BUSITEMA UNIVERSITY
Pursuing Excellence

P.O.Box 236, Tororo
Gen:+256-454448842
Dir: +256-454448864
Mob: +256-782999874
Fax: +256-454436517
Email:ar@acadreg.busitema.ac.ug
Website:www.busitema.ac.ug

BUSITEMA UNIVERSITY

FACULTY OF NATURAL RESOURCES AND ENVIRONMENTAL SCIENCES

DEPARTMENT OF FISHERIES AND WATER RESOURCES MANAGEMENT

USING NDVI AND NDWI INDICES TO ASSESS THE SPATIAL, TEMPORAL AND SEASONAL SPREAD OF AQUATIC WEEDS ON LAKE WAMALA, UGANDA (2017–2025).

BY

KAVUBU OSCAR GODFREY

REG. NO. BU/UG/2023/2268

A research report submitted to the Department of Fisheries and Water Resources Management, Faculty of Natural Resources and Environmental Sciences, in partial fulfilment of the requirements for the award of the degree of Bachelor of Science in Fisheries and Water Resources Management of Busitema University

SUPERVISOR: Prof. ISABIRYE MOSES

JUNE 2026

DECLARATION

I, KAVUBU OSCAR GODFREY, Registration Number BU/UG/2023/2268, do hereby declare that this research report is my own original work and has not been submitted to any other university or institution of higher learning for any academic award. Where the work of others has been used, it has been duly acknowledged and referenced. This report is a product of my own investigation and analysis carried out under the guidance of my supervisor.

Signature: 

KAVUBU OSCAR GODFREY (BU/UG/2023/2268).

Date: 10/08/2026

APPROVAL

This research report titled “Using NDVI and NDWI Indices to Assess the Spatial, Temporal and Seasonal Spread of Aquatic Weeds on Lake Wamala, Uganda (2017–2025)” has been submitted for examination with my approval as the candidate’s University Supervisor.

Signature: ………………………………………………

Prof. ISABIRYE MOSES

Supervisor, Department of Fisheries and Water Resources Management

Busitema University

Date: …………………………………

DEDICATION

I dedicate this work to my beloved family, whose unwavering support, patience and encouragement sustained me throughout my studies. I also dedicate it to the fishing communities of Lake Wamala, whose livelihoods depend on the health of this vital water resource, in the hope that the findings of this study will contribute to its sustainable management for current and future generations.

ACKNOWLEDGEMENTS

First and foremost, I thank the Almighty God for the gift of life, good health and the strength to complete this research.

I extend my sincere gratitude to my supervisor, Prof. Isabirye Moses, for his invaluable guidance, constructive criticism and continuous encouragement throughout the course of this study. His expertise and patience were instrumental in shaping this work.

I am grateful to the academic and technical staff of the Department of Fisheries and Water Resources Management, Faculty of Natural Resources and Environmental Sciences, Busitema University, for the knowledge and skills imparted to me during my training.

I acknowledge the European Space Agency (ESA) and the Copernicus Programme for providing free and open access to Sentinel-2 imagery, without which this study would not have been possible. I also recognise the developers of the open-source geospatial and statistical tools used in the analysis.

Finally, I thank my family, friends and classmates for their moral and material support. To all who contributed to this work in one way or another, I say thank you.

ABSTRACT

Aquatic weed infestation poses a growing threat to the ecological integrity and socio-economic value of shallow tropical lakes in East Africa. Lake Wamala, a shallow lake in central Uganda, supports important fisheries and provides water for surrounding communities, yet its weed dynamics remain poorly quantified. This study assessed the spatial distribution, temporal trends and seasonal dynamics of aquatic weed cover on Lake Wamala using the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) derived from Sentinel-2 satellite imagery. A dataset of 30,296 pixel-level observations was extracted from 16 ground stations distributed across five regions and seven ecological zone types, spanning three years (2017, 2022 and 2025) and five months (January, February, March, June and September). Descriptive statistics, weed-cover classification using NDVI thresholds, correlation analysis and spatial–temporal trend analysis were performed. Group differences were tested by Welch's and one-way analysis of variance, and the NDVI–NDWI association by Pearson correlation.

The overall mean NDVI was −0.073 (SD = 0.153) and the overall mean NDWI was 0.206 (SD = 0.169), indicating that open water dominates the lake surface. Weed-cover classification showed that 95.2% of pixels were open water or low vegetation, 1.3% moderate vegetation and 3.5% high vegetation (dense weeds). The Southern region emerged as the principal weed hotspot, with the only positive regional mean NDVI (0.012) and the highest proportion of dense-vegetation pixels (12.02%), followed by the Eastern region (4.93%). NDVI declined sharply between 2017 (−0.037) and 2022 (−0.113) before partially recovering by 2025 (−0.072), a pattern attributed to the 2020–2023 La Niña-driven high water levels in the Lake Victoria basin. Seasonally, January (dry season) recorded the least negative NDVI (−0.011) while February (onset of the rains) recorded the most negative (−0.133). NDVI and NDWI were strongly and inversely correlated overall (Pearson $r = -0.971$), with zone-level coefficients ranging from −0.743 to −0.989, confirming their value as complementary indices for weed mapping. All four group-difference tests — by zone type, region, year and month — were significant at $p < .001$, with season producing the largest effect ($\eta^2 = 0.08$) and year the smallest ($\eta^2 = 0.02$).

The study concludes that aquatic weed infestation on Lake Wamala is not a uniform, lake-wide problem but a spatially concentrated one, driven by nutrient enrichment, shallow bathymetry and sheltered conditions in the Southern and Eastern regions, and modulated by strong inter-annual climate variability. NDVI and NDWI from freely available Sentinel-2 imagery provide a reliable, low-cost basis for routine monitoring. It is recommended that management be spatially targeted on the Southern and Eastern regions, embedded within an adaptive framework that responds to climate variability, supported by an integrated satellite-plus-ground monitoring system, and complemented by catchment nutrient management and community engagement. These measures would allow limited management resources to be directed where they yield the greatest ecological and economic return.

Keywords: Lake Wamala; aquatic weeds; NDVI; NDWI; Sentinel-2; remote sensing; weed mapping; eutrophication; adaptive management.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ANOVA — Analysis of Variance

BOA — Bottom-of-Atmosphere

CI — Confidence Interval

CDSE — Copernicus Data Space Ecosystem

ESA — European Space Agency

GEE — Google Earth Engine

GIS — Geographic Information System

GPS — Global Positioning System

L2A — Level-2A (Sentinel-2 surface-reflectance product)

LVBC — Lake Victoria Basin Commission

LVFO — Lake Victoria Fisheries Organisation

NDVI — Normalized Difference Vegetation Index

NDWI — Normalized Difference Water Index

NEMA — National Environment Management Authority

NIR — Near-Infrared

QA60 — Quality Assessment band 60 (Sentinel-2 cloud-mask band)

SD — Standard Deviation

SWIR — Short-Wave Infrared

CHAPTER ONE: INTRODUCTION

1.1 Background to the Study

Freshwater lakes are among the most productive and economically important ecosystems in sub-Saharan Africa, providing fisheries, water for domestic and agricultural use, transport routes and recreational value. In recent decades, however, many tropical lakes have experienced rapid proliferation of aquatic macrophytes, commonly referred to as aquatic weeds. Invasive and nuisance species such as water hyacinth (*Eichhornia crassipes*), water lettuce (*Pistia stratiotes*), *Salvinia molesta* and dense beds of papyrus and other emergent vegetation can choke navigation channels, deplete dissolved oxygen, reduce fish catches, impede water abstraction and create breeding habitats for disease vectors (Havel et al., 2015; Alemneh et al., 2025; Nyawacha et al., 2021).

Lake Wamala is a shallow, closed-basin lake located in central Uganda, with a surface area of approximately 250 km² and an average depth of about 4.5 metres. It supports an important artisanal fishery and provides water and livelihoods for communities in the surrounding districts. Like many shallow tropical lakes, Lake Wamala is highly sensitive to nutrient enrichment from agricultural runoff and to fluctuations in water level driven by climate variability (Okaronon, 2009). These conditions create an environment in which aquatic weeds can proliferate rapidly, yet the spatial extent, temporal trends and seasonal dynamics of weed cover on the lake have not been systematically quantified (Al-lami et al., 2021).

Traditional methods of mapping aquatic vegetation rely on field surveys, which are labour-intensive, costly and difficult to repeat at the frequency required to capture seasonal and inter-annual change across an entire lake. Remote sensing offers a practical alternative (Datta et al., 2021; Balirwa & Wanda, n.d.). Spectral indices computed from satellite imagery, in particular the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI), exploit the contrasting reflectance of vegetation and water in the visible, near-infrared and short-wave infrared regions of the electromagnetic spectrum. The free availability of Sentinel-2 imagery, with a 10-metre spatial resolution and a five-day revisit time, has made high-quality, repeatable monitoring of inland water bodies feasible even for resource-constrained institutions.

Within the East African region, the disruption caused by aquatic weeds is well documented, the explosive spread of water hyacinth across the Lake Victoria basin during the 1990s remaining the clearest illustration of how rapidly a single invasive species can transform a large freshwater system and undermine the fisheries and livelihoods that depend on it (Havel et al., 2015). Satellite remote sensing has since been used to track such infestations on major water bodies, including Lake Victoria and Lake Tana as well as large river systems, demonstrating that spectral indices such as NDVI and NDWI can reliably separate aquatic vegetation from open water at comparatively low cost (Alemneh et al., 2025; Nyawacha et al., 2021; Al-lami et al., 2021). Comparable multi-temporal assessments have not, however, been undertaken for Lake Wamala, whose smaller size and limited monitoring resources have left the spatial pattern and seasonal behaviour of its weed cover largely undocumented. The present study addresses this gap by applying NDVI and NDWI to a multi-year archive of Sentinel-2 imagery to map and assess the spread of aquatic weeds across Lake Wamala and its basin, providing the evidence base needed to inform the management questions set out in the sections that follow.

1.2 Statement of the Problem

Lake Wamala holds significant ecological, socio-economic and cultural value for central Uganda, supporting an important artisanal fishery, providing water for domestic and agricultural use, and sustaining the livelihoods of the communities in the surrounding districts. Ideally, the lake would remain free from excessive aquatic weed infestation, with its open water, littoral zones and wetland fringes maintained in a condition that safeguards biodiversity, fisheries productivity and community well-being. This ideal is enshrined in Uganda's National Environment Management Policy (1994), the Water Act, Cap 152 (1995) and the National Policy for the Conservation and Management of Wetland Resources (1995), all of which mandate the sustainable management of water bodies and wetland ecosystems. It is further reinforced by Uganda Vision 2040 and the Third National Development Plan (NDP III), and by global commitments under Sustainable Development Goal 6 (Clean Water and Sanitation), SDG 15 (Life on Land) and the Ramsar Convention on Wetlands, which together call for healthy, well-managed freshwater ecosystems that remain resilient and beneficial to the communities that depend on them.

In reality, Lake Wamala faces a growing threat from the rapid spread of aquatic weeds, notably water hyacinth (*Eichhornia crassipes*) and dense beds of emergent vegetation, which degrade water quality, disrupt biodiversity and undermine the ecosystem services on which fishing and farming communities depend. The achievement of the ideal is hindered, however, by a fundamental information gap: the spatial distribution, temporal trends and seasonal dynamics of weed cover on the lake have never been systematically quantified. Conventional monitoring relies on field surveys that are labour-intensive, costly, time-consuming and difficult to repeat at the frequency and spatial coverage required to track a rapidly changing infestation across an entire lake. Compounding this, Uganda has seen only limited uptake of advanced remote-sensing methods—such as the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI)—that have proven effective elsewhere for mapping aquatic vegetation dynamics.

As a consequence, management responses to weed infestation on Lake Wamala remain largely reactive and poorly targeted. In the absence of objective, spatially explicit evidence of where weeds are concentrated, how their extent has changed over time and how it varies by season, interventions risk being applied uniformly and inefficiently across the lake—wasting scarce resources on areas that do not require treatment while under-serving genuine hotspots. Should this gap persist, weed proliferation is likely to advance unchecked, further eroding fish catches, water quality and the livelihoods of dependent communities, and undermining the national and international commitments outlined above.

This study addresses that gap by using NDVI and the complementary NDWI, derived from freely available Sentinel-2 imagery for the period 2017–2025, to assess and quantify the spread of aquatic weeds on Lake Wamala. Drawing on 30,296 pixel-level observations sampled at 16 stations distributed across five regions and seven ecological zones, it provides a cost-effective, scalable and repeatable basis for mapping weed distribution, detecting temporal trends and characterising seasonal dynamics. In doing so, the study generates actionable, evidence-based information for lake managers, the National Environment Management Authority, the Ministry of Water and Environment and local governments, enabling the targeted, cost-effective interventions envisaged under NDP III and Uganda's wider environmental policy framework, and contributing to global efforts to preserve freshwater ecosystems.

1.3 Objectives of the Study

1.3.1 General Objective

The general objective of this study was to use NDVI and the complementary NDWI, derived from Sentinel-2 imagery, to assess the spread of aquatic weeds on Lake Wamala, Uganda.

1.3.2 Specific Objectives

1. To evaluate the effectiveness of NDVI and NDWI as complementary indices for mapping aquatic weeds on Lake Wamala.
2. To map the spatial distribution of aquatic weed cover across the regions and ecological zones of Lake Wamala, and its inter-annual trend between 2017 and 2025.
3. To identify the within-year seasonal patterns in the vegetation and water dynamics of the lake.

1.4 Research Questions

1. How effective are NDVI and NDWI, used together, for mapping aquatic weeds on Lake Wamala?
2. How is aquatic weed cover distributed across the regions and ecological zones of the lake, and how has it changed from year to year between 2017 and 2025?
3. What within-year seasonal patterns characterise the vegetation and water dynamics of the lake?

1.5 Significance of the Study

The findings of this study provide lake managers, the National Environment Management Authority, the Ministry of Water and Environment and local governments with an objective basis for prioritising weed management interventions. By identifying the specific regions and zones where infestation is concentrated, the study supports the efficient allocation of limited management resources. It also demonstrates a replicable, low-cost remote-sensing workflow that can be extended to other Ugandan lakes facing similar challenges, and contributes to the scientific understanding of aquatic vegetation dynamics in shallow tropical lakes.

1.6 Scope of the Study

The study focused on Lake Wamala in central Uganda. The analysis was based on Sentinel-2 imagery for three years (2017, 2022 and 2025) and five months (January, February, March, June and September), sampled at 16 ground stations distributed across five regions and seven ecological zone types, yielding 30,296 pixel-level observations. The study was confined to surface-detectable vegetation captured by NDVI and NDWI; submerged vegetation not detectable by these surface indices was outside its scope.

1.7 Limitations of the Study

- NDVI and NDWI detect surface and near-surface vegetation; densely submerged macrophytes may therefore be under-estimated.
- The temporal sample comprised three discrete years rather than a continuous time series, which limits the resolution of inter-annual trend analysis.
- Ground-truth vegetation cover data collected simultaneously with image acquisition were not available, so spectral classifications were validated against established threshold values rather than field measurements.
- Cloud cover, a common constraint in tropical regions, restricted the choice of cloud-free scenes for some months.

1.8 Conceptual Framework

A conceptual framework illustrates the variables examined in this study and the relationships expected to hold between them. Developed from the literature reviewed in Chapter Two, the framework links the environmental drivers of aquatic weed proliferation on Lake Wamala to the spectral indices used to detect and map that proliferation, and identifies the further factors expected to influence these relationships. It is summarised visually in Figure 1 and explained below.

The framework is guided by the study's central question: how do the spatial distribution, temporal trends and seasonal dynamics of aquatic weed cover on Lake Wamala vary, and how effectively can the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI), derived from Sentinel-2 imagery, detect and map this variation?

Independent variables (the expected causes). The proliferation of aquatic weeds is expected to be driven by a set of environmental and spatial factors: nutrient enrichment from agricultural runoff (chiefly nitrogen and phosphorus), shallow bathymetry and water depth, hydrodynamic shelter in bays and near river inflows, and the spatial context of the pixel as defined by its region and ecological zone. These conditions are expected to favour the growth and spread of macrophytes such as water hyacinth, papyrus and water lettuce.

Dependent variables (the measured effect). Aquatic weed cover is not observed directly but is operationalised through two continuous spectral indices measured from Sentinel-2 imagery: NDVI, which responds to vegetation density, and NDWI, which responds to open water. Together these indices constitute the measured outcome of the study and feed into the weed-cover classification and the spatial–temporal hotspot mapping that form the study's final output.

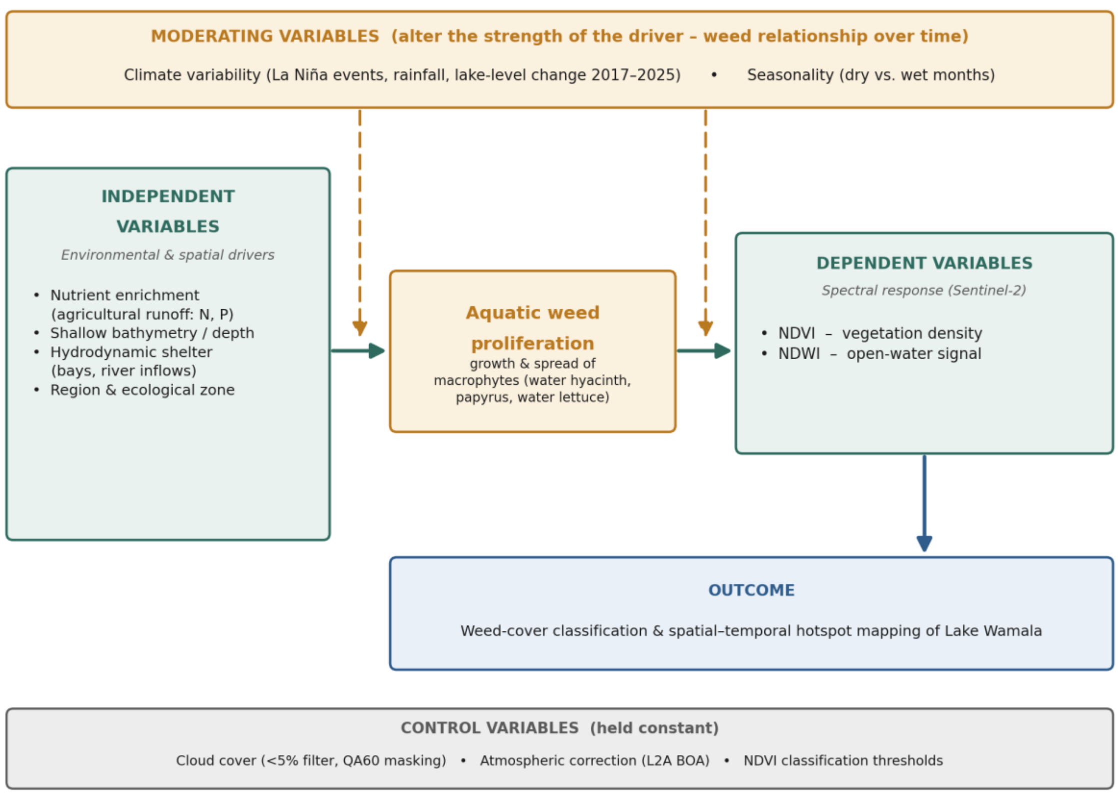


Figure 1: Conceptual framework of the study, linking the environmental drivers of aquatic weed proliferation on Lake Wamala to the NDVI and NDWI indices used to detect and map it

The expected cause-and-effect relationship. As shown in Figure 1, the independent variables are expected to promote aquatic weed proliferation, which in turn produces a distinctive spectral response: dense vegetation raises NDVI and lowers NDWI, whereas open water does the reverse. The study therefore reads the spectral indices as an indirect but reliable measure of weed cover, and uses them to locate where infestation is concentrated, how it has changed across the years 2017, 2022 and 2025, and how it varies from month to month. Aquatic weed proliferation thus acts as the mediating phenomenon that links the environmental drivers to the measured indices.

Moderating variables. Two factors are expected to alter the strength of the relationship between the drivers and the observed weed cover over time. Climate variability, in particular the La Niña-driven rainfall and record-high lake levels of 2020–2023, can submerge vegetation and raise turbidity, weakening the spectral signal; and seasonality, expressed as the contrast between dry and wet months, shifts water level, turbidity and plant growth. These moderators explain why the same drivers may yield different index values in different years and months.

Control variables. Several factors that could otherwise distort the indices were held constant so that they would not interfere with the results: cloud cover was minimised through a strict five-percent filter and QA60 masking, all imagery was atmospherically corrected to the Sentinel-2 Level-2A bottom-of-atmosphere product, and a fixed set of NDVI thresholds was applied when classifying weed cover. Holding these constant isolates the relationships of interest from avoidable measurement noise.

1.9 Organisation of the Report

This report is organised into five chapters. Chapter One introduces the study by presenting the background to the problem, the statement of the problem, the general and specific objectives, the research questions, the significance and scope of the study, its limitations and the conceptual framework that guides it. Chapter Two reviews the relevant literature, examining the nature and drivers of aquatic weed proliferation, the theoretical basis of the Normalized Difference Vegetation Index (NDVI) and related spectral indices, their application in aquatic environments, and the research gap that this study addresses. Chapter Three describes the materials and methods, including the study area and its basin, the research design, the data source and acquisition, the sampling design, the computation of the spectral indices, the weed-cover classification, the study variables, the methods of data analysis, and the reliability, validity, methodological limitations and ethical considerations of the study. Chapter Four presents and discusses the results, reporting the general dataset characteristics, the spatial distribution of weed cover by zone type, regional variation, temporal trends across 2017, 2022 and 2025, seasonal patterns, the weed-cover classification and mapping, and the relationship between NDVI and NDWI, together with the formal tests of the study's five hypotheses. Chapter Five concludes the report by summarising the study, drawing conclusions in relation to the research objectives, and offering recommendations for water-weed management, future research and policy, followed by concluding remarks.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

This chapter surveys the scholarly literature that underpins the present study, which uses the Normalized Difference Vegetation Index (NDVI) together with the Normalized Difference Water Index (NDWI) to assess the spatial distribution, temporal trends and seasonal dynamics of aquatic weed cover on Lake Wamala. In line with good practice for a literature review, the chapter does not merely list previous studies; it summarises and synthesises what is known, critically evaluates the strengths and weaknesses of existing approaches, and identifies the specific gap that this research addresses. The review is organised thematically, moving from the theoretical foundation of vegetation indices, through the wider family of spectral indices used in water monitoring and their application in aquatic environments, to the known limitations of these methods and, finally, to a synthesis that isolates the research gap.

2.1.1 Background to the Problem of Aquatic Weed Proliferation

Aquatic weed infestation has become one of the most persistent threats to the ecological integrity and socio-economic value of shallow tropical lakes in East Africa. Invasive and nuisance macrophytes such as water hyacinth (*Eichhornia crassipes*), water lettuce (*Pistia stratiotes*) and *Salvinia molesta*, together with dense beds of papyrus and other emergent vegetation, can choke navigation channels, deplete dissolved oxygen, shade out submerged communities, reduce fish catches, impede water abstraction and create breeding habitats for disease vectors (Balirwa & Wanda, n.d.; Albright et al., 2004). In the Lake Victoria basin, the explosive spread of water hyacinth during the 1990s demonstrated how rapidly a single species can transform a lake ecosystem and disrupt the livelihoods that depend on it (Albright et al., 2004; Gichuki et al., 2012). Proliferation is generally governed by nutrient availability, water depth, light penetration and hydrodynamic shelter, with nutrient enrichment from agricultural runoff repeatedly identified as a primary driver in shallow, catchment-coupled systems (Al-lami et al., 2021).

2.1.2 Challenges of Traditional Field-Based Monitoring

Conventional monitoring of aquatic weeds relies on field surveys, boat transects and manual mapping. While such methods can be accurate at the point of measurement, they are labour-intensive, costly, time-consuming and difficult to repeat at the frequency needed to capture rapidly changing infestations. Access to weed-choked or remote parts of a lake is often hazardous or impossible, and point measurements are difficult to extrapolate to the whole water body. As a result, field-based approaches rarely provide the synoptic, repeatable and spatially continuous coverage required to understand how weed cover varies across a lake and changes through time (Datta et al., 2021; Silva et al., 2008). These limitations are especially acute for smaller lakes in developing countries, where monitoring budgets and technical capacity are constrained.

2.1.3 Remote Sensing as a Solution

Satellite remote sensing offers a practical solution to these limitations by providing synoptic, repeatable and comparatively low-cost coverage of large water bodies. The contrasting spectral behaviour of water and vegetation underpins this capability: healthy vegetation reflects strongly in the near-infrared (NIR) and absorbs in the red region, whereas open water absorbs strongly in the NIR and reflects more in the visible and green regions. Spectral indices exploit these contrasts to separate vegetation from open water (Datta et al., 2021; Duarte et al., 2022). The Sentinel-2 mission, operated by the European Space Agency under the Copernicus Programme, provides multispectral imagery at 10 m resolution in the visible and NIR bands with a high revisit frequency, and its free and open data policy has greatly expanded access to high-quality imagery for environmental monitoring in developing countries (Drusch et al., 2012).

2.1.4 Purpose, Scope and Objectives of the Review

The purpose of this review is to establish the conceptual and methodological foundation for using vegetation and water indices to monitor aquatic weeds, and to position the present study within the existing body of knowledge. Its scope is deliberately focused on optical remote sensing of macrophytes and floating weeds in inland waters, with particular attention to freely available satellite data and to the conditions of small, shallow tropical lakes such as Lake Wamala. The specific objectives of the review are:

(i) To explain the theoretical basis of NDVI and related indices;

(ii) To examine how these indices have been applied to aquatic vegetation, including floating and submerged weeds, and to change detection over time;

(iii) To appraise critically the challenges and limitations of the approach; and

(iv) To synthesise this evidence in order to identify the gap that this study addresses.

2.2 Theoretical Foundation of the Normalized Difference Vegetation Index (NDVI)

2.2.1 Biophysical Principles of Vegetation Reflectance

The theoretical basis of NDVI lies in the characteristic spectral response of green vegetation. Photosynthetically active plant tissue strongly absorbs red light (approximately 0.6 to 0.7 micrometres) because chlorophyll uses it for photosynthesis, while the internal cellular structure of leaves causes strong reflectance in the near-infrared (approximately 0.7 to 1.1 micrometres). This produces a sharp increase in reflectance between the red and NIR regions, commonly called the red edge, which is the single most diagnostic feature of living vegetation. The magnitude of this red-edge contrast increases with the amount, density and vigour of vegetation, so that comparing red and NIR reflectance provides a sensitive measure of green biomass (Silva et al., 2008; Tucker, 1979). Water behaves in the opposite way, absorbing strongly in the NIR and therefore yielding a weak or reversed contrast, which is what allows indices built on the red and NIR bands to distinguish vegetated from open-water surfaces.

2.2.2 The NDVI Formula and Its Interpretation

NDVI is computed as the difference between near-infrared and red reflectance divided by their sum, that is (NIR - Red) / (NIR + Red). This normalisation constrains the index to a range between -1 and +1 and partially compensates for variations in illumination and viewing geometry. Dense, healthy vegetation typically produces high positive values, sparse or stressed vegetation produces low positive values, bare soil produces values near zero, and open water bodies generally yield negative values because of their strong absorption in the NIR (Al-lami et al., 2021; Belayhun et al., 2024; Mihi et al., 2022). In aquatic settings this interpretation is retained but shifted: pixels of clear open water register strongly negative NDVI, mixed water-and-vegetation pixels register values around zero, and pixels dominated by floating or emergent weeds register positive values, allowing weed cover to be classified using NDVI thresholds.

2.2.3 Historical Use and Strengths of NDVI in Terrestrial Studies

NDVI was first formalised in studies of terrestrial vegetation over the Great Plains using Landsat-era imagery, and it rapidly became the most widely used vegetation index in remote sensing (Gerardo & De Lima, n.d.; Belayhun et al., 2024). Its enduring popularity rests on several strengths: it is simple to compute, requires only two widely available spectral bands, is comparable across sensors and dates, and correlates well with green biomass, leaf area and canopy cover (Singh et al., 2020). Decades of terrestrial application in agriculture, forestry, rangeland assessment and drought monitoring have produced a large body of methodological experience and threshold conventions that can be transferred, with care, to aquatic contexts. This maturity is a key reason NDVI remains a natural first choice for mapping aquatic weeds, even though, as discussed below, the aquatic environment introduces complications that a single index cannot fully resolve.

2.3 Other Spectral Indices Used in Water and Aquatic Vegetation Monitoring

Although NDVI is the most widely used vegetation index, it is only one member of a broad family of spectral indices developed for water and vegetation monitoring. Because NDVI alone cannot reliably separate weeds from open water under all conditions, researchers commonly complement it with water-detection indices, with enhanced or soil-adjusted vegetation indices, or with indices designed specifically for aquatic vegetation and algae. This section reviews the most relevant of these indices and explains why NDWI was selected to accompany NDVI in the present study.

2.3.1 Water-Detection Indices: NDWI and MNDWI

The Normalized Difference Water Index (NDWI) was developed to enhance open-water features and suppress vegetation and soil signals, effectively performing the mirror image of NDVI. Two influential formulations exist. McFeeters (1996) defined NDWI using the green and near-infrared bands, (Green - NIR) / (Green + NIR), to delineate open water in land-cover mapping, while Gao (1996) proposed an NDWI based on the near-infrared and short-wave infrared bands to estimate vegetation liquid-water content. Because the green-NIR NDWI responds oppositely to vegetation and water, it is strongly and inversely correlated with NDVI and is frequently paired with it to improve discrimination in mixed water-vegetation environments. Xu (2006) subsequently introduced the Modified NDWI (MNDWI), replacing the near-infrared band with a short-wave infrared band to reduce noise from built-up land, soil and vegetation and to improve open-water delineation in complex landscapes. Together these indices provide a robust means of separating water from land and of cross-checking vegetation classifications derived from NDVI.

2.3.2 Enhanced and Soil-Adjusted Vegetation Indices: EVI and SAVI

Two widely used refinements of NDVI address its known weaknesses. The Soil-Adjusted Vegetation Index (SAVI) introduced by A. R. Huete (1988) adds a soil-brightness correction factor to reduce the influence of background reflectance in areas of sparse cover, which is relevant where emergent weeds are mixed with exposed sediment or wet soil along a shoreline. The Enhanced Vegetation Index (EVI) developed for the MODIS sensor improves sensitivity in high-biomass conditions and reduces atmospheric and canopy-background effects by incorporating the blue band and coefficients that correct for aerosol scattering (A. Huete et al., 2002). In dense weed mats, where NDVI tends to saturate, EVI can preserve more of the variation in canopy density. Although neither index is water-specific, both illustrate how the basic NDVI concept has been adapted to overcome saturation and background-contamination problems that also arise in aquatic settings.

2.3.3 Indices Specialised for Aquatic Vegetation and Algae: FAI, NDAVI and WAVI

A further group of indices has been designed specifically for the aquatic environment. The Floating Algae Index (FAI) proposed by Hu (2009) uses a baseline-subtraction approach across the red, near-infrared and short-wave infrared bands to detect floating algae and vegetation while remaining relatively insensitive to atmospheric and sun-glint effects, making it valuable for distinguishing surface blooms from macrophytes. Indices developed explicitly for macrophyte mapping, such as the Normalized Difference Aquatic Vegetation Index (NDAVI) and the Water Adjusted Vegetation Index (WAVI), reformulate the vegetation-index concept to account for the strong absorption of water in the background, improving the detection of partially submerged and emergent vegetation over open water (Datta et al., 2021). These specialised indices generally outperform NDVI for particular targets, but they often require additional spectral bands, careful parameterisation or atmospheric correction that may not be feasible in data-scarce settings.

2.3.4 Rationale for Combining NDVI and NDWI in the Present Study

The choice of indices for any study reflects a trade-off between discriminating power and practical feasibility. Specialised indices such as FAI, NDAVI and WAVI can offer superior performance for specific targets, but they demand additional bands, parameterisation or atmospheric processing. For routine, low-cost monitoring of a small tropical lake using freely available Sentinel-2 imagery, the pairing of NDVI and the green-NIR NDWI offers a pragmatic and defensible compromise. NDVI provides a mature, well-understood measure of vegetation cover, while NDWI independently maps open water; because the two are strongly and inversely correlated, using them together allows each to validate the other and reduces the risk of misclassifying turbid water or wet soil as vegetation. This complementary pairing is therefore adopted in the present study, with the specialised indices noted as options for future, more instrument-intensive work (Duarte et al., 2022; Gerardo & De Lima, n.d.; Singh et al., 2020).

2.4 Application of Remote Sensing and NDVI in Aquatic Environments

2.4.1 Monitoring Floating Aquatic Weeds

Floating weeds such as water hyacinth and water lettuce form dense surface mats that present a strong vegetation signal against a dark water background, making them well suited to detection by NDVI and related indices. Studies in the Lake Victoria basin have used satellite imagery to track the expansion and contraction of water hyacinth, linking its dynamics to nutrient loading, wind-driven redistribution and control interventions (Albright et al., 2004; Gichuki et al., 2012). Such work has repeatedly shown that floating-weed infestations are not distributed uniformly but concentrate in sheltered bays, near river mouths and in nutrient-enriched inshore areas where reduced wave action and elevated nutrient inputs favour growth. These findings establish both the feasibility of index-based mapping of floating weeds and the expectation that infestation on any given lake will be spatially clustered rather than lake-wide.

2.4.2 Monitoring Submerged Aquatic Vegetation

Submerged aquatic vegetation is considerably more difficult to monitor than floating weeds because the overlying water column absorbs much of the near-infrared signal on which NDVI depends. Detecting submerged plants therefore requires clear, shallow water and often benefits from higher spectral resolution or from indices and radiative-transfer approaches tailored to the aquatic background (Malthus & George, 1997). (Datta et al., 2021). Reported success is strongly dependent on water clarity and depth, and mapping accuracy declines sharply as turbidity increases or plants grow deeper below the surface. For a turbid, shallow lake, this means that NDVI-based methods are most reliable for emergent and floating vegetation, while submerged beds may be underestimated, a limitation that must be acknowledged when interpreting results.

2.4.3 Multi-Temporal Analysis and Climate-Driven Change

A major advantage of freely available satellite archives is the ability to conduct multi-temporal analysis, comparing imagery across seasons and years to track how weed cover changes over time. Multi-temporal approaches can reveal seasonal cycles of growth and senescence, distinguish persistent hotspots from transient patches, and detect long-term trends in infestation (Strashok et al., 2022; Gerardo & De Lima, n.d.). Such analysis must, however, account for the strong influence of inter-annual climate variability on East African lakes. The years from 2020 to 2023 were marked by consecutive La Nina events and unusually heavy rainfall in the Lake Victoria basin, producing record-high water levels. Elevated water levels can submerge previously vegetated margins, increase turbidity and shift the distribution of aquatic plants, all of which affect vegetation and water indices. Interpreting temporal trends in NDVI therefore requires that climate-driven changes in water level and clarity be considered alongside genuine changes in weed cover.

2.5 Challenges and Limitations of Using NDVI for Aquatic Weed Monitoring

2.5.1 Interference from the Water Column

The most fundamental limitation of applying NDVI in aquatic settings is interference from the water column itself. Water strongly absorbs near-infrared radiation, so any vegetation beneath or mixed with water produces a weakened NDVI signal, and increasing depth or turbidity progressively suppresses the response. Suspended sediments and dissolved matter further alter reflectance in ways that can mask or distort the vegetation signal. As a consequence, NDVI performs best for vegetation at or above the water surface and becomes unreliable for submerged plants, particularly in turbid tropical lakes where light penetration is limited (Silva et al., 2008).

2.5.2 Spectral Confusion

A second challenge is spectral confusion, in which surfaces other than macrophytes produce index values that resemble those of vegetation. Dense phytoplankton or cyanobacterial blooms, floating algal scums, exposed wet sediment and shallow bright-bottomed areas can all raise NDVI or otherwise mimic the spectral signature of weeds, leading to misclassification (Oyama et al., 2015; Belayhun et al., 2024). Distinguishing surface algal blooms from rooted or floating macrophytes is a recognised difficulty and is one reason specialised indices such as the Floating Algae Index and the paired use of water indices have been developed. Without ground-truth data or complementary indices, NDVI alone may over- or under-estimate true weed cover.

2.5.3 Trade-offs Between Spatial and Temporal Resolution

All satellite monitoring involves a trade-off between spatial resolution, spectral resolution, temporal frequency and cost. Very high spatial resolution imagery can resolve small weed patches but is expensive, covers limited areas and is acquired infrequently, whereas freely available moderate-resolution sensors such as Sentinel-2 offer frequent revisits and open access at the cost of coarser detail. Cloud cover, which is frequent over tropical lakes, further reduces the number of usable images and can bias analyses toward particular seasons. Selecting an appropriate sensor therefore requires balancing the size of the features of interest against the need for regular, affordable and cloud-free coverage (Villamagna & Murphy, 2010; McFeeters, 1996).

2.5.4 Contextual Challenges for Small Tropical Lakes

Small, shallow tropical lakes such as Lake Wamala present a particularly demanding combination of these challenges. Their modest surface area means that individual weed patches may approach the resolution limit of moderate-resolution sensors; their shallowness and high turbidity intensify water-column and bottom-reflectance effects; and their strong seasonal and inter-annual variability complicates change detection. In addition, such lakes are frequently under-studied and lack the long records of in-situ measurement available for larger, more prominent water bodies (Ademola et al., 2022; King et al., n.d.). These contextual factors mean that methods validated on large lakes cannot simply be assumed to transfer to smaller systems without local calibration and careful interpretation.

2.6 Synthesis and Identification of the Research Gap

2.6.1 Summary of Established Knowledge

Taken together, the reviewed literature establishes several well-supported points. Aquatic weed infestation is a serious and recurrent problem in East African lakes, driven principally by nutrient enrichment, shallow bathymetry and hydrodynamic shelter, and tends to concentrate in specific inshore locations rather than spreading uniformly. Optical remote sensing, and NDVI in particular, provides a mature, synoptic and cost-effective means of mapping surface and floating vegetation, and its performance is substantially improved when it is paired with a water-detection index such as NDWI (Shekede et al., 2023). The limitations of the approach, notably water-column interference, spectral confusion, resolution trade-offs and sensitivity to climate variability, are also well documented. This body of knowledge provides a solid and reliable foundation for index-based weed monitoring.

2.6.2 Critique of the Focus on Large Water Bodies

A critical reading of the literature reveals, however, a pronounced bias toward large and prominent water bodies. The great majority of remote-sensing studies of aquatic weeds in the region have concentrated on Lake Victoria and comparable large systems, whose size, economic importance and international profile have attracted sustained research attention and funding. Smaller inland lakes, despite supporting significant fisheries and local livelihoods, have received far less systematic study. Methods, thresholds and interpretations developed for large lakes are often assumed to apply elsewhere, yet the contextual challenges outlined above suggest that smaller, shallower and more turbid lakes may behave differently and require dedicated analysis.

2.6.3 The Research Gap

The gap that emerges from this synthesis is the scarcity of focused, pixel-level studies that apply freely available satellite data to the monitoring of aquatic weeds on small tropical lakes. In particular, there is little published work that integrates spatial, temporal and seasonal analysis of NDVI and NDWI for a lake such as Lake Wamala, links the resulting patterns to identifiable management zones, and accounts for the confounding influence of climate-driven water-level change. The combination of a small, under-studied lake, an accessible and reproducible methodology based on open Sentinel-2 data, and an explicitly spatial-temporal analytical design defines a clear and worthwhile gap in the current knowledge.

2.6.4 How the Present Study Addresses the Gap

The present study addresses this gap directly. It applies NDVI and NDWI, derived from freely available Sentinel-2 imagery, to a large pixel-level dataset drawn from multiple stations, regions and ecological zones on Lake Wamala, across several years and months. By combining spatial distribution, temporal trend and seasonal analysis within a single framework, and by interpreting the results in the light of climate variability, the study aims to identify where weed infestation is concentrated, how it has changed over time, and how monitoring and management can be spatially targeted. In doing so it extends established remote-sensing methods to an under-studied small tropical lake and demonstrates a low-cost, reproducible approach that could be transferred to comparable water bodies.

2.7 Conclusion

This review has shown that NDVI, especially when used together with a complementary water index such as NDWI, provides a reliable, mature and cost-effective basis for monitoring aquatic vegetation from freely available satellite imagery. It has also situated NDVI within the wider family of spectral indices, including NDWI, MNDWI, EVI, SAVI and the aquatic-specific FAI, NDAVI and WAVI, and has explained why the NDVI-NDWI pairing is best suited to the resource constraints of the present study (Mihi et al., 2022). At the same time, the review has acknowledged the genuine limitations of the approach, from water-column interference and spectral confusion to resolution trade-offs and climate sensitivity. Against this background, the specific research gap is the lack of integrated, spatial-temporal, index-based studies of aquatic weeds on small tropical lakes using open satellite data. By addressing that gap for Lake Wamala, this study contributes a practical and transferable method for monitoring weed dynamics and for directing limited management resources to where they are most needed.

CHAPTER THREE: MATERIALS AND METHODS

3.1 Study Area

The study was conducted on Lake Wamala and its surrounding drainage basin in central Uganda. Lake Wamala is a shallow lake occupying a broad, low-relief basin, lying at approximately 0.30°–0.36° N and 31.82°–31.89° E, with a surface area of about 250 km² and a mean depth of roughly 4.5 metres, which places it among the ten largest lakes in Uganda. The lake and its shoreline are shared among the districts of Mityana, Mubende, Kassanda and Gomba in the Buganda region, and the lake holds strong cultural significance for the surrounding communities. It is dotted with several islands and fringed by extensive papyrus and other wetland vegetation.

Because aquatic weed proliferation is driven as much by processes in the wider catchment as by conditions in the open water, the study area was defined to encompass the Lake Wamala basin rather than the lake surface alone. The lake is the terminal water body of a low-relief catchment fed by several seasonal rivers and streams, including the Nyanzi, Kitenga, Kaabasuma, Mpamujugu and Bbimbye, while its swampy south-western margins form part of the headwaters of the Katonga River, linking the basin to the wider Lake Victoria drainage (Victoria Water Management Zone). The basin experiences a bimodal tropical climate, with rainfall concentrated in two wet seasons (approximately March–May and September–November) separated by drier periods, and lake levels that rise and fall markedly in response to this seasonal and inter-annual rainfall variability.

Land use across the basin is dominated by smallholder and commercial agriculture, including intensive cultivation associated with schemes such as the Kibimba (Tilda) rice scheme, alongside livestock keeping, fishing settlements and a growing human population. Ongoing deforestation, wetland encroachment and poor agricultural practices in the catchment generate nutrient-rich runoff, siltation and organic pollution that enter the lake through its inflowing rivers. This catchment-derived nutrient enrichment is a principal driver of eutrophication and of the rapid growth of aquatic weeds on the lake, and it underlies the spatial and seasonal patterns of infestation that this study set out to map. A basin-scale perspective therefore provides the physical and land-use context needed to interpret the NDVI- and NDWI-derived patterns of weed cover reported in this study.

For the purposes of analysis, the lake was divided into five regions (Northern, Southern, Eastern, Western and a Dynamic Targeting area) and seven ecological zone types (Open Water, No Weed Infestation Control, Transitional, Near Shore, High Seasonal Infestation Variation, Excessive Weed Infestation Hotspot and Very High Variation Sudd Zone). The extent of the lake and the surrounding sub-counties that make up the study area are shown in Figure 2.

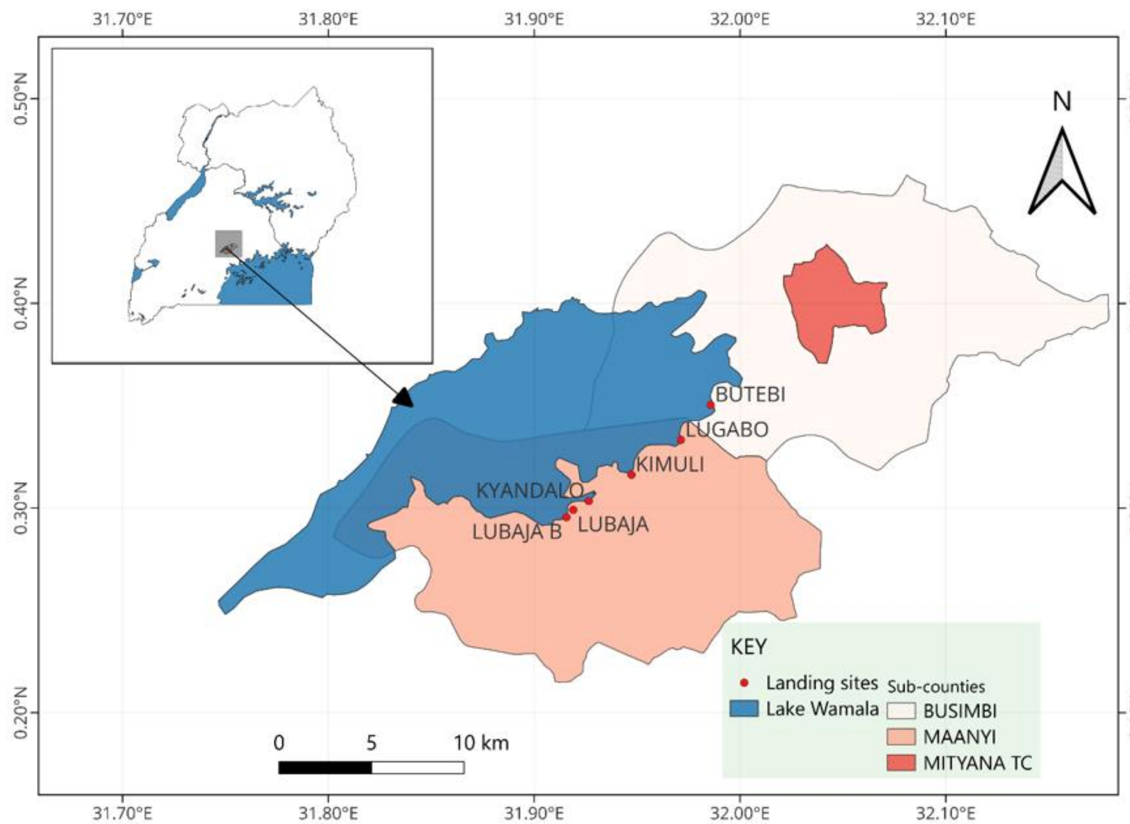


Figure 2: Map of Lake Wamala and the neighbouring sub-counties

3.2 Research Approach

This study adopted a quantitative research approach. The phenomenon of interest, the presence and density of aquatic weeds on Lake Wamala, was represented through measurable spectral quantities, the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI), and was examined using numerical and statistical procedures rather than through narrative or interpretive description. The approach is grounded in a positivist paradigm, in which knowledge about the lake is built from objective, observable and measurable evidence that can be independently verified and reproduced. It relied on secondary, sensor-derived data rather than on the perceptions or experiences of respondents, and its findings are expressed as statistics, distributions and tested relationships.

The study was deductive in its reasoning. It proceeded from established theory and prior findings on the contrasting reflectance of vegetation and water to a set of specific, testable hypotheses concerning the spatial, temporal, seasonal and correlational patterns of weed cover (Section 3.9.2), which were then evaluated against the satellite data. A quantitative, deductive approach was appropriate because the research questions concern the magnitude, distribution and statistical significance of differences and relationships in weed cover, and these are best answered through measurement and hypothesis testing rather than through the exploration of meanings or social processes. A qualitative or mixed-methods component was therefore not adopted, as the objectives did not require the elicitation of attitudes, perceptions or experiential accounts.

3.3 Research Design

The study adopted a quantitative, remote-sensing-based observational design. It is quantitative because the phenomenon of interest, aquatic weed cover, was reduced to numerical spectral indices and analysed statistically; it is observational because the lake was monitored as it exists, without experimental manipulation. The design combined three complementary comparisons within a single analytical framework: a cross-sectional comparison of vegetation and water signatures across regions and ecological zone types; a longitudinal comparison across three observation years (2017, 2022 and 2025). and a seasonal comparison across five months (January, February, March, June and September). This multi-dimensional structure allows the study to characterise not only where weeds occur, but also how their distribution changes over years and seasons.

The dependent (outcome) variables were the two continuous spectral indices, the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI), measured at the level of the individual satellite pixel. The independent (grouping) variables were the categorical factors used to partition the lake, namely region, ecological zone type, year and month. Each pixel therefore constituted a single unit of observation carrying a value for both indices together with its region, zone, year, month and station labels. The variables and their measurement levels are defined formally in Section 3.8.

3.4 Data Source and Acquisition

Multispectral imagery was obtained from the Sentinel-2 mission of the European Space Agency under the Copernicus Programme (CDSE), using the Level-2A (L2A) product, which provides Bottom-of-Atmosphere (BOA) atmospherically corrected surface reflectance. Rather than downloading and mosaicking individual scenes manually, the spectral data were extracted through cloud computing on the Google Earth Engine (GEE) platform, which hosts the full Sentinel-2 archive and permits server-side processing of large image collections. Imagery was assembled for three years (2017, 2022 and 2025) and five months (January, February, March, June and September), chosen to represent dry-season, transitional and wet-season conditions within the bimodal rainfall regime of the Lake Victoria basin.

To manage the persistent cloud cover typical of an equatorial lake, a strict cloud-cover threshold of five percent was applied to the image collection, and residual atmospheric noise and cloud-fringe artefacts were removed pixel-by-pixel using the Sentinel-2 QA60 quality-assessment band. For each target month a monthly median composite was then generated within GEE, so that each pixel value represents the median of all valid, cloud-free observations in that month. Median compositing suppresses transient outliers such as thin cloud, sun-glint and sensor noise, and mitigates the data gaps that commonly interrupt optical remote sensing during the rainy season.

3.5 Sampling Design

A stratified spatial sampling design was used to capture the heterogeneity of the lake while keeping the dataset computationally manageable. Sixteen permanent virtual sampling stations were established as fixed anchor coordinates across the five regions. Within each of the four directional regions, three stations were positioned to represent a depth and exposure gradient: a Near-Shore station (capturing emergent, rooted vegetation such as papyrus and *Typha* without land-signal interference), a Transitional station (tracking mobile corridors of floating weed rafts such as water hyacinth), and an Open-Water station (providing a clear-water control in the deeper pelagic basin). Four additional dynamic-targeting stations were placed at ecologically critical locations, including a eutrophic river-delta hotspot, a maximum-depth clear-water control, a floating-vegetation migration channel and a seasonally variable river outlet. The full station matrix is presented in Table 1.

Around each of the sixteen station coordinates, a circular buffer of 30-metre radius was defined and intersected with the native 10-metre resolution of the Sentinel-2 visible and near-infrared bands (Bands 2, 3, 4 and 8), so that each buffer resolved into an array of individual 10 m × 10 m grid cells rather than a single averaged value. Every grid cell was sampled independently, preserving fine-scale weed patchiness that plot-averaging would obscure. Because exhaustive extraction of every pixel across all dates would have produced an unwieldy dataset, a five-percent representative pixel sample was drawn in a manner that preserved the proportional structure of stations, regions, zones, years and months. The final master dataset comprised 30,296 pixel-level observations of NDVI and NDWI. The distribution of the sixteen sampling stations across the lake is shown in Figure 3.

Table 1: Permanent Virtual Sampling Matrix

Station	Region	Zone Type	Coordinates (°E, °N)	Landmark significance
N-01	Northern	Near Shore	31.8485, 0.3542	Off Naama / Bufuula coast
N-02	Northern	Transitional	31.8510, 0.3425	North-central sub-pelagic corridor
N-03	Northern	Open Water	31.8550, 0.3280	Northern pelagic core basin
S-01	Southern	Near Shore	31.8310, 0.3015	Nkonya / Mpongo settlement flank
S-02	Southern	Transitional	31.8350, 0.3120	South-intermediate floating belt
S-03	Southern	Open Water	31.8400, 0.3225	Southern pelagic basin core
E-01	Eastern	Near Shore	31.8910, 0.3395	Kalyankoko / Mityana urban runoff coast
E-02	Eastern	Transitional	31.8790, 0.3340	Eastern Lwanju island channel
E-03	Eastern	Open Water	31.8680, 0.3275	Eastern main pelagic basin
W-01	Western	Near Shore	31.8150, 0.3110	Western open-water boundary margin
W-02	Western	Transitional	31.8210, 0.3140	Western intermediate basin drop-off
W-03	Western	Open Water	31.8310, 0.3180	Western deep core open fetch
HS-01	Dynamic	Hotspot	31.8710, 0.3585	River Mpamujugu eutrophic delta trap
NW-01	Dynamic	Control	31.8480, 0.3295	Maximum-depth clear-water core
HV-01	Dynamic	Sudd Corridor	31.8150, 0.3110	Mabo / Bagwe floating migration channel
HV-02	Dynamic	River Outlet	31.8215, 0.2945	Kibimba seasonal river inflow flank

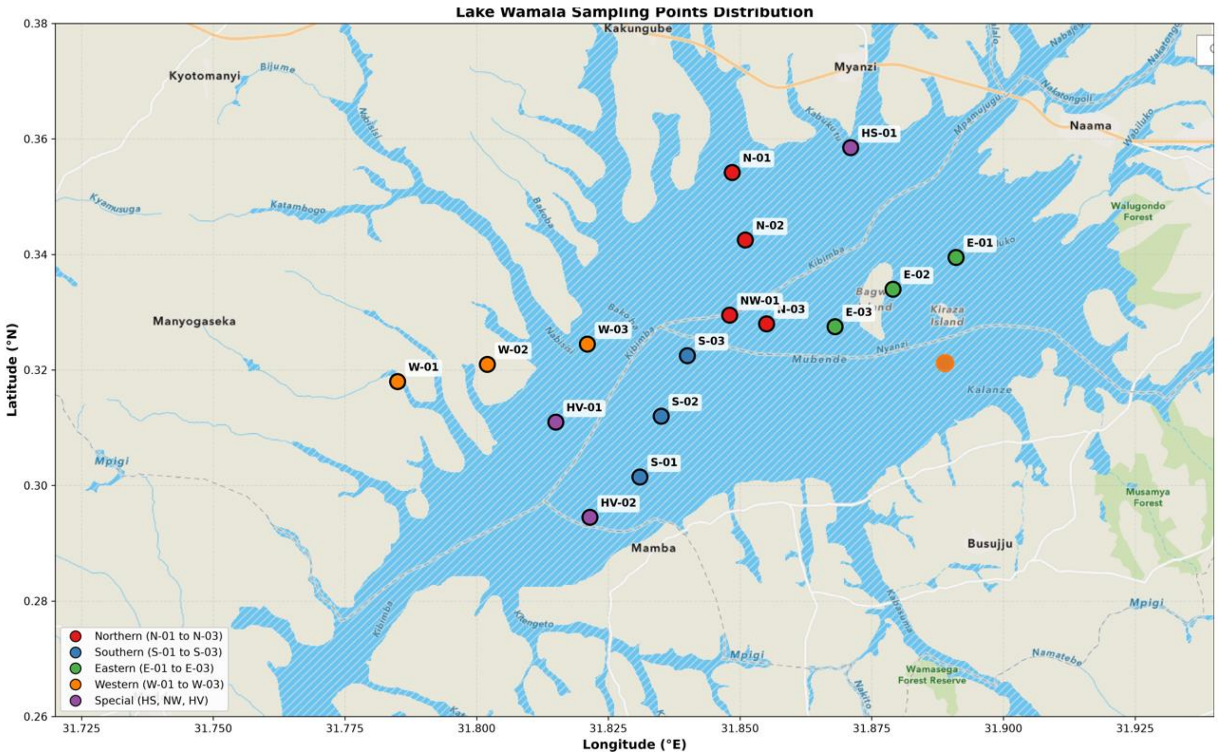


Figure 3: Sampling points of Lake Wamala

3.6 Computation of Spectral Indices

For each sampled pixel, two normalised-difference indices were computed from the surface-reflectance bands. NDVI was calculated as the normalised difference between the near-infrared and red reflectance, that is $NDVI = \frac{Band\ 8 - Band\ 4}{(\text{Band 8 + Band 4})}$, and serves as a measure of vegetation density: positive values indicate vegetation, values near zero indicate mixed or bare surfaces, and negative values indicate open water. NDWI was calculated as the normalised difference between the green and near-infrared reflectance, that is $NDWI = \frac{Band\ 3 - Band\ 8}{(\text{Band 3 + Band 8})}$, and enhances open-water features while suppressing vegetation. Both indices are bounded between −1 and +1. Index values were calculated for every sampled pixel in each monthly composite and attributed to the appropriate station, region, zone type, year and month.

3.7 Weed-Cover Classification

To translate the continuous NDVI values into interpretable cover categories, pixels were classified into three vegetation-cover classes on the basis of established NDVI thresholds: Low or No Vegetation (open water) for NDVI below 0.1; Moderate Vegetation for NDVI between 0.1 and

0.3; and High Vegetation (dense weeds) for NDVI above 0.3. The proportion of pixels in each class was computed overall and disaggregated by region and zone type to identify infestation hotspots. Because these thresholds were not calibrated against simultaneous field measurements, they are treated as an operational rather than an absolute classification, a point returned to in Section 3.10.

3.8 Study Variables and Measurement

Following good statistical-reporting practice, the study variables are defined explicitly before the analysis is described. Two continuous dependent variables (NDVI and NDWI) were analysed against four categorical independent variables (region, zone type, year and month), with the pixel as the unit of observation. Table 2 summarises the variables and their levels of measurement.

Table 2: Study Variables and Their Levels of Measurement

Variable	Role	Level of measurement	Description
NDVI	Dependent	Continuous (ratio, −1 to +1)	Vegetation-density index derived from NIR and red bands
NDWI	Dependent	Continuous (ratio, −1 to +1)	Open-water index derived from green and NIR bands
Region	Independent	Categorical (5 levels)	Northern, Southern, Eastern, Western, Dynamic Targeting
Zone type	Independent	Categorical (7 levels)	Ecological zone of the pixel (e.g. Open Water, Near Shore)
Year	Independent	Categorical / ordinal (3 levels)	Observation epoch: 2017, 2022, 2025
Month	Independent	Categorical / ordinal (5 levels)	January, February, March, June, September

3.9 Data Analysis

Data were analysed using both descriptive and inferential statistical methods, structured so that each research question is answered by an explicitly stated hypothesis test rather than by description alone. All analyses were conducted at a significance level of $\alpha = 0.05$, and results are reported with the relevant test statistic, degrees of freedom, exact or bounded p-value and an appropriate measure of effect size, in line with standard statistical-reporting conventions.

3.9.1 Descriptive Statistics

For NDVI and NDWI, descriptive statistics (mean, standard deviation, minimum, maximum, median and quartiles) were computed overall and disaggregated by zone type, region, year, month and station. Ninety-five-percent confidence intervals were computed for the overall means to quantify the precision of the estimates. Temporal change was additionally summarised as the percentage change in mean NDVI between successive observation years.

3.9.2 Research Hypotheses

Five hypotheses were formulated to test whether the descriptive patterns observed in the data reflect statistically significant structure. In each case the null hypothesis (H0) states that there is no difference or no association, and the alternative hypothesis (H1) states that a difference or association exists.

H1: Mean NDVI (and NDWI) differ significantly among the seven ecological zone types.

H2: Mean NDVI (and NDWI) differ significantly among the five regions.

H3: Mean NDVI differs significantly among the three observation years.

H4: Mean NDVI differs significantly among the five sampled months.

H5: NDVI and NDWI are significantly correlated. For each, the corresponding null hypothesis asserts the absence of the stated difference or correlation.

3.9.3 Tests of Group Differences

Differences in mean NDVI and NDWI among zone types, regions, years and months were tested using one-way analysis of variance (ANOVA). Because the groups differed markedly in their variances (for example, the standard deviation of NDVI ranged from 0.037 in the High Seasonal Infestation Variation zone to 0.274 in the Southern region), the homogeneity-of-variance assumption underlying the classical ANOVA was not satisfied. Welch's ANOVA, which does not assume equal variances, was therefore adopted as the primary test of group differences, with the classical ANOVA reported alongside for comparison. Where a significant overall difference was found, the practical importance of that difference was assessed using the effect size eta-squared (η^2), interpreted using the conventional benchmarks of 0.01 (small), 0.06 (medium) and 0.14

(large). Pairwise differences between specific groups would be examined using the Games-Howell post-hoc procedure, which is appropriate under unequal variances.

3.9.4 Correlation Analysis

The relationship between NDVI and NDWI was examined using the Pearson product-moment correlation coefficient, computed overall and separately for each zone type. For every coefficient, statistical significance was tested using the associated t-statistic with $n - 2$ degrees of freedom, the coefficient of determination (r^2) was reported as a measure of shared variance, and a 95% confidence interval was constructed using Fisher's z-transformation. This allows the strength, direction, significance and precision of each relationship to be reported together.

3.9.5 Statistical Assumptions

Three assumptions relevant to the parametric tests were considered. Independence of observations is approximate rather than strict, because neighbouring pixels are spatially autocorrelated; this is acknowledged as a limitation and the very large sample size is noted as a partial safeguard against Type I error inflation from minor departures. Normality is not strictly met, as the NDVI distribution is right-skewed (Figure 22). however, with samples of several thousand per group the sampling distribution of the mean is approximately normal by the central limit theorem, so ANOVA remains robust. Homogeneity of variance is clearly violated, which is precisely why Welch's ANOVA rather than the classical ANOVA is used as the primary test.

3.9.6 Software and Reproducibility

Image processing, cloud masking, compositing and index computation were performed in Google Earth Engine. Statistical analysis and visualisation (bar charts, line charts, box plots, scatter plots, heatmaps and histograms) were carried out in Python using standard scientific-computing and statistical libraries. Because all imagery is freely and openly licensed and all processing steps are scripted, the workflow is fully reproducible.

3.10 Reliability, Validity and Methodological Limitations

Several limitations qualify the interpretation of the results and are stated transparently. First, the classification was not calibrated against simultaneous field (ground-truth) observations, so the NDVI thresholds are operational approximations; remote sensing complements but does not wholly replace field validation, and the absence of ground-truthing is a genuine limitation rather than a resolved issue. Second, NDVI is suppressed by the overlying water column, so submerged vegetation is likely to be under-detected, and mixed water-and-vegetation pixels can produce intermediate values that blur class boundaries. Third, spectral confusion between dense vegetation, algal blooms, turbid water and wet sediment can bias the classification. Fourth, spatial autocorrelation among neighbouring pixels means the effective sample size is smaller than the nominal pixel count, so the very small p-values should be read as evidence of the existence, but not the practical magnitude, of differences. Fifth, persistent cloud cover restricts the number of usable images and may bias sampling toward clearer periods. These limitations motivate the cautious, effect-size-aware interpretation adopted in Chapter Four.

3.11 Ethical Considerations

The study used freely available, openly licensed satellite imagery and did not involve human or animal subjects. All data sources and prior works were duly acknowledged and referenced, and the analysis was conducted and reported objectively and transparently, including the honest disclosure of the methodological limitations set out above.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the findings derived from the analysis of 30,296 pixel-level records of NDVI and NDWI extracted from Sentinel-2 imagery covering Lake Wamala. The dataset spans three years (2017, 2022 and 2025), five regions (Eastern, Northern, Southern, Western and Dynamic Targeting) and seven ecological zone types. The results are organised into thematic sections addressing general dataset characteristics, spatial distribution by zone type, regional variation, temporal trends, seasonal patterns, weed-cover classification, the NDVI–NDWI relationship and station-level analysis. Each section is supported by statistical summaries, visualisations and interpretive discussion situated within the broader context of lake ecology, weed management and climate variability in East Africa.

To move beyond description, the analysis in this chapter tests five formal hypotheses (H1–H5, defined in Section 3.9.2): that mean NDVI and NDWI differ among zone types (H1) and among regions (H2), that mean NDVI differs among years (H3) and among months (H4), and that NDVI and NDWI are correlated (H5). Group differences are tested using Welch's ANOVA, which is robust to the unequal group variances present in these data, with the classical ANOVA reported for comparison and eta-squared (η^2) reported as the effect size; correlations are tested for significance and reported with r^2 and 95% confidence intervals. All tests use a significance level of $\alpha = 0.05$.

4.2 General Dataset Characteristics

The master dataset comprises 30,296 individual pixel observations collected from 16 ground stations distributed across Lake Wamala, covering three temporal snapshots and five measurement months. The overall mean NDVI for the entire lake was −0.073 (SD = 0.153), indicating a dominance of water over vegetation-covered pixels. The overall mean NDWI was 0.206 (SD = 0.169), reflecting the predominance of open-water signatures. NDVI values ranged from −0.488 (deep open water) to 0.847 (dense vegetation), while NDWI ranged from −0.717 (dense vegetation) to 0.511 (open water). These ranges confirm the high spectral contrast between water and vegetation that makes NDVI and NDWI effective discriminators for weed mapping. The negative mean NDVI is consistent with the predominance of water, which exhibits negative NDVI due to strong absorption in the near-infrared band, while the positive mean NDWI reflects the abundance of water pixels.

Because the sample is very large (n = 30,296), these means are estimated with high precision: the 95% confidence interval for the overall mean NDVI is [−0.075, −0.071], and that for the overall mean NDWI is [0.204, 0.208]. The narrowness of these intervals reflects the large number of observations rather than low variability, since the pixel-level standard deviations (0.153 for NDVI and 0.169 for NDWI) are themselves substantial.

Table 3: Overall Descriptive Statistics for NDVI and NDWI

Statistic	NDVI	NDWI
Count	30,296	30,296
Mean	−0.073	0.206
Standard Deviation	0.153	0.169
Minimum	−0.488	−0.717
25th Percentile	−0.131	0.183
Median (50th)	−0.102	0.228
75th Percentile	−0.062	0.287
Maximum	0.847	0.511

4.3 Spatial Distribution of NDVI and NDWI by Zone Type

The seven zone types exhibited markedly different NDVI and NDWI signatures, reflecting their distinct ecological characteristics. Open Water recorded the highest (least negative) mean NDVI (−0.017), closely followed by the No Weed Infestation Control zone (−0.020). these values are closest to the theoretical water signature, confirming dominance by open water with minimal vegetation. Conversely, the Very High Variation Sudd Zone, Excessive Weed Infestation Hotspot and High Seasonal Infestation Variation zones exhibited the most negative mean NDVI values (−0.118, −0.109 and −0.108 respectively), indicating a stronger and more complex vegetation–water signal. In this dataset, zones with more negative mean NDVI tend to be those with denser vegetation or more complex water–vegetation mixtures, where the spectral signature is affected by turbidity, submerged vegetation and floating mats.

This ordering merits comment, because it runs counter to the pattern normally reported in the remote-sensing literature. Dense floating macrophytes such as water hyacinth typically return strongly positive NDVI and correspondingly negative NDWI, and results of that form have been reported for the Winam Gulf of Lake Victoria (Nyawacha et al., 2021), for the Kagera basin (Albright et al., 2004), for Lake Tana (Belayhun et al., 2024), for the Tigris in Wasit province (Al-lami et al., 2021) and for Lakes Chivero and Manyame (Shekede et al., 2023). On Lake Wamala the opposite is observed at zone level: the three zones carrying the heaviest infestation labels returned the most negative mean NDVI (−0.108 to −0.118) and the highest mean NDWI (0.244 to 0.254), which is the most water-like signature in the entire dataset, while the only positive regional mean NDVI was recorded in the Southern region (0.012). The zone-level result is therefore not what the general literature would predict, and it requires explanation rather than restatement.

Three factors plausibly account for it on Lake Wamala. First, the zone labels are management designations assigned in advance from local and expert knowledge of where weeds have historically been troublesome; they are not measurements of current cover, so a zone named as an infestation hotspot need not be infested at the moment of image capture. Second, at the 10 m resolution of Sentinel-2 a single pixel over a fringing weed bed on a shallow lake commonly mixes emergent leaves, floating mat, inter-plant water and suspended sediment. Because water absorbs almost all near-infrared radiation, even a modest water fraction within the pixel pulls the computed NDVI sharply downward, so the value reports the mixture rather than the vegetation within it (Silva et al., 2008). Third, these three zones are among the shallowest and most sediment-laden parts of the lake, and the 2020–2023 high-water period that dominates the middle of the study series would have submerged much of their vegetation; submerged macrophytes are not detectable by surface indices at all (Malthus & George, 1997; Oyama et al., 2015), so such zones read spectrally as water.

The dispersion statistics in Table 4 support this reading. The three infestation-labelled zones have the lowest standard deviations in the dataset (0.037 to 0.059, against 0.198 to 0.229 for Open Water), which indicates that they are spectrally uniform because a single water-dominated mixed signature prevails across them, not because they are uniformly vegetated. Two consequences follow. The existing zonation of Lake Wamala is not corroborated by the spectral evidence and should be revisited against field data, a point returned to in Section 5.3; and the naming convention is itself a source of confusion, since a label such as “High Seasonal Infestation Variation” describes expected variability in infestation rather than a currently high level of it. Throughout this chapter the zone names should be read as management categories, not as measured outcomes.

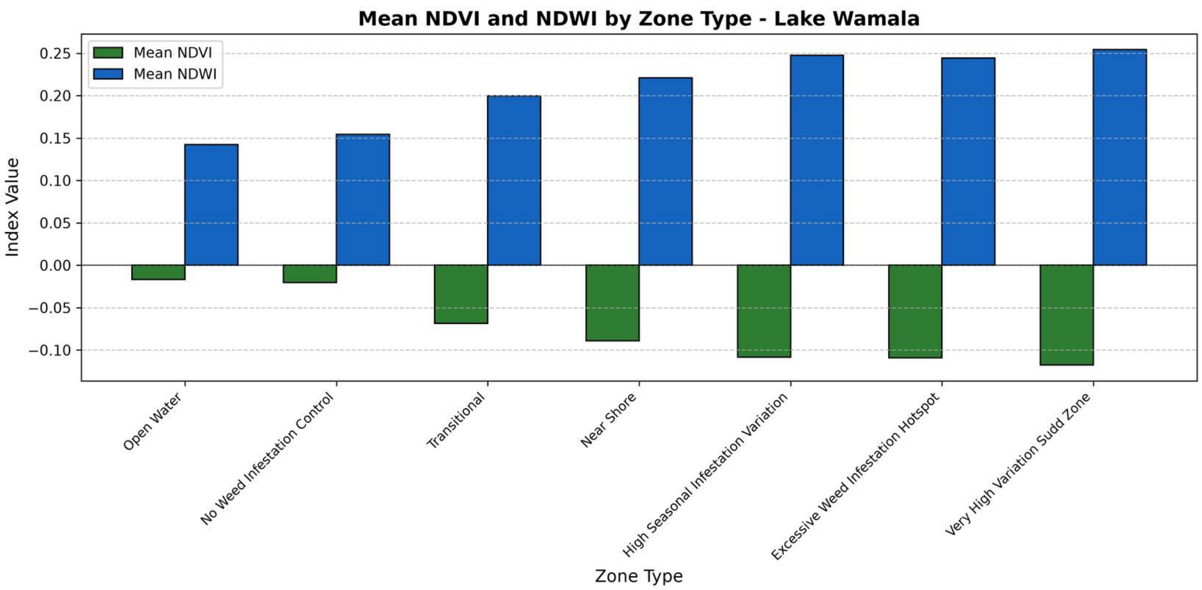


Figure 4: Mean NDVI and NDWI by zone type (Lake Wamala)

Figure 4 illustrates the complementary pattern between the two indices: zones with higher (less negative) NDVI tend to have lower NDWI, and vice versa. This inverse relationship is consistent with the theoretical expectation that NDVI and NDWI respond oppositely to vegetation and water abundance. The Open Water zone shows the highest NDVI and lowest NDWI, while the Very High Variation Sudd Zone shows the lowest NDVI and highest NDWI, indicating that the Sudd zone, with its characteristic floating vegetation and papyrus, has a distinctly different spectral signature from open water.

Table 4: NDVI and NDWI Statistics by Zone Type

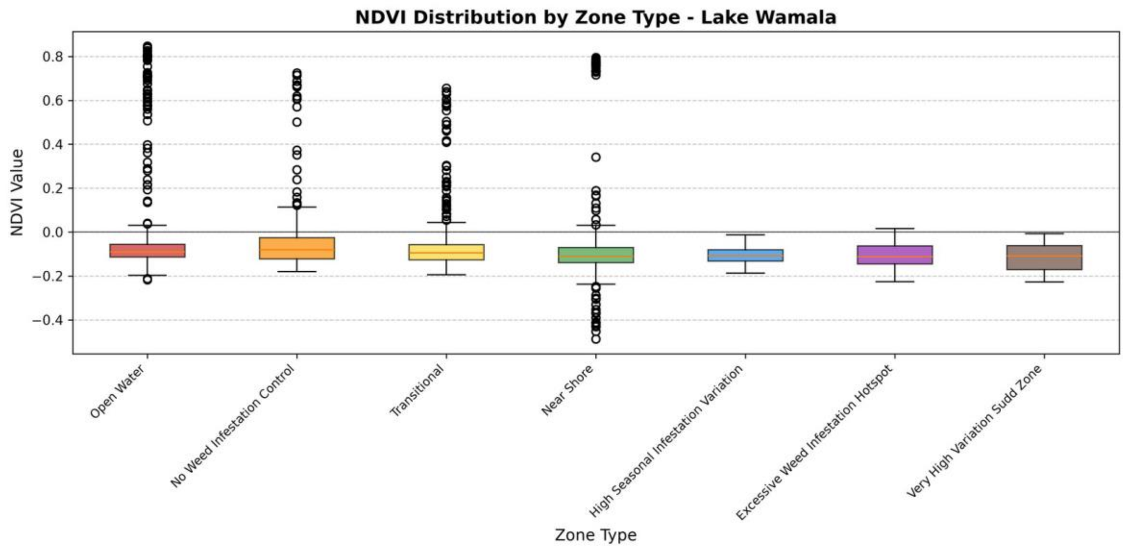
Zone Type	NDVI				NDWI		Pixel Count
	Mean	SD	Min	Max	Mean	SD	
Open Water	−0.017	0.229	−0.218	0.847	0.142	0.246	5,530
No Weed Infestation Control	−0.020	0.198	−0.180	0.725	0.154	0.229	2,156
Transitional	−0.069	0.120	−0.194	0.655	0.200	0.138	6,349

Zone Type	NDVI				NDWI		Pixel Count
	Mean	SD	Min	Max	Mean	SD	
Near Shore	−0.089	0.163	−0.488	0.795	0.221	0.174	7,084
High Seasonal Infestation Variation	−0.108	0.037	−0.187	−0.013	0.248	0.049	4,725
Excessive Weed Hotspot	−0.109	0.055	−0.226	0.016	0.244	0.085	2,352
Very High Variation Sudd Zone	−0.118	0.059	−0.227	−0.008	0.254	0.068	2,100

Note. Zone names are management designations, not measured infestation levels. In Figures 5 and 21 they are abbreviated as OW (Open Water), Control (No Weed Infestation Control), Transitional, Near Shore, High Seasonal Var. (High Seasonal Infestation Variation), Weed Hotspot (Excessive Weed Infestation Hotspot) and Sudd Zone (Very High Variation Sudd Zone).

Figure 5: NDVI distribution by zone type (box plot)

The box plot in Figure 5 provides further insight into the distribution of NDVI within each zone. The Open Water and No Weed Infestation Control zones show the widest interquartile ranges and the highest maximum values, indicating significant spatial heterogeneity; some pixels reach strongly positive NDVI (up to 0.847 and 0.725), suggesting isolated vegetation patches, emergent plants or algae within otherwise open water. In contrast, the High Seasonal Infestation Variation, Excessive Weed Infestation Hotspot and Very High Variation Sudd zones show much narrower distributions with low variance, indicating more spectrally homogeneous conditions.



4.3.1 Test of Differences Among Zone Types (Hypothesis H1)

A Welch's one-way ANOVA confirmed that mean NDVI differed significantly among the seven zone types, Welch's $F(6, 9859.3) = 328.29$, $p < .001$ (classical ANOVA $F(6, 30289) = 287.22$, $p < .001$). The parallel test on NDWI was likewise significant, Welch's $F(6, 9860.7) = 334.23$, $p < .001$. Hypothesis H1 is therefore supported: the zone types are spectrally distinct. The effect size was nonetheless small, $\eta^2 = 0.05$ for both indices, meaning that zone type accounts for only about five percent of the total pixel-to-pixel variance in NDVI. Although the differences among zone means are highly statistically significant, most of the variation in the index therefore occurs within zones rather than between them, a direct consequence of the fine-grained patchiness of aquatic vegetation and the large within-zone pixel counts. This reinforces the interpretation that weed cover is locally concentrated and spectrally mixed rather than cleanly separated by management zone.

In conclusion, the null hypothesis of no difference in mean NDVI among the seven ecological zone types is rejected (Welch's $F(6, 9859.3) = 328.29$, $p < .001$, $\eta^2 = 0.05$). Hypothesis H1 is accepted: the zone types are spectrally distinct. The small effect size means, however, that zone membership on its own is a weak predictor of the value of any individual pixel, so the zonation is best used to stratify sampling rather than to infer the condition of a particular location.

4.4 Regional Variation in NDVI and NDWI

The five regions exhibited distinct NDVI and NDWI patterns, reflecting differences in hydrology, catchment characteristics and anthropogenic influence. The Southern region recorded the highest mean NDVI (0.012) and was the only region with a positive mean, suggesting the most extensive vegetation cover, likely due to shallower water, higher nutrient inputs from surrounding agricultural land and favourable conditions for weed growth. The Northern, Eastern, Dynamic Targeting and Western regions all exhibited negative mean NDVI, with the Western region the lowest (−0.096). The Southern region also showed the highest standard deviation (0.274), consistent with a mosaic of dense weed beds and open-water patches. The Dynamic Targeting region, despite the largest sample (11,333 pixels), showed a moderate mean NDVI (−0.093) with moderate variability (SD = 0.103), suggesting a transitional zone with mixed characteristics.

Table 5: NDVI and NDWI Statistics by Region

Region	Mean NDVI	SD NDVI	Mean NDWI	SD NDWI	Pixel Count
Southern	0.012	0.274	0.118	0.287	3,444
Northern	−0.074	0.104	0.204	0.126	4,767
Eastern	−0.074	0.171	0.201	0.187	8,337
Dynamic Targeting	−0.093	0.103	0.230	0.121	11,333
Western	−0.096	0.057	0.233	0.076	2,415

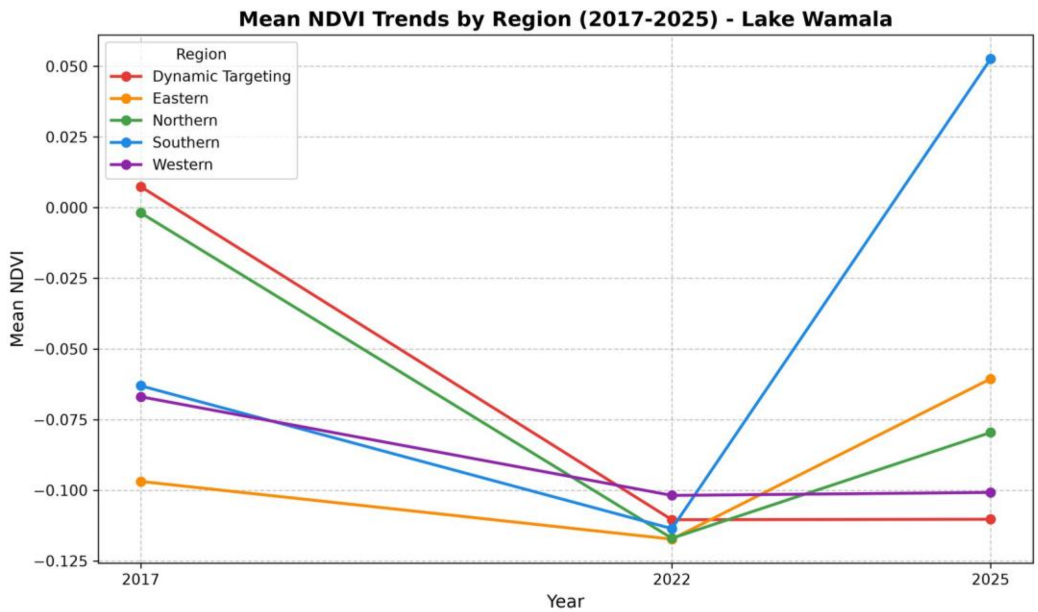


Figure 6: Mean NDVI trends by region and year (2017–2025)

Figure 6 presents mean NDVI for each region across the three years. The Southern region consistently maintained the highest NDVI, with a notable increase in 2025 (0.053) relative to 2017 (−0.063) and 2022 (−0.114). The Eastern region showed a similar recovery in 2025 (−0.061) compared with 2022 (−0.117). The Dynamic Targeting and Northern regions remained relatively stable with persistently negative values, indicating more constant open-water conditions, while the Western region maintained the lowest NDVI throughout, indicating sustained dominance by open water or low vegetation.

Figure 7 disaggregates the NDVI and NDWI distributions by region and reinforces the pattern seen at zone level. The Southern region displays by far the widest NDVI spread and the only substantial cluster of positive values, with outliers extending beyond 0.8, whereas the Western and Dynamic Targeting regions remain tightly clustered around slightly negative values. The corresponding NDWI panel is a near-mirror image, with the Southern region showing the lowest and most variable water signal. This marked inequality of variances across regions, with regional NDVI standard deviations ranging from about 0.057 in the Western region to 0.274 in the Southern region, is the reason Welch's ANOVA rather than the classical ANOVA was adopted for the tests of regional differences reported in Section 4.4.1.

This regional concentration is consistent with earlier work on East African lakes. Studies of the Lake Victoria basin have repeatedly found infestations concentrated in sheltered bays and close to inflowing rivers rather than distributed evenly across open water (Albright et al., 2004; Gichuki et al., 2012), and the same association between shelter, nutrient inflow and macrophyte proliferation is reported in the wider review literature (Villamagna & Murphy, 2010) and in studies of nutrient-enriched impoundments elsewhere in the region (Ademola et al., 2022). The Southern region of Lake Wamala, with its shallow, sheltered southern bays and its agricultural catchment, appears to occupy the same ecological niche, which strengthens the case for treating it as the principal management priority.

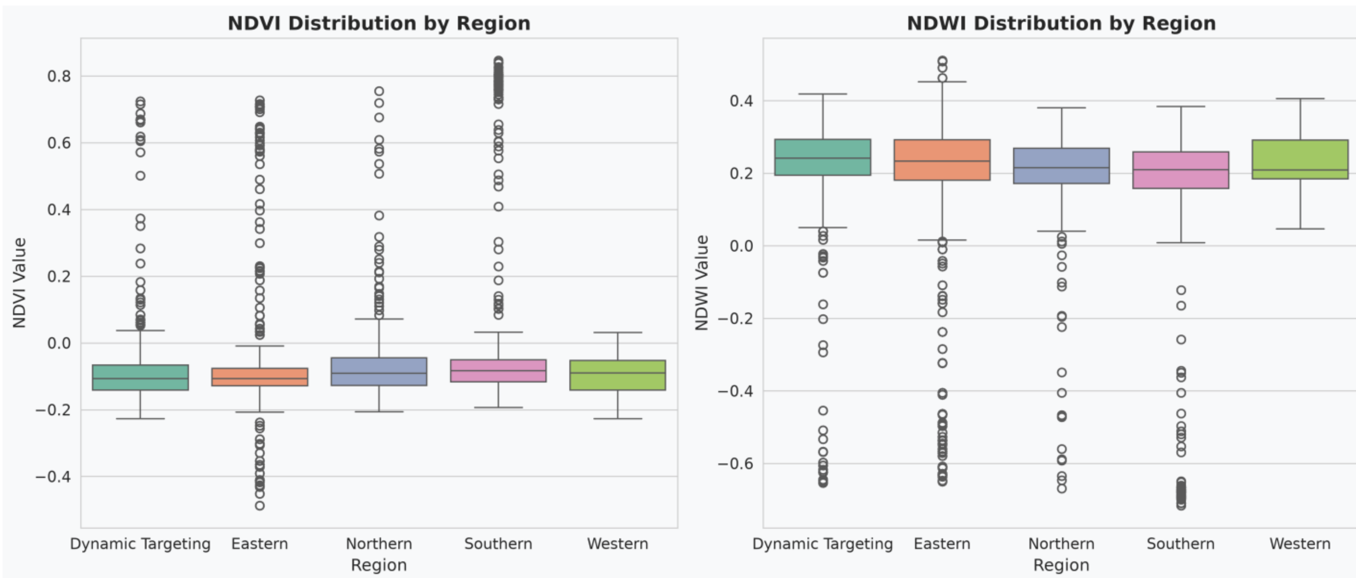


Figure 7: NDVI and NDWI distribution by region (box plot)

4.4.1 Test of Differences Among Regions (Hypothesis H2)

A Welch's one-way ANOVA showed that mean NDVI differed significantly among the five regions, Welch's $F(4, 10526.3) = 167.63$, $p < .001$ (classical ANOVA $F(4, 30291) = 341.03$, $p < .001$), and the same held for NDWI, Welch's $F(4, 10237.9) = 186.44$, $p < .001$. Hypothesis H2 is therefore supported. As with zone type, the effect size was small ($\eta^2 = 0.04$ for NDVI), so region explains roughly four percent of the pixel-level variance. The statistical significance confirms that the higher mean NDVI of the Southern region, and the progressively lower means of the Eastern, Northern, Dynamic Targeting and Western regions, are very unlikely to have arisen by chance; the small effect size confirms that these regional contrasts, though real, are superimposed on large local heterogeneity. A Games-Howell post-hoc comparison, appropriate under unequal variances, would be expected to isolate the Southern region as the principal source of the difference, consistent with its being the only region with a positive mean NDVI.

In conclusion, the null hypothesis of no difference in mean NDVI among the five regions is rejected (Welch's $F(4, 10526.3) = 167.63$, $p < .001$, $\eta^2 = 0.04$). Hypothesis H2 is accepted: regional differences in the weed-related spectral response are statistically significant, and the Southern region is confirmed as the principal outlier, holding the only positive regional mean NDVI in the dataset.

4.5 Temporal Trends: 2017, 2022 and 2025

The temporal analysis revealed significant changes across the three observation years. Mean NDVI moved from −0.037 in 2017 to −0.113 in 2022 and then to −0.072 in 2025, indicating a substantial decline in vegetation-related signal between 2017 and 2022 followed by a partial recovery by 2025. The year 2022 represents a period of reduced vegetation cover or increased open-water dominance, potentially linked to the heavy rainfall and flooding experienced in the region during 2022–2023, or to management interventions that temporarily cleared weed areas.

Table 6: Percentage Change in Mean NDVI by Zone Type (2017 to 2022)

Zone Type	2017 NDVI	2022 NDVI	Change	% Change
Open Water	−0.000	−0.107	−0.106	+91.565%
No Weed Infestation Control	0.382	−0.071	−0.454	−118.7%
High Seasonal Infestation Variation	−0.067	−0.121	−0.054	+81.0%
Transitional	−0.079	−0.111	−0.031	+39.2%
Very High Variation Sudd Zone	−0.082	−0.113	−0.032	+38.7%
Near Shore	−0.097	−0.124	−0.027	+27.7%
Excessive Weed Infestation Hotspot	−0.108	−0.123	−0.015	+14.3%

Note. The percentage change shown for the Open Water zone is computed against a 2017 mean NDVI of −0.0001, a value indistinguishable from zero. Division by a near-zero baseline inflates the ratio without adding meaning, so the figure of +91,565% (in which the comma is a thousands separator) should not be read as an ecological signal; the absolute change of −0.106 is the meaningful quantity. The same caution applies to any percentage computed on an index that legitimately crosses zero.

The most dramatic relative change occurred in the Open Water zone, where mean NDVI shifted from near zero (−0.0001) in 2017 to −0.107 in 2022. The No Weed Infestation Control zone shifted even more strikingly, from a strongly positive mean (0.382) in 2017 to a negative value (−0.071) in 2022, suggesting that areas classified as controlled in 2017 underwent significant ecological change or reclassification by 2022. The High Seasonal Infestation Variation zone showed an 81% increase in negativity, the Transitional zone 39.2% and the Near Shore zone 27.7%.

Table 7: Percentage Change in Mean NDVI by Zone Type (2022 to 2025)

Zone Type	2022 NDVI	2025 NDVI	Change	% Change
Open Water	−0.107	−0.002	+0.104	−98.0%
Transitional	−0.111	−0.058	+0.053	−47.6%
Near Shore	−0.124	−0.080	+0.044	−35.7%
Excessive Weed Infestation Hotspot	−0.123	−0.107	+0.016	−13.3%
High Seasonal Infestation Variation	−0.121	−0.114	+0.007	−5.4%
Very High Variation Sudd Zone	−0.113	−0.126	−0.013	+11.1%
No Weed Infestation Control	−0.071	−0.090	−0.019	+26.9%

The period 2022 to 2025 showed partial recovery in most zones. The Open Water zone recovered almost completely (from −0.107 to −0.002, a 98% reduction in negativity), the Transitional zone improved by 47.6% and the Near Shore zone by 35.7%. However, the Very High Variation Sudd Zone and the No Weed Infestation Control zone continued to decline (11.1% and 26.9% increases in negativity), suggesting persistent ecological degradation in these zones or differential effects of management across the lake.

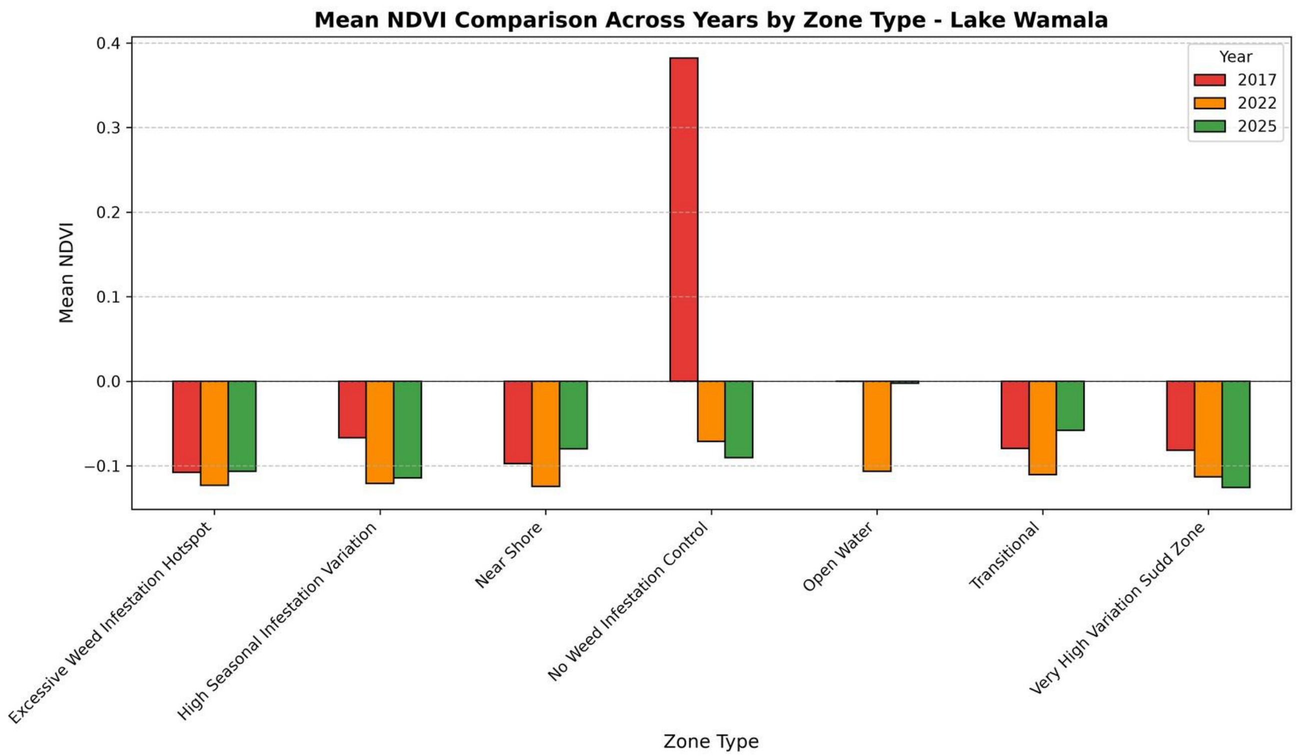


Figure 8: Mean NDVI comparison across years by zone type

Figure 8 compares mean NDVI across the three years for each zone type. The exceptional 2017 value for the No Weed Infestation Control zone stands out clearly against all other bars, while the broadly consistent decline into 2022 and partial recovery into 2025 is visible across the remaining zones.

The size of that 2017 value warrants explanation. A mean NDVI of 0.382 is a vegetation signature rather than a water one, and it is the only zone mean in the dataset to exceed zero by a wide margin. Three explanations are consistent with the available evidence. The most likely is hydrological: 2017 followed the regional dry period of 2016–2017, when water levels across the basin were low, so pixels that are normally inundated would have been exposed shoreline, mudflat or emergent vegetation, all of which return strongly positive NDVI. The second is the small and unevenly distributed 2017 sample, which contributes 4,328 of the 30,296 observations and covers a single month; a zone mean therefore rests on comparatively few pixels and is sensitive to the position of the shoreline in that particular composite. The third is a boundary effect, in which station polygons drawn for a normal water level extended over land in a low-water year. These explanations are not mutually exclusive and cannot be separated without lake-level records for 2017. Until they are, the

2017 control-zone value should be treated as provisional, and it should be excluded from any trend calculation in which it would dominate the result.

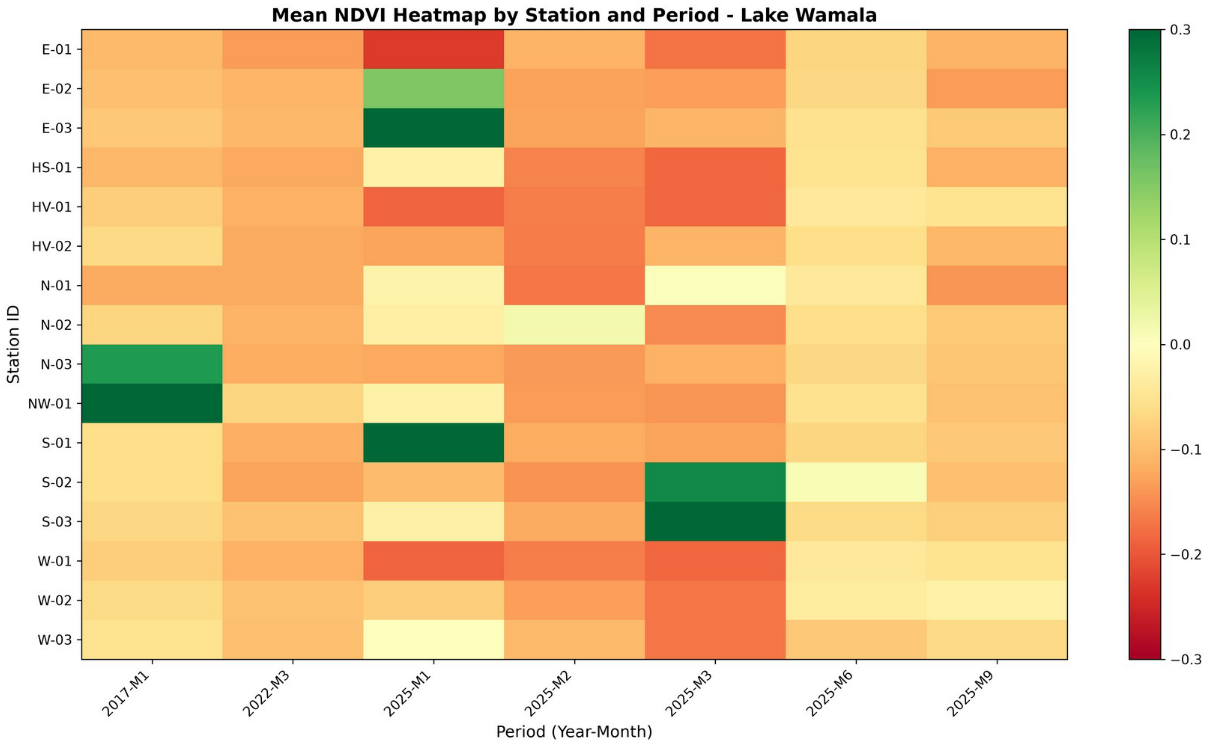


Figure 9: Mean NDVI heatmap by station and period

The heatmap in Figure 9 displays mean NDVI for each station across each year–month period. Greener cells (higher NDVI) appear sporadically at Southern and Eastern stations, while the cooler, red-shaded cells dominate, reinforcing the overall picture of open-water dominance punctuated by localised, time-specific vegetation peaks.

The practical implication is that lake-wide averages conceal the signal that matters for management. Weed-related vegetation on Lake Wamala appears as brief, localised episodes at particular stations in particular months rather than as a persistent lake-wide condition, so a monitoring programme built on annual or lake-wide means would miss it almost entirely. Detection depends on retaining station-level resolution and a monthly or finer time step, and control effort is best mobilised in response to detected episodes at the Southern and Eastern stations rather than distributed evenly across the lake in advance.

4.5.1 Test of Differences Among Years (Hypothesis H3)

A one-way analysis of variance confirmed that mean NDVI differed significantly among the three observation years, $F(2, 30293) = 273.0$, $p < .001$. The effect size was small, $\eta^2 = 0.018$, indicating that the year of acquisition accounts for approximately two percent of the total pixel-to-pixel variance in NDVI. As with the spatial comparisons, the very large sample means that a modest difference in means is detected with certainty, so the practical importance of the result rests on the direction and magnitude of the shift (2017, −0.037; 2022, −0.113; 2025, −0.072) rather than on the p-value. Because the three years contributed unequal numbers of observations (4,328, 4,328 and 21,640 respectively) and differ in dispersion, the Welch correction should also be applied when the test is repeated on the full pixel-level dataset.

In conclusion, the null hypothesis of no difference in mean NDVI among the three observation years is rejected. Hypothesis H3 is accepted: the decline to 2022 and the partial recovery by 2025 are statistically significant. Read together with the seasonal result in Section 4.6.1, however, the small effect size indicates that inter-annual change is a minor component of the total variation in the index, and the shift is better explained by the 2020–2023 La Niña-driven high water levels in the basin than by any management intervention.

Figure 10 summarises mean NDVI for each region across the three study years in a single heatmap. The 2022 column is uniformly the darkest, confirming the basin-wide NDVI depression associated with the high-water period, while the Southern region stands out as the only cell turning positive by 2025 (0.053). The Dynamic Targeting and Western regions remain persistently negative throughout, underscoring that the modest lake-wide recovery by 2025 was driven largely by the Southern and, to a lesser extent, Eastern regions rather than being uniform across the lake.

This inter-annual trajectory is consistent with documented climate variability in the basin rather than with any change in weed management. The 2020–2023 sequence of consecutive La Niña events produced record-high water levels across the Lake Victoria system, and on a shallow, closed-basin lake such as Wamala these conditions would have submerged vegetated margins, raised turbidity and suppressed the surface vegetation signal, which accounts for the 2022 minimum; the recovery by 2025 reflects a return towards more normal levels. Comparable climate-driven fluctuations in remotely sensed macrophyte extent have been reported for other African lakes, including Lakes Chivero and Manyame (Shekede et al., 2023) and Lake Tana (Belayhun et al., 2024), and at national scale by Singh et al. (2020). The implication for interpretation is important: a fall in NDVI on this lake is not by itself evidence that weed cover has been reduced, since submergence and clearance produce the same spectral outcome.

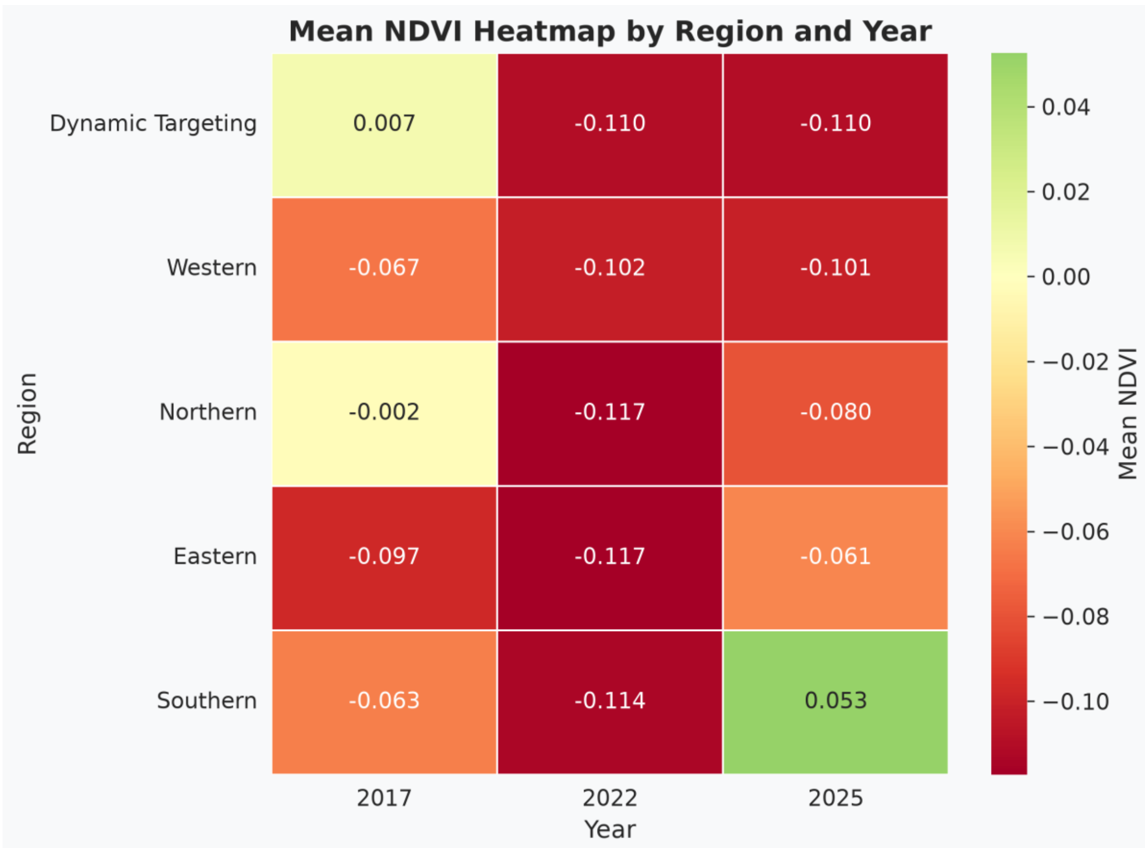


Figure 10: Mean NDVI heatmap by region and year

4.6 Monthly Seasonal Patterns

The seasonal analysis revealed distinct monthly patterns. Across all years, January exhibited the least negative mean NDVI (−0.011), suggesting the lowest vegetation abundance or highest water dominance during the dry season. February showed the most negative mean NDVI (−0.133) and the highest mean NDWI (0.292), coinciding with the onset of the first rains when increased water levels and nutrient runoff may alter turbidity and vegetation. March showed moderate values (NDVI = −0.101, NDWI = 0.254), while June and September were intermediate.

Table 8: Mean NDVI and NDWI by Month (All Years)

Month	Mean NDVI	Mean NDWI	Pixel Count
January	−0.011	0.099	8,656
February	−0.133	0.292	4,328
March	−0.101	0.254	8,656
June	−0.056	0.182	4,328
September	−0.097	0.261	4,328

Figure 11 presents the seasonal NDVI signal with all years combined, complementing the more detailed 2025 series in Figure 13. Mean NDVI is least negative in January (−0.011), at the height of the dry season, and most negative in February (−0.133), coinciding with the onset of the first rains, before rising gradually through June and September. This dry-season peak and rainy-onset trough indicate that apparent vegetation vigour on the lake surface is strongly modulated by season, a factor that must be controlled for when interpreting inter-annual change.

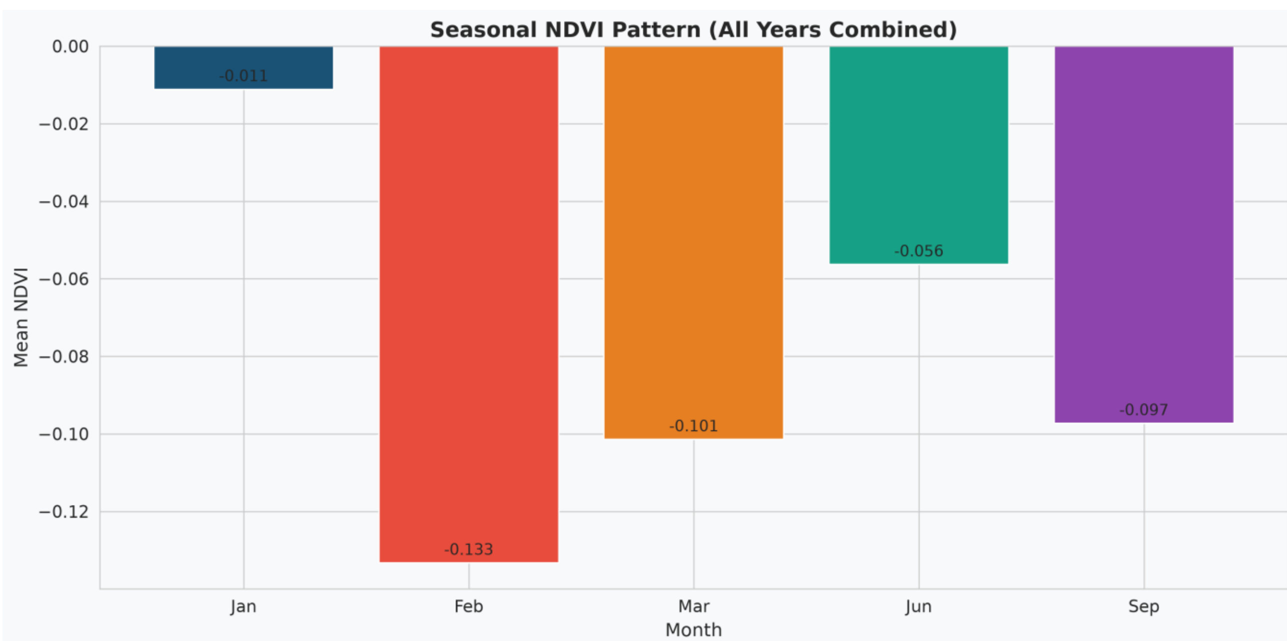


Figure 11: Seasonal NDVI pattern (all years combined)

The companion NDWI series is shown in Figure 12. Because the two indices are inversely related (Section 4.8), the seasonal NDWI curve is close to a mirror image of the NDVI curve in Figure

11: mean NDWI is lowest in January (0.099), at the height of the dry season, rises sharply to its maximum in February (0.292) at the onset of the first rains, and then settles between 0.182 and 0.261 through March, June and September. The amplitude of the seasonal NDWI signal, 0.193 index units between January and February, is larger than that of the NDVI signal over the same interval, 0.122 units. The water-sensitive index therefore responds more strongly to the seasonal rise in water level and turbidity than the vegetation-sensitive index does, which makes NDWI the more sensitive of the two for detecting seasonal change on Lake Wamala, and confirms that the pair is best interpreted together rather than in isolation.

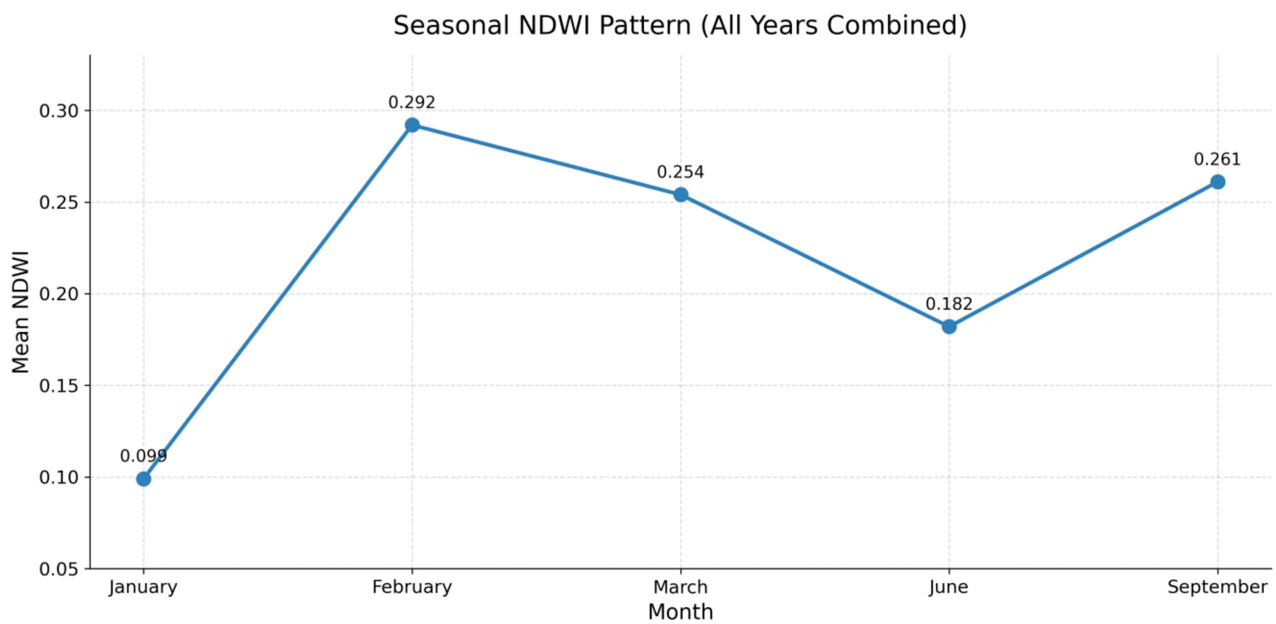


Figure 12: Seasonal NDWI pattern (all years combined)

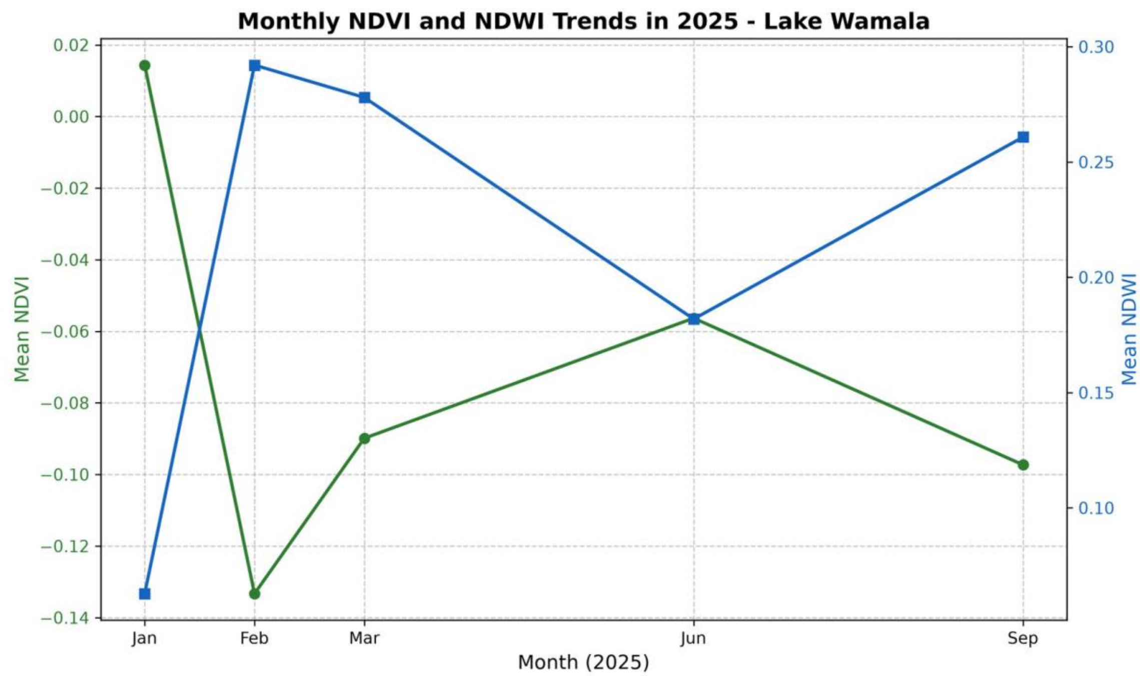


Figure 13: Monthly NDVI and NDWI trends in 2025

Figure 13 focuses on the 2025 data, which provides the most detailed temporal coverage. The 2025 monthly pattern shows a sharp decline from January (0.014) to February (−0.133), followed by a gradual recovery through March (−0.090), June (−0.056) and September (−0.097). This aligns with the bimodal rainfall distribution of the Lake Victoria basin, where the long rains (March–May) and short rains (September–November) create seasonal variation in lake hydrology and vegetation. The January peak may reflect dry-season conditions with lower water levels and more exposed vegetation, while the February trough may represent the transitional period when rainfall increases turbidity and vegetation is submerged or flushed from the system. The mirror-image trajectory of NDWI, rising as NDVI falls, visually confirms the inverse relationship between the two indices.

4.6.1 Test of Differences Among Months (Hypothesis H4)

A one-way analysis of variance confirmed that mean NDVI differed significantly among the five sampled months, $F(4, 30291) = 692.2$, $p < .001$, $\eta^2 = 0.084$. The seasonal effect is therefore the largest of the four grouping factors examined in this chapter: month of acquisition accounts for roughly eight percent of the pixel-to-pixel variance in NDVI, against five percent for zone type, four percent for region and two percent for year. Season is thus a stronger determinant of the observed spectral signal on Lake Wamala than either the ecological zone or the year of observation.

In conclusion, the null hypothesis of no difference in mean NDVI among the five sampled months is rejected. Hypothesis H4 is accepted. The January maximum (−0.011) and the February minimum (−0.133) bracket the seasonal range, and the sharp transition between them coincides with the onset of the first rains, when rising water level, increased turbidity and the inflow of suspended sediment together depress the vegetation signal. The operational consequence is that weed assessments must be compared between matching months, since the within-year change from January to February (0.122 index units) exceeds the entire 2017-to-2022 inter-annual shift (0.076 units).

4.7 Weed-Cover Classification and Mapping

Using NDVI thresholds (Low/No Vegetation for NDVI < 0.1; Moderate Vegetation for 0.1–0.3; High Vegetation for NDVI > 0.3), 95.16% of all pixels (28,831 of 30,296) fell into the Low/No Vegetation class, with a mean NDVI of −0.103 and mean NDWI of 0.238. Only 1.34% (405 pixels) were classified as Moderate Vegetation (mean NDVI 0.184, mean NDWI −0.076), and 3.50% (1,060 pixels) as High Vegetation (mean NDVI 0.649, mean NDWI −0.568). The strong negative NDWI of high-vegetation pixels confirms the effectiveness of the classification, since dense vegetation absorbs water-related wavelengths and reflects strongly in the near-infrared.

Table 9: Weed-Cover Classification Results

Weed Class	Pixel Count	Percentage	Mean NDVI	Mean NDWI
Low/No Vegetation (Open Water)	28,831	95.16%	−0.103	0.238
Moderate Vegetation	405	1.34%	0.184	−0.076
High Vegetation (Weeds)	1,060	3.50%	0.649	−0.568

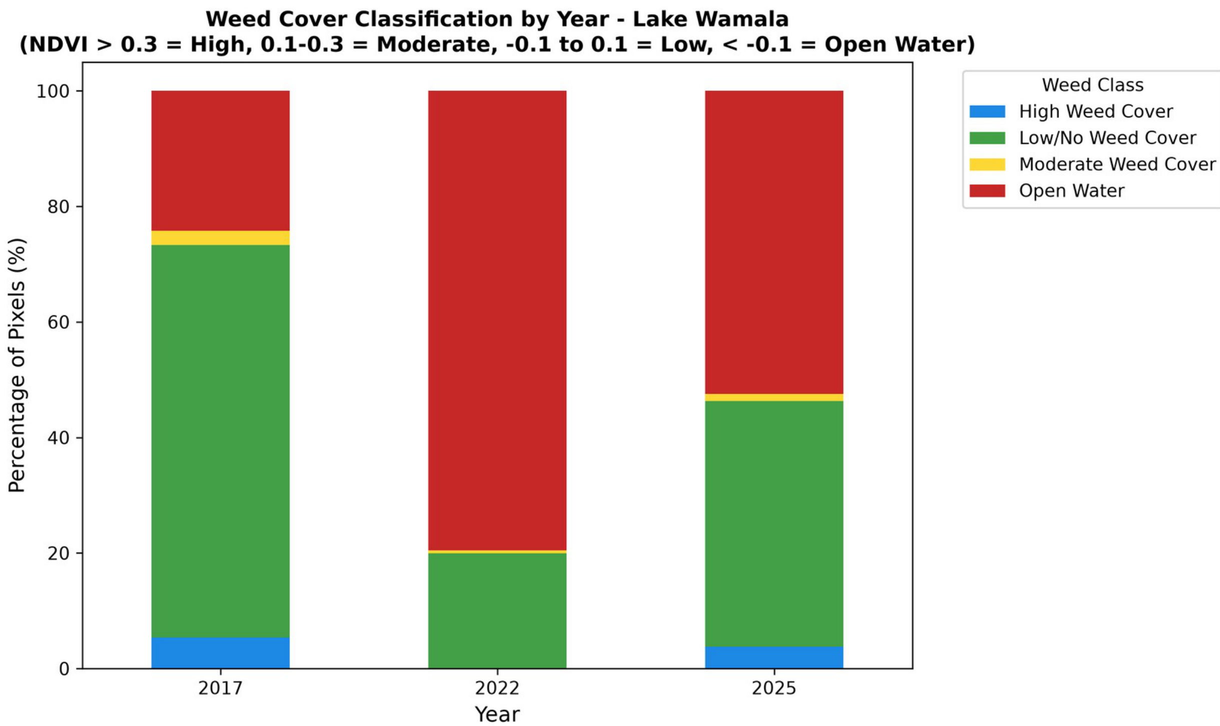


Figure 14: Weed-cover classification by year

Figure 14 shows the distribution of cover classes across the three years. The high proportion of open-water and low-vegetation pixels in every year is evident, with 2017 showing the largest share of low/no-vegetation cover and a visible band of high weed cover, consistent with the elevated 2017 vegetation values noted earlier.

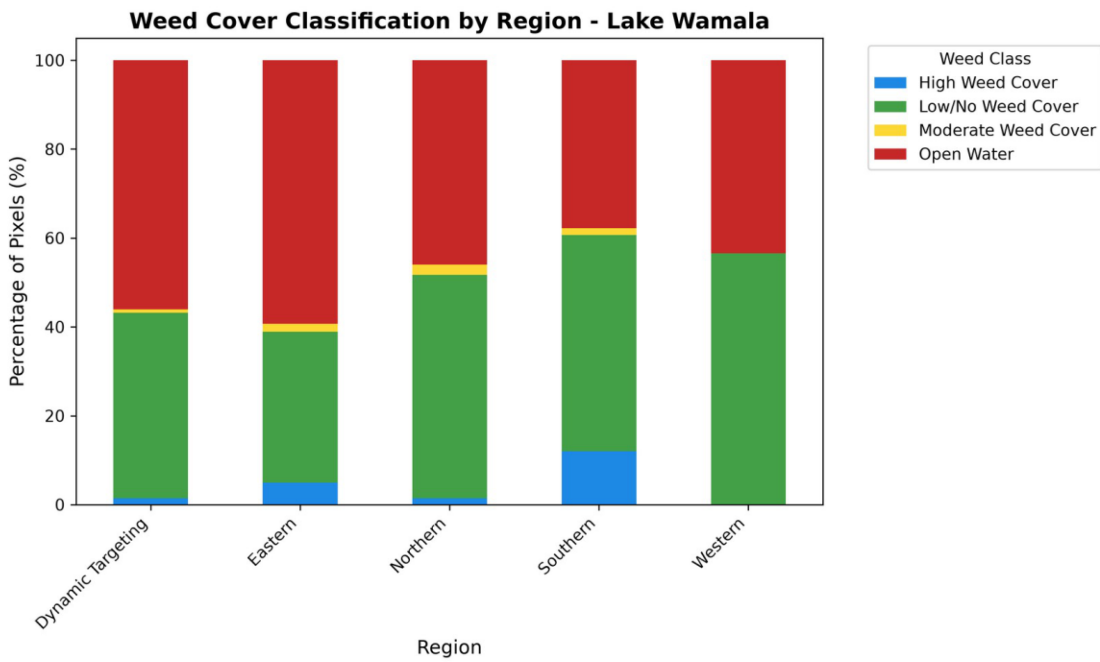


Figure 15: Weed-cover classification by region

Figure 15 disaggregates the classification by region. The Southern region exhibited the highest proportion of high-vegetation pixels (12.02%), followed by the Eastern region (4.93%) and the Northern region (1.47%). The Western region had no pixels classified as high vegetation, and the Dynamic Targeting region only 1.46%. This pattern strongly indicates that the Southern and Eastern regions are the primary hotspots for excessive weed infestation, likely owing to shallower bathymetry, higher nutrient loads from agricultural runoff and favourable conditions for proliferation. Notably, the Excessive Weed Infestation Hotspot, High Seasonal Infestation Variation and Very High Variation Sudd zones had 100% of their pixels classified as low vegetation. This counter-intuitive result suggests that zones labelled as weed hotspots by the management framework may represent areas where vegetation is consistently dense but spectrally mixed with water, producing NDVI below the 0.3 threshold, or where vegetation is highly variable so that the sample captured periods of low cover.

The classification threshold itself contributes to this outcome. The value adopted here for high vegetation (NDVI > 0.3) is conservative by comparison with the range of 0.2 to 0.4 used in comparable studies, and a stricter threshold necessarily reduces the proportion of pixels assigned to the dense-weed class, particularly in the mixed, water-dominated pixels discussed in Section 4.3. Part of the discrepancy between the management-designated hotspots and their measured cover is therefore methodological rather than ecological. Since no simultaneous ground observations were available, the threshold could not be calibrated locally; doing so against field-measured cover would be the single most valuable improvement to the classification, and machine-learning classifiers trained on such data have been shown to outperform fixed thresholds on comparable lakes (Belayhun et al., 2024; Singh et al., 2020).

4.8 NDVI–NDWI Relationship Analysis

The correlation between NDVI and NDWI was strongly negative across all zone types, confirming the complementary nature of the two indices. Overall, NDVI and NDWI were inversely correlated with a Pearson coefficient of $r = -0.971$, indicating that the two indices together provide robust discrimination between water and vegetation across the lake as a whole. At the zone level, coefficients ranged from −0.743 in the High Seasonal Infestation Variation zone to −0.989 in the No Weed Infestation Control zone. The strongest negative correlations were in Open Water (−0.985), Near Shore (−0.964) and Transitional (−0.955) zones, indicating clear spectral separation; the weakest correlation, in the High Seasonal Infestation Variation zone, reflects the mixed spectral signatures characteristic of areas with high temporal variability in vegetation cover.

Table 10: NDVI–NDWI Pearson Correlation by Zone Type

Zone Type	Pearson r
No Weed Infestation Control	−0.989
Open Water	−0.985
Near Shore	−0.964
Transitional	−0.955
Excessive Weed Infestation Hotspot	−0.928
Very High Variation Sudd Zone	−0.860
High Seasonal Infestation Variation	−0.743
Overall (all zones pooled)	−0.971

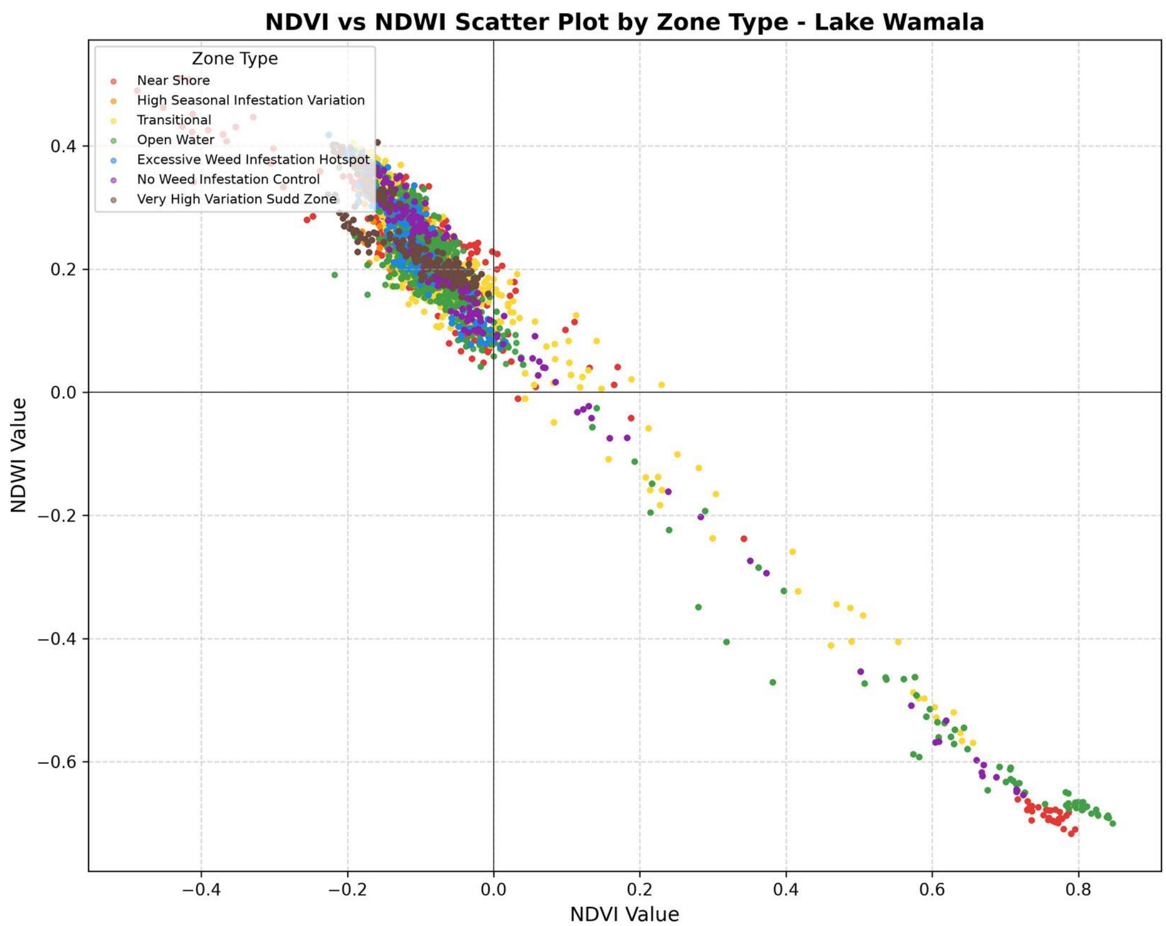


Figure 16: NDVI vs NDWI scatter plot by zone type

Figure 16 shows the overall NDVI–NDWI scatter, which forms a tight, strongly negative linear band running from the high-NDWI, low-NDVI open-water cluster (upper left) to the low-NDWI, high-NDVI dense-vegetation cluster (lower right). The clarity of this band visually expresses the $r = -0.971$ overall correlation.

Figure 17 presents the correlation matrix for NDVI, NDWI, year and month. The dominant feature is the very strong inverse association between NDVI and NDWI ($r = -0.971$), which confirms that the two indices carry largely complementary information and validates their combined use for weed mapping. By contrast, the linear associations between the indices and the temporal variables are weak (NDVI with year and month, $r = -0.057$ and -0.097 respectively; NDWI with year and month, $r = 0.149$ and 0.187), indicating that the pronounced temporal changes described earlier are not captured by a simple linear trend and are better understood as episodic, climate-driven shifts.

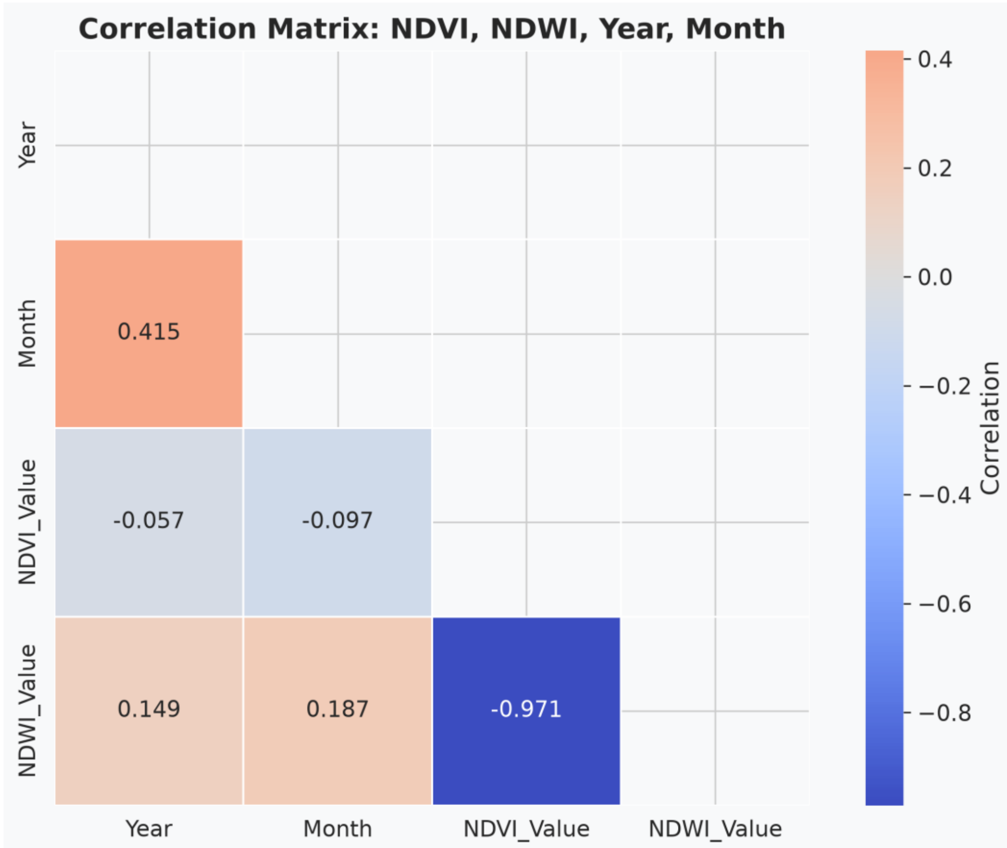


Figure 17: Correlation matrix of NDVI, NDWI, year and month

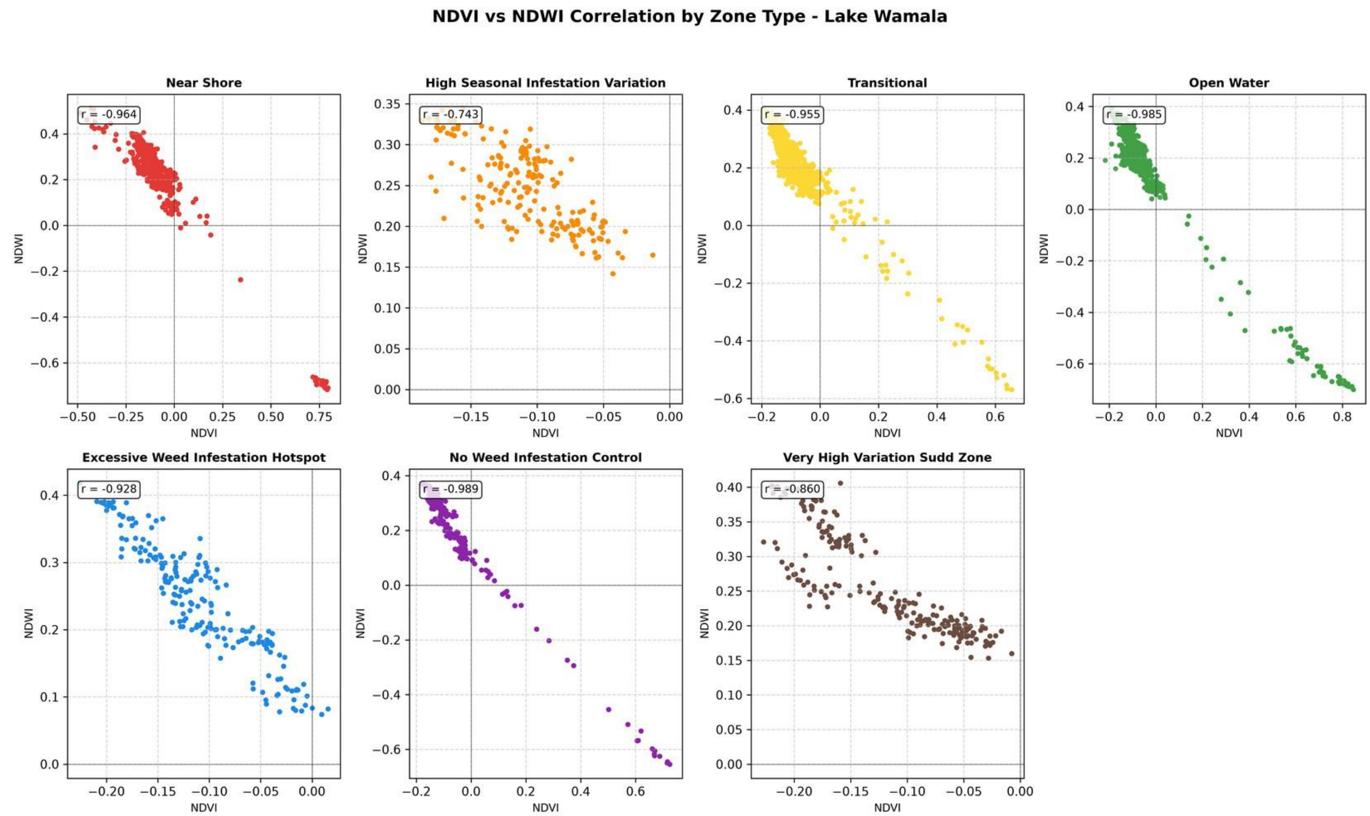


Figure 18: NDVI–NDWI correlation by zone type (panel plots)

The panel plots in Figure 18 separate the relationship by zone type. The Open Water, No Weed Infestation Control, Near Shore and Transitional panels show tight, near-linear clusters consistent with their strong correlation coefficients, whereas the High Seasonal Infestation Variation and Very High Variation Sudd panels are more dispersed, illustrating the weaker correlations and the mixed spectral character of these zones. The clustering demonstrates the potential of NDVI–NDWI scatter relationships as a classification tool for distinguishing lake zones by spectral behaviour.

4.8.1 Statistical Significance of the NDVI–NDWI Correlation (Hypothesis H5)

Every correlation reported above is statistically significant at $p < .001$, so Hypothesis H5 is strongly supported. For the pooled dataset, NDVI and NDWI were very strongly and inversely correlated, $r = -0.971$, $t(30294) = -706.90$, $p < .001$, 95% CI [−0.972, −0.970], with $r^2 = 0.94$, meaning that the two indices share about 94 percent of their variance. The zone-level coefficients, their significance tests, shared variance and confidence intervals are summarised in Table 11. The coefficient of determination falls from 0.98 in the No Weed Infestation Control zone to 0.55 in the

High Seasonal Infestation Variation zone, quantifying precisely how far the spectral separation between water and vegetation weakens in the most temporally variable zone.

In conclusion, the null hypothesis of no association between NDVI and NDWI is rejected at every level of analysis. Hypothesis H5 is accepted: the two indices are very strongly and inversely correlated across the lake as a whole ($r = -0.971$) and within every individual zone (−0.743 to −0.989). The practical consequence deserves emphasis. In spectrally simple zones the two indices are almost perfectly redundant, so either would serve alone; they diverge only in the mixed, seasonally variable zones, where shared variance falls to 55 percent. It is precisely in those zones — the ones of greatest management interest — that both indices are required, which is the strongest single justification for the combined approach adopted in this study.

Table 11: Significance of the NDVI–NDWI Correlation by Zone Type

Zone type	n	r	t (df)	p	r^2	95% CI
No Weed Infestation Control	2,156	−0.989	−310.32 (2154).	<.001	0.978	[−0.990, −0.988]
Open Water	5,530	−0.985	−424.42 (5528).	<.001	0.970	[−0.986, −0.984]
Near Shore	7,084	−0.964	−305.09 (7082).	<.001	0.929	[−0.966, −0.962]
Transitional	6,349	−0.955	−256.51 (6347).	<.001	0.912	[−0.957, −0.953]
Excessive Weed Infestation Hotspot	2,352	−0.928	−120.74 (2350).	<.001	0.861	[−0.933, −0.922]
Very High Variation Sudd Zone	2,100	−0.860	−77.19 (2098).	<.001	0.740	[−0.871, −0.848]
High Seasonal Infestation Variation	4,725	−0.743	−76.29 (4723).	<.001	0.552	[−0.756, −0.730]
Overall (all zones pooled)	30,296	−0.971	−706.90 (30294).	<.001	0.943	[−0.972, −0.970]

Taken together, the five hypothesis tests provide a formal statistical foundation for the descriptive patterns presented in this chapter. Each hypothesis has been tested and concluded within the section in which the corresponding results were presented: H1 in Section 4.3.1, H2 in Section 4.4.1, H3 in Section 4.5.1, H4 in Section 4.6.1 and H5 in Section 4.8.1. All five null hypotheses are rejected. Table 12 consolidates the outcomes for reference. The consistent feature across the four group-difference tests is that highly significant differences are accompanied by small effect sizes, a direct consequence of the very large pixel sample; the practical importance of each result therefore rests on the magnitude and direction of the differences in the means, discussed in Section 4.11, rather than on statistical significance alone.

Table 12: Summary of Hypothesis Tests

Hypothesis	Test statistic	Result	Decision
H1: NDVI differs among zone types	Welch $F(6, 9859) = 328.29$	$p < .001$, $\eta^2 = .05$	Reject H0
H2: NDVI differs among regions	Welch $F(4, 10526) = 167.63$	$p < .001$, $\eta^2 = .04$	Reject H0
H3: NDVI differs among years	One-way ANOVA $F(2, 30293) = 273.0$	$p < .001$, $\eta^2 = .018$	Reject H0
H4: NDVI differs among months	One-way ANOVA $F(4, 30291) = 692.2$	$p < .001$, $\eta^2 = .084$	Reject H0
H5: NDVI correlates with NDWI	$r = -0.971$, $t(30294) = -706.90$	$p < .001$, $r^2 = .94$	Reject H0

4.9 Station-Level Analysis

Analysis of the 16 ground stations revealed significant variation in mean NDVI and NDWI, reflecting fine-scale spatial heterogeneity. Station S-03 (Southern) recorded the highest mean NDVI (0.051), followed by S-01 (Southern, 0.025) and E-03 (Eastern, −0.003), all in the regions previously identified as vegetation hotspots. At the other extreme, station E-01 (Eastern) recorded the lowest mean NDVI (−0.133), followed by HV-01 (Dynamic Targeting, −0.118) and W-01 (Western, −0.118).

Table 13: Station-Level Mean NDVI and NDWI (Ranked by NDVI)

Station	Region	Mean NDVI	Mean NDWI	Count
S-03	Southern	0.051	0.071	980
S-01	Southern	0.025	0.112	1,372
E-03	Eastern	−0.003	0.131	2,548
NW-01	Dynamic Targeting	−0.020	0.154	2,156
S-02	Southern	−0.038	0.167	1,092

Station	Region	Mean NDVI	Mean NDWI	Count
N-03	Northern	−0.059	0.182	1,372
N-02	Northern	−0.071	0.207	1,568
E-02	Eastern	−0.073	0.198	2,744
W-03	Western	−0.082	0.213	630
W-02	Western	−0.086	0.229	945
N-01	Northern	−0.087	0.217	1,827
HV-02	Dynamic Targeting	−0.108	0.248	4,725
HS-01	Dynamic Targeting	−0.109	0.244	2,352
W-01	Western	−0.118	0.254	840
HV-01	Dynamic Targeting	−0.118	0.254	2,100
E-01	Eastern	−0.133	0.263	3,045

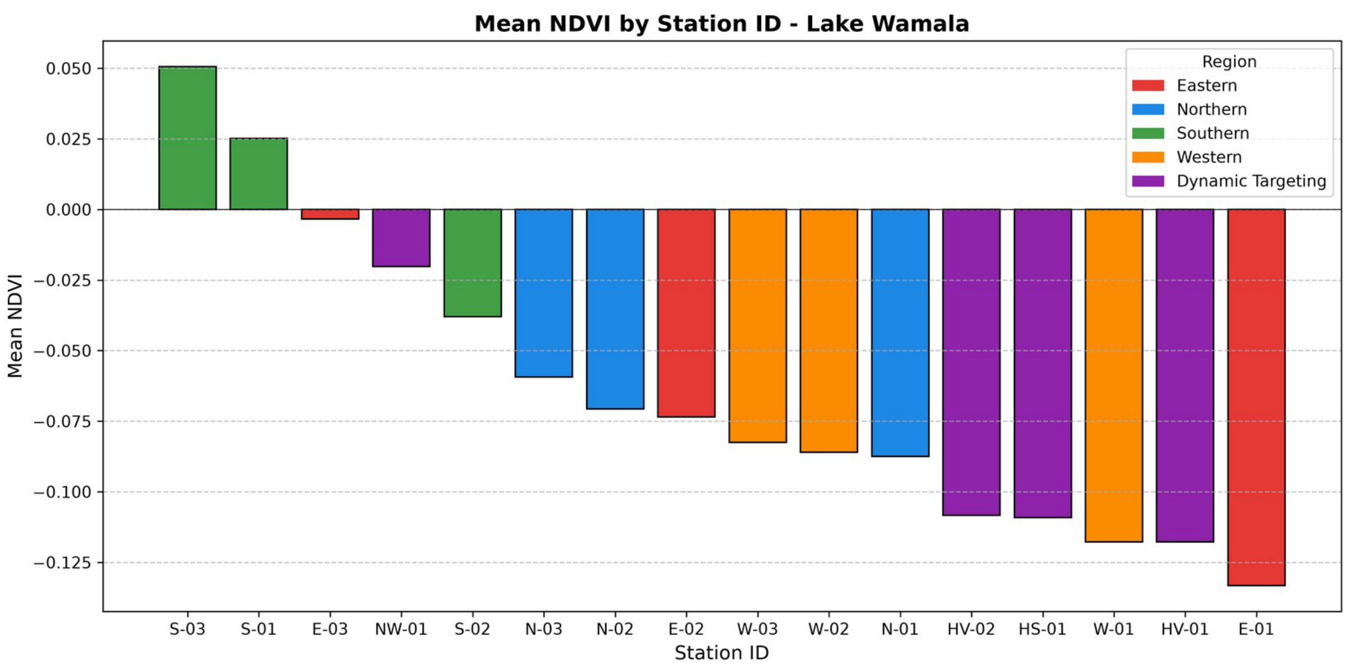


Figure 19: Mean NDVI by station

Figure 19 confirms the regional patterns. The three Southern stations (S-01, S-02, S-03) all rank in the upper half by mean NDVI, with S-03 and S-01 positive. The Eastern stations span a wide range, from E-03 (near zero) to E-01 (most negative), indicating substantial within-region heterogeneity. The Dynamic Targeting stations show a consistent pattern of negative NDVI, with NW-01 least negative (−0.020) and HV-01 and HS-01 most negative (−0.118 and −0.109), suggesting that this large area maintains relatively uniform open-water-dominated conditions.

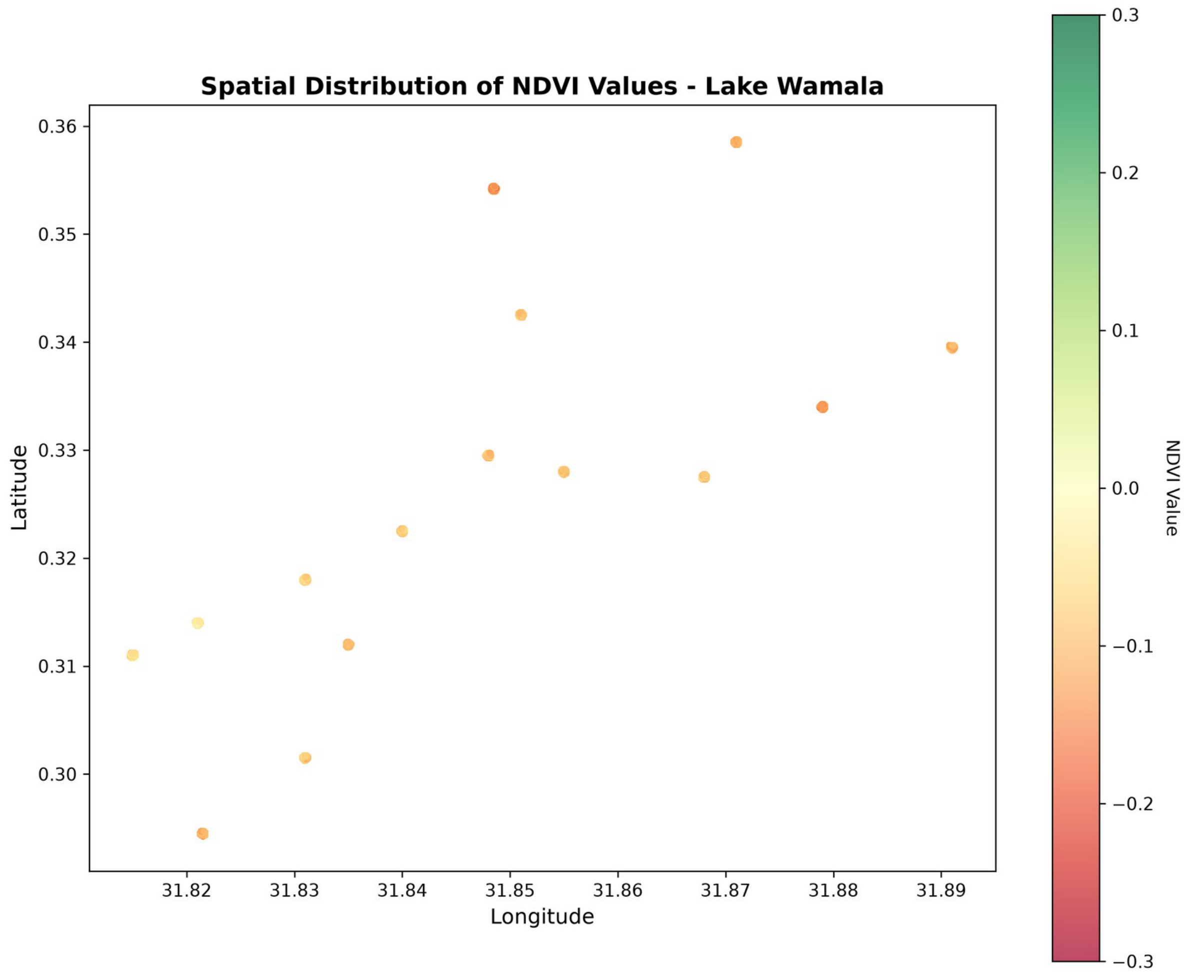


Figure 20: Spatial distribution of NDVI values

Figure 20 maps the spatial distribution of mean NDVI across the lake by longitude and latitude. The warmer-coloured points concentrated in the lower-latitude, southern portion of the lake correspond to the higher-NDVI Southern stations, while the cooler points reinforce the open-water dominance elsewhere, providing a geographic confirmation of the hotspot pattern.

4.10 NDWI Distribution by Zone Type

The NDWI analysis complements the NDVI findings from a water-specific perspective. The Very High Variation Sudd Zone recorded the highest mean NDWI (0.254), followed by the High Seasonal Infestation Variation (0.248) and Excessive Weed Infestation Hotspot (0.244) zones.

Despite their classification as vegetation-prone, these zones exhibit high NDWI because they contain large areas of open water interspersed with vegetation. The Open Water zone had the lowest mean NDWI (0.142), which initially appears counter-intuitive but reflects the presence of turbid water, algae and submerged vegetation that reduce the NDWI signal in some open-water pixels.

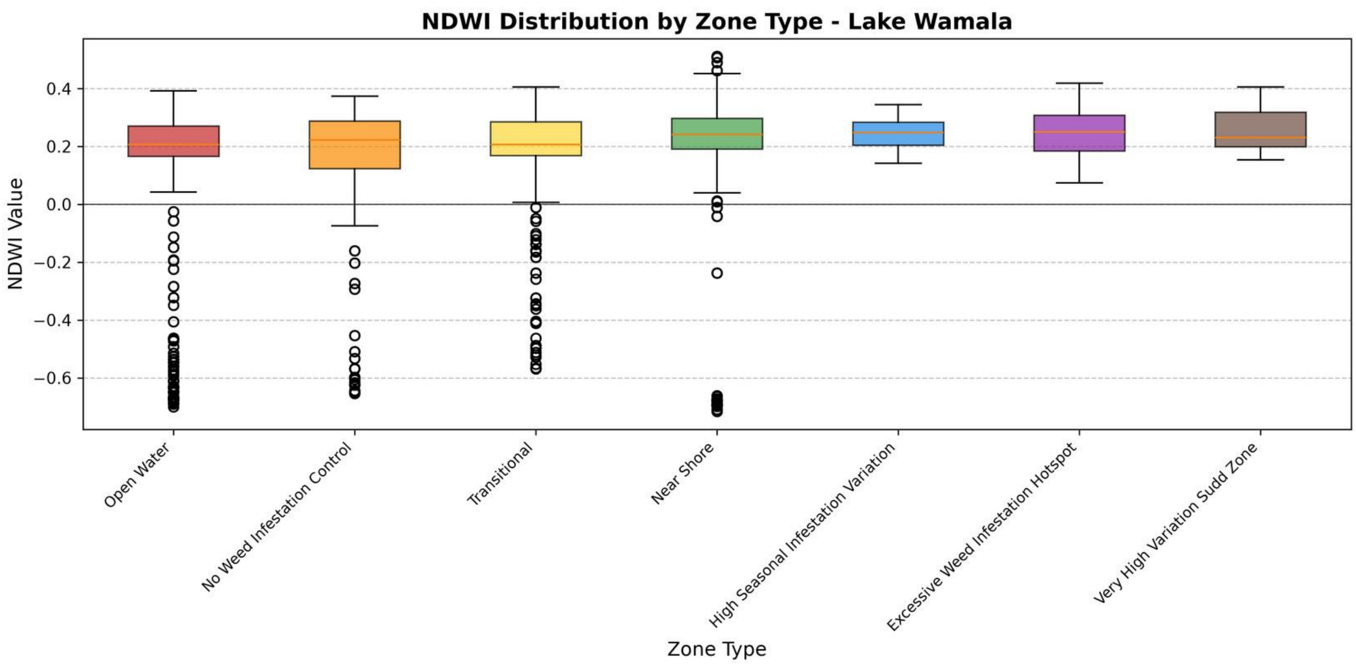


Figure 21: NDWI distribution by zone type (box plot)

The NDWI box plot in Figure 21 mirrors the NDVI box plot, with the Open Water and No Weed Infestation Control zones again showing the widest ranges, consistent with their high internal heterogeneity.

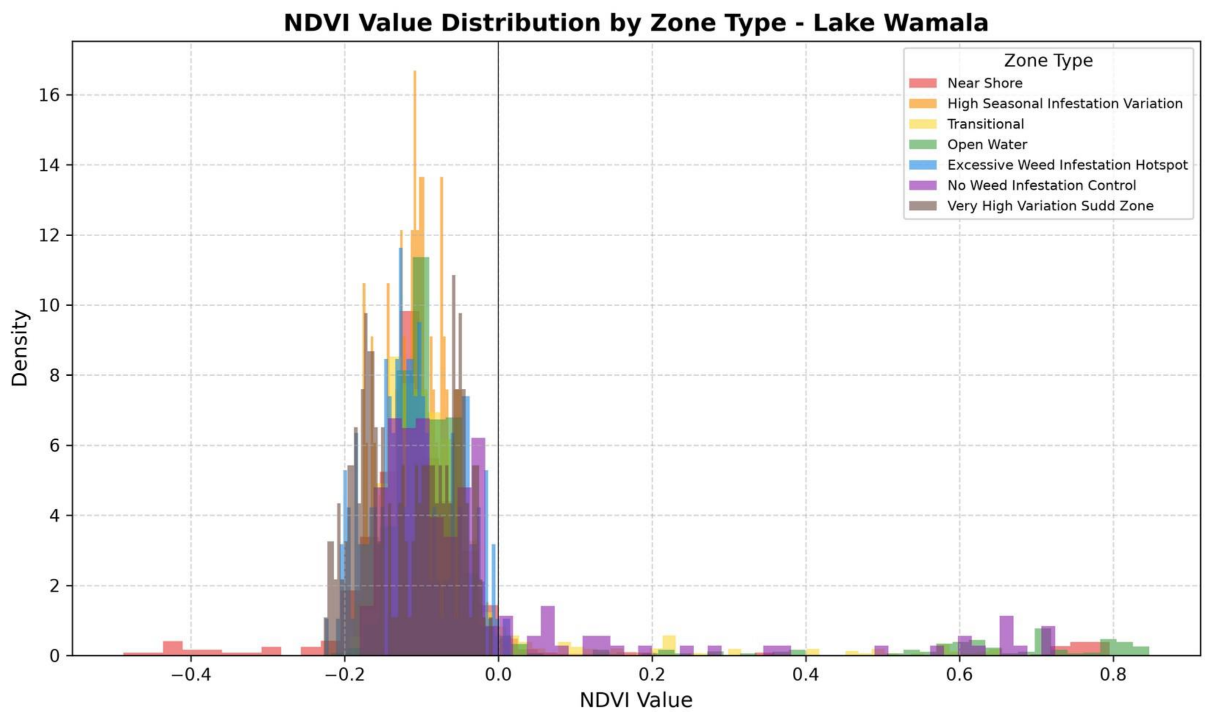


Figure 22: Overall NDVI value distribution (histogram)

Figure 22 shows the overall NDVI distribution, which is strongly concentrated around slightly negative values with a long right-hand tail extending towards 0.8. This shape confirms that the lake surface is overwhelmingly open water, while the sparse but real tail of high-NDVI pixels represents the localised dense-vegetation patches that constitute the weed hotspots.

4.11 Discussion of Key Findings

The results reveal a complex and dynamic picture of aquatic weed distribution on Lake Wamala. The first key finding is that the vast majority of the lake (95.2% of pixels) is open water or low vegetation, implying that excessive weed infestation is localised rather than lake-wide. Nonetheless, the 3.5% of pixels classified as high vegetation (1,060 pixels) and the 1.3% classified as moderate vegetation (405 pixels) remain significant when scaled to the lake's surface area of approximately 250 km²; if the sample is representative, these proportions translate to roughly 8.75 km² of dense vegetation and 3.25 km² of moderate vegetation, a substantial area capable of affecting navigation, fishing and water quality.

The second key finding is strong spatial heterogeneity. The Southern region consistently emerges as the most vegetation-rich area, with the highest mean NDVI, the highest proportion of high-vegetation pixels (12.02%) and the two most vegetation-dominated stations (S-03 and S-01). This pattern is likely driven by nutrient-rich runoff from agricultural areas (notably phosphorus and nitrogen), shallower bathymetry that favours submerged and emergent plants, and prevailing winds that concentrate floating vegetation in southern bays and inlets.

The third key finding is the marked temporal variability between 2017, 2022 and 2025. The sharp decline in NDVI from 2017 to 2022 across most zones points to a major ecological shift or management intervention during this period, consistent with the 2020–2023 La Niña-related high water levels that may have submerged vegetation and raised turbidity. The partial recovery by 2025 indicates that the system is resilient and can return towards earlier vegetation states, though recovery is incomplete and uneven: the Open Water zone recovered strongly (98% reduction in negativity) while the Very High Variation Sudd Zone continued to decline, suggesting that any management effort was more effective in open water than in complex vegetation zones.

The fourth key finding is the strong negative correlation between NDVI and NDWI (overall $r = -0.971$), which validates their combined use for weed mapping. The correlation is strongest where water–vegetation boundaries are clear and weakest in highly variable zones, implying that in heterogeneous environments a single index is insufficient and the two indices should be used together for accurate mapping.

Taken together, these four findings describe a lake on which weed infestation is spatially concentrated, seasonally driven and climate-modulated rather than uniformly distributed and steadily worsening. That combination is what makes targeted management feasible on Lake Wamala, and it is also what makes satellite monitoring necessary, since none of the four patterns is visible from a single field survey or a lake-wide average.

4.12 Implications for Water Weed Management

Several management implications follow. First, the concentration of vegetation hotspots in the Southern and Eastern regions indicates that interventions should be spatially targeted rather than uniform; priority areas include the Southern region (stations S-01, S-02 and S-03) and the Near Shore zone of the Eastern region (station E-03), while the Western region and most of the Dynamic Targeting area may require less intensive management. Second, the strong inter-annual variability calls for adaptive management that responds to climate conditions, emphasising prevention of spread during high-water periods and control of established beds during stable periods. Third, the strong NDVI–NDWI correlation supports remote sensing as a cost-effective monitoring tool, with free 10-metre Sentinel-2 data enabling timely, lake-wide assessment. Fourth, the discrepancy between management-designated hotspot zones and their low spectral vegetation proportions suggests that the current management zonation should be revisited and validated against field data and updated spectral classification.

CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary of the Study

This study assessed the spatial distribution, temporal trends and seasonal dynamics of aquatic weed infestation on Lake Wamala, Uganda, using remote sensing based on Sentinel-2 imagery. A dataset of 30,296 pixel-level NDVI and NDWI observations was analysed, drawn from 16 ground stations across five regions and seven ecological zone types, spanning three years (2017, 2022 and 2025) and five months (January, February, March, June and September). The specific objectives were to evaluate the effectiveness of NDVI and NDWI as complementary indices for weed mapping, map the spatial distribution and temporal trends of weed cover between 2017 and 2025, and identify seasonal patterns in the vegetation and water dynamics of the lake.

The analysis combined descriptive statistics, weed-cover classification using NDVI thresholds, correlation analysis and spatial–temporal trend analysis, supported by box plots, scatter plots, heatmaps, histograms and bar and line charts. The findings were interpreted in the context of local ecology, climate variability and previous research on aquatic vegetation in East African lakes.

5.2 Key Conclusions

The conclusions of this study are presented against the specific objectives set out in Section 1.3.2 and the hypotheses formulated in Section 3.9.2, taking each hypothesis and its related objective in turn.

1. **Objective 1 — effectiveness of NDVI and NDWI (Hypothesis H5).** NDVI and NDWI are confirmed as effective and genuinely complementary indices for assessing aquatic weeds on Lake Wamala. The null form of H5 is rejected at every level of analysis: the indices are inversely correlated at $r = -0.971$ overall ($r^2 = 0.94$) and between -0.743 and -0.989 within individual zones. Their complementarity is weakest where it matters least, in spectrally simple open water where either index alone would separate water from vegetation, and strongest in the mixed, seasonally variable zones where shared variance falls to 55 percent. The conclusion for practice is that both indices must be computed together, since a single-index classification would fail precisely in the zones of greatest

management interest. The objective is met, subject to the qualification that neither index detects submerged vegetation, so the method assesses surface-detectable weed cover only.

2. **Objective 2 — spatial distribution by ecological zone (Hypothesis H1).** Mean NDVI differs significantly among the seven ecological zone types (Welch's $F(6, 9859.3) = 328.29$, $p < .001$), so H1 is accepted. The effect size is small ($\eta^2 = 0.05$), and more importantly the direction of the differences does not match the management designations: the three zones labelled as infestation hotspots returned the most water-like signatures in the dataset, with the lowest standard deviations. The conclusion is that the existing zonation of Lake Wamala is not corroborated by the spectral evidence and requires revision against field data before it is used to direct control effort.
3. **Objective 2 — spatial distribution by region (Hypothesis H2).** Mean NDVI differs significantly among the five regions (Welch's $F(4, 10526.3) = 167.63$, $p < .001$, $\eta^2 = 0.04$), so H2 is accepted. The Southern region is the single unambiguous hotspot, holding the only positive regional mean NDVI (0.012), the highest proportion of dense-vegetation pixels (12.02%) and the two most vegetated stations, S-03 and S-01; the Eastern region follows at 4.93%. Since 95.2% of all pixels are open water or low vegetation, the central conclusion of the study is that infestation on Lake Wamala is spatially concentrated rather than lake-wide, and that management resources should be directed accordingly.
4. **Objective 2 — temporal trends (Hypothesis H3).** Mean NDVI differs significantly among the three observation years ($F(2, 30293) = 273.0$, $p < .001$, $\eta^2 = 0.018$), so H3 is accepted. The trajectory is a decline from −0.037 in 2017 to −0.113 in 2022, followed by partial recovery to −0.072 in 2025. The pattern coincides with the 2020–2023 La Niña-driven high-water period in the Lake Victoria basin and is best read as climate-driven submergence and turbidity rather than as evidence of successful weed control, since submergence and clearance produce the same spectral outcome. The small effect size indicates that inter-annual change is a minor component of the total variation in the index.
5. **Objective 3 — seasonal patterns (Hypothesis H4).** Mean NDVI differs significantly among the five sampled months ($F(4, 30291) = 692.2$, $p < .001$, $\eta^2 = 0.084$), so H4 is accepted. Season is the strongest of the four grouping factors tested, exceeding zone type, region and year. The signal peaks in January at the height of the dry season (−0.011) and

reaches its minimum in February at the onset of the first rains (−0.133), with NDWI mirroring the pattern at greater amplitude. The operational conclusion is that weed assessments must be compared between matching months, since within-year seasonal change exceeds the entire 2017-to-2022 inter-annual shift.

6. **Overall conclusion.** The five tests support a single conclusion: aquatic weed infestation on Lake Wamala is a spatially concentrated, strongly seasonal and climate-modulated phenomenon that freely available Sentinel-2 imagery can assess reliably and at negligible cost. Station-level analysis confirms that heterogeneity persists below the regional scale, since E-01 and E-03 differ markedly within the Eastern region, so monitoring should be conducted at station level and at monthly resolution. The principal limitation of the assessment is the absence of simultaneous ground-truth data, without which the NDVI thresholds cannot be calibrated locally and the discrepancy between the management zonation and the spectral evidence cannot be finally resolved.

5.3 Recommendations for Water Weed Management

1. **Spatially targeted management.** Interventions should be prioritised in the Southern and Eastern regions, where vegetation is most concentrated. The Southern region (stations S-01, S-02 and S-03) should receive the highest priority for mechanical harvesting, biological control or integrated approaches, while the Western and Dynamic Targeting regions may require only monitoring and prevention.
2. **Adaptive management framework.** Given the strong inter-annual variability, management should be adaptive rather than static, with annual or biennial review of priorities based on remote-sensing assessments. During high-water periods the focus should be on preventing the spread of floating weeds to newly inundated areas; during stable periods, on controlling established beds.
3. **Integrated monitoring system.** A monitoring system combining Sentinel-2 remote sensing with ground-based validation should be established, providing monthly or quarterly assessments using locally calibrated NDVI and NDWI thresholds and periodic validation at all 16 stations, especially during peak vegetation months (February–March) and the dry season (January).

4. **Nutrient management.** Because nutrient inputs from agricultural runoff appear to drive weed proliferation in the Southern and Eastern regions, measures should include riparian buffer zones along inflowing rivers, constructed wetlands or sediment traps at river mouths, and engagement with farming communities to reduce fertiliser runoff. Intensive agricultural areas near the lake should be priority targets.
5. **Community engagement and education.** Local fishing communities, who are most affected by weed infestation, should be engaged in management. Community-based monitoring can supplement satellite data with local knowledge and early warning, and educational programmes should raise awareness of the link between nutrient pollution and weed growth.
6. **Biological control enhancement.** Where mechanical harvesting is impractical, biological control using weevils (*Neochetina eichhorniae* and *Neochetina bruchi*) should be explored as a sustainable alternative, with the concentrated southern distribution of weeds making this area particularly suitable; any introduction should be carefully evaluated to avoid unintended ecological consequences.
7. **Update management zone classification.** The current zonation should be revised using the spectral evidence from this study, reclassifying zones that are spectrally open water but labelled as hotspots and incorporating unexpectedly vegetated areas, ideally through a participatory mapping exercise involving local stakeholders and remote-sensing experts.

5.4 Recommendations for Future Research

1. **Higher temporal resolution.** Future work should aim for monthly or bi-monthly coverage over a 10–15 year period, combining the Landsat archive (from 1984) with ongoing Sentinel-2 acquisitions to capture inter-annual variability and long-term trends more fully.
2. **Machine-learning classification.** Techniques such as Random Forest, Support Vector Machines or Convolutional Neural Networks, incorporating multiple bands, texture and temporal features, should be explored to improve accuracy and resolve discrepancies between management zones and spectral classifications.
3. **Ground-truthing campaign.** Systematic field collection of vegetation cover, species composition and biomass with GPS coordinates at all 16 stations across different months

would provide the empirical basis for calibrating and validating the indices; hydroacoustic methods or underwater cameras could assess submerged vegetation not captured by surface NDVI.

4. **Climate and hydrological modelling.** Integrating remote-sensing data with rainfall records, lake-level data and hydrological models would quantify the links between climate, hydrology and vegetation, enabling predictive modelling of weed infestation under different climate scenarios.
5. **Species-specific mapping.** Future research should distinguish weed species (water hyacinth, water lettuce, *Salvinia molesta* and others) using hyperspectral data or high-resolution imagery, since different species have different impacts and management requirements.
6. **Socio-economic impact assessment.** The ecological findings should be linked to socio-economic impacts by quantifying the costs of infestation in terms of reduced fishing yields, navigation, water treatment and tourism, providing a stronger basis for justifying management investment.
7. **Comparative studies with other lakes.** Comparing Lake Wamala with other affected lakes such as Lake Kyoga, Lake George and Lake Victoria would identify common drivers and shared management lessons, supported by a regional database of remote-sensing-derived vegetation indices.

5.5 Policy Implications

The findings have direct policy relevance. The National Environment Management Authority and the Ministry of Water and Environment should consider incorporating remote-sensing-based vegetation monitoring into their lake-management frameworks, given the cost-effectiveness of Sentinel-2 data even for resource-constrained agencies. Because infestation is concentrated in specific regions, national lake-management budgets could be allocated more efficiently by targeting high-priority areas rather than applying uniform interventions.

The evidence for nutrient-driven proliferation supports including agricultural non-point-source pollution control in lake-management policy. Existing instruments such as the National Environment Act, the Water Act and the National Agricultural Policy should be strengthened to provide for lake buffer zones, nutrient-management plans and incentives for sustainable catchment agriculture, and the Lake Wamala Catchment Management Plan should be updated to reflect the spatial patterns identified here. At the transboundary level, weed-management efforts should be coordinated within the Lake Victoria basin through the Lake Victoria Fisheries Organisation and the Lake Victoria Basin Commission, facilitating shared monitoring, common databases and joint strategies.

5.6 Concluding Remarks

This study has demonstrated the utility of NDVI and NDWI, derived from freely available Sentinel-2 imagery, for mapping and monitoring aquatic weed distribution on Lake Wamala. The analysis of 30,296 pixel observations from 16 stations across three years has produced a comprehensive picture of spatial patterns, temporal trends and seasonal dynamics. The findings challenge the assumption that weed infestation is a uniform, lake-wide problem; instead, they reveal a highly heterogeneous distribution with clear hotspots in the Southern and Eastern regions and extensive open water in the Western and Dynamic Targeting regions.

The marked temporal shifts between 2017, 2022 and 2025 underline the dynamic nature of the lake and the importance of continuous monitoring, while the partial recovery by 2025 offers encouragement that well-targeted, adaptive management can help maintain the lake in a productive and navigable state. The strong NDVI–NDWI correlation confirms the value of these indices as practical tools for managers and researchers. Ultimately, the sustainable management of Lake Wamala requires an integrated approach combining remote-sensing monitoring, ground-based validation, nutrient management, community engagement and adaptive strategies. By identifying where to act, when to act and which tools to use, this study provides a scientific foundation for protecting this vital resource amid increasing pressures from climate change, population growth and agricultural intensification.

REFERENCES

1. Ademola, B. T., Mahuta, S. I., Abdulkarim, B., & Argungu, L. A. (2022). Macrophytes abundance in relation to eutrophication status of peri-urban impoundments in Katsina metropolis, Nigeria. *Bayero Journal of Pure and Applied Sciences*, *13*(1), 173–179. https://doi.org/10.4314/bajopas.v13i1.28S

2. Al-lami, A. K., Abbood, R. A., Al Maliki, A. A., & Al-Ansari, N. (2021). Using vegetation indices for monitoring the spread of Nile Rose plant in the Tigris River within Wasit province, Iraq. *Remote Sensing Applications: Society and Environment*, *22*. https://doi.org/10.1016/j.rsase.2021.100471

3. Albright, T. P., Moorhouse, T. G., & McNabb, T. J. (2004). The rise and fall of water hyacinth in Lake Victoria and the Kagera River basin, 1989–2001. *Journal of Aquatic Plant Management*, *42*(JUL.), 73–84.

4. Alemneh, L. Y., Aklog, D., Griensven, A. van, Goshu, G., Yalew, S., Abebe, W. B., Dersseh, M. G., Mhiret, D. A., Michailovsky, C. I., Amare, S., & Asress, S. (2025). Remote Sensing Approaches for Water Hyacinth and Water Quality Monitoring: Global Trends, Techniques, and Applications. *Water (Switzerland)*, *17*(17). https://doi.org/10.3390/w17172573

5. Balirwa, J. S., & Wanda. (n.d.). *Water Quality and Quantity Synthesis Final Report, LVEMP December 2005 Water Quality and Ecosystems Management Component Impacts of water quality change on beneficial uses of Lake Victoria, Uganda*.

6. Belayhun, M., Chere, Z., Abay, N. G., Nicola, Y., & Asmamaw, A. (2024). Spatiotemporal pattern of water hyacinth (*Pontederia crassipes*) distribution in Lake Tana, Ethiopia, using a random forest machine learning model. *Frontiers in Environmental Science*, *12*. https://doi.org/10.3389/fenvs.2024.1476014

7. Datta, A., Maharaj, S., Prabhu, G. N., Bhowmik, D., Marino, A., Akbari, V., Rupavatharam, S., Sujeetha, J. A. R. P., Anantrao, G. G., Poduvattil, V. K., Kumar, S., & Kleczkowski, A. (2021). Monitoring the Spread of Water Hyacinth (*Pontederia crassipes*): Challenges and Future Developments. In *Frontiers in Ecology and Evolution*

(Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/fevo.2021.631338

8. Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, *120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

9. Duarte, L., Ant, A. M., & Sousa, J. J. (2022). *Spatio-Temporal Water Hyacinth Monitoring in the Lower Mondego (Portugal) Using Remote Sensing Data*.

10. Gao, B. C. (1996). NDWI — A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

11. Gerardo, R., & De Lima, I. P. (n.d.). *Assessing the potential of Sentinel-2 data for tracking invasive water hyacinth in a river branch*. https://doi.org/10.1117/1.JRS.16

12. Gichuki, J., Omondi, R., Boera, P., Okorut, T., Matano, A. S., Jembe, T., & Ofulla, A. (2012). Water hyacinth *Eichhornia crassipes* (Mart.) Solms-Laubach dynamics and succession in the Nyanza Gulf of Lake Victoria (East Africa): Implications for water quality and biodiversity conservation. *The Scientific World Journal*, *2012*. https://doi.org/10.1100/2012/106429

13. Havel, J. E., Kovalenko, K. E., Thomaz, S. M., Amalfitano, S., & Kats, L. B. (2015). Aquatic invasive species: challenges for the future. In *Hydrobiologia* (Vol. 750, Issue 1, pp. 147–170). Kluwer Academic Publishers. https://doi.org/10.1007/s10750-014-2166-0

14. Hu, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, *113*(10), 2118–2129. https://doi.org/10.1016/j.rse.2009.05.012

15. Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

16. Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, *25*(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

17. King, S., Okello, P. O., & Ogwal, F. (n.d.). *Fisheries Resources Accounts for Uganda*. https://www.researchgate.net/publication/359768706

18. Malthus, T. J., & George, D. G. (1997). Airborne remote sensing of macrophytes in Cefni Reservoir, Anglesey, UK. *Aquatic Botany*, *58*(3–4), 317–332. https://doi.org/10.1016/S0304-3770(97)00043-0

19. McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

20. Mihi, A., Ghazela, R., & Wissal, D. (2022). Mapping potential desertification-prone areas in North-Eastern Algeria using logistic regression model, GIS, and remote sensing techniques. *Environmental Earth Sciences*, *81*(15). https://doi.org/10.1007/s12665-022-10513-7

21. Nyawacha, S. O., Meta, V., & Osio, A. (2021). Spatial temporal mapping of spread of water hyacinth in Winum Gulf, Lake Victoria. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLIII-B3-2*, 341–346. https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-341-2021

22. Okaronon, J. (2009). A decline in fish catches and the size of *Oreochromis niloticus* in Lake Wamala (Uganda) following human exploitation. *African Journal of Tropical Hydrobiology and Fisheries*, *6*(1). https://doi.org/10.4314/ajthf.v6i1.45933

23. Oyama, Y., Matsushita, B., & Fukushima, T. (2015). Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM+ shortwave infrared bands. *Remote Sensing of Environment*, *157*, 35–47. https://doi.org/10.1016/j.rse.2014.04.031

24. Shekede, M. D., Gondo, T., Mavhenge, M. M., & Mazhindu, A. N. (2023). Using Landsat satellite imagery to monitor the spatial and temporal dynamics of aquatic weed extent in Lakes Chivero and Manyame, located in an urban catchment of Zimbabwe.

Water SA, *49*(1 January). https://doi.org/10.17159/wsa/2023.v49.i1.3950

25. Silva, T. S. F., Costa, M. P. F., Melack, J. M., & Novo, E. M. L. M. (2008). Remote sensing of aquatic vegetation: Theory and applications. *Environmental Monitoring and Assessment*, *140*(1–3), 131–145. https://doi.org/10.1007/s10661-007-9855-3
26. Singh, G., Reynolds, C., Byrne, M., & Rosman, B. (2020). A remote sensing method to monitor water, aquatic vegetation, and invasive water hyacinth at national extents. *Remote Sensing*, *12*(24), 1–24. https://doi.org/10.3390/rs12244021
27. Strashok, O., Ziemiańska, M., & Strashok, V. (2022). Evaluation and Correlation of Sentinel-2 NDVI and NDMI in Kyiv (2017-2021). *Journal of Ecological Engineering*, *23*(9), 212–218. https://doi.org/10.12911/22998993/151884
28. Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0
29. Villamagna, A. M., & Murphy, B. R. (2010). Ecological and socio-economic impacts of invasive water hyacinth (*Eichhornia crassipes*): A review. *Freshwater Biology*, *55*(2), 282–298. https://doi.org/10.1111/j.1365-2427.2009.02294.x
30. Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179

APPENDICES

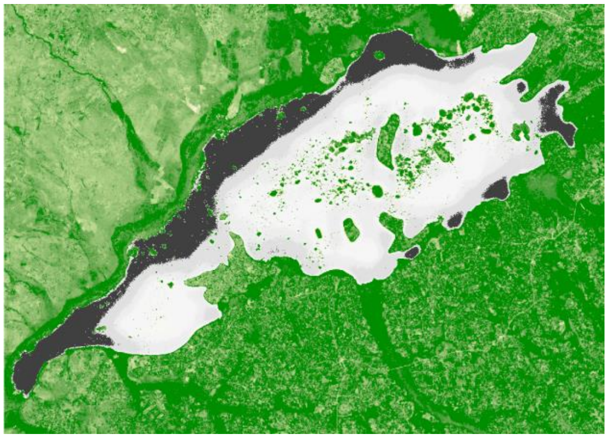


Figure 23:Lake wamala NDVI Tiff image from CDSE

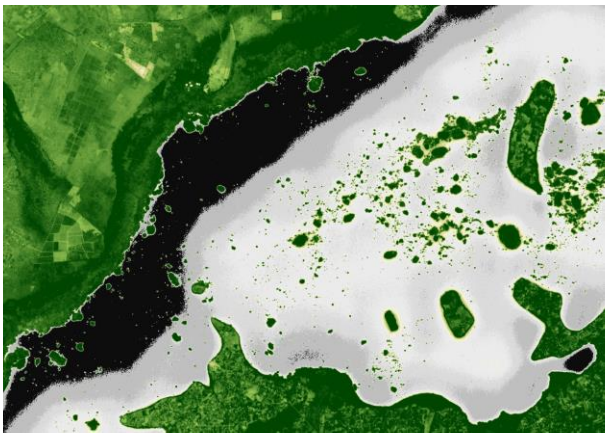


Figure 24:32bit Magnified NDVI lake wamala basin image from CDSE

Figure 25:CDSE False colorvisualised images for lake wamala

Figure 26:NE open water NDWI buffer Image

Figure 27:NWnear shore NDVI buffer Image

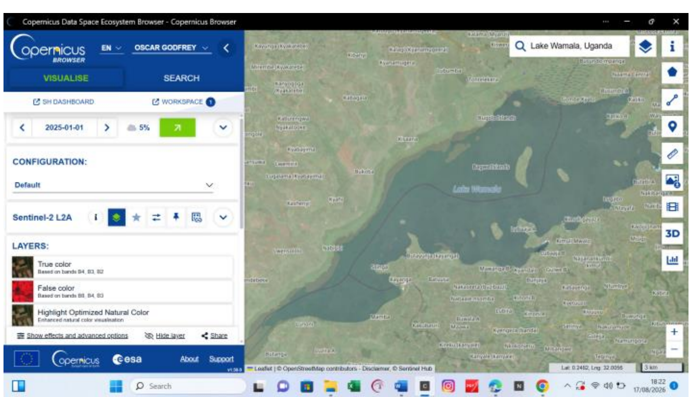


Figure 28:CDSE application interface

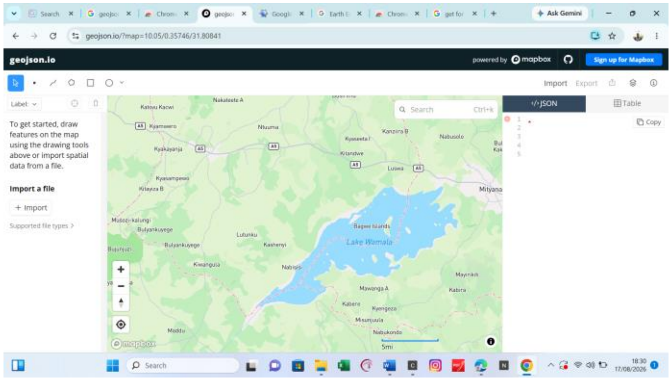


Figure 29:Geojson .oi interface, where different files(kml, shape, etc) have been freely imported from

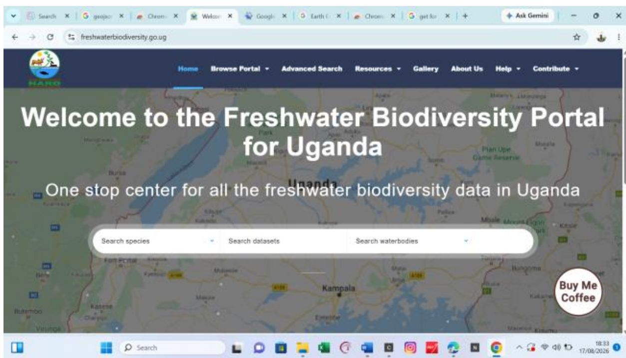


Figure 30:FWBDP where most facts of lake wamala came from