the data used. In this study, the cloud-free images that were used were mostly available for the dry season, and semi-arid wetlands are mostly seasonal and tend to merge with the surrounding terrestrial ecosystems during the dry season (Day et al., 2010), which makes them difficult to monitor when using dry season images as some of the wetlands features, such as the surface inundation, are not clearly visible and some wetland vegetation species are not healthy.

In addition, an imbalance in the training and validation data, which was caused by the low spatial coverage of wetlands and sparse vegetation, could have resulted in high commission and omission errors, as well as low producer's and user's accuracy. Imbalanced training data tend to introduce biasness during classification towards the class that is mostly represented in the dataset resulting in a high accuracy for that particular class and low accuracies for the under-represented classes. Studies by Millard and Richardson (2013), Ustuner et al., (2016) and Amani et al., (2021) have shown that an imbalance in the training data tends to introduce biasness in the classification outputs, where the under-represented classes are likely to be classified more inaccurately than the represented classes. In this study, wetlands and sparse vegetation had the lowest number of training and validation points, resulting from the limited spatial coverage of these classes in the Limpopo Transboundary River Basin (LTRB). This could therefore have introduced a bias towards the other classes that had a high number of training and validation points, thus causing these classes to be classified more accurately than the wetland and sparse vegetation class.

Overall, this chapter provides new insights about the diminishing state of semi-arid wetlands in southern Africa, particularly in the LTRB. These findings are critical for environmental planners and conservation specialists, as they provide the baseline information that is required to develop proper strategies and that will assist in curbing the negative impacts that these LULC cover changes have on wetlands. In particular, the conservation status of the small and seasonally-flooded wetlands is currently overlooked because of their size and their intermittent nature, despite the fact that they provide significant eco-hydrological services to the surrounding communities. Proper and inclusive wetland management strategies on a regional scale, such as those in transboundary basins, require adequate integrated data collection, reliable information and reviews. On a larger scale, for example in transboundary basins, the collection of proper data that produce reliable information and that will inform proper LULC management policies, is hindered by the limited resources. This chapter therefore provides robust, cost-effective and efficient methodologies that can assist in the generation of the required knowledge that is acceptable and precise, which can inform the LULC management policies, as well as wetland conservation and management strategies, on a larger scale.

6.5 Conclusions

This chapter analysed the impacts of large-scale LULC changes impacts on the semi-arid wetlands systems in southern Africa over a 20-year period (2000-2005), using the Limpopo Transboundary River Basin (LTRB) as a case study. The findings showed that 9 land-cover classes were characterised. These included savanna grasslands, shrublands, croplands, grasslands, wetlands, sparse vegetation, bare surface and built-up areas. Shrublands were found to be the most dominating class throughout the study periods and covered proportion ranging between 76% and 82% of the study area, while wetlands and sparse vegetation class were the least dominating classes, covering between 0.9% and 2%, and 0.3% and 0.04% of the study area, respectively. It was also observed that the wetlands and sparse vegetation continued to decline at an average rate of 19% and 44%, while shrublands, croplands and savanna grasslands continued to increase at an average rate of 0.4%, 12.4% and 4.25%, respectively. Most of the wetland area (40%) was replaced by built-up area, which indicates that urbanisation is a major driving factor in the loss of wetlands in the area. The accuracy analysis results showed that the OA was within an acceptable range (77%-86%), although some classes, such as wetlands and sparse vegetation, reported a lower accuracy, which is attributed to the unbalanced training and validation data that was caused by the low spatial coverage of these classes in the LTRB. This study provides new insights into the diminishing state of the semi-arid wetlands in southern

Africa, particularly in the LTRB. These findings are critical for environmental planners and conservation specialists, as they provide baseline information for the development of proper strategies that will help to curb the negative impact that these LULC cover changes have on the wetlands. Moreover, the chapter provides robust, cost-effective and efficient ways to assist with the generation of the required, acceptable and precise knowledge for informing the LULC management policies, wetlands conservation and management strategies on a larger scale.



CHAPTER 7

AN INTEGRATED REMOTE-SENSING FRAMEWORK FOR THE DETECTION AND MONITORING OF SMALL SEASONALLY- FLOODED WETLANDS IN SEMI-ARID SOUTHERN AFRICA: A SYNTHESIS

7.1 Introduction

This work aimed at developing an integrated remotely-sensed framework that will assist in the detection and monitoring of small and seasonally-flooded wetlands in semi-arid southern Africa. The data, methodologies and results in this study have direct implications for conservation planning and for the management and wiser use of wetlands in semi-arid environments, particularly in data-scarce regions like southern Africa. They will provide cost-effective and feasible methodologies for generating information on the condition and extent of the wetlands. Moreover, the study demonstrates how the improved remote sensing approaches, such as introduction of cloud-computing platforms like GEE, are cost-effective and efficient tools for assessing and monitoring the changes in small and seasonally-flooded wetlands located in semi-arid areas, which was previously a daunting task when using the traditional classification approaches because of the intermittent nature and size of these wetlands. The realisation of the overall goal of this study was achieved by means of a series of interlinking manuscripts and publications that address the main objectives of the study outlined in Chapter One. These publications and manuscripts are outlined as follows:

- 1) A review on the progress, challenges and future research direction of the multispectral remote sensing of wetlands in semi-arid and arid areas.
- 2) Leveraging the recently-available GEE cloud-computing platform to characterise and map small and seasonally-flooded wetlands in the semi-arid environments of South Africa. This publication addresses objective one of the studies.
- 3) Available remotely-sensed data in the GEE catalogue, which are capable of monitoring the long-term variations in the eco-hydrological dynamics of small and seasonally-flooded wetlands in semi-arid South Africa. This manuscript addresses objective two of the study.
- 4) Using cloud-computing techniques to monitor the long-term variations in eco-hydrological dynamics of small and seasonally-flooded wetlands in semi-arid South Africa. This publication addresses objective three of the study.

- 5) Exploring the capabilities of the GEE cloud-computing platform and its advanced machine-learning algorithms for the long-term monitoring of the impacts of LULC changes on semi-arid wetland systems, on a regional scale. This manuscript addresses objective four of the study.

7.2 A Review of the Progress, Challenges and Future Research Directions of the Multispectral Remote Sensing of Wetlands in Semi-arid and Arid Areas

This section focused on providing a comprehensive review on the progress, challenges and the future research directions in the remote sensing of small seasonally-flooded wetlands (<10 ha –2 500 ha) in semi-arid and arid areas, using freely-available multispectral remotely-sensed data. Although a significant number of reviews have been published on the use of remote sensing for the monitoring and assessment of wetlands in different parts of the world and in different climatic zones, none of them have focused only on semi-arid and arid systems (Adam et al., 2010; Dronova, 2015; Guo et al., 2017; Klemas, 2014). This review therefore focused on the multispectral remote sensing of wetlands found within semi-arid and arid areas. A systematic analysis of the papers published between 2000-2020, which is within the scope of this review, showed an increasing trend ($R^2 = 0.76$) in the number of studies published during that period, with most of them increasing significantly after the year 2008. This was attributed to the fact that most remotely-sensed data, such as the Landsat series, became freely-available to all users after 2008 (Zhu et al., 2019). The characterisation (mapping and classification), inundation and vegetation of wetlands were the most-studied aspects, while their extent and the land use land cover changes were the least-studied aspects. The spatial distribution analyses in this study revealed that most of them were conducted in semi-arid China, India, Australia and Canada. Although significant research strides have been made with regard to the use of freely-available multispectral data for studying small and seasonally-flooded wetlands in semi-arid areas, the review noted that challenges, such as mapping these wetlands with the highest precision, is still an issue of concern, because some of them have been missed or confused with other land-cover features. This has resulted in the conservation status of these systems being overlooked, as the information on their eco-hydrological dynamics at an appropriate spatial scale is limited. The review also noted that improvements in remote sensing techniques, such as the introduction of big data analytic tools such as the Google Earth Engine (GEE), offer unique opportunities to address the issues relating to the assessment and monitoring of small and seasonally-flooded wetlands. With their advanced processing power, data analysis tools and algorithms offer the benefit of improving the assessment and monitoring

of these systems. However, the application of such tools in the mapping, monitoring and understanding of small and seasonally-flooded wetlands is still in its infancy in semi-arid areas, particularly in the sub-Saharan African region. This study therefore took advantage of the readily-available GEE cloud-computing platform and its benefits to assess and monitor the small and seasonally-flooded in semi-arid southern Africa, with the aim of improving the detection and monitoring of these systems by using freely-available, remotely-sensed data.

7.3 The Characterisation and Mapping of Small Seasonally-flooded Wetlands in the Semi-arid Environments of South Africa, using the Google Earth Engine

Although the Google Earth Engine offers the benefit of improving the detection and monitoring of small and seasonally-flooded wetlands in semi-arid areas, its application to these systems and other wetland types is still in its infancy. This section is based on publication two which addresses objective one of the study, and was informed by the literature review papers of Tamiminia et al., (2020) and Kumar and Mutanga (2018), which indicated that the use of the GEE platform wetlands studies is still in its infancy. Based on that, it therefore explored the use of this platform in the characterisation and mapping of small and seasonally-flooded wetlands in the semi-arid environment of South Africa, using two seasonal systems, namely the Nylsvley floodplain and the Lindani valley bottom, as case study wetlands. In order to achieve its main goal, this study assessed the spectral separability of different wetland-cover classes that were detected in the field surveys, by using the GEE and multi-year Sentinel-2 composite derivatives, coupled with the Jeffries Matusita (JM) distances, to evaluate the capabilities of the GEE cloud-computing platform in producing customised wetland cover maps at a reasonable accuracy, using the high-resolution Sentinel-2 data and its advanced machine-learning algorithms (Random Forest (RF), Support Vector Machine (SVM), Classification Regression Tree (CART), as well as Naïve Bayes (NB)). It also identified a suitable GEE machine-learning algorithm for accurately detecting and mapping the characteristics of semi-arid seasonally-flooded wetlands, using the multi-year Sentinel 2 composite data.

The classification results yielded overall accuracies that were within an acceptable range, for both systems (69%-78%), although some classes, such as bare surface long grass and water, reported low class accuracies for some algorithms, such as NB. Moreover, the spectral separability results showed that the synergic use of spectral data (Visible Red, Green and Blue, as well as the Near-infrared), coupled with the vegetation and water indices (NDWI, NDVI

and MSAVI2), improved the separability of different wetland classes with JM distances of around 2. The comparative analysis of the overall accuracies, as well as the class accuracies, showed that RF outperformed the other algorithms for the Lindani valley bottom, while RF, CART and SVM outperformed the NB in the case of the Nylsvley floodplain. Based on these findings this section concluded that, these results underscore the relevance of the GEE with its advanced machine-learning algorithms (RF, SVM and CART), coupled with the synergy of visible bands Near-infrared, as well as vegetation and water indices in improving the characterisation and mapping of small and seasonally-flooded wetlands in semi-arid environments, which was a challenge based on traditional mapping techniques.

7.4 Available Remotely-sensed Data in the GEE Catalogue, which are capable of Monitoring the Long-term Variations in the Eco-hydrological Dynamics of Small and Seasonally-flooded Wetlands in Semi-arid South Africa

Freely-available, higher-resolution multispectral data offer unprecedented opportunities to monitor the long-term changes in the eco-hydrological dynamics of small and seasonally-flooded wetlands in semi-arid areas. However, obtaining time-sensitive data and the generation of precise and accurate information on the wetland dynamics still presents some challenges. The availability of cloud-computing platforms, such as the GEE, offers an opportunity to address the issues associated with data availability, as it hosts 40 years of peta bytes of freely-available, remotely-sensed data, ranging from coarse to finer resolution data (Gorelick et al., 2017). Moreover, these platforms offer specialised data processing, filtering and integration features, which provide an opportunity to improve the classification of wetlands from remotely-sensed data. Although these platforms offer such benefits, uncertainties with respect to data continuity still exist (Wang and Yésou, 2018). Moreover, the capabilities of the available data, in terms of the image quality (cloud-free) for monitoring small and seasonally-flooded wetlands is unknown, particularly for cloud-prone areas (Wua et al., 2019). This section was based on manuscript one and addressed objective two of the study which was aimed at evaluating the available remotely-sensed data in the GEE catalogue that are capable of monitoring the long-term (2000-2020) variations in the eco-hydrological dynamics of small and seasonally-flooded wetlands in semi-arid southern Africa, using the Nylsvley floodplain as a case study. In order to realise this goal, the chapter established the number of scenes available, with different cloud-cover ranges, for various remotely-sensed products on the GEE, using specialised filtering algorithms. The study also determined how slight variations in the cloud-cover ranges (0%, 1-10%, 11-25%) affected the monitoring of the wetland eco-

hydrological dynamics across different seasons, by using multi-year Landsat-7 and Sentinel-2 composite images representing each season, as well as the Random Forest machine-learning algorithm.

The results showed that approximately 1 376 images were found for the period under study. These included 492 images for Landsat series data, 394 images for Sentinel-1 and 490 images for Sentinel-2. The Moderate Resolution Imaging Spectroradiometer (MODIS) was not included in the analysis because of the spatial resolution of the data, which makes it impossible to detect small and seasonally-flooded wetlands. Sentinel-2 and Landsat-7 had the highest number of images (69% and 76%) with a cloud-cover percentage range of 0-20%, while other products, such as Landsat 4, 5 and 8, had the lowest number of images. In determining the effects and impacts of cloud cover on wetlands detection, only images with a cloud cover of less than 26% and products with the most scene were used (Landsat-7 and Sentinel-2). The use of satellite images with 0% cloud-cover yielded an Overall Accuracy (OA) of between 69-72%, 1-10% had an OA ranging between 68-70%, and 11-25% had an OA ranging between 69-80.55%, for both the dry and wet seasons. Generally, the classification results indicated satisfactory overall accuracies (68-80%) for all the scenes, although there were some inaccuracies for some classes (e.g. bare surfaces and long grass), particularly when using Landsat-7 scenes, and this was associated with the sensing characteristics of the data used. Based on these findings this chapter concluded that there is a reasonable amount of satellite data available that are capable of monitoring small and seasonally-flooded wetlands, which will provide useful insights into the eco-hydrological dynamics of these ecosystems. In addition, the use of cloud-computing platforms, such as the Google Earth Engine (GEE), provide a unique opportunity for addressing the problems associated with big data filtering, processing and analytics, in order to improve the monitoring and assessment of these critical ecosystems.

7.5 Monitoring the Long-term Variations in the Eco-hydrological Dynamics of Small and Seasonally-flooded Wetlands in Semi-arid South Africa, using Cloud-computing Techniques

This section was based on the publication four which addressed objective three of the study focusing on assessing the potential of cloud-computing techniques, particularly the Google Earth Engine, in monitoring the long-term variations in the eco-hydrological dynamics of small and seasonally-flooded wetlands in semi-arid South Africa. In order to realise this goal, the

chapter evaluated the spatio-temporal variations in the eco-hydrological dynamics of wetlands by using time-series Landsat composite data, coupled with the Random Forest machine-learning algorithm, and established the major drivers of the eco-hydrological changes in the studied systems by using remotely-sensed metrics (NDVI, NDWI and NDPI), coupled with climate data. The chapter was motivated by the lack of critical data on wetland eco-hydrological dynamics at an appropriate spatial and temporal scale, which hinders the setting up of proper management plans for the wiser use and conservation of wetlands in semi-arid and arid areas, particularly in the sub-Saharan African region. Moreover, although the availability of cloud-computing platforms, such as the Google Earth Engine, offer numerous advantages for addressing the issues of limited and inconsistent data for wetland monitoring, the use of this platform for monitoring the long-term changes in the eco-hydrological dynamics of semi-arid wetlands requires further evaluation (Wua et al., 2019). Studies that have utilised this platform for the long-term monitoring of environmental change, include those on forest mapping, crop mapping and open water mapping (Kumar and Mutanga, 2018; Tamiminia et al., 2020). The most recent studies that have utilised the GEE in wetland mapping have demonstrated the potential value of the platform in wetland science (Gxokwe et al., 2021; Zhou et al., 2019). Although these studies underscore the relevance of the GEE for investigating small seasonally-flooded wetlands, they did not consider the longer-term changes that can be assessed by using the time-series analysis in the GEE platform. This study therefore assessed the use of the platform for monitoring the long-term variations in the eco-hydrological dynamics of small and seasonally-flooded wetlands that are found in semi-arid southern Africa.

The results showed that the eco-hydrology of the studied system changed over time; where some wetland-cover classes increased twice as much as in the previous measurements, and others decreased significantly during the study period. Notably, the bare surface class increased at a rate of 230% and 350% between 2006-2010 and 2016-2020, respectively. The indices also showed similar trends throughout the 20-year period, with NDWI having the lowest values of less than zero in all cases, which implies that there was no surface inundation; however, the presence of some wetland vegetation indicates seasonal to semi-permanent soil saturation conditions. A comparative analysis of the climate data and remotely-sensed indices revealed that annual changes in the precipitation and evapotranspiration were the main drivers of wetland eco-hydrological variations. This section, concluded that the findings demonstrated the usefulness of the GEE and its advanced machine-learning algorithms for monitoring the eco-hydrological dynamics of small seasonally-flooded wetlands in semi-arid environments

over a long period of time. This information is critical for setting up appropriate measures for the conservation and wiser use of these wetlands, as it provides robust methodologies that will provide an understanding of their status.

7.6 Assessing Large-scale LULC Changes and Impacts on Wetlands using Cloud-computing and Multisource Remotely-sensed Data

An analysis of the impacts of LULC changes on the semi-arid wetlands on a larger scale, over a 20-year period, were conducted in the Limpopo Transboundary River Basin in southern Africa. The study was motivated by the limited number of studies that have focused on understanding the impacts of LULC changes on semi-arid wetlands, on a larger scale (Thamaga et al., 2022). It was noted that the existing studies focus either on one aspect of LULC change, or on its impact on large and protected systems, but they have overlooked the small and seasonally-flooded wetlands, despite the fact that these provide critical eco-hydrological services to the surrounding communities (Thamaga et al., 2022). The study utilised the Google Earth Engine platform and its advanced machine learning techniques, coupled with multisource remotely-sensed data (Sentinel-1, Landsat 5 and 8), to examine the LULC changes between 2000-2020 and how they relate to the extent of the wetlands in the LTRB. This chapter was based on manuscript three, which addressed objective four of the study outlined in chapter one.

This section identified 9 landcover classes, namely, shrublands, croplands, bare surfaces, wetlands, sparse vegetation, grasslands, built-up areas and savanna grasslands. Shrubland was the most dominating class throughout the study period, covering between 76% and 82% of the LTRB surface area, while wetlands and sparse vegetation were the least dominating, covering between 0.9% and 2% and 0.3% and 0.04%, respectively. The overall accuracies were within an acceptable range (77%-78%), although some classes, such as wetlands and sparse vegetation, presented low class accuracies for some periods. This was attributed to the unbalanced training and validation data caused by the low spatial coverage of these classes in the studied area. It was also observed that wetlands and sparse vegetation continued to decline at an average rate of 19% and 44%, respectively, while shrublands, croplands and savanna grasslands continued to increase at an average rate of 0.4%, 12.4% and 4.25%, respectively. Most of the wetland area (40%) was replaced by built-up area, which indicates that urbanisation is a major driving factor in the loss of wetlands in the area. This section concluded that the findings provide new insights into the diminishing state of small and seasonally-flooded wetlands in semi-arid areas, which are critical for conservation planners, as they provide them

with the baseline information required to develop proper strategies that will assist in curbing the negative impacts of these LULC cover changes in wetlands. Moreover, the study demonstrated the benefits of data integration on the GEE platform, which provides robust and efficient ways of generating the required knowledge with acceptable precision, in order to properly inform the LULC management policies, as well as the wetlands conservation and management strategies.

7.7 Overall Implications of the Results of Semi-arid Wetland Detection and Monitoring, using Remotely-sensed Data

The study demonstrated how integrated remote sensing and the cloud-computing platform can assist in improving the detection and monitoring of small seasonally-flooded wetlands in the semi-arid environments of southern Africa, which was previously a daunting task when using the traditional mapping techniques. The overall implications of the investigation into the detection and monitoring of semi-arid seasonal flooded wetlands using remote sensing based on the findings of the study are as follows:

The first section on the characterisation and mapping of small and seasonally-flooded wetlands demonstrated the feasibility of the GEE and its advanced data processing algorithms for mapping the wetlands in these regions. More specifically, it showed that the use of multi-data images composited from the median algorithm allowed for the integration of the strengths from different scenes, in order to better enhance the wetlands features, rather than relying on a single data image. It also assisted in the minimisation of the effects of clouds and haze, and it demonstrated that an object-based analysis and the synergic use of spectral bands (Visible Red, Green and Blue as well as Near-infrared band), as well as water and vegetation indices, greatly enhanced the detection of different wetland features. It also proved that the Random Forest (RF), the Support Vector Machine (SVM), as well as the Classification and Regression Tree (CART) machine-learning algorithms that are implemented in an object-based analysis, will most likely improve the detection and monitoring of small and seasonally-flooded wetlands.

The second section of this thesis demonstrated the available remotely-sensed data on the GEE cloud-computing platform that can be used to monitor the eco-hydrological changes of small and seasonally-flooded wetlands. Based on these findings, it was noted that a reasonable amount of remotely-sensed Landsat and Sentinel-1 and 2 good-quality data are available for the long-term monitoring of small and seasonally-flooded wetlands. It was also demonstrated that careful attention should be given, when selecting images for analysing these wetlands, as

some wetland features become more visible than others in certain seasons. In addition, the cloud cover of the selected images should be clear (less than 20%), as this is one of the major sources of error during the classification of these wetlands.

The third section demonstrated the feasibility of the GEE and its advanced machine-learning algorithms, particularly RF, in the monitoring of long-term eco-hydrological variations and their drivers. The findings of this section revealed that the addition of the Normalized Difference Phenology Index (NDPI) to the synergy of Visible Red, Green, Blue, Near-infrared bands, NDVI and NDWI, improved the detection and monitoring of the eco-hydrological dynamics of these wetlands. In this study, NDPI allowed for the better detection of moisture and vegetation water content, as well as the separation of the bare soil from the unhealthy short wetland vegetation. This also allowed for the inference whether the wetlands are really seasonally, semi-permanent or permanently saturated.

The last section of this study demonstrated how the integration of multisource data on the Google Earth Engine is beneficial for monitoring the large-scale impacts of LULC changes on these wetlands. The findings revealed that multi-source and multitemporal data integrations, such as the combination of Synthetic Aperture Radar data (Sentinel-1) and optical products like Landsat-8, improved the monitoring of the LULC changes and their impact on wetlands in semi-arid areas. Overall this study provides an overall guiding framework, for the improved detection and monitoring wetlands using cloud computing approaches, in particular GEE coupled with freely available remotely-sensed data- a previously daunting from traditional image processing analysis.

7.8 Overall Recommendations

Although the study noted that the GEE platform offers new ways for improving the detection and monitoring of small and seasonally-flooded wetlands, there are still challenges associated with the use of the GEE, for example the computational time costs, particularly for a large-scale analysis, which can result in computational time-out errors. The study recommends the evaluation of the splits and parallel processing of the images, which has not been applied much, although researchers like Shafizadeh-Moghadam et al., (2021) who used the GEE platform for large-scale mapping and recommended the approach. It also has implications for the use of classified thematic layers after the parallel process split images have been merged. It is therefore necessary to consider the following future research:

- An exploration into the application of methodologies, such as the deep Artificial Neural Networks (ANN) and Convolutional Neural Network (CNN) on the GEE cloud-computing platform for the detection and monitoring of semi-arid wetlands.
- An analysis of the impacts of the groundwater level changes and the impacts of the eco-hydrological dynamics of semi-arid wetlands, on a regional scale, using GEE, as some of these systems are groundwater fed.
- The study also recommends an exploration into the use of Synthetic Aperture Radar (SAR) data for the assessment and monitoring of small and seasonally-flooded wetlands, as it not influenced by cloud coverage and it is able to penetrate through forested vegetation cover, because of its wavelengths.



REFERENCES

- Abdi, A.M., 2020. Land cover and land use classification performance of machine-learning algorithms in a boreal landscape, using Sentinel-2 data. *GIScience Remote Sens.* 57, 1-20. <https://doi.org/10.1080/15481603.2019.1650447>
- Achanta, R., Süsstrunk, S., 2017. Superpixels and polygons using simple non-iterative clustering, in: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. pp. 4895–4904. <https://doi.org/10.1109/CVPR.2017.520>
- 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 the random forest and support vector machines classifiers. *Int. J. Remote Sens.* 35, 3440–3458. <https://doi.org/10.1080/01431161.2014.903435>
- Adam, E., Mutanga, O., Rugege, D., 2010. Multispectral and hyperspectral remote sensing for the identification and mapping of wetland vegetation: A review. *Wetl. Ecol. Manag.* 18, 281–296. <https://doi.org/10.1007/s11273-009-9169-z>
- Adeeyo, A.O., Ndlovu, S.S., Ngwagwe, L.M., Mudau, M., Alabi, M.A., Edokpayi, J.N., 2022. Wetland resources in South Africa: Threats and metadata study. *Resources* 11, 54. <https://doi.org/10.3390/resources11060054>
- Adeli, S., Salehi, B., Mahdianpari, M., Quackenbush, L.J., Brisco, B., Tamiminia, H., Shaw, S., 2020. Wetland monitoring using SAR Data: A meta-analysis and comprehensive review. *Remote Sens.* 12. <https://doi.org/10.3390/rs12142190>
- Al-Shehhi, M.R., Saffarini, R., Farhat, A., Al-Meqbali, N.K., Ghedira, H., 2011. Evaluating the effect of soil moisture, surface temperature, and humidity variations on MODIS-derived NDVI values, in: *International Geoscience and Remote Sensing Symposium (IGARSS)*. pp. 3160–3163. <https://doi.org/10.1109/IGARSS.2011.6049889>
- Alam, A., Rashid, S.M., Bhat, M.S., Sheikh, A.H., 2011. Impact of land use / land cover dynamics on Himalayan wetland ecosystem. *J. Exp. Sci.* 2, 60–64.
- Amani, M., Brisco, B., Afshar, M., Mirmazloumi, S.M., Mahdavi, S., Mirzadeh, S.M.J., Huang, W., Granger, J., 2019. A generalized supervised classification scheme to produce provincial wetland inventory maps: an application of Google Earth Engine for big geo data processing. *Big Earth Data* 3, 378–394.

<https://doi.org/10.1080/20964471.2019.1690404>

Amani, M., Brisco, B., Mahdavi, S., Ghorbanian, A., Moghimi, A., Delancey, E.R., Merchant, M., Jahncke, R., Fedorchuk, L., Mui, A., Fisette, T., Kakooei, M., Ahmadi, S.A., Leblon, B., Larocque, A., 2021. Evaluation of the Landsat-based Canadian Wetland Inventory Map using multiple sources: Challenges of large-scale wetland classification using remote sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 14, 32–52. <https://doi.org/10.1109/JSTARS.2020.3036802>

Amani, M., Ghorbanian, A., Ahmadi, S.A., Kakooei, M., Moghimi, A., Mirmazloumi, S.M., Moghaddam, S.H.A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., Brisco, B., 2020. Google Earth Engine cloud-computing platform for remote sensing big data applications: A comprehensive review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 13, 5326–5350. <https://doi.org/10.1109/JSTARS.2020.3021052>

Amani, M., Salehi, B., Mahdavi, S., Brisco, B., 2018. Spectral analysis of wetlands using multi-source optical satellite imagery. *ISPRS J. Photogramm. Remote Sens.* 144, 119–136. <https://doi.org/10.1016/j.isprsjprs.2018.07.005>

Amani, M., Salehi, B., Mahdavi, S., Granger, J.E., Brisco, B., Hanson, A., 2017. Wetland classification using multi-source and multi-temporal optical remote sensing data in Newfoundland and Labrador, Canada. *Can. J. Remote Sens.* 43, 360–373. <https://doi.org/10.1080/07038992.2017.1346468>

Amler, E., Schmidt, M., Menz, G., 2015. Definitions and mapping of East African wetlands: A Review. *Remote Sens.* 7, 5256–5282. <https://doi.org/10.3390/rs70505256>

Ao, Y., Li, H., Zhu, L., Ali, S., Yang, Z., 2019. The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. *J. Pet. Sci. Eng.* 174, 776–789. <https://doi.org/10.1016/j.petrol.2018.11.067>

Bahilu, B., Tadesse, M., 2017. Review on distribution, importance, threats and consequences of wetland degradation in Ethiopia. *Int. J. Water Resour. Environ. Eng.* 9, 64–71. <https://doi.org/10.5897/ijwree2016.0697>

Barbier, E.B., 1993. Valuing tropical wetland benefits: economic methodologies and applications. *Geogr. J.* 159, 22–32.

Berberoglu, S., Yilmaz, K.T., Özkan, C., 2004. Mapping and monitoring of coastal wetlands

- of Çukurova Delta in the Eastern Mediterranean region. *Biodivers. Conserv.* 13, 615–633. <https://doi.org/10.1023/B:BIOC.0000009493.34669.ec>
- Berhane, T.M., Lane, C.R., Wu, Q., Anenkhonov, O.A., Chepinoga, V. V., Autrey, B.C., Liu, H., 2018. Comparing pixel- and object-based approaches in effectively classifying wetland-dominated landscapes. *Remote Sens.* 10. <https://doi.org/10.3390/rs10010046>
- Bhatnagar, S., Gill, L., Regan, S., Naughton, O., Johnston, P., Waldren, S., Ghosh, B., 2020. Mapping vegetation communities inside wetlands using Sentinel-2 Imagery in Ireland. *Int. J. Appl. Earth Obs. Geoinf.* 88, 102083. <https://doi.org/10.1016/j.jag.2020.102083>
- Birkhead, A.L., James, C.S., Kleynhans, M.T., 2007. Hydrological and hydraulic modelling of the Nyl River floodplain Part 2: Modelling hydraulic behaviour. *Water SA* 33, 9–20. <https://doi.org/10.4314/wsa.v33i1.47866>
- Blanckenberg, M., Mlambo, M.C., Parker, D., Motitsoe, S.N., Reed, C., 2020. Protected and unprotected urban wetlands have similar aquatic macroinvertebrate communities: A case study from the Cape Flats Sand Fynbos region of southern Africa. *PLoS One* 15, 1–18. <https://doi.org/10.1371/journal.pone.0233889>
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* 65, 2–16. <https://doi.org/10.1016/j.isprsjprs.2009.06.004>
- Bowman, M., 2002. The Ramsar Convention on Wetlands: Has it Made a Difference? *Ybiced* 61–68.
- Butt, A., Shabbir, R., Ahmad, S.S., Aziz, N., 2015. Land use change mapping and analysis using remote sensing and GIS: A case study of Simly watershed, Islamabad, Pakistan. *Egypt. J. Remote Sens. Sp. Sci.* 18, 251–259. <https://doi.org/10.1016/j.ejrs.2015.07.003>
- Camacho Valdez, V., Ruiz-Luna, A., 2016. Effects of land use changes on ecosystem services value provided by coastal wetlands: Recent and future landscape scenarios. *J. Coast. Zo. Manag.* 19, 1–7. <https://doi.org/10.4172/2473-3350.1000418>
- Cao, L., Pan, J., Li, R., Li, J., Li, Z., 2018. Integrating airborne LiDAR and optical data to estimate forest aboveground biomass in arid and semi-arid regions of China. *Remote Sens.* 10. <https://doi.org/10.3390/rs10040532>
- Cape, W., Africa, S., Thomas, A., Ellery, F., Africa, S., 2015. Wetlands in drylands : ‘ Hotspots ’ of ecosystem services in marginal environments 1–4. available at:

<https://sustainabledevelopment.un.org>, accessed on: 20/02/2022

- Carreño, M.F., Esteve, M.A., Martinez, J., Palazón, J.A., Pardo, M.T., 2008. Habitat changes in coastal wetlands associated to hydrological changes in the watershed. *Estuar. Coast. Shelf Sci.* 77, 475–483. <https://doi.org/10.1016/j.ecss.2007.10.026>
- Caturegli, L., Matteoli, S., Gaetani, M., Grossi, N., Magni, S., Minelli, A., Corsini, G., Remorini, D., Volterrani, M., 2020. Effects of water stress on spectral reflectance of bermuda grass. *Sci. Rep.* 1–12. <https://doi.org/10.1038/s41598-020-72006-6>
- Chapman, A., 2017. *Transboundary Water Climate change and development impacts*. Cape Town.
- Chen, L., Jin, Z., Michishita, R., Cai, J., Yue, T., Chen, B., Xu, B., 2014. Dynamic monitoring of wetland cover changes using time-series remote sensing imagery. *Ecol. Inform.* 24, 17–26. <https://doi.org/10.1016/j.ecoinf.2014.06.007>
- Chen, M., Liu, J., 2015. Historical trends of wetland areas in the agriculture and pasture interlaced zone: A case study of the Huangqihai Lake Basin in northern China. *Ecol. Modell.* 318, 168–176. <https://doi.org/10.1016/j.ecolmodel.2014.12.012>
- Chen, Y., Huang, C., Ticehurst, C., Merrin, L., Thew, P., 2013. An evaluation of MODIS daily and 8-day composite products for floodplain and wetland inundation mapping. *Wetlands* 33, 823–835. <https://doi.org/10.1007/s13157-013-0439-4>
- Chen, Y., Qiao, S., Zhang, G., Xu, Y.J., Chen, L., Wu, L., 2020. Investigating the potential use of Sentinel-1 data for monitoring wetland water level changes in China's Momoge National Nature Reserve. *PeerJ* 2020, 1–24. <https://doi.org/10.7717/peerj.8616>
- Chikodzi, D., Mufori, R.C., 2018. Wetland fragmentation and key drivers: A case of the Murewa District of Zimbabwe. *IOSR J. Environ. Sci.* 12, 49–61. <https://doi.org/10.9790/2402-1209014961>
- Collins, S.D., Heintzman, L.J., Starr, S.M., Wright, C.K., Henebry, G.M., McIntyre, N.E., 2014. Hydrological dynamics of temporary wetlands in the southern Great Plains as a function of surrounding land use. *J. Arid Environ.* 109, 6–14. <https://doi.org/10.1016/j.jaridenv.2014.05.006>

- Corcoran, J., Knight, J., Pelletier, K., Rampi, L., Wang, Y., 2015. The effects of point or polygon based training data on random forest classification accuracy of wetlands. *Remote Sens.* 7, 4002–4025. <https://doi.org/10.3390/rs70404002>
- Corcoran, J.M., Knight, J.F., Gallant, A.L., 2013. Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota. *Remote Sens.* 5, 3212–3238. <https://doi.org/10.3390/rs5073212>
- Cuba, N., 2015. Research note: Sankey diagrams for visualizing land cover dynamics. *Landsc. Urban Plan.* 139, 163–167. <https://doi.org/10.1016/j.landurbplan.2015.03.010>
- Dabboor, M., Howell, S., Shokr, M., Yackel, J., 2014. The Jeffries–Matusita distance for the case of complex Wishart distribution as a separability criterion for fully polarimetric SAR data. *Int. J. Remote Sens.* 35, 6859–6873. <https://doi.org/10.1080/01431161.2014.960614>
- Davidson, N.C., 2014. How much wetland has the world lost? Long-term and recent trends in global wetland area. *Mar. Freshw. Res.* 65, 934–941. <https://doi.org/10.1071/MF14173>
- Day, J., Day, E., Ross-Gillespie, V., Ketley, A., 2010. The assessment of temporary wetlands during dry conditions, *Wetland Health and Importance Research Programmeme*. Pretoria.
- De Alban, J.D.T., Connette, G.M., Oswald, P., Webb, E.L., 2018. Combined Landsat and L-band SAR data improves land cover classification and change detection in dynamic tropical landscapes. *Remote Sens.* 10. <https://doi.org/10.3390/rs10020306>
- De Vries, B., Huang, C., Lang, M.W., Jones, J.W., Huang, W., Creed, I.F., Carroll, M.L., 2017. Automated quantification of surface water inundation in wetlands using optical satellite imagery. *Remote Sens.* 9. <https://doi.org/10.3390/rs9080807>
- Ding, C.S., Hsieh, C.T., Wu, Q., Pedram, M., 1998. Stratified random sampling for power estimation, in: *Low-Power CMOS Design*. pp. 501–507. <https://doi.org/10.1109/9780470545058.sect13>
- Dini, J., Bahadur, U., 2016. South Africa’s National Wetland Rehabilitation Programmeme: Working for Wetlands, in: *The Wetland Book*. Springer Netherlands, pp. 1–7. https://doi.org/10.1007/978-94-007-6172-8_145-2
- Dlamini, M., Adam, E., Chirima, G., Hamandawana, H., 2021a. A remote sensing-based approach to investigate changes in land use and land cover in the lower uMfolozi

- floodplain system, South Africa. *Trans. R. Soc. South Africa* 0, 1–13. <https://doi.org/10.1080/0035919X.2020.1858365>
- Dlamini, M., Chirima, G., Sibanda, M., Adam, E., Dube, T., 2021b. Characterising leaf nutrients of wetland plants and agricultural crops with non-parametric approach using Sentinel-2 imagery data. *Remote Sens.* 13, 4249. <https://doi.org/10.3390/rs13214249>
- Dronova, I., 2015. Object-based image analysis in wetland research: A review. *Remote Sens.* 7, 6380–6413. <https://doi.org/10.3390/rs70506380>
- Dube, T., Mutanga, O., 2015. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in the uMgeni Catchment, South Africa. *ISPRS J. Photogramm. Remote Sens.* 101, 36–46. <https://doi.org/10.1016/j.isprsjprs.2014.11.001>
- Dubeau, P., King, D.J., Unbushe, D.G., Rebelo, L.M., 2017. Mapping the Dabus Wetlands, Ethiopia, using random forest classification of Landsat, PALSAR and topographic data. *Remote Sens.* 9, 1–23. <https://doi.org/10.3390/rs9101056>
- Dvoretz, D., Davis, C., Papeş, M., 2016. Mapping and hydrologic attribution of temporary wetlands using recurrent Landsat imagery. *Wetlands* 36, 431–443. <https://doi.org/10.1007/s13157-016-0752-9>
- Dzurume, T., 2021. The use of remote sensing data for assessing water quality in wetlands within the Limpopo River Basin. University of the Western Cape, Cape Town.
- Dzurume, T., Dube, T., Thamaga, K.H., Shoko, C., Mazvimavi, D., 2021. Use of multispectral satellite data to assess impacts of land management practices on wetlands in the Limpopo Transfrontier River Basin, South Africa. *South African Geogr. J.* 00, 1–20. <https://doi.org/10.1080/03736245.2021.1941220>
- Ewart-smith, J., Ollis, D., Day, J., Malan, H., 2006. National Wetland Inventory: Development of a Wetland Classification System for South Africa Final Report March 2006 Prepared for : Water Research Commission
- Fang, C., Wen, Z., Li, L., Du, J., Liu, G., Wang, X., Song, K., 2019. Agricultural development and implications for wetland sustainability: A case from Baoqing County, North-east China. *Chinese Geogr. Sci.* 29, 231–244. <https://doi.org/10.1007/s11769-019-1019-1>
- Federal Geographic Data Committee (FGDC), 1992. Application of satellite data for mapping

and monitoring wetlands - fact-finding report, Technical Report 1. Wetlands Subcommittee.

Frohn, R.C., Autrey, B.C., Lane, C.R., Reif, M., 2011. Segmentation and object-oriented classification of wetlands in a karst Florida landscape using multi-season Landsat-7 ETM+ imagery. *Int. J. Remote Sens.* 32, 1471–1489. <https://doi.org/10.1080/01431160903559762>

Fu, B., Wang, Y., Campbell, A., Li, Y., Zhang, B., Yin, S., Xing, Z., Jin, X., 2017. Comparison of object-based and pixel-based Random Forest algorithm for wetland vegetation mapping using high spatial resolution GF-1 and SAR data. *Ecol. Indic.* 73, 105–117. <https://doi.org/10.1016/j.ecolind.2016.09.029>

Gallant, A.L., 2015. The challenges of remote monitoring of wetlands. *Remote Sens.* 7, 10938–10950. <https://doi.org/10.3390/rs70810938>

Gardner, R.C., Connolly, K.D., Bamba, A., Diana Connolly, K., 2009. African wetlands of international importance: Assessment of African wetlands of international importance: Assessment of benefits associated with designations under the Ramsar Convention Afr, in: 21 *Geo. Int'l. Env'tl. L. Rev.* 257 (2009). pp. 1–39.

Gebreslassie, H., Gashaw, T., Mehari, A., 2014. Wetland degradation in Ethiopia : Causes , Consequences and Remedies. *J. Environ. Earch Sci.* 4, 40–49.

George H. Hargreaves, Zohrab A. Samani, 1985. Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* 1, 96–99. <https://doi.org/10.13031/2013.26773>

Ghobadi, Y., Pradhan, B., Kabiri, K., Pirasteh, S., Shafri, H.Z.M., Sayyad, G.A., 2012. Use of multi-temporal remote sensing data and GIS for wetland change monitoring and degradation. *CHUSER 2012 - 2012 IEEE Colloq. Humanit. Sci. Eng. Res.* 103–108. <https://doi.org/10.1109/CHUSER.2012.6504290>

Gómez-Rodríguez, C., Bustamante, J., Díaz-Paniagua, C., 2010. Evidence of hydro-period shortening in a preserved system of temporary ponds. *Remote Sens.* 2, 1439–1462. <https://doi.org/10.3390/rs2061439>

Gong, Z., Cui, T., Gong, H., 2014. Estimating wetland vegetation abundance based on spectral mixture analysis: A comparison between LSMA and FCM classification methods. *Int. J.*

- Remote Sens. 35, 189–203. <https://doi.org/10.1080/01431161.2013.866292>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Guo, M., Li, J., Sheng, C., Xu, J., Wu, L., 2017. A review of wetland remote sensing. Sensors (Switzerland) 17, 1–36. <https://doi.org/10.3390/s17040777>
- Gxokwe, S., Dube, T., Mazvimavi, D., 2021. Leveraging Google Earth Engine platform to characterise and map small seasonal wetlands in the semi-arid environments of South Africa. Sci. Total Environ. 803, 150139. <https://doi.org/10.1016/j.scitotenv.2021.150139>
- Gxokwe, S., Dube, T., Mazvimavi, D., 2020. Multispectral remote sensing of wetlands in semi-arid and arid areas: A review on applications, challenges and possible future research directions. Remote Sens. 12, 1–19. <https://doi.org/10.3390/rs12244190>
- Gxokwe, S., Dube, T., Mazvimavi, D., Grenfell, M., 2022. Using cloud-computing techniques to monitor long-term variations in eco-hydrological dynamics of small seasonally-flooded wetlands in semi-arid South Africa. J. Hydrol. 612, 128080. <https://doi.org/10.1016/j.jhydrol.2022.128080>
- Hakdaoui, S., Emran, A., Pradhan, B., Lee, C.W., Fils, S.C.N., 2019. A collaborative change detection approach on multi-sensor spatial imagery for desertwetland monitoring after a flash flood in Southern Morocco. Remote Sens. 11. <https://doi.org/10.3390/rs11091042>
- Halabisky, M., 2011. Object-based classification of semi-arid wetlands. J. Appl. Remote Sens. 5, 053511. <https://doi.org/10.1117/1.3563569>
- Halabisky, M., Babcock, C., Moskal, L.M., 2018. Harnessing the temporal dimension to improve object-based image analysis classification of wetlands. Remote Sens. 10. <https://doi.org/10.3390/rs10091467>
- Hansen, M.C., 2013. High-resolution global maps of 850, 850–854. <https://doi.org/10.1126/science.1244693>
- Hayri Kesikoglu, M., Haluk Atasever, U., Dadaser-Celik, F., Ozkan, C., 2019. Performance of ANN, SVM and MLH techniques for land use/cover change detection at Sultan Marshes wetland, Turkey. Water Sci. Technol. 80, 466–477. <https://doi.org/10.2166/wst.2019.290>

- Hird, J.N., DeLancey, E.R., McDermid, G.J., Kariyeva, J., 2017. Google Earth Engine, open-access satellite data, and machine learning in support of large-area probabilistic wetland mapping. *Remote Sens.* 9. <https://doi.org/10.3390/rs9121315>
- Hollis, G.E., 1990. Les effets sur l'environnement de la mise en valeur des zones humides dans Ses regions arides et semi arides. *Hydrol. Sci. J.* 35, 411–428. <https://doi.org/10.1080/02626669009492443>
- Hu, S., Niu, Z., Chen, Y., Li, L., Zhang, H., 2017. Global wetlands: Potential distribution, wetland loss, and status. *Sci. Total Environ.* 586, 319–327. <https://doi.org/10.1016/j.scitotenv.2017.02.001>
- Huang, C., Chen, Y., Zhang, S., Wu, J., 2018. Detecting, extracting and monitoring surface water from space using optical sensors: A Review. *Rev. Geophys.* 56, 333–360. <https://doi.org/10.1029/2018RG000598>
- Huang, S., Dahal, D., Young, C., Chander, G., Liu, S., 2011. Integration of Palmer Drought Severity Index and remote sensing data to simulate wetland water surface from 1910 to 2009 in Cottonwood Lake area, North Dakota. *Remote Sens. Environ.* 115, 3377–3389. <https://doi.org/10.1016/j.rse.2011.08.002>
- Inman, V.L., Lyons, M.B., 2020. Automated inundation mapping over large areas using landsat data and google earth engine. *Remote Sens.* 12, 1–12. <https://doi.org/10.3390/RS12081348>
- Jenkins, K.M., Boulton, A.J., Ryder, D.S., 2005. A common parched future? Research and management of Australian arid-zone floodplain wetlands. *Hydrobiologia* 552, 57–73. <https://doi.org/10.1007/s10750-005-1505-6>
- Ji, H., Li, X., Wei, X., Liu, W., Zhang, L., Wang, L., 2020. Mapping 10-m resolution rural settlements using multi-source remote sensing datasets with the google earth engine platform. *Remote Sens.* 12, 1–23. <https://doi.org/10.3390/rs12172832>
- John Mwita, E., 2013. Land cover and land use dynamics of semi-arid wetlands: A case of Rumuruti (Kenya) and Malinda (Tanzania). *J. Remote Sens. GIS* s1. <https://doi.org/10.4172/2469-4134.s1-001>
- Jones, J.W., 2015. Efficient wetland surface water detection and monitoring via landsat: Comparison with in-situ data from the everglades depth estimation network. *Remote Sens.*

- 7, 12503–12538. <https://doi.org/10.3390/rs70912503>
- Judah, A., Hu, B., 2019. The integration of multi-source remotely-sensed data in support of the classification of wetlands. *Remote Sens.* 11. <https://doi.org/10.3390/rs11131537>
- Kabii, T, Kabii, Tom, 2005. An overview of African wetlands item type working paper. An Overview of African Wetlands.
- Kamal, M., Phinn, S., 2011. Hyperspectral data for mangrove species mapping: A comparison of pixel-based and object-based approach. *Remote Sens.* 3, 2222–2242. <https://doi.org/10.3390/rs3102222>
- Kandus, P., Minotti, P.G., Morandeira, N.S., Grimson, R., Trilla, G.G., González, E.B., Martín, L.S., Gayol, M.P., 2018. Remote sensing of wetlands in South America: Status and challenges. *Int. J. Remote Sens.* 39, 993–1016. <https://doi.org/10.1080/01431161.2017.1395971>
- Kaplan, G., Avdan, U., 2017. Mapping and monitoring wetlands using Sentinel-2 satellite imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 4, 271–277. <https://doi.org/10.5194/isprs-annals-IV-4-W4-271-2017>
- Karanam, H.K., BabuNeela, V., 2018. Study of normalized difference built-up (NDBI) index in automatically mapping urban areas from Landsat TM imagery. *Int. J. Eng. Sci. Math.* 6, 239–248.
- Kates, R.W., Parris, T.M., Leiserowitz, A.A., 2005. What is sustainable development? Goals, indicators, values., and practice. *Environment* 47, 8–21. <https://doi.org/10.1080/00139157.2005.10524444>
- Klein, I., Dietz, A.J., Gessner, U., Galayeva, A., Myrzakhmetov, A., Kuenzer, C., 2014. Evaluation of seasonal water body extents in Central Asia over the past 27 years derived from medium-resolution remote sensing data. *Int. J. Appl. Earth Obs. Geoinf.* 26, 335–349. <https://doi.org/10.1016/j.jag.2013.08.004>
- Klemas, V., 2014. Remote sensing of riparian and wetland buffers: An overview. *J. Coast. Res.* 297, 869–880. <https://doi.org/10.2112/jcoastres-d-14-00013.1>
- Klemas, V., 2013. Remote sensing of emergent and submerged wetlands: an overview. *Int. J. Remote Sens.* 34, 6286–6320. <https://doi.org/10.1080/01431161.2013.800656>

- Kotze, D.C., O'Connor, T.G., 2000. Vegetation variation within and among Palustrine wetlands along an altitudinal gradient. *Source Plant Ecol.* 146, 77–96.
- Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T.V., Dech, S., 2011. Remote sensing of mangrove ecosystems: A review, *Remote Sensing* 12,1-17. <https://doi.org/10.3390/rs3050878>
- Kulithalai Shiyam Sundar, P., Deka, P.C., 2021. Spatio-temporal classification and prediction of land use and land cover change for the Vembanad Lake system, Kerala: a machine learning approach. *Environ. Sci. Pollut. Res.* <https://doi.org/10.1007/s11356-021-17257-0>
- Kumar, L., Mutanga, O., 2018. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* 10, 1–15. <https://doi.org/10.3390/rs10101509>
- Landmann, T., Schramm, M., Colditz, R.R., Dietz, A., Dech, S., 2010. Wide area wetland mapping in semi-arid Africa using 250-meter MODIS metrics and topographic variables. *Remote Sens.* 2, 1751–1766. <https://doi.org/10.3390/rs2071751>
- Landmann, T., Schramm, M., Huettich, C., Dech, S., 2013. MODIS-based change vector analysis for assessing wetland dynamics in Southern Africa. *Remote Sens. Lett.* 4, 104–113. <https://doi.org/10.1080/2150704X.2012.699201>
- Li, J., Wang, S., 2018. Using SAR-derived vegetation descriptors in a water cloud model to improve soil moisture retrieval. *Remote Sens.* 10, 11–14. <https://doi.org/10.3390/rs10091370>
- Li, L., Chen, Y., Xu, T., Shi, K., Liu, R., Huang, C., Lu, B., Meng, L., 2019. Remote sensing of wetland flooding at a sub-pixel scale based on random forests and spatial attraction models. *Remote Sens.* 11, 1–15. <https://doi.org/10.3390/rs11101231>
- Li, L., Vrieling, A., Skidmore, A., Wang, T., Muñoz, A.R., Turak, E., 2015. Evaluation of MODIS spectral indices for monitoring hydrological dynamics of a small, seasonally-flooded wetland in southern Spain. *Wetlands* 35, 851–864. <https://doi.org/10.1007/s13157-015-0676-9>
- Liu, D., Cao, C., Chen, W., Ni, X., Tian, R., Xing, X., 2017. Monitoring and predicting the degradation of a semi-arid wetland due to climate change and water abstraction in the Ordos Larus relictus National Nature Reserve, China. *Geomatics, Nat. Hazards Risk* 8,

367–383. <https://doi.org/10.1080/19475705.2016.1220024>

- Liu, H.Q., Huete, A., 1995. A feedback-based modification of the NDVI to minimise canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.* 33, 457. <https://doi.org/doi:10.1109/TGRS.1995.8746027>.
- Macarthur, R., 1975. Geographic Information Systems and their use for environmental monitoring, in: Artiola, J., Pepper, I., Brusseau, M. (Eds.), *Environmental Monitoring and Characterization*. Elsevier Science and Technology Books, Amsterdam, pp. 86–100.
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Brisco, B., Homayouni, S., Gill, E., DeLancey, E.R., Bourgeau-Chavez, L., 2020. Big data for a big vountry: The first generation of the Canadian Wetland Inventory Map at a spatial resolution of 10-m using Sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud-computing platform. *Can. J. Remote Sens.* 46, 15–33. <https://doi.org/10.1080/07038992.2019.1711366>
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Homayouni, S., Gill, E., 2019. The first wetland inventory map of newfoundland at a spatial resolution of 10 m using Sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud-computing platform. *Remote Sens.* 11. <https://doi.org/10.3390/rs11010043>
- Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J.M., McNairn, H., Rao, Y.S., 2020. Dual polarimetric radar vegetation index for crop growth monitoring using Sentinel-1 SAR data. *Remote Sens. Environ.* 247, 111954. <https://doi.org/10.1016/j.rse.2020.111954>
- Marambanyika, T., Beckedahl, H., Ngetar, N.S., Dube, T., 2017. Assessing the environmental sustainability of cultivation systems in wetlands using the WET-Health framework in Zimbabwe. *Phys. Geogr.* 38, 62–82. <https://doi.org/10.1080/02723646.2016.1251751>
- Martínez-López, J., Carreño, M.F., Palazón-Ferrando, J.A., Martínez-Fernández, J., Esteve, M.A., 2014. Remote sensing of plant communities as a tool for assessing the condition of semi-arid Mediterranean saline wetlands in agricultural catchments. *Int. J. Appl. Earth Obs. Geoinf.* 26, 193–204. <https://doi.org/10.1016/j.jag.2013.07.005>
- Maswanganye, S.E., Dube, T., Jovanovic, N., Mazvimavi, D., 2022. Use of multi-source remotely-sensed data in monitoring the spatial distribution of pools and pool dynamics along non-perennial rivers in semi-arid environments, South Africa. *Geocarto Int.* 0, 1–

20. <https://doi.org/10.1080/10106049.2022.2043453>

Mazzarino, M.M., 2014. Agro-Pastoralist Livelihood in the Andes of Peru. PhD thesis, School of Geoscience, University of Massachusetts Amherst

McCarthy, J., Gumbricht, T., McCarthy, T.S., 2005. Eco-region classification in the Okavango Delta, Botswana, from multi-temporal remote sensing. *Int. J. Remote Sens.* 26, 4339–4357. <https://doi.org/10.1080/01431160500113583>

Mccartney, M.P., Houghton-Carr, H.A., 2009. Working wetland potential: An index to guide the sustainable development of African wetlands. *Nat. Resour. Forum* 33, 99–110. <https://doi.org/10.1111/j.1477-8947.2009.01214.x>

McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. <https://doi.org/10.1080/01431169608948714>

Millard, K., Richardson, M., 2013. Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR-SAR fusion using a random forest classifier. *Can. J. Remote Sens.* 39, 290–307. <https://doi.org/10.5589/m13-038>

Millennium Ecosystem Assessment (Programme), 2005. Ecosystems and human well-being: wetlands and water synthesis: a report of the Millennium Ecosystem Assessment. World Resources Institute.

Minckley, T.A., Turner, D.S., Weinstein, S.R., 2013. The relevance of wetland conservation in arid regions: A re-examination of vanishing communities in the American South-west. *J. Arid Environ.* 88, 213–221. <https://doi.org/10.1016/j.jaridenv.2012.09.001>

Mishra, N.B., 2014. Wetlands: Remote Sensing. *Encycl. Nat. Resour. L.* 15; 566–574. <https://doi.org/10.1081/e-enrl-120049156>

Mohammadimanesh, F., Salehi, B., Mahdianpari, M., Brisco, B., Motagh, M., 2018. Multi-temporal, multi-frequency and multi-polarization coherence and SAR back-scatter analysis of wetlands. *ISPRS J. Photogramm. Remote Sens.* 142, 78–93. <https://doi.org/10.1016/j.isprsjprs.2018.05.009>

Mosase, E., Ahiablame, L., 2018. Rainfall and temperature in the Limpopo River Basin, southern Africa: Means, variations and trends from 1979 to 2013. *Water (Switzerland)* 10:1-29). <https://doi.org/10.3390/w10040364>

- Mosase, E., Ahiablame, L., Srinivasan, R., 2019. Spatial and temporal distribution of blue water in the Limpopo River Basin, southern Africa: A case study. *Ecohydrol. Hydrobiol.* 19, 252–265. <https://doi.org/10.1016/j.ecohyd.2018.12.002>
- Moser, L., Voigt, S., Schoepfer, E., Palmer, S., 2014. Multitemporal wetland monitoring in sub-Saharan West Africa using medium resolution optical satellite data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 3402–3415. <https://doi.org/10.1109/JSTARS.2014.2336875>
- Mtengwana, B., Dube, T., Mkunyana, Y.P., Mazvimavi, D., 2020. Use of multispectral satellite datasets to improve ecological understanding of the distribution of invasive alien plants in a water-limited catchment, South Africa. *Afr. J. Ecol.* 1–10. <https://doi.org/10.1111/aje.12751>
- Mudereri, B.T., Abdel-Rahman, E.M., Dube, T., Niassy, S., Khan, Z., Tonnang, H.E.Z., Landmann, T., 2021. A two-step approach for detecting *Striga* in a complex agro-ecological system using Sentinel-2 data. *Sci. Total Environ.* 762, 143151. <https://doi.org/10.1016/j.scitotenv.2020.143151>
- Mwita, E., Menz, G., Misana, S., Nienkemper, P., 2012. Detection of small wetlands with multi-sensor data in East Africa. *Adv. Remote Sens.* 01, 64–73. <https://doi.org/10.4236/ars.2012.13007>
- Na, X., Zhang, S., Li, X., Yu, H., Liu, C., 2010. Improved land cover mapping using random forests combined with Landsat thematic mapper imagery and ancillary geographic data. *Photogramm. Eng. Remote Sensing* 76, 833–840. <https://doi.org/10.14358/PERS.76.7.833>
- Nhamo, L., Magidi, J., Dickens, C., 2017. Determining wetland spatial extent and seasonal variations of the inundated area using multispectral remote sensing. *Water SA* 43, 543–552. <https://doi.org/10.4314/wsa.v43i4.02>
- Nicolau, A.P., Flores-Anderson, A., Griffin, R., Herndon, K., Meyer, F.J., 2021. Assessing SAR C-band data to effectively distinguish modified land uses in a heavily-disturbed Amazon forest. *Int. J. Appl. Earth Obs. Geoinf.* 94, 102214. <https://doi.org/10.1016/j.jag.2020.102214>
- Noi Phan, T., Kuch, V., Lehnert, L.W., 2020. Land cover classification using Google Earth

- Engine and random forest classifier - the role of image composition. *Remote Sens.* 12. <https://doi.org/10.3390/RS12152411>
- Oommen, T., Misra, D., Twarakavi, N.K.C., Prakash, A., Sahoo, B., Bandopadhyay, S., 2008. An objective analysis of support vector machine-based classification for remote sensing. *Math. Geosci.* 40, 409–424. <https://doi.org/10.1007/s11004-008-9156-6>
- Otukey, J.R., Blaschke, T., 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *Int. J. Appl. Earth Obs. Geoinf.* 12, 27–31. <https://doi.org/10.1016/j.jag.2009.11.002>
- Ozesmi, S.L., Bauer, M.E., 2002. Satellite remote sensing of wetlands. *Wetl. Ecol. Manag.* 10, 381–402. <https://doi.org/10.1023/A:1020908432489>
- Pal, M., 2005. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* 26, 217–222. <https://doi.org/10.1080/01431160412331269698>
- Pande-Chhetri, R., Abd-Elrahman, A., Liu, T., Morton, J., Wilhelm, V.L., 2017. Object-based classification of wetland vegetation using very high-resolution unmanned air system imagery. *Eur. J. Remote Sens.* 50, 564–576. <https://doi.org/10.1080/22797254.2017.1373602>
- Pekel, J.F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418–422. <https://doi.org/10.1038/nature20584>
- Pereira, M., 2004. Land cover change detection in the Limpopo River Basin, Mozambique. MSC, International Institute for Geo-Information and Earth Observation Enschede.
- Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: An attempt on Sentinel-1. *Remote Sens. Environ.* 217, 537–549. <https://doi.org/10.1016/j.rse.2018.09.003>
- Peter, K.H., Nnko, H.J., Mubako, S., 2020. Impacts of anthropogenic and climate variation on spatiotemporal pattern of water resources: a case study of Lake Babati, Tanzania. *Sustain. Water Resour. Manag.* 6, 1–12. <https://doi.org/10.1007/s40899-020-00400-z>
- Powell, M., Hodgins, G., Danaher, T., Ling, J., Hughes, M., Wen, L., 2019. Mapping wetland types in semi-arid floodplains: A statistical learning approach. *Remote Sens.* 11, 609. <https://doi.org/10.3390/rs11060609>

- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil-adjusted Vegetation Index. *Remote Sens. Environ.* 48, 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)
- Ramachandra, T. V., Kumar, U., 2008. Wetlands of Greater Bangalore, India: Automatic delineation through pattern classifiers. *Electron. Green J.* <https://doi.org/10.5070/g312610729>
- Rana, V.K., Venkata Suryanarayana, T.M., 2020. Performance evaluation of MLE, RF and SVM classification algorithms for watershed scale land use/land cover mapping using Sentinel 2 bands. *Remote Sens. Appl. Soc. Environ.* 19, 100351. <https://doi.org/10.1016/j.rsase.2020.100351>
- Rapinel, S., Fabre, E., Dufour, S., Arvor, D., Mony, C., Hubert-Moy, L., 2019. Mapping potential, existing and efficient wetlands using free remote sensing data. *J. Environ. Manage.* 247, 829–839. <https://doi.org/10.1016/j.jenvman.2019.06.098>
- Rebelo, L.M., 2010. Eco-hydrological characterization of inland wetlands in Africa using L-Band SAR. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 3, 554–559. <https://doi.org/10.1109/JSTARS.2010.2070060>
- Riddell, E.S., Lorentz, S.A., Kotze, D.C., 2012. The hydrodynamic response of a semi-arid headwater wetland to technical rehabilitation interventions. *Water SA* 38, 55–66. <https://doi.org/10.4314/wsa.v38i1.8>
- Rocheftort, L., Strack, M., Poulin, M., Price, J.S., Graf, M., Desrochers, A., Lavoie, C., 2012. Northern Peatlands, in: Darold P. Batzer and Andrew (Ed.), *Wetland habitats of North America: Ecology and conservation concerns*. University of California Press, Ltd, California, pp. 119–134.
- 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 J. Photogramm. Remote Sens.* 67, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Ruiz, E., 2008. Management of Natura 2000 habitats * Mediterranean temporary ponds 3170. *Ecosystems* 20.
- Ryan, M.E., Palen, W.J., Adams, M.J., Rocheftort, R.M., 2014. Amphibians in the climate vise:

- Loss and restoration of resilience of montane wetland ecosystems in the western US. *Front. Ecol. Environ.* 12, 232–240. <https://doi.org/10.1890/130145>
- Sakané, N., Alvarez, M., Becker, M., Böhme, B., Handa, C., Kamiri, H.W., Langensiepen, M., Menz, G., Misana, S., Mogha, N.G., Mösel, B.M., J. Mwita, E., Oyieke, H., van Wijk, M.T., 2011. Classification, characterisation and use of small wetlands in East Africa. *Wetlands* 31, 1103–1116. <https://doi.org/10.1007/s13157-011-0221-4>
- Scanlon, B., Keese, K., Flint, A., Gaye, C., Edmunds, M., Simmers, I., 2010. Advanced Bash-Scripting Guide: An in-depth exploration of the art of shell scripting Table of Contents. *Hydrol. Process.* 2274, 2267–2274. <https://doi.org/10.1002/hyp>
- Schmid, T., Koch, M., Gumuzzio, J., 2005. Multisensor approach to determine changes of wetland characteristics in semiarid environments (Central Spain). *IEEE Trans. Geosci. Remote Sens.* 43, 2516–2525. <https://doi.org/10.1109/TGRS.2005.852082>
- Schmid, T., Koch, M., Gumuzzio, J., Mather, P.M., 2004. A spectral library for a semi-arid wetland and its application to studies of wetland degradation using hyperspectral and multispectral data. *Int. J. Remote Sens.* 25, 2485–2496. <https://doi.org/10.1080/0143116031000117001>
- Shafizadeh-Moghadam, H., Khazaee, M., Alavipanah, S.K., Weng, Q., 2021. Google Earth Engine for large-scale land use and land cover mapping: an object-based classification approach using spectral, textural and topographical factors. *GIScience Remote Sens.* 58, 914–928. <https://doi.org/10.1080/15481603.2021.1947623>
- Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S., 2017a. Exploring the Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Front. Earth Sci.* 5, 1–10. <https://doi.org/10.3389/feart.2017.00017>
- Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., Skakun, S., 2017b. Large-scale crop classification using the Google Earth Engine platform, in: *International Geoscience and Remote Sensing Symposium (IGARSS)*. pp. 3696–3699. <https://doi.org/10.1109/IGARSS.2017.8127801>
- Shen, G., Yang, X., Jin, Y., Xu, B., Zhou, Q., 2019. Remote sensing and evaluation of the wetland ecological degradation process of the Zoige Plateau Wetland in China. *Ecol.*

Indic. 104, 48–58. <https://doi.org/10.1016/j.ecolind.2019.04.063>

Shine, C., Klemm, C. De, 1999. Wetlands, water and the law. Using law to advance wetland conservation and wise use. Available at: <https://portals.iucn.org/library/efiles/documents/eplp-038.pdf>, Date accessed: 20/05/2021

Shoko, C., Mutanga, O., Dube, T., 2016. Progress in the remote sensing of C3 and C4 grass species aboveground biomass over time and space. *ISPRS J. Photogramm. Remote Sens.* 120, 13–24. <https://doi.org/10.1016/j.isprsjprs.2016.08.001>

Sibanda, S., Ahmed, F., 2021. Modelling historic and future land use/land cover changes and their impact on wetland area in Shashe sub-catchment, Zimbabwe. *Model. Earth Syst. Environ.* 7, 57–70. <https://doi.org/10.1007/s40808-020-00963-y>

Sieben, E.J.J., Collins, N.B., Corry, F.T.J., Kotze, D.C., Job, N., Muasya, A.M., Venter, C.E., Mtshali, H., Zondo, S.A., Janks, M., Pretorius, L., 2016. The vegetation of grass lawn wetlands of floodplains and pans in semi-arid regions of South Africa: Description, classification and explanatory environmental factors. *South African J. Bot.* 104, 215–224. <https://doi.org/10.1016/j.sajb.2015.11.003>

Simioni, J.P.D., Guasselli, L.A., de Oliveira, G.G., Ruiz, L.F.C., de Oliveira, G., 2020. A comparison of data mining techniques and multi-sensor analysis for inland marshes delineation. *Wetl. Ecol. Manag.* 28, 577–594. <https://doi.org/10.1007/s11273-020-09731-2>

Singh, A., Lin, J., 2015. Microbiological, coliphages and physico-chemical assessments of the Umgeni River, South Africa. *Int. J. Environ. Health Res.* 25, 33–51. <https://doi.org/10.1080/09603123.2014.893567>

Slagter, B., Tsendbazar, N.-E., Vollrath, A., Reiche, J., 2020. Mapping wetland characteristics using temporally-dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *Int. J. Appl. Earth Obs. Geoinf.* 86, 102009. <https://doi.org/10.1016/j.jag.2019.102009>

Soltani, K., Amiri, A., Zeynoddin, M., Ebtehaj, I., Gharabaghi, B., Bonakdari, H., 2020. Forecasting monthly fluctuations of lake surface areas using remote sensing techniques and novel machine learning methods. *Theor. Appl. Climatol.* <https://doi.org/10.1007/s00704-020-03419-6>

- Spruce, J., Bolten, J., Mohammed, I.N., Srinivasan, R., Lakshmi, V., 2020. Mapping land use land cover change in the Lower Mekong Basin from 1997 to 2010. *Front. Environ. Sci.* 8. <https://doi.org/10.3389/fenvs.2020.00021>
- Story, M., Congalton, R.G., 1986. Remote sensing brief accuracy assessment: A user's perspective. *Photogramm. Eng. Remote Sensing* 52, 397–399.
- Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A., Blaschke, T., 2020. Big earth data: disruptive changes in earth observation data management and analysis? *Int. J. Digit. Earth* 13, 832–850. <https://doi.org/10.1080/17538947.2019.1585976>
- Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., Brisco, B., 2020. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* 164, 152–170. <https://doi.org/10.1016/j.isprsjprs.2020.04.001>
- Thakur, J.K., Srivastava, P.K., Singh, S.K., Vekerdy, Z., 2012. Ecological monitoring of wetlands in semi-arid region of the Konya closed Basin, Turkey. *Reg. Environ. Chang.* 12, 133–144. <https://doi.org/10.1007/s10113-011-0241-x>
- Thamaga, K.H., Dube, T., 2018. Testing two methods for mapping water hyacinth (*Eichhornia crassipes*) in the Greater Letaba River system, South Africa: discrimination and mapping potential of the polar-orbiting Sentinel-2 MSI and Landsat 8 OLI sensors. *Int. J. Remote Sens.* 39, 8041–8059. <https://doi.org/10.1080/01431161.2018.1479796>
- Thamaga, K.H., Dube, T., Shoko, C., 2022. Evaluating the impact of land use and land cover change on unprotected wetland ecosystems in the arid-tropical areas of South Africa using the Landsat dataset and support vector machine. *Geocarto Int.* 0, 1–22. <https://doi.org/10.1080/10106049.2022.2034986>
- Thamaga, K.H., Dube, T., Shoko, C., 2021. Advances in satellite remote sensing of the wetland ecosystems in sub-Saharan Africa. *Geocarto Int.* 0, 1–22. <https://doi.org/10.1080/10106049.2021.1926552>
- The Wetlands in Dry Lands Research Network, 2014. Parys Declaration on the Importance of Wetlands in Drylands.
- Tian, S., Zhang, X., Tian, J., Sun, Q., 2016. Random forest classification of wetland landcovers

- from multi-sensor data in the arid region of Xinjiang, China. *Remote Sens.* 8, 1–14. <https://doi.org/10.3390/rs8110954>
- Tooth, S., McCarthy, T.S., Brandt, D., Hancox, P.J., Morris, R., 2002. Geological controls on the formation of alluvial meanders and floodplain wetlands: The example of the Klip River, eastern Free State, South Africa. *Earth Surf. Process. Landforms* 27, 797–815. <https://doi.org/10.1002/esp.353>
- Trudel, M., Charbonneau, F., Leconte, R., 2012. Using RADARSAT-2 polarimetric and ENVISAT-ASAR dual-polarization data for estimating soil moisture over agricultural fields. *Can. J. Remote Sens.* 38, 514–527. <https://doi.org/10.5589/m12-043>
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- Turpie, J., Malan, H., 2010. Wetland valuation. Vol III - a tool for the assessment of the livelihood value of wetlands., WRC Report. Cape Town.
- Uluocha, N.O., Okeke, I.C., 2004. Implications of wetlands degradation for water resources management: Lessons from Nigeria. *GeoJournal* 61, 151–154. <https://doi.org/10.1007/s10708-004-2868-3>
- Ustuner, M., Sanli, F.B., Abdikan, S., 2016. Balanced vs imbalanced training data: Classifying RapidEye data with support vector machines, in: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. pp. 379–384. <https://doi.org/10.5194/isprsarchives-XLI-B7-379-2016>
- van Deventer, H., Cho, M.A., Mutanga, O., 2019. Multi-season RapidEye imagery improves the classification of wetland and dryland communities in a subtropical coastal region. *ISPRS J. Photogramm. Remote Sens.* 157, 171–187. <https://doi.org/10.1016/j.isprsjprs.2019.09.007>
- Van Deventer, H., Smith-Adao, L., Collins, N., Grenfell, M., Grundling, A., Grundling, P.-L., Dean, I., Job, N., Dean, O., Petersen, C., Patsy, S., Erwin, S., Snaddon, K., Tererai, F., Lotter, M., van der Collf, D., 2019. Volume 2b: Inland Aquatic (Freshwater) Realm.
- Vinciková, H., Hanuš, J., Pechar, L., 2015. Spectral reflectance is a reliable water-quality estimator for small, highly turbid wetlands. *Wetl. Ecol. Manag.* 23, 933–946.

<https://doi.org/10.1007/s11273-015-9431-5>

Wang, C., Chen, J., Wu, J., Tang, Y., Shi, P., Black, T.A., Zhu, K., 2017. A snow-free vegetation index for improved monitoring of vegetation spring green - update in deciduous ecosystems. *Remote Sens. Environ.* 196, 1–12. <https://doi.org/10.1016/j.rse.2017.04.031>

Wang, J.W., Li, M., Zhang, G.Y., Zhang, H.R., Yu, C.Q., 2020. Growing season precipitation rather than growing season length predominates maximum normalized difference vegetation index in alpine grasslands on the Tibetan Plateau. *Sustain.* 12: 1-16. <https://doi.org/10.3390/su12030968>

Wang, Y., Qi, Q., Liu, Y., 2018. Unsupervised segmentation evaluation using area-weighted variance and Jeffries-Matusita distance for remote sensing images. *Remote Sens.* 10. <https://doi.org/10.3390/rs10081193>

Wang, Y., Yésou, H., 2018. Remote sensing of floodpath lakes and wetlands: A challenging frontier in the monitoring of changing environments. *Remote Sens.* 10. <https://doi.org/10.3390/rs10121955>

Wang, Z., Huang, N., Luo, L., Li, X., Ren, C., Song, K., Chen, J.M., 2011. Shrinkage and fragmentation of marshes in the West Songnen Plain, China, from 1954 to 2008 and its possible causes. *Int. J. Appl. Earth Obs. Geoinf.* 13, 477–486. <https://doi.org/10.1016/j.jag.2010.10.003>

West, H., Quinn, N., Horswell, M., White, P., 2018. Assessing vegetation response to soil moisture fluctuation under extreme drought using Sentinel-2. *Water (Switzerland)* 10, 1–22. <https://doi.org/10.3390/w10070838>

Westra, T., de Wulf, R.R., 2007. Monitoring Sahelian floodplains using Fourier analysis of MODIS time-series data and artificial neural networks. *Int. J. Remote Sens.* 28, 1595–1610. <https://doi.org/10.1080/01431160600887698>

White, D.C., Lewis, M.M., Green, G., Gotch, T.B., 2016. A generalizable NDVI-based wetland delineation indicator for remote monitoring of groundwater flows in the Australian Great Artesian Basin. *Ecol. Indic.* 60, 1309–1320. <https://doi.org/10.1016/j.ecolind.2015.01.032>

Wohlfart, C., Winkler, K., Wendleder, A., Roth, A., 2018. TerraSAR-X and wetlands: A review. *Remote Sens.* 10, 1–24. <https://doi.org/10.3390/rs10060916>

- Wua, Q., Lane, C.R., Li, X., Zhou, K., Zhou, Y., Clinton, N., DeVries, B., Golden, H.E., Lang, M.W., 2019. Integrating LiDAR data and multi-temporal aerial imagery to map wetland inundation dynamics using Google Earth Engine. *Remote Sens. Environ.* 60, 1–24. <https://doi.org/10.1016/j.rse.2019.04.015>. Integrating
- Xia, W., Ma, C., Liu, J., Liu, S., Chen, F., Yang, Z., Duan, J., 2019. High-resolution remote sensing imagery classification of imbalanced data using multi-stage sampling method and deep neural networks. *Remote Sens.* 11, 1-17. <https://doi.org/10.3390/rs11212523>
- Xie, Z., Huete, A., Ma, X., Restrepo-Coupe, N., Devadas, R., Clarke, K., Lewis, M., 2016. Landsat and GRACE observations of arid wetland dynamics in a dryland river system under multi-decadal hydroclimatic extremes. *J. Hydrol.* 543, 818–831. <https://doi.org/10.1016/j.jhydrol.2016.11.001>
- Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely-sensed imagery. *Int. J. Remote Sens.* 27, 3025–3033. <https://doi.org/10.1080/01431160600589179>
- Yamazaki, D., Trigg, M.A., 2016. The dynamics of Earth's surface water. *Nature* 540, 348.
- Zaitunah, A., Samsuri, S., Ahmad, A.G., Safitri, R.A., 2018. Normalized Difference Vegetation Index (NDVI) analysis for land cover types using Landsat 8 oli in Besitang watershed, Indonesia, in: IOP Conference Series: Earth and Environmental Science. <https://doi.org/10.1088/1755-1315/126/1/012112>
- Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* 24, 583–594. <https://doi.org/10.1080/01431160304987>
- Zhang, R., Tang, C., Ma, S., Yuan, H., Gao, L., Fan, W., 2011. Using Markov chains to analyze changes in wetland trends in arid Yinchuan Plain, China. *Math. Comput. Model.* 54, 924–930. <https://doi.org/10.1016/j.mcm.2010.11.017>
- Zhen, Z., Quackenbush, L.J., Stehman, S. V., Zhang, L., 2013. Impact of training and validation sample selection on classification accuracy and accuracy assessment when using reference polygons in object-based classification. *Int. J. Remote Sens.* 34, 6914–6930. <https://doi.org/10.1080/01431161.2013.810822>

- Zhou, Y., Dong, J., Xiao, X., Liu, R., Zou, Z., Zhao, G., Ge, Q., 2019. Continuous monitoring of lake dynamics on the Mongolian Plateau using all available Landsat imagery and Google Earth Engine. *Sci. Total Environ.* 689, 366–380. <https://doi.org/10.1016/j.scitotenv.2019.06.341>
- Zhu, Z., Wulder, M.A., Roy, D.P., Woodcock, C.E., Hansen, M.C., Radeloff, V.C., Healey, S.P., Schaaf, C., Hostert, P., Strobl, P., Pekel, J.F., Lyburner, L., Pahlevan, N., Scambos, T.A., 2019. Benefits of the free and open Landsat data policy. *Remote Sens. Environ.* 224, 382–385. <https://doi.org/10.1016/j.rse.2019.02.016>
- Ziter, C.D., Pedersen, E.J., Kucharik, C.J., Turner, M.G., 2019. Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. *Proc. Natl. Acad. Sci. U. S. A.* 116, 7575–7580. <https://doi.org/10.1073/pnas.1817561116>



APPENDICES

Chapter 5 appendix

Table A. Indices extracted from the time series Landsat-7 data

Data	Data extracted	Formula	Reference
	NDVI	$\frac{NIR - Red}{NIR + Red}$	Tucker, (1979)
Landsat-7	NDWI	$\frac{NIR - Green}{NIR + Green}$	McFeeters, (1996)
	NDPI	$\frac{NIR - (0.74 \times Red + 0.26 \times SWIR1)}{NIR + (0.74 \times Red + 0.26 \times SWIR1)}$	Wang et al. (2017)

**NDVI: Normalised Difference Vegetation Index, NDWI: Normalised Difference Water Index, NDPI: Normalised Difference Phenology Index*



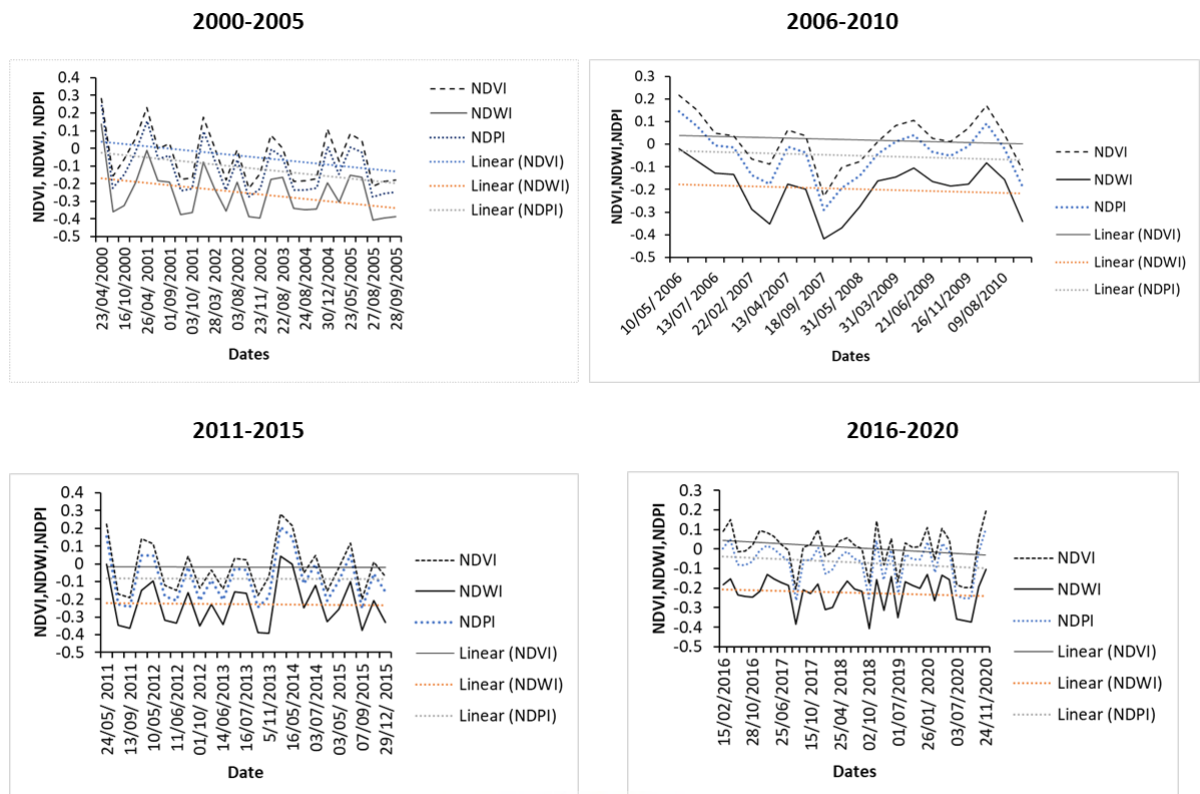
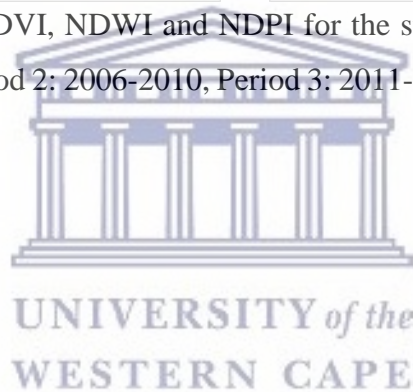


Figure A. Variations in NDVI, NDWI and NDPI for the studies time periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4:2016-2020)



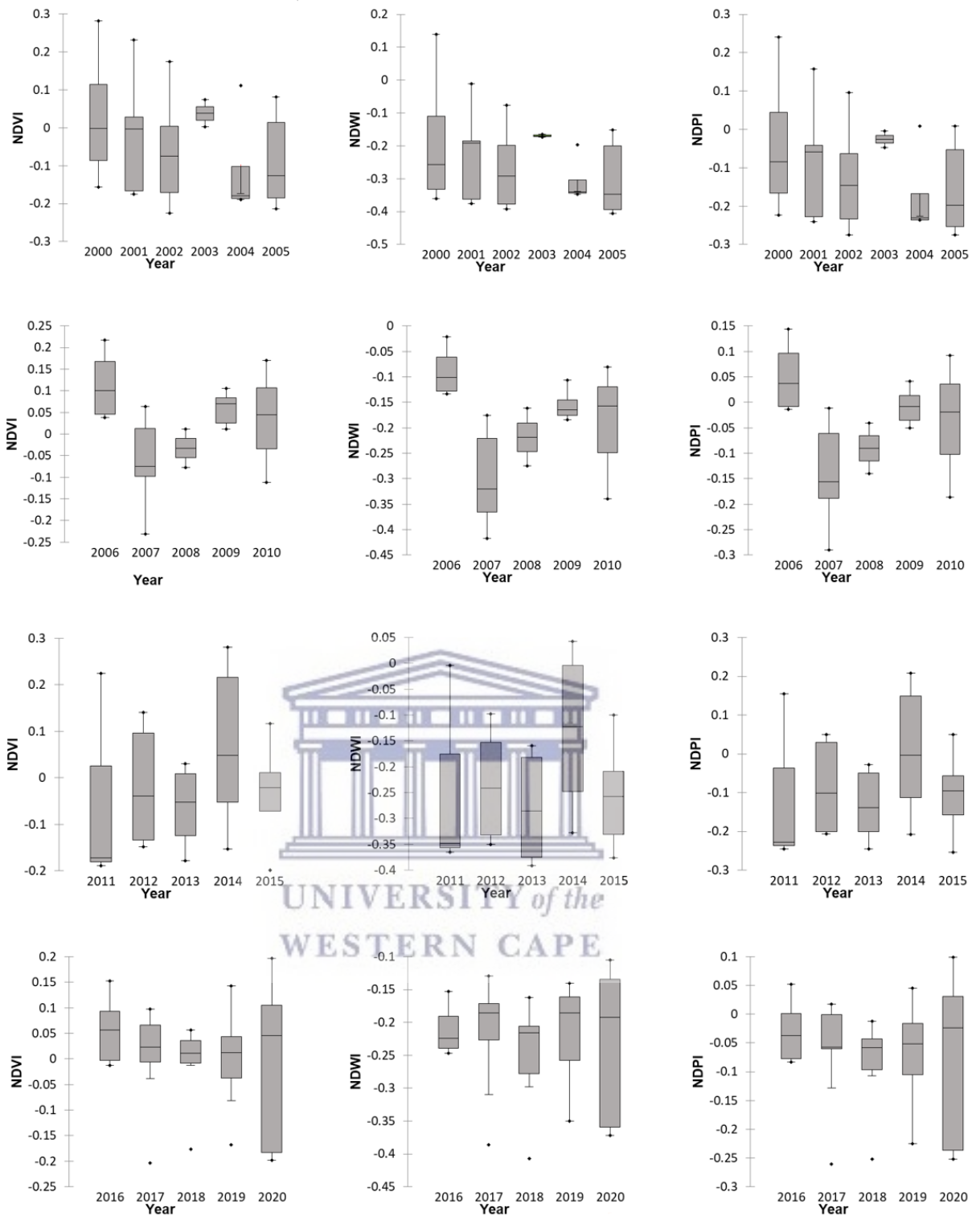


Figure B. Statistical summaries of NDVI, NDWI and NDPI extracted from the Landsat image stacks representing the time periods (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4: 2016-2020)

Table B. Dates of images obtained and used in this study (Period 1: 2000-2005, Period 2: 2006-2010, Period 3: 2011-2015 and Period 4:2016-2020)

2000-2005	2006-2010	2011-2015	2016-2020
23/04/2000	10/05/ 2006	24/05/ 2011	15/02/2016
14/09/2000	26/05/ 2006	12/08/ 2011	02/03/ 2016
16/10/2000	13/07/ 2006	13/09/ 2011	09/08/ 2016
03/12/2000	29/07/ 2006	03/01/ 2012	25/08/2016
26/04/ 2001	22/02/ 2007	10/05/2012	28/10/2016
16/08/2001	11/04/2007	26/05/ 2012	29/11/2016
01/09/2001	13/04/2007	11/06/2012	24/05/ 2017
17/09/ 2001	29/05/ 2007	27/06/ 2012	9/06/ 2017
03/10/ 2001	18/09/ 2007	01/10/ 2012	25/06/2017
07/01/2002	05/11/ 2007	27/04/ 2013	27/07/2017
28/03/ 2002	31/05/ 2008	14/06/2013	12/08/ 2017
16/06/ 2002	03/10/ 2008	30/06/ 2013	13/09/ 2017
03/08/2002	31/03/2009	16/07/2013	15/10/ 2017
22/10/ 2002	03/06/ 2009	17/10/2013	31/10/ 2017
23/11/ 2002	21/06/2009	5/11/2013	18/12/ 2017
15/03/2003	06/09/2009	30/04/2014	19/01/2018
22/08/ 2003	26/11/2009	16/05/2014	25/04/ 2018
08/08/2004	02/03/2010	01/06/2014	28/06/2018
24/08/2004	09/08/2010	03/07/2014	30/07/ 2018
09/09/2004	26/09/2010	19/07/ 2014	16/09/ 2018
30/12/2004		03/05/ 2015	02/10/ 2018
21/04/2005		19/05/ 2015	22/01/ 2019
23/05/2005		07/09/2015	14/05/2019
08/06/2005		24/09/2015	15/06/ 2019
27/08/2005		29/12/ 2015	01/07/2019
12/09/2005			17/07/ 2019
28/09/2005			03/09/ 2019
			06/11/ 2019
			26/01/ 2020

30/04/ 2020

16/05/ 2020

01/06/2020

03/07/2020

19/07/2020

04/09 2020

23/10/2020

24/11/2020



UNIVERSITY *of the*
WESTERN CAPE