University of Zurich[UZH]

Effects of Livelihood Choices on Forest Dynamics in Protected Areas in Mozambique

ESS 511 Master's Thesis

Author
Lino Jastrowski
16-917-957

Supervised by
Prof. Dr. Maria J. Santos

Faculty representative
Prof. Dr. Maria J. Santos

31.01.2025
Department of Geography, University of Zurich

University of Zurich
Departement of Geography

Effects of Livelihood Choices on Forest Dynamics in Protected Areas in Mozambique

A Case Study in Niassa Special Reserve, Mozambique

ESS511 Master's Thesis

Lino Jastrowski
16-917-957

Supervised by
Maria J. Santos, Prof. Dr.

Faculty Representative
Maria J. Santos, Prof. Dr.

31.01.2025

Contact

Author: Lino Jastrowski
Student ID: 16-917-957
Institution: University of Zurich
Email: lino.jastrowski@uzh.ch

Supervision: Maria J. Santos, Prof. Dr.
Institution: University of Zurich
Email: maria.j.santos@geo.uzh.ch

Abstract

Forests provide essential ecosystem services, including carbon sequestration, biodiversity conservation, and hydrological regulation. However, they are increasingly threatened by deforestation, land-use change, and socio-economic pressures. This is especially in conservation areas where ecological integrity must be balanced with human livelihoods the case. Aim of this study is to analyze how village infrastructure, household wealth, water security, and agricultural diversification influence vegetation health indicators such as NDVI and NDWI. This was achieved by combining remote sensing indices, statistical modeling, and Structural Equation Modeling (SEM). The regression and correlation analyses reveal weak to moderate relationships between socio-economic drivers and forest functioning, with Agricultural Livelihood Diversification (ALD) positively influencing NDVI, while Village Infrastructure shows a weak negative association with forest health. The SEM results further emphasize the role of Water Security as a mediating factor, suggesting that infrastructure development indirectly supports ecological resilience when linked with resource access. These findings highlight the complex trade-offs between socio-economic development and conservation. While livelihood diversification appears to support forest sustainability, increased infrastructure and wealth do not necessarily translate into improved ecological outcomes. This research emphasizes the importance of integrated land management policies that consider both direct and indirect socio-economic impacts on forests. The study provides insights for conservation planning in sub-Saharan Africa and beyond, advocating for strategies that align local livelihoods with ecological sustainability.

Acknowledgments

Thanks to Maria J. Santos for guiding me through this process and always be there when I did not know how to continue. Also, I want to thank Maria's group as a whole for helping me in the beginning setting up my thesis and sharing their expertise when I needed it. A big thank you also goes out to Joana G. Pereira, who let me use her survey data, and also answered my question even though just having finished her own PhD (Congratulations!).
Additionally, I want to thank the Copernicus Programme for making Sentinel-2 imagery available freely. OpenAI and their chatbot ChatGPT-4o were also a great help to find adequate formulations and speeding up the writing process.
Finally, I also want to thank my friends and family for supporting me and making this possible.

Declaration of Independent Work

I, Lino Jastrowski, declare that this thesis was conducted independently and only the cited sources were used.

Contents

List of Figures

List of Tables

1 Introduction

Forests are key components of the Earth's biosphere, providing essential ecosystem services for both natural and human systems. The services they provide include regulating climate through carbon sequestration, maintaining biodiversity by providing habitat for millions of species, and contributing to hydrological and nutrient cycles (Trumbore et al., 2015; Chazdon and Guariguata, 2018; Bonan, 2008). Nevertheless, forests globally continue to be degraded due to socio-economic pressures, land-use changes, and resource exploitation, particularly in conserved areas where the integrity provided by ecological services or functions is needs to be balanced against the need for human livelihoods (FAO, 2020; Hansen et al., 2013). Given the significant human impact on forests, the interaction of socio-economic factors and forest functioning are fundamental to the sustainable management of forests that prevents environmental degradation while simultaneously supporting local communities (Fisher and Shively, 2005). In protected areas that overlap human settlements, such as Mozambique's Niassa Special Reserve, conservation must balance complex trade-offs between ecological sustainability and socio-economic development. Using an integrative approach (i.e., combining remote sensing, statistical modeling, and Structure Equation Modeling-SEM), this study analyzed these interactions, focusing on the extent to which forest dynamics is influenced by livelihood strategies, infrastructure, and socio-economic conditions.

1.1 Forests and their Importance

Forests represent an essential driver of carbon storage, preservation of biodiversity, climate regulation and the maintenance of water resources around the world (Worm et al., 2006; Trumbore et al., 2015; Lewis et al., 2015). Forests contribute around 75% of terrestrial gross primary production, and with an estimated 861 gigatonnes of carbon stored, and absorbing 4.9 gigatonnes of CO_2e per year, they remain a major component of climate change mitigation strategies (Ruiz, 2024; Harris et al., 2021). Yet, deforestation and degradation re-emit much of this stored carbon, and therefore play a major role in global greenhouse gas emissions (Harris et al., 2021). The rates of deforestation are increasing and are mainly linked to agricultural expansion, logging and infrastructure development, threatening carbon storage as well as hydrological cycles and biodiversity (Hansen et al., 2013; FAO, 2020). This alarming trend emphasizes the necessity of sustainable forest management approaches to mitigate environmental degradation and secure beneficial and vital ecosystem services (Chazdon and Guariguata, 2018).

Forests offer the richest terrestrial biodiversity, with approximately 60'000 tree species, 80%, 75% and 68% of amphibians, birds and mammals, respectively(FAO, 2020). They support critical ecosystem services like soil stabilization, pollination, and nutrient cycling (Chazdon and Guariguata, 2018; Foley et al., 2005). Forests shape hydrological cycles, and thus water availability and quality, which are essential for both human and ecological systems (Brauman, 2015). In places like Mozambique's Niassa Special Reserve, the biotic pump mechanism, where atmospheric moisture is generated in forests and then transported around the planet via evapotranspiration, is especially crucial (Sheil, 2018; Makarieva et al., 2007). Regular precipitation maintains ecosystems and livelihoods dependent on forest resources for agriculture, water and biodiversity (Sheil and Murdiyarso, 2009). These services highlight the essential role of forests in supplying functions that sustain ecological and social systems at both the local and global scale.

Forests are extremely important but are under pressure from anthropogenic activities and natural disturbances. Approximately 420 million hectares were lost globally between 1990 to 2020, mainly as results of deforestation and land-use changes (FAO, 2020; Hansen et al., 2013). Between 2000 and 2020 alone, an estimated additional 231 million hectares of trees were cut down, with commodity-driven deforestation and urban expansion responsible for 22% of this deforestation (Global Forest Watch, 2024; Turubanova et al., 2018). Fires were also involved in this loss and accounted for 28% of tree cover loss between 2001 and 2023 (Global Forest Watch, 2024). These losses are especially pronounced in tropical regions, such as Mozambique, where the key drivers of deforestation and forest degradation are agricultural expansion, logging and infrastructure development (Cumming and Allen, 2017).

1.1.1 Forests in Africa

Africa represents some of the most ecologically significant forest ecosystems, from the humid tropical rain forests of the Congo Basin to the dry woodlands of the Miombo and mangrove forests in the coastal regions (FAO, 2020; Mayaux et al., 2004; Lewis et al., 2015). Forests represent around 624 million hectares and approximately 21% of Africa's lands, being important providers of biodiversity maintenance, climate regulatory systems, and livelihoods for millions of people (Bastin et al., 2019). Home to almost 10'000 plant species and the second-largest rain forest in the world, the Congo Basin acts as a carbon sink, absorbing an estimated 600 million tonnes of CO_2 per year (Harris et al., 2021). On the other hand, savanna woodlands including the Miombo harbor rich and biodiverse ecosystems and offer critical resources for subsistence agriculture, energy generation and grazing (Campbell et al., 2007).

Nonetheless, the forests of Africa are increasingly threatened by deforestation and degradation, mainly due to agriculture, logging, expansion of infrastructure and population growth (Tyukavina et al., 2018; FAO, 2020). According to estimates, the continent lost 74.3 million hectares of forest between 1990 and 2020, with an annual deforestation rate of 3.9 million hectares, one of the highest in the world (FAO, 2020; Global Forest Watch, 2024; Turubanova et al., 2018). The Democratic Republic of the Congo (DRC) alone contributed to more than 40% of the continent-wide total forest loss in recent decades driven by the combined processes of shifting cultivation, commercial logging and land tenure insecurity (Turubanova et al., 2018). In Mozambique, forests are being rapidly degraded by the production of charcoal, slash-and-burn agriculture, and infrastructure projects , contributing to land degradation, and biodiversity loss (Cumming and Allen, 2017).

However, African forests are indispensable for climate resilience, water regulation, and food security despite these threats. Non-Timber Forest Products (NTFPs) like wild fruits, honey, and medicinal plants are an important source of food and income for many rural communities (Shackleton et al., 2007; Kassa et al., 2017). Populations in the Sahel region that depend on forests use agroforestry to reverse desertification, and mangroves in West and East Africa act as natural barriers against coastal erosion and climate-induced sea-level rise (Herr and Pidgeon, 2015).

Understanding the importance of conservation and the need to act fast, both governments and international organizations have undertaken large-scale schemes to protect and restore forests. The AF110 (African Forest Landscape Restoration Initiative) seeks to restore 100 Million hectares by 2030, which are valued in terms of increased carbon sequestration and providing ecosystem services (AFR100, 2020). The Great Green Wall Project, which covers 11countries in the Sahel, aims to prevent desertification through restoring native vegetation and practices of sustainable land-use (Gnacadja and Wiese, 2013). Community-Based Natural Resource Management (CBNRM) programs have been implemented in countries such as Namibia, Kenya, and Tanzania, which link conservation actions to local livelihoods in an effort to curb deforestation and enhance forest governance (Knight et al., 2019; Cumming and Allen, 2017).

A greater understanding of these complex dynamics is critical for devising strategies that protect Africa's forests without compromising long-term benefits for people and the environment. This study explores the interplay between socio-economic and ecological variables in the Niassa Special Reserve (NSR) to inform sustainable conservation planning.

1.2 Livelihoods and Conservation

Livelihoods include all of the activities that people take up in order to earn a living, utilizing their skills, assets and opportunities to satisfy their needs and enhance their quality of life (Kapur, 2019; Tambe, 2022a; Acharya, 2006). However, where conservation targets and human communities share the same landscapes, livelihoods go beyond purely economic and physical dimensions to become embedded in ecological and governance processes (Rai et al., 2021; Cumming and Allen, 2017; Adams, 2004). In continents like Africa, for example in the Niassa Special Reserve in Mozambique the socio-ecological landscape is characterized by competing conservation goals, poverty, and resource scarcity, which makes it imperative to find integrated solutions that address the needs and lifestyles of people (Adams, 2004; Cumming and Allen, 2017). This understanding is essential to not only addressing poverty and food security but also how conservation can equitably meet the needs of people and biodiversity (Bernstein et al., 1992; Kalyanram et al., 2014).

Agricultural and natural resource use are a cornerstone of daily life for local communities, e.g. in Niassa (Acharya, 2006; Tambe, 2022b; Shackleton et al., 2007). People depend on forests as a source of timber, fuelwood, medicinal plants and other non-timber forest products, while practicing subsistence agriculture and livestock husbandry (Bernstein et al., 1992; Angelsen et al., 2014; B. Fisher and Burgess, 2011). These livelihood strategies are not fixed, and adapt to risks including climate variability, changing precipitation patterns and wildlife conflicts (IPCC, 2019; Pereira et al., 2021; Shackleton et al., 2015). Households diversify their activities—through petty commodity production and migration—to buffer shocks and respond to changing contexts of (Scoones, 1998; Kapur, 2019; Ellis, 2000). However, if access to important resources is limited with the establishment of protected areas, traditional practices may be disrupted, creating conflict between conservation goals and community needs (Rai et al., 2021; Cumming and Allen, 2017; Nelson et al., 2016). This quagmire raises questions about who conservation benefits, what objectives protected areas crystallize and by what means to sustain livelihoods under such conditions.

The establishment of protected areas frequently restricts for land-use practices, with repercussions for social and economic processes in impacted communities (West et al., 2006; Adams, 2004; Cumming and Allen, 2017). Conservation initiatives aim to protect biodiversity and ecosystems but face the difficulties of local poverty, resource dependencies and governance forms (Adams, 2004; West et al., 2006). In Niassa, resource scarcity as well as formal credit, education, healthcare and infrastructure constraints further complicate these interactions (Ellis, 1999; Baumgartner and Högger, 2004).

Helping rural households thrive in a world of uncertainty is where the Sustainable Livelihoods Framework (SLF) comes in handy for diagnosing and addressing these challenges. Originally developed for analysis of rural livelihoods, the SLF has been adapted to incorporate the wider socio-ecological systems in which people are embedded (Chambers et al., 1989; Sallu et al., 2010; Scoones, 1998). It distinguishes five capitals (natural, physical, human, social, and financial) that households draw on to achieve desired outcomes and emphasizes resilience as the ability to recover from shocks and sustain livelihoods (Moench, 2009; Proag, 2014; Adger, 2006). For Niassa it is useful to apply the SLF to understand the interplay of conservation policies, access to resources for adaptation and local adaptation strategies. For instance, natural capital (e.g., forests) serves as a source of livelihood and a target of conservation, which introduces tensions over resource use (Shackleton et al., 2015). Social and human capital, such as community networks and local knowledge, can likewise strengthen resilience, but potentially be eroded by exclusionary conservation practices (Rai et al., 2021; Knight et al., 2019; Nelson et al., 2016).

Understanding how conservation and broader environmental changes impact livelihoods primarily centers around the concepts of vulnerability and resilience. Vulnerability emphasizes the vulnerability of communities to negative pressures, such as resource demands or climate shocks, while resilience emphasizes the ability of communities to rebuild and adapt (Twigg et al., 2001; Adger, 2006). In Niassa Special Reserve, communities are confronted with intersecting vulnerabilities such as wildlife conflicts, low economic diversification, and the effects of climate variability (Rai et al., 2021; Shackleton et al., 2015). Yet, building resilience through Participatory Governance Models (PGMs) e.g., Community-Based Natural Resource Management (CBNRM) that allow communities to more actively engage in conservation, while providing for their livelihoods can enhance resilience (Knight et al., 2019; Palomo et al., 2014).

By utilizing these concepts, this research seeks to fill gaps in understanding how conservation interacts with local livelihoods in socio-Ecological systems such as the Niassa Special Reserve. Although much has been published on the SLF and resilience, there is less understanding of how the SLF and resilience frameworks function where there is widespread poverty, resource deprivation and competing land-use agendas (Scoones, 1998). This paper looks at how protected areas can reconcile both biodiversity conservation and community needs and potentially provide new insights on how integrated management approaches can simultaneously pursue both ecological and social goals (Cumming and Allen, 2017; Shackleton et al., 2007).

1.2.1 Livelihoods in Africa: Challenges and Opportunities

Africa's rural and urban populations depend heavily on forest, grassland and water bodies for their livelihoods (FAO, 2020; Shackleton et al., 2007; Angelsen et al., 2014). The majority (more than 60%) of Africa's human population depend on agriculture, forestry and fisheries for their sustenance and livelihoods, thus, natural ecosystems are critical to achieving food security and economic stability (Kassa et al., 2017). Forests, specifically, are crucial for rural livelihoods, supplying goods like timber, fuelwood, medicinal plants, and non-timber forest products (NTFPs), including honey, wild fruits and nuts (Chidumayo and Gumbo, 2010; Angelsen et al., 2014).

In the African context, on the other hand, challenges are overlapping, which shapes the dynamics of livelihoods. Resource scarcity is made worse by high levels of poverty, high rates of population growth, and land-use pressures, while climate change adds further vulnerabilities (IPCC, 2019). Communities in areas such as the Sahel and East Africa are adapting in response to changes like prolonged droughts, desertification, and changes in rainfall patterns through strategies like agroforestry, migration, and diversification of livelihoods (Herr and Pidgeon, 2015; Kassa et al., 2017).

Novel approaches like Community-Based Natural Resource Management (CBNRM) have been shown to have the potential for tackling these issues by giving local populations a participatory role in the management of resources (Knight et al., 2019; Cumming and Allen, 2017; Shackleton et al., 2015). Nonetheless, the success of these initiatives tends to depend on several factors, including secure land tenure, mechanisms for equitable benefit-sharing, and inclusion of indigenous knowledge in management practices (Shackleton et al., 2007).

Nevertheless, much remains to be known about the impact of the complex interactions between socio-economics and institutions, on ecological processes and multiple livelihoods across the continent (Adger, 2006; Twigg et al., 2001). This challenge is approached through an integrated approach (ecological and socio-economic), preserving Africa's livelihoods in a sustainable manner as much as possible (Scoones, 1998; Shackleton et al., 2015).

1.3 Framework for Study

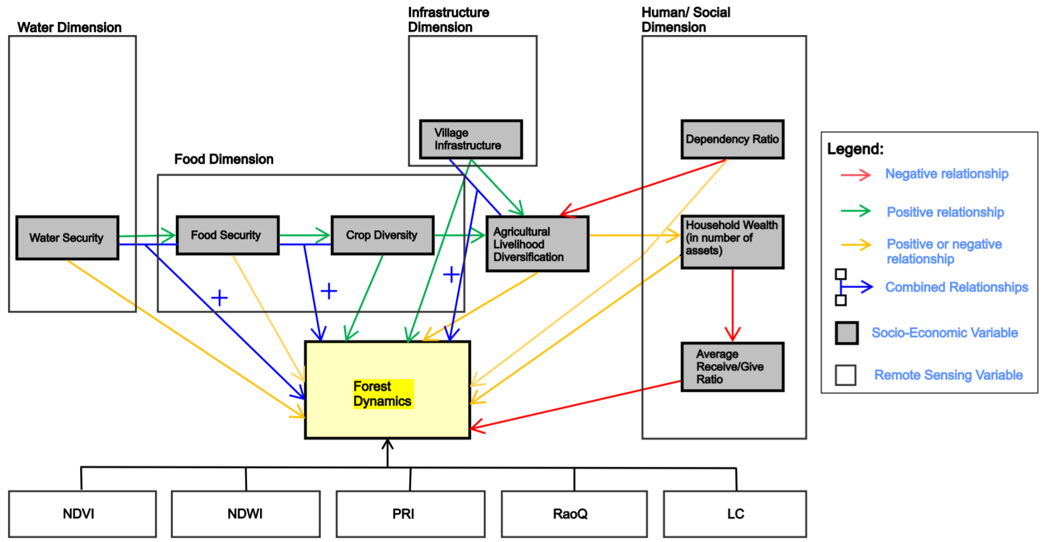


Figure 1: The diagram illustrates the relationships between socio-economic and ecological variables in the Niassa Special Reserve. Socio-economic variables (gray background) were based on survey data by Pereira et al. (2021) and ecological variables (white background) were obtained by remote sensing (European Space Agency (ESA), 2024). This framework demonstrates the interrelationships between systems of food, water, infrastructure and society and emphasizes the need for integrating diverse data sources to better understand forest trajectories and their drivers

Forest dynamics monitoring is key to understanding interactions between ecological processes and socio-economic systems. Forest health and productivity, major attributes of forest dynamics, are measures of a forest's capacity to preserve biodiversity, withstand disturbances, and sequester carbon (Lausch et al., 2016; Trumbore et al., 2015). Forest health encompasses the status of canopies, leaf pigments and water availability, whereas productivity reflects rates of photosynthesis, biomass accumulation and growth (Running et al., 2004). These facets are closely linked to socio-economic contextual dimensions induced by agricultural livelihoods, water security and household wealth, particularly in the context of Mozambique's Niassa Special Reserve.

This work employs a conceptual framework (Figure 1) that describes the crucial variables affecting forest dynamics and their respective relationships, with which to analyze these relationships. This framework incorporates ecological indicators obtained by remote sensing, e.g. NDVI, PRI, NDWI with socio-economic parameters, e.g. agricultural diversity, dependency ratios, wealth of households. These variables span food, water, infrastructure, and social dimensions, reflecting their wide-ranging impacts on forests and communities. This conceptual framework emphasizes data needs for characterizing forest dynamics and makes a distinction between variables available from remote sensing and those that require ground-based socio-economic surveys.

1.4 Remote Sensing

Remote sensing has been widely used in multiple global programs to track forest dynamics such as deforestation, biodiversity and productivity. For example, global databases such as the Global Forest Watch and the Forest Resources Assessment (FRA), have been provided at scales sufficient to quantify forest cover, carbon stocks, and degradation of forests on the same land (Hansen et al., 2013; FAO, 2020; Turubanova et al., 2018). For example, satellite platforms such as Landsat, MODIS, and Sentinel have provided global, consistent, repeatable, and high-frequency datasets for forest monitoring (Silveira et al., 2021). These datasets also allow us to accurately compute vegetation indices such as the Normalized Difference Vegetation Index (NDVI), the Enhanced Vegetation Index (EVI), and Solar-Induced Chlorophyll Fluorescence (SIF) to evaluate photosynthesis, biomass production, and canopy health (Li et al., 2020; Running et al., 2004).

For instance, there are metrics such as Gross Primary Productivity (GPP) and Net Primary Productivity (NPP) used to evaluate forest productivity and have greatly increased the extent of our knowledge of ecosystem functions (Beer et al., 2010; Zhao and Running, 2010). Long-term monitoring of forests benefits from using traditional indices such as NDVI and EVI, while newer metrics (e.g SIF) highlight the direct and sensitive measurements of photosynthesis and early response relating forest hemispheric function (Wu et al., 2011). Nonetheless, these methods exhibit some limitations such as difficulties in estimating understory vegetation, atmospheric effects, and requirements for ground validation to improve accuracy (Chrysafis et al., 2020).

One specific area where new technologies like LiDAR and radar imaging have expanded the horizons of forest tracking. Light detection and ranging (LiDAR) offers high-resolution, 3D data, permitting the exact assessment of tree height, canopy structure and biomass (Jastrowski, 2022; Somers and Asner, 2012). These systems, e.g. Sentinel-1 radar, perform well in cloudy conditions, which is advantageous for tropical countries like Mozambique (Chrysafis et al., 2020; Shimada et al., 2014). The combination of these datasets with satellite imagery provides a multi-dimensional perspective on geographic and temporal trends in forest ecosystems, elucidating significant relationships among forest structure, productivity, and socio-economic variables defined in the conceptual framework. As an example, LiDAR data paired with NDVI and PRI metrics can be used to evaluate how water security mediates forest productivity, and radar imagery can be used to monitor forest resilience to agricultural intensification (Watmough et al., 2019).

In addition, machine learning methods are also increasingly becoming used in processing and analyzing large-scale remote sensing datasets. Such methods have demonstrated capabilities for forest type classification, biomass prediction, and disturbance detection, allowing for better connections between ecological indicators and socio-economic dynamics (Silveira et al., 2021).

Remote sensing has limitations, however, despite its transformative potential. It also cannot directly assess ecological processes such as soil respiration or biodiversity indices, or fully capture socio-economic drivers of forest change (Schimel et al., 2015; Pettorelli et al., 2014). These gaps can only be filled through ground-based surveys and socio-economic data. Fieldwork is indispensable for conditions like household wealth, agricultural diversification and village infrastructure which are essential for comprehending livelihoods in Niassa are poorly observed from remote sensing observations (Goetz et al., 2012; Tyukavina et al., 2018).

A conceptual framework (Figure 1) was devised to illustrate the need for these integrated approaches by viably categorizing variables requiring remote sensing, socio-economic surveys, or both. For instance, although forest productivity can be estimated based on NDVI, it is necessary to link its measurement to other socio-economic data in order to analyse household wealth or food security. This mixed method approach facilitates a comprehensive understanding of the dynamics between forest and human systems (B. Fisher and Burgess, 2011; Shackleton et al., 2007).

Remote sensing has emerged as an essential tool for forest ecosystem monitoring that can provide large-scale, high-resolution information on forest dynamics. Another component of this research is the data linkage with ground-based socio-economic surveys to obtain a full understanding of forest-people interactions. This integrated approach focuses on the interconnections between conservation and the livelihoods and is especially relevant for the Niassa Special Reserve, where sustainable forest management needs to balance ecological integrity with those of local livelihoods (Chazdon and Guariguata, 2018; Cumming and Allen, 2017). This research will contribute to a science-policy dialogue by linking advanced remote sensing technologies with socio-economic frameworks that can help shape strategies to achieve both conservation and development objectives (Watmough et al., 2019).

1.5 Goals and Reasoning

In this study, I aim to address the gap in understanding the relationships between socio-economic factors and forest functioning in the Niassa Special Reserve, Mozambique. Specifically, I investigate how livelihood strategies influence forest functioning and, in turn, how improved forest functioning supports sustainable livelihoods. Forest functioning in this context refers to the capacity of forest ecosystems to deliver key ecosystem services, including carbon sequestration, biodiversity support, and water regulation (Chapin et al., 2011; Mori et al., 2017).

The Niassa Special Reserve is an ideal site for studying these dynamics. It is one of Africa's largest conservation areas, encompassing diverse ecosystems that support rich biodiversity and provide essential ecosystem services (Lindsey et al., 2017; Harris et al., 2021). However, this region faces significant challenges, including high poverty rates, resource scarcity, and pressures from unsustainable resource use and climate change. Conservation areas like Niassa are critical for global biodiversity goals, yet their success depends on balancing ecological integrity with the needs of local communities who depend on forest resources for their livelihoods (Bennett and Roth, 2015; Shackleton et al., 2015). Despite the global focus on African conservation areas, research remains limited on how specific livelihood strategies, such as agricultural diversification and access to water, interact with forest functioning in regions like Niassa (Angelsen et al., 2014; Shackleton et al., 2007).

The African conservation landscape presents a unique context for exploring socio-ecological systems. Protected areas often overlap with regions of high poverty and resource dependency, where communities rely on forests for fuelwood, non-timber forest products, and water resources (Kassa et al., 2017). Livelihood strategies like diversified agriculture have been shown to reduce forest dependency in some regions by providing alternative income sources and reducing the need for land conversion (Scoones, 1998; Ellis, 2000). However, findings on the relationship between agricultural diversification and forest functioning are mixed and context-dependent, requiring further exploration in areas like Niassa (FAO, 2020; Shackleton et al., 2007).

Water security is another critical factor influencing both livelihoods and forests. Access to clean, reliable water supports agricultural productivity and reduces direct pressures on forest water sources. Studies in similar African regions highlight that water scarcity can exacerbate forest degradation, as communities extract resources to compensate for insufficient water supplies (IPCC, 2019; Goetz et al., 2012; Kassa et al., 2017). Despite this, there is limited understanding of how water security specifically interacts with forest functioning in conservation areas.

1.6 Research Questions and Hypotheses

This study is guided by **two hypotheses**:

1. Livelihood strategies and forest functioning: Improved livelihood strategies—such as diversified agriculture and enhanced water security—reduce dependency on forest resources, leading to improved forest functioning. Diversified agriculture refers to incorporating a variety of crops and practices to reduce reliance on forest products. Enhanced water security means reliable access to sufficient, clean water for household and agricultural use, which can reduce pressures on forest ecosystems for water extraction. Improved forest functioning is defined as enhanced carbon storage, greater biodiversity, and stable hydrological cycles. Not all forest functions may improve equally; the impacts are expected to vary by the specific livelihood strategy employed (Cumming and Allen, 2017; B. Fisher and Burgess, 2011).

2. Forest functioning and sustainable livelihoods: Enhanced forest functioning supports sustainable livelihoods by providing vital ecosystem services, such as water availability, soil fertility, and climate regulation, which promote community well-being and reduce resource extraction pressures (Adams, 2004; Shackleton et al., 2015).

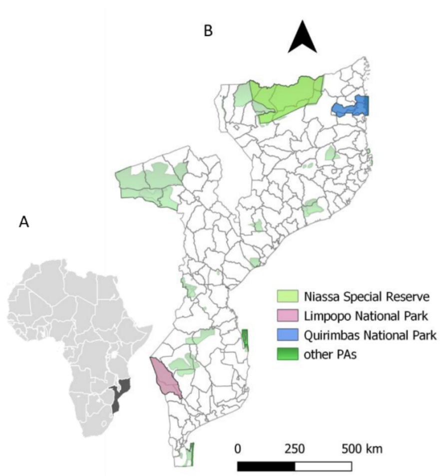
To explore these hypotheses, I address the following specific **research questions**:

1. How do socio-economic dimensions, including village infrastructure and household wealth, influence forest functioning in protected areas like the Niassa Special Reserve?

2. How do livelihood strategies contribute to forest functionality and mediate the pressures between socio-economic development and ecological sustainability?

2 Data

2.1 Description of study site



(a)

(b)

Figure 2:
a.A) Location of the Republic of Mozambique on the African continent marked in black.
a.B) Map of Mozambique with the locations of Protected Areas, Niassa Special Reserve in light green and Quirimbas National Park in blue (Pereira et al., 2021).
b) Sampled villages in Niassa Special Reserve (NSR) (Pereira et al., 2021).

2.1.1 Mozambique

Human Dimension

Mozambique is a country located in the south-east of the African continent, bordering Tanzania, Malawi, Zimbabwe, Eswatini, South Africa, and the Indian Ocean to the east (Fig. 2a). The population of is roughly 35 million people with diverse religious beliefs (e.g. Christian, Muslim), ethnics (e.g. Makua, Sena, Tsong), languages (Makua, Portuguese, Tsonga), and cultures (Bantu, Sawahili, Portuguese) (United Nations Department of Economic and Social Affairs, 2024). The Republic of Mozambique is a poor country where almost 75% of the population lives under poverty levels (The World Bank, 2019) and 30% no access to electricity (The World Bank, 2021). Over 60% of the population lives in rural areas (The World Bank, 2022), where employment chances are sparse, and therefore much of the population depends on the exploitation of land and forest resources (Pereira et al., 2021). Primary activities include subsistence agriculture, forestry, livestock rearing, and fishing (World Food Programme, 2021). To fulfill their resource needs communities living next to or within secondary forests commonly use them to fulfill their resource needs (World Food Programme, 2021).

Nature Dimension

Mozambique has a mostly tropical climate with a wet and dry season. From October to May, the wet season brings 80% of annual rain, a key determinant of agriculture and forest ecosystems (Bank, 2024). However, the country faces high inter-annual variability in rainfall, resulting in periodic droughts and floods, mostly in the southern and central parts of the country. In contrast, Northern Mozambique (such as Niassa province) is wetter and better suited to growing dense forest and agriculture. Its ecosystems, which range from miombo woodlands and coastal mangroves to grasslands, provide habitats for numerous endemic and migratory species. With climate change temperatures are rising, patterns of precipitation are changing, and extreme weather events are becoming more common, which exacerbates already existing challenges (Bank, 2024; World Food Programme, 2021).

2.1.2 Niassa Special Reserve

Human Dimension

The Niassa Special Reserve, located in northern Mozambique, encompasses 58 villages with approximately 40,000 inhabitants. Of these, 14 villages in two provinces were surveyed as part of this study (Fig. 2b). Communities within the reserve primarily rely on subsistence agriculture, fishing, and non-timber forest products for their livelihoods (Pereira et al., 2021). High poverty levels and limited access to education, healthcare, and infrastructure exacerbate dependency on natural resources. Despite this, local knowledge and community-based conservation initiatives play a vital role in resource management and biodiversity conservation (Wildlife Conservation Society, 2021).

Nature Dimension

The Niassa Special Reserve covers 42,300 km^2, making it one of the largest protected areas in Africa. Via a natural corridor of forestry concessions, it links with the Quirimbas National Park, allowing for a vast and contiguous land-space for wildlife (Wildlife Conservation Society, 2021). Miombo woodlands are the dominant habitat type in the reserve, broken up by savanna and riverine forests, and it is home to some of the most iconic species in the world, including African elephants, lions,and leopards (World Food Programme, 2021). The area is extremely vulnerable to climate change, with declining early-season rainfall and increasing incidence of drought threatening its ecosystems. The reserve is still an important area for biodiversity conservation and provides essential ecosystem services such as water regulation and carbon storage, despite these threats (Pereira et al., 2021).

2.2 Data Collection

2.2.1 Sentinel-2

Sentinel-2 is a satellite constellation, funded by the European Space Agency's Copernicus programme, comprised of two satellites with improved spatial, temporal, and spectral resolution compared to previous missions (Perez-Priego et al., 2015). The spatial resolution is either 10m, 20m or 60m (ESA, 2024). The spectral resolution contains 13 bands in the visible, red-edge, near infrared, and shortwave infrared, split over the different resolutions (see Figure 3). Helfenstein et al. (2022, (Helfenstein et al., 2022)) analyzed Sentinel-2 data and concluded that meaningful analysis requires a minimum area with a radius of 60m. Thus, using a dataset with a 20m pixel size should yield significant results. The temporal resolution (3-5 days revisit time) makes it very fitting for ecological observations as a patch of land can be observed over the entire year with its seasons (Drusch et al., 2012). Also, the spectral resolution of the satellite constellation is very useful for observing plant species from space as the red-edge is very important to determine leaf chemistry, PRI, NDVI, and NDWI (Papale et al., 2008; Perez-Priego et al., 2015) (see Table 2 for more information).

Band number	S2A		S2B	
	Central wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)	Bandwidth (nm)
2	492.7	65	492.3	65
3	559.8	35	558.9	35
4	664.6	30	664.9	31
8	832.8	105	832.9	104

(a)

Band number	S2A		S2B	
	Central wavelength (nm)	Bandwidth (nm)	Central wavelength (nm)	Bandwidth (nm)
5	704.1	14	703.8	15
6	740.5	14	739.1	13
7	782.8	19	779.7	19
8a	864.7	21	864.0	21
11	1613.7	90	1610.4	94
12	2202.4	174	2185.7	184

(b)

Figure 3:
a) Spectral Resolution for bands with spatial resolution of 10m for both Sentinel-2 satellites (ESA, 2024)
b) Spectral Resolution for bands with spatial resolution of 20m for both Sentinel-2 satellites (ESA, 2024)

2.2.2 Questionnaire from Pereira et al. (2021)

The study applied the questionnaire of Pereira et al. (2021), which was a study devised to examine the links between socio-economic aspects and Human-Wildlife dynamics in the Niassa Special Reserve, Mozambique. It draws on the Sustainable Livelihoods Framework (SLF) (Chambers et al., 1989; Ellis, 2000) and the Social-Ecological Systems (SES) Theory (Ostrom, 2009). These frameworks provide a foundation for analyzing how human activities and ecological systems interact.

The questionnaire includes several sections to collect comprehensive data about households and their livelihoods. It begins by addressing basic demographic information, such as household size, age, education levels, and dependency ratios. This kind of data provides a baseline for understanding the socio-economic conditions of the interviewees. The next section focuses on livelihood strategies, including primary and secondary income sources, agricultural practices, and the use of forest resources (Pereira et al., 2021). The questionnaire also collects information on people's perceptions of conservation policies and their experiences with wildlife interactions, such as how frequently these interactions occur and their impacts on daily life (Wildlife Conservation Society, 2021; Pereira et al., 2021).

The questionnaire also seeks to measure vulnerability and how well households can adapt to challenges. It analyzes the availability of natural, physical, human, social, and financial capital for households (Chambers et al., 1989; Sallu et al., 2010). These are then assessed in the context of how households cope with risks such as scarcity of resources and conflicts with wildlife and how rainfall patterns change (IPCC, 2019; Pereira et al., 2021).

The theory behind the questionnaire is based around the five types of capital mentioned in the SLF and their enhancement in improving livelihoods. SES theory provides a framework to link these social and economic variables to forest function, for example, how household strategies influence ecosystem services such as carbon storage or water regulation (Ellis, 2000; Ostrom, 2009).

Data collected from the questionnaire are used in this thesis to investigate how livelihood strategies impact forest dynamics. The study shows which socio-economic factors favor or harm forest dynamics by exploiting determinants like income sources, resource utilization and adaptive capacity. Results from this line of inquiry contribute to answering the research questions and testing the observational hypotheses of relationships between livelihoods and forest functioning.

2.3 Socio-Economic Variables and Their Dimensions

This study evaluates socio-economic variables influencing livelihoods and forest functioning in the Niassa Special Reserve, Mozambique. These variables are selected for their relevance to household resilience, forest dependency, and well-being. The calculations were adapted from Pereira et al. (2021).

2.3.1 Water Dimension

Forests contribute significantly to water security by maintaining groundwater levels, enhancing water filtration, and regulating the hydrological cycle (IPCC, 2019). This study uses the *Water Security* variable, which measures the proportion of households in a village that obtain water from wells or fountains, indicating better infrastructure and reduced dependence on forest water sources (IPCC, 2019; Pereira et al., 2021):

$$Water\ Security = 1 - \frac{\text{HH using rivers or lakes}}{\text{Total HH}}$$

2.3.2 Food Dimension

Forests support food security by providing essential ecosystem services such as soil stabilization, nutrient cycling, and water regulation (Ellis, 2000; Pereira et al., 2021). Two key variables are used:

Food Security Index: This variable measures the months households struggle to find food, normalized on a scale of 0 to 1. Fewer months of food scarcity indicate better food security (Pereira et al., 2021):

$$Food\ Security\ Index = 1 - \frac{\text{\#Food scarcity months}}{12}$$

However, the data for this variable was unavailable in the final dataset, as the column for Food Security Index was completely empty. It remains unclear whether this omission occurred during data collection, was the result of an error in data entry, or was inadvertently removed during pre-processing. Unfortunately, this oversight went unnoticed during the planning and design of the study, and as a result, the Food Security Index could not be included in the analysis. Instead, other variables, such as Crop Diversity and Agricultural Livelihood Diversification, were selected to indirectly capture aspects of food security.

Crop Diversity: This variable reflects the number of crops grown by a household. Higher crop diversity reduces risks from pests, diseases, and climatic shocks (Ellis, 2000; Pereira et al., 2021):

$$Crop\ Diversity = 1 - \frac{\#\text{Crops}}{\#\text{Crops} + 1}$$

2.3.3 Agricultural Livelihood Diversification

Diversified agricultural livelihoods provide economic stability and reduce reliance on forest resources by increasing food security (Waha et al., 2022; Sallu et al., 2010; Pereira et al., 2021). The *Agricultural Livelihood Diversification Index* measures the number of distinct agricultural activities practiced by households (Sallu et al., 2010; Pereira et al., 2021):

$$Agricultural\ Livelihood\ Diversification\ Index = 1 - \frac{1}{\#\text{Agricultural activities} + 1}$$

2.3.4 Infrastructure Dimension

Village infrastructure increases access to education, healthcare, and markets, reducing forest dependency over time (Wildlife Conservation Society, 2021; Pereira et al., 2021). The *Village Infrastructure Index* measures the presence of key infrastructure components like schools, healthcare facilities, and water sources (Pereira et al., 2021; Wildlife Conservation Society, 2021):

$$Village\ Infrastructure\ Index = \frac{\#\text{Infrastructure components}}{8}$$

2.3.5 Human and Social Dimensions

Dependency Ratio: This variable measures the proportion of non-working individuals (children, younger than 16, and elderly, older than 64 years old) to the working-age population. A higher ratio is hypothesized to increase reliance on forest resources for survival (Sapkota and Odén, 2008; Pereira et al., 2021):

$$Dependency\ Ratio = \frac{\#\text{Children} + \#\text{Elderly}}{\#\text{Household members}}$$

Household Wealth: Wealthier households are more resilient and less dependent on forest resources (Ellis, 2000). The *Household Wealth* variable is calculated based on the number of assets owned by a household:

$$Household\ Wealth = \frac{\#\text{Assets}}{7}$$

Average Receive/Give Ratio: This variable reflects the balance between resources given to and received from the community, capturing social capital and economic stability (Pereira et al., 2021):

$$Average\ Receive/Give\ Ratio = \frac{\#\text{Received}}{\#\text{Given}}$$

Table 1: Socio-Economic Variables in Niassa Special Reserve

Variable Name	Definition	Formula	Range
Water Security	Proportion of HH accessing wells/fountains	$1 - \frac{\text{HH using rivers/lakes}}{\text{Total HH}}$	[0-1]
Food Security Index	Months of food scarcity	$1 - \frac{\text{\#Food scarcity months}}{12}$	[0-1]
Crop Diversity	Number of crops grown	$1 - \frac{\text{\#Crops}}{\text{\#Crops+1}}$	[0-1]
Agricultural Livelihood Diversification Index	Agricultural activities per HH	$1 - \frac{1}{\text{\#Activities+1}}$	[0-1]
Village Infrastructure Index	Presence of key amenities	$\frac{\text{\#Amenities}}{8}$	[0-1]
Dependency Ratio	Non-working to working population ratio	$\frac{\text{\#Children + Elderly}}{\text{\#Working}}$	[0-1]
Household Wealth	Number of household assets	$\frac{\text{\#Assets}}{7}$	[0-1]
Average Receive/Give Ratio	Resource exchange balance	$\frac{\text{\#Received}}{\text{\#Given}}$	[0-∞]

2.4 Remote Sensing Variables

Using variables from the Sentinel-2 dataset, this study analyses forest functioning, water availability, nutrient and biological diversity in the Niassa Special Reserve. The high spatial resolution (10m, 20m or 60m) of Sentinel-2 images, along with the 13 spectral bands acquired in the visible, red-edge, near-infrared and shortwave infrared regions, makes it well-suited for ecological studies (Perez-Priego et al., 2015; ESA, 2024). The high temporal resolution (3–5 days) translates to consistent observations and accurate representation of ecosystem states over time (Drusch et al., 2012). In order to merge variables with differing spatial resolutions, all data were resampled to a resolution of 20m (Helfenstein et al., 2022) An Overview of the Remote Sensing Variables and their Formulas are given in Table 2.

2.4.1 Normalized Difference Vegetation Index (NDVI)

The NDVI is a widely used index for estimating forest productivity. It measures the difference between near-infrared (strongly reflected by healthy vegetation) and red light (absorbed during photosynthesis):

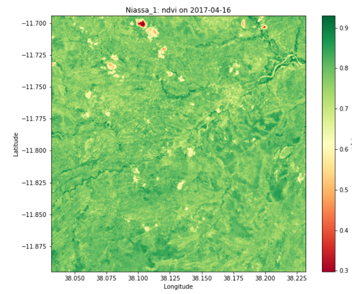
$$\text{NDVI} = \frac{\rho\text{Band 8} \;-\; \rho\text{Band 4}}{\rho\text{Band 8} \;+\; \rho\text{Band 4}}$$

NDVI values range from -1 to 1, where higher values indicate more vigorous vegetation. This index has been validated as a robust measure of forest productivity in numerous studies, including Wang et al. (2017); Horion et al. (2014), and is used here to quantify the functional state of forest ecosystems.

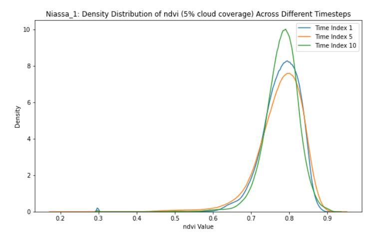
Figure 4 provides an overview of NDVI variations in the Niassa Special Reserve. The spatial distribution of NDVI values at a single time step is shown in Figure 4a, where higher values indicate dense vegetation, while lower values represent sparsely vegetated or degraded areas.

To further analyze temporal changes, Figure 4b presents the density distribution of NDVI values across multiple time steps. Seasonal fluctuations in vegetation health can be observed, likely influenced by precipitation patterns and human activities.

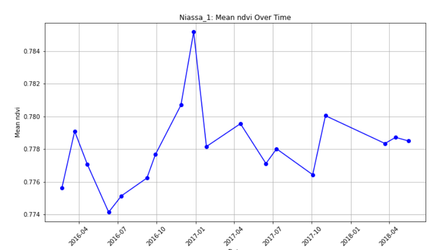
Finally, Figure 4c depicts the mean NDVI trend over time, revealing long-term changes in forest productivity. This visualization helps assess vegetation resilience and potential degradation patterns in response to environmental and anthropogenic factors.



(a) NDVI at a single time step.



(b) NDVI distribution over multiple time steps.



(c) Mean NDVI over the observed period.

Figure 4:
Overview of NDVI dynamics for Necuti village:
(a) NDVI values at a single time step, showing spatial variation. (b) NDVI distribution across multiple time steps, highlighting seasonal fluctuations. (c) Mean NDVI trends, indicating long-term changes in forest productivity.

2.4.2 Normalized Difference Water Index (NDWI)

The NDWI measures water content in vegetation or surface water, helping to monitor water availability for plants:

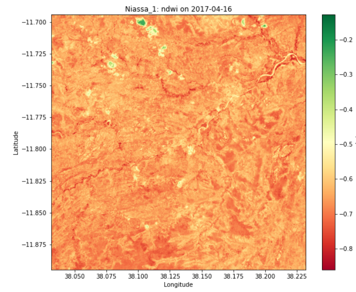
$$\text{NDWI} = \frac{\rho\text{Band } 3 \ - \ \rho\text{Band } 8}{\rho\text{Band } 3 \ + \ \rho\text{Band } 8}$$

Higher NDWI values suggest greater water content, which is critical for understanding plant health and resilience to drought. Studies such as Marusig et al. (2020) have demonstrated NDWI's utility in tracking changes in vegetation water content and hydrological conditions. Unlike NDVI, which primarily measures vegetation health, NDWI is particularly useful for identifying water stress, monitoring seasonal fluctuations in hydrology, and detecting drought impacts on ecosystems.

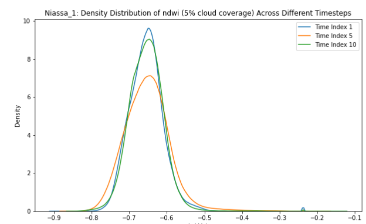
Figure 5 provides an overview of NDWI variations in the Niassa Special Reserve. The spatial distribution of NDWI values at a single time step is shown in Figure 5a, where higher values indicate areas with more water, while lower values suggest drier conditions. This helps to identify locations where water resources are concentrated and areas that may be prone to moisture stress.

To analyze temporal changes, Figure 5b presents the density distribution of NDWI values across multiple time steps, highlighting fluctuations in water availability that may be influenced by seasonal variation and environmental factors. These distributions reflect shifts in water content, which could result from rainfall patterns, evaporation rates, or land use changes.

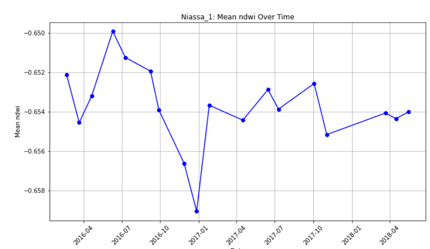
Lastly, Figure 5c depicts the mean NDWI trend over time, allowing us to assess how water availability in vegetation changes over the observed period.



(a) NDWI at a single time step.



(b) NDWI distribution over multiple time steps.



(c) Mean NDWI over the observed period.

Figure 5:
Overview of NDWI dynamics for Necuti village:
(a) NDWI values at a single time step, showing spatial variation and highlighting areas with high water content. (b) NDWI distribution across multiple time steps, illustrating seasonal fluctuations in vegetation water availability. (c) Mean NDWI trends, indicating long-term changes in water content and hydrological conditions.

2.4.3 Leaf Chemistry ($\mathrm{CI}_{\text{red edge}}$)

Leaf Chemistry, particularly nitrogen (N) and phosphorus (P) concentrations, is estimated using the red-edge chlorophyll index ($\mathrm{CI}_{\text{red edge}}$):

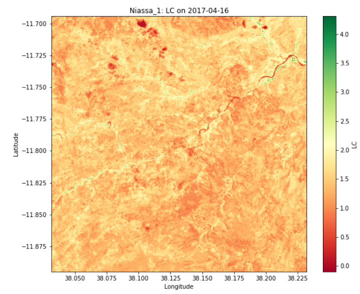
$$\mathrm{CI}_{\text{red edge}} = \frac{\rho\text{Band 7}}{\rho\text{Band 5}} - 1$$

This index reflects nutrient availability in leaves, influencing photosynthetic efficiency and productivity. The $\mathrm{CI}_{\text{red edge}}$ has been validated for ecological applications by Schlemmera et al. (2013); Jay et al. (2014) and is particularly relevant for understanding nutrient limitations in forest functioning. Leaf chemistry plays a crucial role in defining ecosystem productivity, as nitrogen and phosphorus availability directly impact carbon fixation and primary production.

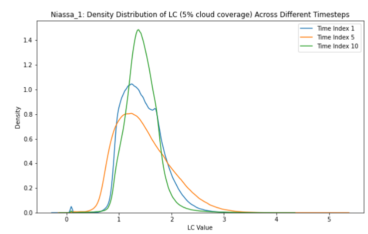
Figure 6 provides an overview of leaf chemistry variations across the Niassa Special Reserve. The spatial distribution of $\mathrm{CI}_{\text{red edge}}$ at a single time step is displayed in Figure 6a, highlighting variations in nutrient content across the landscape. Higher values indicate nutrient-rich regions, while lower values may suggest areas experiencing nutrient deficiencies.

To examine temporal patterns, Figure 6b presents the density distribution of $\mathrm{CI}_{\text{red edge}}$ values at different time steps, allowing us to assess how nutrient availability shifts seasonally. These variations may be linked to leaf senescence, climatic influences, or land-use changes.

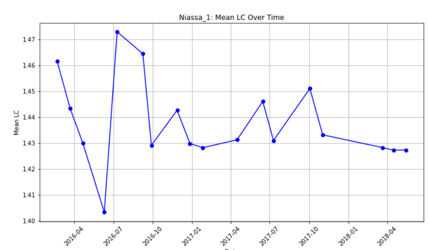
Finally, Figure 6c illustrates the mean $\mathrm{CI}_{\text{red edge}}$ trend over time, offering insight into long-term nutrient availability in the ecosystem. A declining trend could indicate soil degradation or shifts in plant composition, whereas stable or increasing trends might suggest resilient nutrient cycles. Understanding these dynamics is essential for assessing forest health and predicting responses to environmental change.



(a) Leaf chemistry at a single time step.



(b) Leaf chemistry distribution over multiple time steps.



(c) Mean leaf chemistry index over the observed period.

Figure 6:
Overview of leaf chemistry variations for Necuti village:
(a) $\mathrm{CI}_{\text{red edge}}$ values at a single time step, showing spatial variations in nutrient availability. (b) $\mathrm{CI}_{\text{red edge}}$ distribution across multiple time steps, highlighting seasonal fluctuations. (c) Mean $\mathrm{CI}_{\text{red edge}}$ trends, indicating long-term changes in nutrient availability.

2.4.4 Photochemical Reflectance Index (PRI)

The PRI measures photosynthetic efficiency by quantifying the light absorbed for carbon fixation:

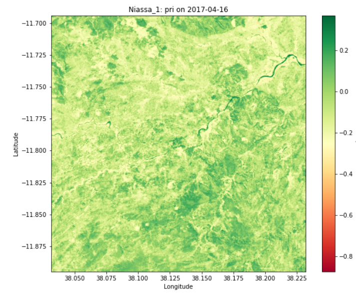
$$\text{PRI} = \frac{\rho\text{Band } 2 - \rho\text{Band } 3}{\rho\text{Band } 2 + \rho\text{Band } 3}$$

This index provides insights into light-use efficiency and stress conditions in vegetation, offering a finer understanding of photosynthetic performance compared to NDVI. Studies such as Nelson et al. (2016); WEI Nan et al. (2017) highlight its role in assessing forest productivity and detecting early signs of plant stress before structural changes become visible.

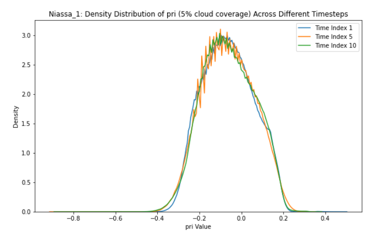
Figure 7 provides an overview of PRI variations in the Niassa Special Reserve. The spatial distribution of PRI at a single time step is shown in Figure 7a, where higher values indicate more efficient light use, while lower values suggest photosynthetic stress or suboptimal conditions.

To examine temporal patterns, Figure 7b presents the density distribution of PRI values over multiple time steps, highlighting seasonal variations in photosynthetic activity. These changes may be linked to fluctuations in water availability, temperature stress, or phenological shifts in vegetation.

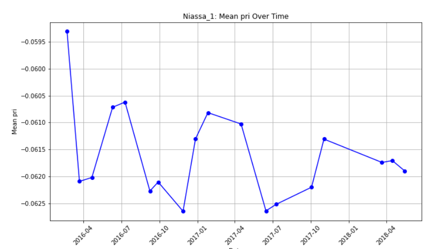
Finally, Figure 7c illustrates the mean PRI trend over time, offering insights into long-term patterns in photosynthetic efficiency. A declining trend could indicate increasing environmental stressors, whereas a stable trend suggests resilient vegetation. Monitoring PRI dynamics is essential for understanding forest health, as it enables the detection of physiological stress before structural degradation becomes apparent.



(a) PRI at a single time step.



(b) PRI distribution over multiple time steps.



(c) Mean PRI over the observed period.

Figure 7:
Overview of PRI variations for Necuti village:
(a) PRI values at a single time step, showing spatial variations in photosynthetic efficiency. (b) PRI distribution across multiple time steps, illustrating seasonal fluctuations in light-use efficiency. (c) Mean PRI trends, indicating long-term changes in photosynthetic performance and stress conditions.

2.4.5 Spectral Diversity (RaoQ)

Spectral diversity serves as a proxy for biodiversity and ecosystem resilience, capturing variability in spectral reflectance across different wavelengths. It is calculated using Rao's Quadratic Entropy (Rao, 1982):

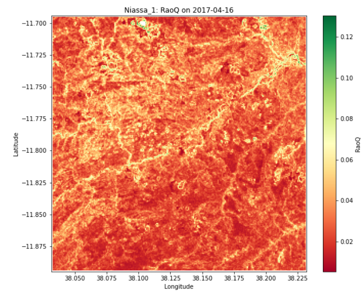
$$RaoQ = \sum_{i=1}^{N} \sum_{j=1}^{N} p_i p_j d_{ij}$$

where p_i and p_j represent the proportions of reflectance in spectral bands i and j, and d_{ij} is the spectral distance between them. Higher RaoQ values indicate greater spectral heterogeneity, which has been shown to correlate with biodiversity and functional diversity in forests (Rocchini et al., 2019; Rao, 1982; Rossi et al., 2021).

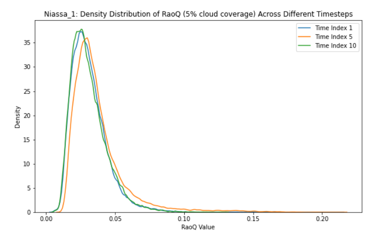
Figure 8 provides an overview of spectral diversity in the Niassa Special Reserve. The spatial distribution of RaoQ values at a single time step is shown in Figure 8a, where higher values suggest increased spectral heterogeneity and, potentially, greater ecosystem complexity.

To examine temporal dynamics, Figure 8b presents the density distribution of RaoQ values across multiple time steps, capturing seasonal or disturbance-driven fluctuations.

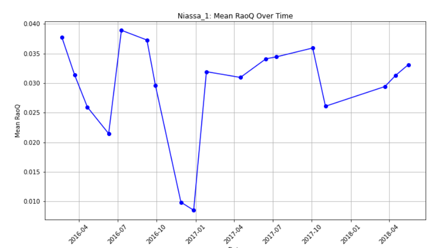
Lastly, Figure 8c shows the mean RaoQ trend over time, revealing shifts in spectral heterogeneity that may reflect changes in biodiversity, habitat structure, or forest resilience. By incorporating this index, the study aims to assess the ecological health and stability of forested landscapes under varying environmental pressures.



(a) RaoQ at a single time step.



(b) RaoQ distribution over multiple time steps.



(c) Mean RaoQ over the observed period.

Figure 8:
Overview of spectral diversity (RaoQ) for Necuti village:
(a) RaoQ values at a single time step, illustrating spatial heterogeneity. (b) RaoQ distribution across multiple time steps, highlighting variability in spectral diversity. (c) Mean RaoQ trends over time, providing insights into ecosystem complexity and resilience.

Table 2: Variables extracted from Sentinel-2 dataset (Band numbers refer to Figure 3)

Variable Name	Definition	Formula
Normalized Difference Vegetation Index (NDVI)	Measure for forest productivity	$\text{NDVI} = \frac{\rho\text{Band 8} - \rho\text{Band 4}}{\rho\text{Band 8} + \rho\text{Band 4}}$
Normalized Difference Water Index (NDWI)	Index for monitoring surface water content	$\text{NDWI} = \frac{\rho\text{Band 3} - \rho\text{Band 8}}{\rho\text{Band 3} + \rho\text{Band 8}}$
Leaf Chemistry ($\text{CI}_{\text{red edge}}$)	Estimates N and P concentration in leaves	$\text{CI}_{\text{red edge}} = \frac{\rho\text{Band 7}}{\rho\text{Band 5}} - 1$
Photochemical Reflectance Index (PRI)	Measures photosynthetic efficiency	$\text{PRI} = \frac{\rho\text{Band 2} - \rho\text{Band 3}}{\rho\text{Band 2} + \rho\text{Band 3}}$
Spectral Diversity (RaoQ)	Estimate of biodiversity	$RaoQ = \sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j d_{ij}$

3 Methods

This section describes the workflow and analytical steps taken in the study. The methods include data acquisition, pre-processing, and analysis, ensuring reproducibility and scientific rigor.

3.1 General Workflow

The general workflow consists of several key steps: (1) data acquisition from Sentinel-2 satellite imagery using an API, (2) cloud filtering to remove low-quality observations, (3) forest extraction to isolate relevant pixels, (4) radiometric normalization for consistent reflectance values across images, and (5) structural equation modeling (SEM) to assess relationships among variables. Figure 9 provides a visual summary of the workflow.

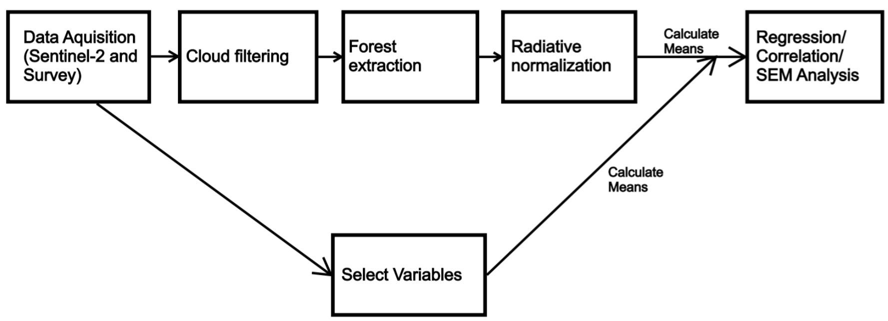


Figure 9: Figure to help visualizing the workflow. First we needed some data, then we had to pre-process it, and then combine the two different datasets using means. Lastly, we were able to conduct three different analysis types.

3.2 Data Acquisition

Sentinel-2 Level-2A data was obtained using the Copernicus Open Access Hub API. A Python script (see Appendix C.1) automated data download for specific coordinates and time intervals based on defined polygons (buffer zones around villages). The script employed a temporal resolution of 28 days, covering the period from January 2016 to December 2023, ensuring seasonal and inter-annual variability was captured (ESA, 2024).

The imagery was first stored as GeoTIFF files and then merged into a NetCDF stack for easier handling during subsequent processing steps.

3.3 Cloud Filtering

Cloud contamination in satellite imagery is a significant challenge for time-series analysis and other applications that demand high-quality data (Meraner et al., 2020). To address this, a multi-step cloud filtering process was implemented, leveraging spectral thresholds, temporal consistency, and seasonal constraints to minimize cloud-related interference. Below, the theoretical foundation, implementation details, and justification of the approach are described. The script used for the cloud filtering and calculation of remote sensing indices can be found in the Appendix C.2.

3.3.1 Theory Behind Cloud Filtering

Clouds and their shadows affect satellite imagery by obstructing ground features, leading to inaccuracies in analyses such as vegetation indices or land cover classification (Stucker et al., 2023). To mitigate these effects, cloud filtering employs techniques such as:

1. **Spectral Thresholding**: Clouds exhibit high reflectance in visible bands (e.g., Blue, Green, and Red) and can be detected using predefined thresholds (Li et al., 2021). This approach is widely adopted due to its simplicity and effectiveness.

2. **Temporal Consistency**: By analyzing pixel stability across time, cloud-free areas can be identified. Pseudo-invariant features (PIFs), characterized by low temporal variability, serve as reliable markers for cloud-free data (Li et al., 2021).

3. **Seasonal Filtering**: Restricting analysis to specific seasons (e.g., dry season) reduces variability introduced by atmospheric conditions and cloud cover, which has been shown to enhance data quality in studies focusing on agricultural and ecological applications.

These methods collectively enhance the quality of data by ensuring that only cloud-free and temporally stable imagery is retained.

3.3.2 Implementation of Cloud Filtering

The filtering process was implemented as follows:

1. **Initial Cloud Masking**: High-reflectance areas in spectral bands (Blue, Green, and Red) were masked using a threshold value. Pixels exceeding the defined threshold were flagged as cloudy. This approach leverages the distinctive spectral signature of clouds, as noted in prior research (Li et al., 2021; Zhu and Woodcock, 2015).

2. **Seasonal Filtering**: Only images captured during the dry season (May to October) were retained as shown in Figure 10, ensuring reduced atmospheric interference and consistent seasonal conditions. Seasonal filtering has been particularly effective in regions with distinct wet and dry seasons (Detsch et al., 2016; Hansen et al., 2003).

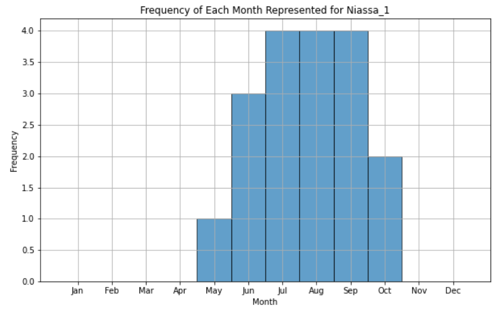


Figure 10:
Frequency distribution of months represented in the dataset for Necuti village (Niassa_1). After filtering only observations between May and October remain in the dataset, corresponding to the dry season, reducing interference from cloud cover.

3.3.3 Justification for the Approach

The approach described here provides a balance between computational efficiency and data quality, leveraging both established and modern techniques to achieve robust cloud filtering. This process is based on spectral thresholding which is a simple and effective approach for identifying high-reflectance clouds in optical imagery (Li et al., 2021; Zhu et al., 2015; Saeed et al., 2021). In this way, the most heavily contaminated pixels are filtered out first in the processing pipeline—a crucial step for reducing the number of clouds present in datasets utilized in global ecological (Sun et al., 2021) or agricultural (Young et al., 2021) studies.

In addition, incorporating seasonal filtering that targets the dry season in which the data is collected reduces the influence of variable phenological states and atmospheric conditions, which can degrade the quality of the data during wet seasons (Li et al., 2021; Guerschman et al., 2009). Restriction of analyses to dry seasons increase the overall homogeneity and reliability of satellite-derived datasets, especially in regions with strong climatic patterns (Zhu and Woodcock, 2017; Tucker et al., 2019). This filtering makes the dataset more uniform and reduces seasonal artifacts (Cai and Du, 2021) by restricting the analysis to this time range.

Lastly, radiometric normalization by histogram matching allows keeping imagery which is cloud-free and radiometrically consistent (Gómez et al., 2016; Lu et al., 2020). This step accounts for possible bias in sensors due to variability in sensor calibrations and atmospheric interactions, which can present themselves in unnatural colors (see Figure 11a and 11b) (Liang and Wang, 2021; Young et al., 2021). The process of radiometric normalization has been utilized extensively in the remote-sensing community, enabling harmonization of multi-temporal datasets for subsequent use in applications, from time-series analysis to ecological monitoring (Ghamisi et al., 2019; Rodrigues et al., 2019).

This multi-pass filtration pipeline has been developed to maximize the extent of preserved high quality imagery, whilst minimizing the impact of clouds, while sustaining the overall consistency of the temporal stack, enabling onward applications such as time-series analysis and ecological monitoring (Stucker et al., 2023; Zhu et al., 2015; Roy et al., 2020).

3.4 Forest Extraction

The identification and classification of forested areas in satellite imagery are essential for monitoring ecosystem health, assessing land use changes, and managing natural resources. This process relies on spectral indices and classification techniques to delineate forested regions with high accuracy. Below, the theoretical background and implementation of forest extraction in this study are described.

3.4.1 Theory Behind Forest Extraction

Forest extraction is generally analyzed by means of vegetation indices and spectral characteristics obtained from satellite images. One of the most widely utilized indexes in vegetation analysis has been the Normalized Difference Vegetation Index (NDVI) that can be used to separate vegetation from the non-vegetation using reflectance of vegetation in near-infrared (NIR) and red bands (Tucker, 1979). Very high NDVI values reflect dense, healthy vegetation, and low NDVI values are indicative of sparse or non-vegetated areas.

A further index is the Enhanced Vegetation Index (EVI), which corrects for atmospheric effects and for the influence of soil background reflectance, which can be important in cases of dense vegetation cover (Huete et al., 2002). The aforementioned indices are typically supplemented with additional spectral bands or indices, such as the Normalized Difference Water Index (NDWI), with the aim of removing water bodies, thus ensuring the accuracy of forest delineation (Gao, 1996).

This is mostly done by applying a classification method - either by thresholding or machine learning models - to classify the imagery into segments of forest and non-forest. Including temporal analysis is also possible to search for stable forested areas by excluding seasonal changes (Potapov et al., 2012).

3.4.2 Implementation in This Study

In this study, forest extraction was performed using NDVI and other spectral indices (see Table 2) derived from Sentinel-2 satellite imagery. The process was designed to minimize noise from non-vegetation areas, water bodies, and seasonal changes. The key steps were as follows:

1. **NDVI Calculation and Filtering**: The NDVI was computed using the near-infrared (Band 8) and red (Band 4) spectral bands:

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \tag{1}$$

   NDVI values were calculated for all images in the temporal stack, and a preliminary threshold of 0.3 was applied to exclude non-vegetated areas. This threshold was chosen based on literature and visual inspection of sample areas, ensuring that forests were correctly delineated (Zhang et al., 2020).

2. **Exclusion of Water Bodies Using NDWI**: To exclude water bodies, the Normalized Difference Water Index (NDWI) was calculated:

$$\text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}} \tag{2}$$

Pixels with high NDWI values were flagged and removed from the forest classification. This step is crucial for ensuring that water bodies, which may sometimes exhibit NDVI-like values due to atmospheric reflections, are not misclassified as forested areas (Gao, 1996).

3. **Temporal Stability Filtering**: Temporal analysis was performed to identify stable forested regions. For each pixel, the standard deviation of NDVI values across the temporal stack was calculated. Pixels with high variability were discarded, as they likely represented transient vegetation, croplands, or disturbed areas. This step ensured that only consistent, forested areas were retained (Potapov et al., 2012; Stucker et al., 2023).

4. **Dry Season Filtering**: The analysis was restricted to images captured during the dry season (May to October) to minimize the influence of phenological variability and cloud cover. This temporal filtering reduced noise and ensured the reliability of NDVI-based classification (Tucker et al., 2019).

5. **Classification and Refinement**: A final forest mask was generated by combining the results of the NDVI, NDWI, and temporal stability analyses. Manual refinement of the classification was performed in areas where misclassification was likely, such as near water bodies or urban regions, based on auxiliary datasets and visual validation (Zhang et al., 2020).

3.4.3 Justification for the Methods

This multi-step approach, which this study took, was to overcome the forest extraction unique challenges in tropical and seasonal places. NDVI sensitivity of vegetation is one of several well-established grounds for its application as the primary index to perform forest delineation (Tucker, 1979). NDVI combined with NDWI added robustness through the effective exclusion of water bodies (Gao, 1996), a common source of error in pixel-level classification based on remote sensing (geo)spatial data.

It improved the precision of forest extraction through filtering the transient or seasonal vegetation. This step is especially critical in tropical regions, where agricultural cycles and phenological variation can render noise (Potapov et al., 2012; Stucker et al., 2023). Likewise limiting the analysis to the dry seasons reduced atmospheric effects that are common in wet seasons (Tucker et al., 2019).

Lastly, manual fine-tuning to enable context-based corrections, e.g. erroneous mix-up of pixels in corridor/urban/peri-urban regions. The balance between this aggregate automated approach combined with manually coding methodology led to a relatively high degree of accuracy that was computationally manageable and feasible.

3.5 Radiometric Normalization

To adjust the discrepancies in the pixel values in multi-temporal or multi-sensor imagery, radiometric normalization is one of the most crucial pre-processing steps taken in remote sensing (Qin, 2012; Pflugmacher et al., 2012). Here, radiometric normalization is undertaken specifically for rectifying color variation across images, enhancing a visually coherent analytical and interpretation basis (Furby and Campbell, 1995b). Despite heterogeneity like sensor calibration differences (Schott et al., 1988), varying intervention contexts (**?**Roy et al., 2016) and illumination geometry differences (Chavez, 1996; Flood, 2014), this process would not have sought to achieve radiometrically calibrated reflectance values, but more likely a common saturated color representation. The rationale, theory, and implementation of this are explained below.

3.5.1 Theory Behind Radiometric Normalization

Radiometric normalization adjusts the pixel intensity values of one image (the "target") to match those of another (the "reference") so that both images share similar visual and radiometric characteristics. It can be full absolute methods requiring external calibration data or pure relative methods that rely on using internal object features as references to align images.

Relative radiometric normalization (RRN) was performed for this study to account for visually detectable discrepancies. In particular, we performed histogram matching to normalize the pixel value distributions of target images to those of a reference image, enabling colors to be consistently represented throughout the dataset. For this, we used pseudo-invariant features (PIFs), which means equal and stable areas across images that serve as a reliable point of correlation for normalization (**?**Furby and Campbell, 1995a).

3.5.2 Implementation in This Study

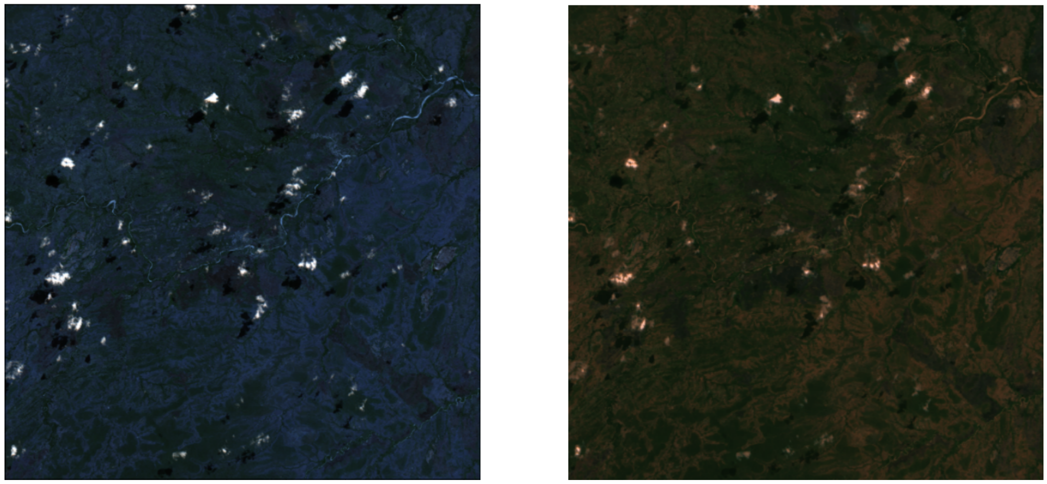
The implementation of radiometric normalization in this study focused on correcting inconsistent color representation (see Figure 11a and 11b) rather than achieving absolute radiometric calibration. The steps are as follows:

1. **Selection of Reference Image**: A single image from the dry season was chosen as the reference image based on its minimal cloud cover and visually consistent colors. Selecting a dry-season image ensured reduced atmospheric variability and better seasonal alignment across the dataset (Furby and Campbell, 1995a).

2. **Identification of Pseudo-Invariant Features (PIFs)**: PIFs were identified by analyzing the temporal stack of images and selecting the 100 pixels with the lowest standard deviation. These pixels, representing stable features such as urban structures or bare soil, provided a robust basis for the normalization process. Their consistent properties helped minimize the impact of transient elements or atmospheric interference (Schott et al., 1988).

3. **Histogram Matching**: The histograms of spectral bands in the target images were adjusted to align with those of the reference image using Python's `match_histograms` function from the `skimage.exposure` library (van der Walt et al., 2014).

4. **Validation of Normalization**: The success of the normalization process was assessed through visual inspection and statistical comparison of histograms between the reference and normalized images. This step confirmed that the adjustments successfully mitigated color inconsistencies without introducing significant artifacts.



(a) (b)

Figure 11:
a) Village Necuti on the 7^{th} of October in 2016 before Histogram matching as RGB visualization.
b) Village Necuti on the 7^{th} of October in 2016 after Histogram matching as RGB visualization.

3.5.3 Justification for the Methods

The purpose of the radiometric normalization implemented in this work was to correct the visual inconsistency in color representation of images rather than to attain absolute radiometric accuracy. Because of this objective, the most appropriate approach was relative radiometric normalization by histogram matching. Meanwhile, this method presented a simple and effective solution without dependency on atmospheric data as well as in situ measurements.

As such, PIFs were a stable and non-varying feature associated with the images while transient or dynamic factors were reduced in their effect during the normalization process (Furby and Campbell, 1995a). In particular, histogram matching was well suited for the task since it alters the discretized color distribution to equalize visual properties between target and reference images (Gómez et al., 2016). By choosing a reference image during the dry season, the method also reduced variability due to atmospheric conditions or seasonal differences.

This process resolved the strange color mismatches, and resulted in visually interpretable images, like the one in Figure 11, that can be used for further analysis like forest delineation or ecological modeling.

3.6 Theoretical Framework for Correlation and Regression Analysis

3.6.1 Correlation Analysis

Correlation analysis is a statistical tool used to identify the strength and direction of linear relationships between variables (Pearson, 1909). In this study, it served as an exploratory method to determine potential associations between socio-economic factors and forest dynamic indicators. The Pearson correlation coefficient (ρ) was used, where values range from

-1 to 1, with:

- **Positive Values:** Indicating a direct relationship (e.g., as Village Infrastructure improves, Water Security increases).
- **Negative Values:** Indicating an inverse relationship (e.g., higher Dependency Ratios correlate with reduced vegetation greenness).
- **Values Near Zero:** Indicating weak or no linear relationship.

This method provided a foundational understanding of variable relationships, informing variable selection for subsequent regression and SEM analyses (Schumacker and Lomax, 2010a; Kline, 2015a; Bollen, 1989a).

3.6.2 Regression Analysis

Regression analysis is a predictive tool that models the relationship between dependent and independent variables (Montgomery et al., 2012). In this study, it was used to quantify the direct effects of socio-economic factors on forest functioning indicators such as NDVI and NDWI. Several types of regression models were employed:

- **Linear Regression:** Assumes a direct proportional relationship between independent and dependent variables.
- **Exponential Regression:** Models situations where changes in the dependent variable occur at an increasing or decreasing rate.
- **Quadratic Regression:** Captures non-linear relationships by including a squared term for the independent variable.

The regression analysis provided insight into both the strength of these relationships and their statistical significance, guiding the inclusion of variables in the SEM (Wooldridge, 2019).

3.7 Structural Equation Modeling (SEM)

Structural Equation Modeling (SEM) is a multivariate statistical analysis technique that includes many statistical techniques such as factor analysis and multiple regression used to analyze item responses for complex relationships among observed and latent variables (Bollen, 1989a; Kline, 2015a). SEM is commonly applied in socio-ecological investigations to study causal relationships among forest functioning, socio-economic, and livelihood strategy pursuits (McCord et al., 2020; Pretty and Bharucha, 2018; Cumming and Allen, 2017). It allows researchers to incorporate measurement error and evaluate direct and indirect effects, making it a powerful tool for understanding multi-dimensional socio-environmental interactions (Kaplan, 2009; Schumacker and Lomax, 2010a). The remainder of this is organized to outline the theoretical basis, implementation, and justification for SEM in this study.

3.7.1 Theory Behind Structural Equation Modeling

SEM enables us to specify, estimate, and test models that describe relationships among observed variables and unobserved latent constructs. Unlike traditional regression techniques, SEM explicitly incorporates measurement error and allows for the simultaneous analysis of multiple dependent relationships (Kaplan, 2009).

SEM models consist of two main components:

1. **Measurement Model**: Describes the relationships between latent variables (unobservable constructs) and their associated observed indicators. For example, forest dynamics can be treated as a latent variable measured by observed indicators such as NDVI, NDWI, and LC (Leaf Chemistry) (Bollen, 1989b).

2. **Structural Model**: Specifies the relationships among latent variables, reflecting hypothesized causal paths based on theoretical or empirical evidence (Kline, 2015b).

SEM provides several advantages, including the ability to test indirect effects, mediation, and feedback loops, making it well-suited for analyzing the complex interplay between socio-economic factors, livelihood strategies, and ecological outcomes (Schumacker and Lomax, 2010b).

3.8 Variable Selection for Structural Equation Modeling

The selection of variables for SEM was informed by the results of regression and correlation analyses as well as theoretical considerations drawn from the Sustainable Livelihoods Framework (SLF). This section outlines the rationale for including specific socio-economic and forest functioning variables and describes the process by which the model structure was determined.

3.8.1 Regression and Correlation Analysis as a Basis

The regression and correlation analyses provided critical insights into the relationships between socio-economic factors and forest functionality. Significant or theoretically relevant predictors identified through these analyses were prioritized for inclusion in the SEM. For example:

- **Village Infrastructure** was consistently associated with NDVI and NDWI, though the relationships were weak to moderate. Its inclusion was justified by its theoretical role in influencing land use and resource access.

- **Dependency Ratio** emerged as a variable with potential indirect effects on NDWI, highlighted by its negative correlations with NDVI and LC_y.

- **Agricultural Livelihood Diversification (ALD)** showed a positive relationship with NDVI, aligning with the hypothesis that diversified livelihoods can enhance ecological outcomes.

- **Water Security and Household Wealth** were identified as potential mediators, given their positive correlations with Village Infrastructure and their potential indirect effects on forest functioning indicators.

3.8.2 Theoretical Rationale

The Sustainable Livelihoods Framework (SLF) served as the theoretical foundation for variable selection. This framework emphasizes the importance of socio-economic factors, resource access, and livelihood strategies in shaping ecological outcomes. Variables were chosen to reflect key dimensions of this framework:

- **Socio-Economic Drivers:** Variables like Village Infrastructure and Household Wealth represent structural and financial capital, which are central to the SLF.
- **Livelihood Strategies:** ALD was included as a representation of how households diversify their income sources to manage risks and enhance sustainability.
- **Environmental Outcomes:** NDVI and NDWI were selected as indicators of vegetation health and water availability, capturing the ecological impacts of socio-economic and livelihood factors.

3.8.3 Progression from Initial to Final SEM Model

The SEM development was guided by theoretical considerations and empirical data from correlation and regression analyses. Figures 12 to 22 illustrate the stepwise refinement process.

- **Initial SEM Model:** The base model (Figure 12) established direct relationships between socio-economic factors, livelihood strategies, and forest functioning indicators. This initial framework aimed to capture wide-scale interactions but required refinement to account for indirect effects and mediation pathways.

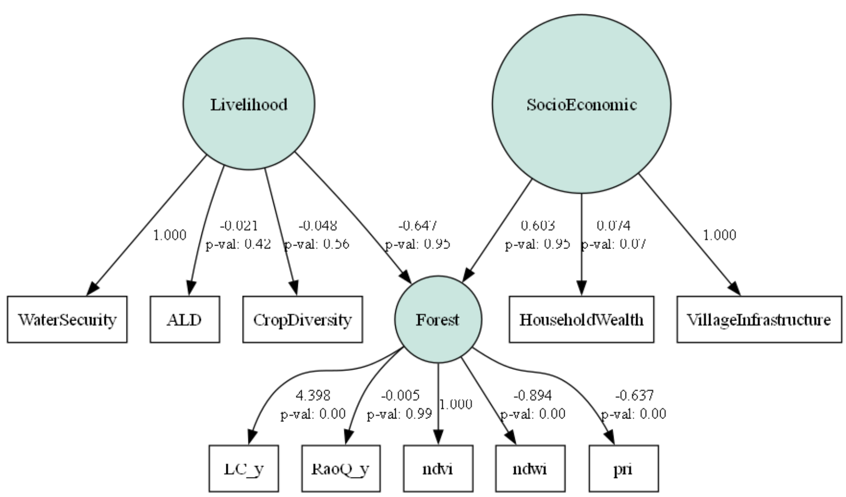


Figure 12: Initial SEM model structure establishing direct relationships between socio-economic factors, livelihood strategies, and forest functioning.

- **Incorporation of Mediators:** Subsequent iterations introduced mediating variables such as Water Security and Household Wealth (Figure 14), hypothesizing their roles in linking socio-economic structures to forest functionality.

- **Final Model Differentiation:** The modeling process culminated in two refined models— a Simple SEM Model (Figure 13) focusing on essential direct pathways and a Complex SEM Model (Figure 14) incorporating additional socio-ecological interactions.

3.8.4 Final Simple and Complex SEM Models

- **Simple SEM Model** (Figure 13): This model retained only the most statistically and theoretically significant relationships, focusing on direct effects of socio-economic and livelihood variables on forest functioning. It prioritizes parsimony while capturing key interactions.

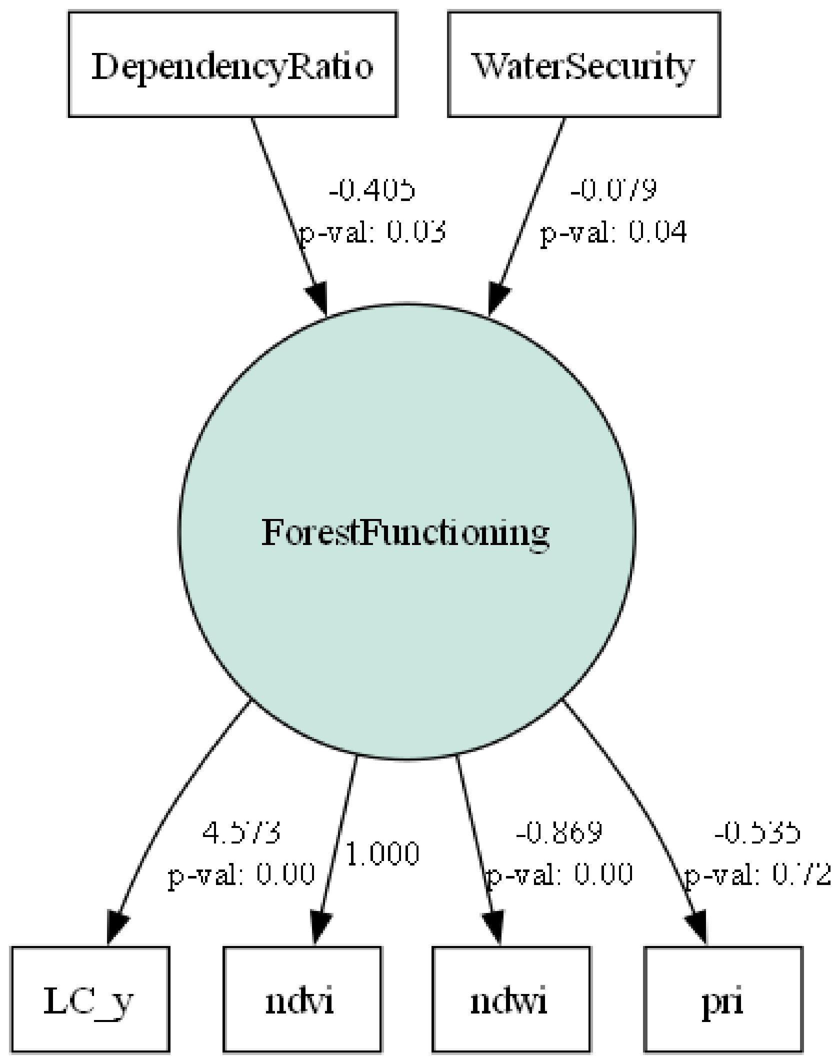


Figure 13: Final Simple SEM Model emphasizing direct pathways between socio-economic structures, livelihood strategies, and forest functioning indicators.

- **Complex SEM Model** (Figure 14): This model includes additional pathways to reflect indirect relationships and mediation effects. It provides a more comprehensive representation of socio-ecological linkages by considering interactions among socio-economic drivers, dependency structures, and livelihood diversification.

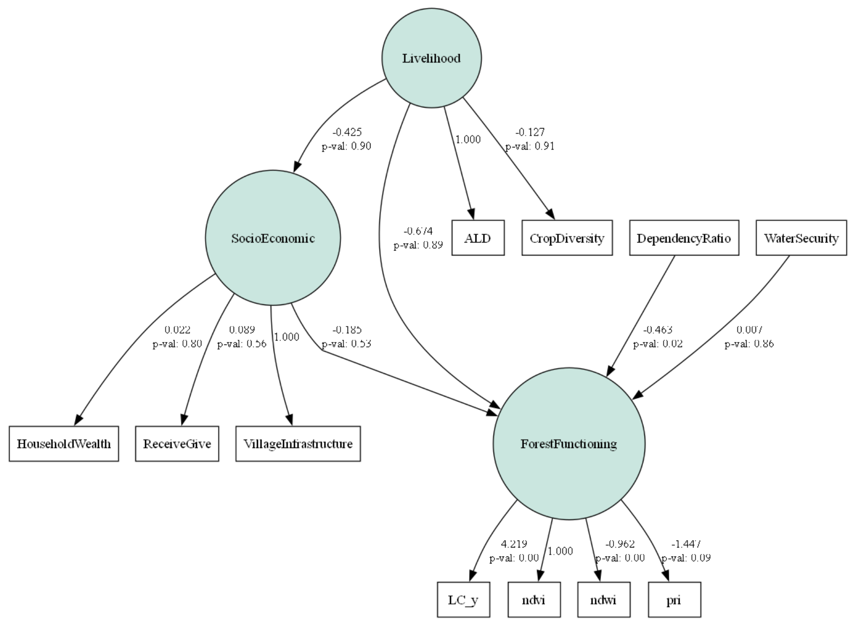


Figure 14: Final Complex SEM Model incorporating additional mediating variables and socio-ecological interactions to better capture systemic dependencies.

3.8.5 Justification for Using SEM

We adopted SEM for this study as it allows for modeling complex relationships among socio-economic, ecological and livelihoods variables simultaneously. The flexibility of SEM allows the inclusion of latent variables like *Forest Functioning*, which cannot be directly measured but may be estimated through several indicators, e.g. NDVI, NDWI and Leaf Chemistry (Kaplan, 2009).

Additionally, SEM allows for the quantification of indirect effects, thus identifying the pathways of mediation. This ability is especially relevant in socio-ecological systems due to the fact that in these systems livelihood strategies frequently mediate the relationship between socio-economic factors and the functioning of forests (Bollen, 1989b). By explicitly accounting for measurement error, SEM enhances the robustness of findings relative to classical regression approaches (Little and Rubin, 2002).

Using Python's `semopy` library allowed SEM to be easily integrated with existing data processing workflows, providing strong model estimation, versatile visualization and simple report generation capabilities. Using a correlation matrix also adds an extra layer of ensuring that the relationships in the model make sense, both in theory and statistically (Rosseel, 2012).

4 Results

The table used for all the following analyses can be seen in the Appendix A. In the Appendix B there are also all linear regression plots that were calculated including their R-squared scored summarized.

4.1 Regression Analysis

4.1.1 Introduction

The regression analysis explores the relationships between socio-economic variables and forest functioning indicators, establishing key patterns and interactions that operate in the Niassa Special Reserve using regression analysis. The dependent variables are forest functioning indicators, such as the NDVI and NDWI. Independent variables captured socio-economic dimensions like Village Infrastructure, Household Wealth, Dependency Ratio and Agricultural Livelihood Diversification (ALD). These analyses give insights into how human development and consumption patterns are influencing ecological dynamics. All linear regression plots are summarized and can be found in the Appendix B.

4.1.2 Methodology

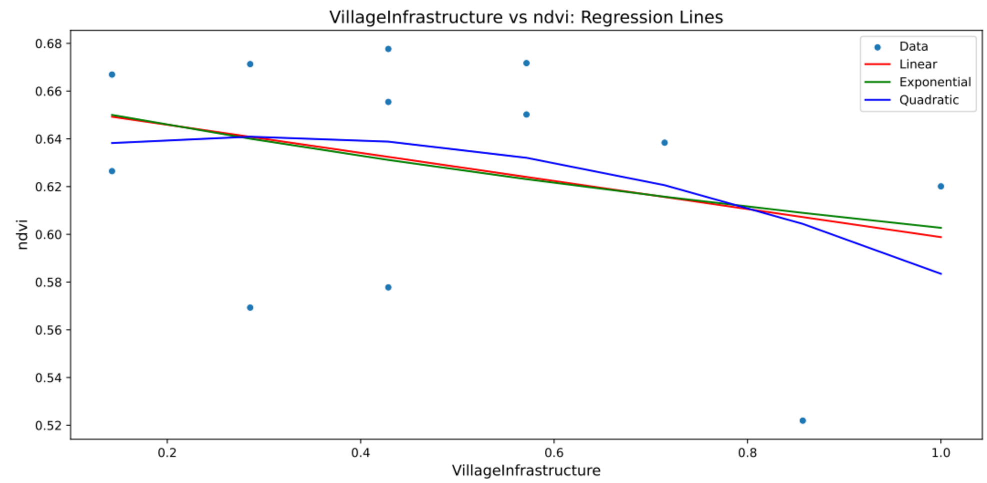
In this study, we utilize three different regression models including linear, exponential, and quadratic regressions to conduct our analysis to investigate linear and non-linear relationships, respectively. Common metrics for each model included R-squared, F-statistics, and p-values to assess their goodness-of-fit and statistical significance. The assumptions of regression analysis were confirmed by diagnostic tests, such as the Durbin-Watson statistic and variance inflation factors (VIF). Although the study had a small sample size (n = 12), limiting statistical power, exploratory insights into these relationships are still possible.

4.1.3 Results

Village Infrastructure NDVI The linear, exponential, and quadratic models were used for regression analysis of Village Infrastructure and NDVI. Proximity to infrastructure explained less than 1% of spatio-temporal variation in vegetation greenness in response to climate variations. With an R-squared value of 0.094, the exponential regression model has a coefficient of -0.0845 and p-value of 0.333. The linear term (p = 0.782) and the squared term (p = 0.597) of the quadratic regression model had a coefficient of 0.0681 and -0.1155, respectively, and its R-squared was 0.133. According to these results, the connection between infrastructure cover and vegetation cover is likely to have some non-linearity, with a general tendency towards very little vegetation cover in most categories, but more data is needed to validate the trends.

Table 3: Regression summary for Village Infrastructure and NDVI.

Model	Coefficient (Linear)	Coefficient (Quadratic)	R-squared
Linear	-0.0589 (p = 0.306)	-	0.104
Exponential	-0.0845 (p = 0.333)	-	0.094
Quadratic	0.0681 (p = 0.782)	-0.1155 (p = 0.597)	0.133



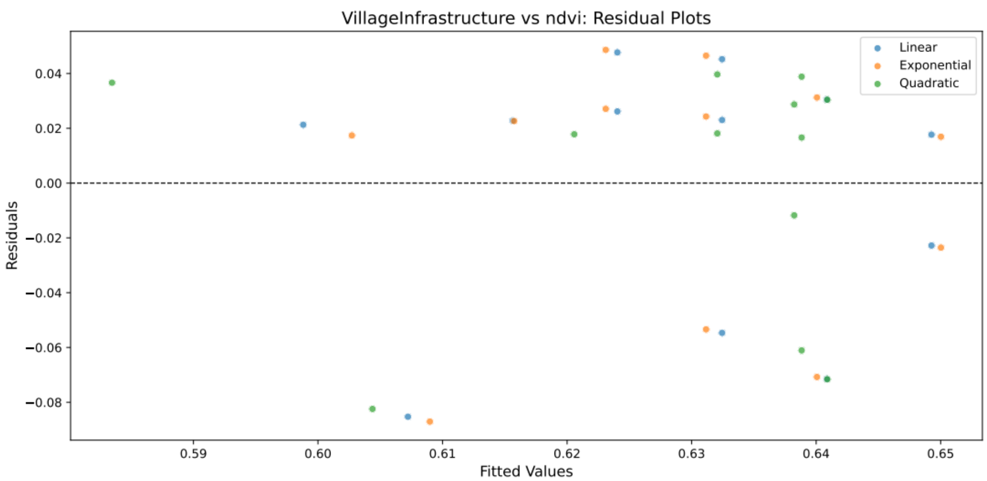


Figure 15: Regression analysis for Village Infrastructure and NDVI with linear, exponential, and quadratic models.

NDWI and Dependent Population In the same manner, three regression models were used to analyze Dependency Ratio and NDWI. The linear regression yielded a coefficient of 0.0440 (p-value = 0.392) of interest and an R-squared of 0.074, suggesting a weak positive relationship. The 2nd order polynomial regression yielded p-value 0.352, r-squared value 0.336 and coefficient 0.0627. The linear and squared terms of the quadratic regression model obtained coefficients of -0.0640 (p = 0.773) and 0.0982 (p = 0.618), respectively, and had an R-squared of 0.101.

Table 4: Regression summary for Dependency Ratio and NDWI.

Model	Coefficient (Linear)	Coefficient (Quadratic)	R-squared
Linear	0.0440 (p = 0.392)	-	0.074
Exponential	0.0627 (p = 0.421)	-	0.066
Quadratic	-0.0640 (p = 0.773)	0.0982 (p = 0.618)	0.101

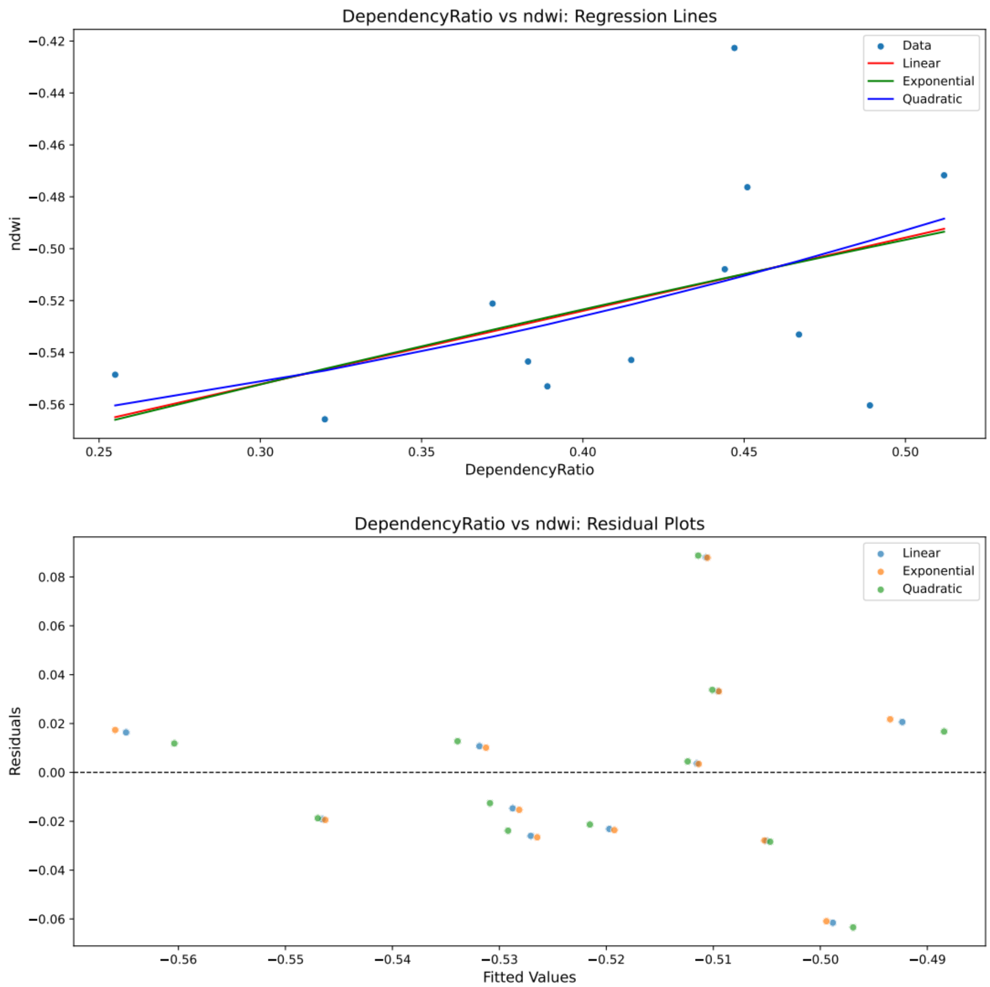


Figure 16: Regression analysis for Dependency Ratio and NDWI with linear, exponential, and quadratic models.

Household Wealth and NDVI Regression models examining the relationship between Household Wealth and NDVI revealed no statistically significant results across all tested models. The linear regression model reported a coefficient of 0.0321 with a p-value of 0.479 and an R-squared value of 0.043. The exponential and quadratic models produced similarly weak relationships, with no substantial improvements in fit. The results indicate that Household Wealth does not have a strong or direct impact on forest greenness.

Table 5: Regression summary for Household Wealth and NDVI.

Model	Coefficient	R-squared
Linear	0.0321 ($p = 0.479$)	0.043
Exponential	0.0456 ($p = 0.502$)	0.041
Quadratic	0.0213 ($p = 0.532$)	0.045

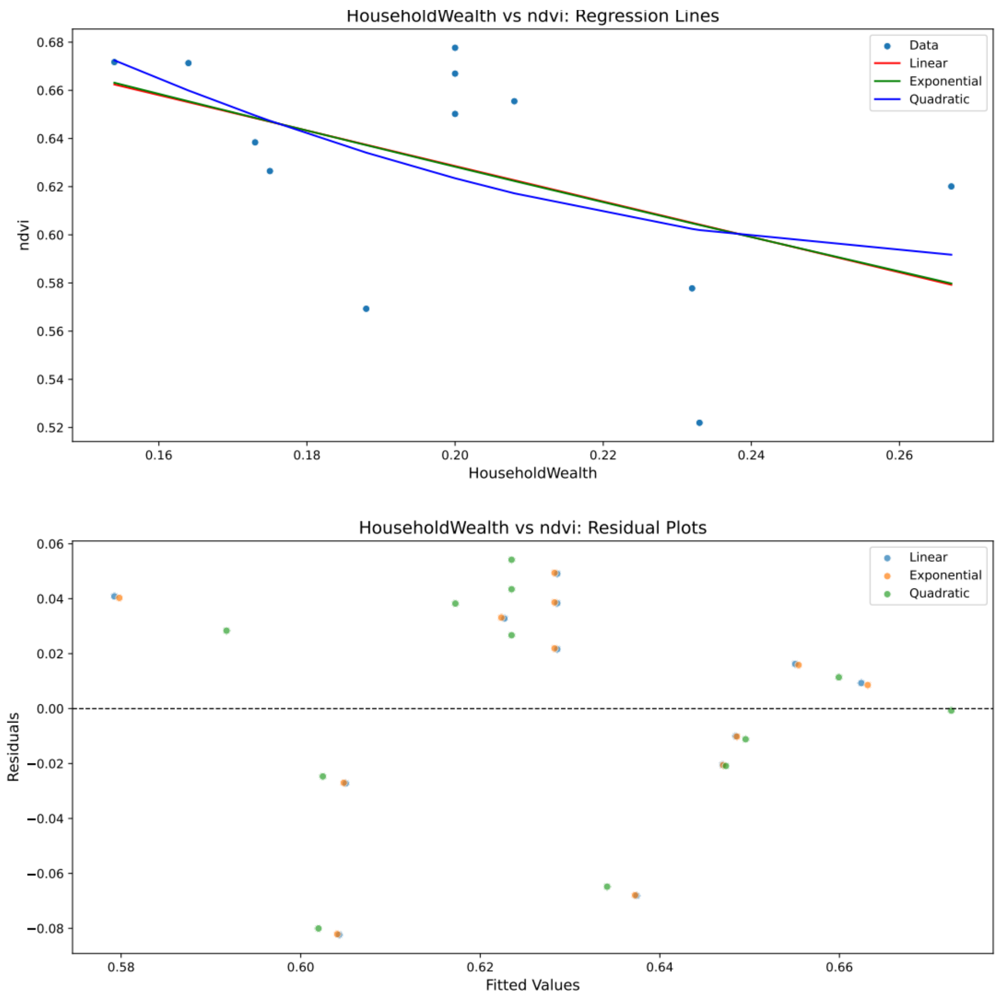


Figure 17: Regression analysis for Household Wealth and NDVI with linear, exponential, and quadratic models.

Agricultural Livelihood Diversification and NDVI Agricultural Livelihood Diversification (ALD) showed a weak positive relationship with NDVI in the linear model, with a coefficient of 0.112 (p = 0.410) and an R-squared value of 0.067. Exponential and quadratic models yielded similar patterns, though none achieved statistical significance. These findings suggest that ALD has a minimal but positive influence on vegetation greenness.

Table 6: Regression summary for Agricultural Livelihood Diversification and NDVI.

Model	Coefficient	R-squared
Linear	0.112 (p = 0.410)	0.067
Exponential	0.105 (p = 0.432)	0.062
Quadratic	0.095 (p = 0.480)	0.070

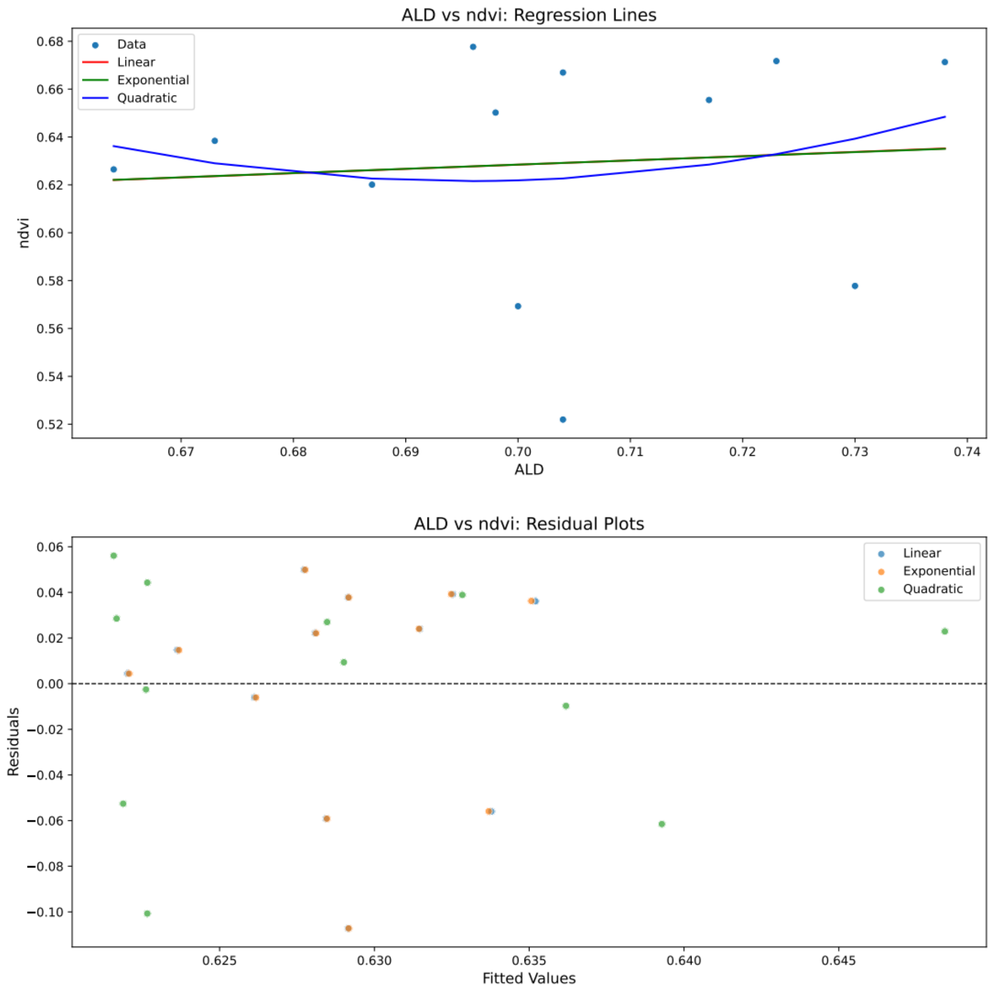


Figure 18: Regression analysis for Agricultural Livelihood Diversification and NDVI with linear, exponential, and quadratic models.

Combined Relationships The combined impact of socio-economic variables such as Village Infrastructure and Household Wealth was tested using linear models. The combined predictors did not yield statistically significant results for any of the forest functioning indicators. For instance, when combined, Village Infrastructure and Household Wealth produced coefficients of -0.041 ($p = 0.505$) and 0.032 ($p = 0.482$), respectively, with an R-squared value of 0.052 for NDVI. Similar non-significant results were found for NDWI and RaoQ_y.

Table 7: Combined regression summary for Village Infrastructure and Household Wealth on NDVI.

Predictor	Coefficient	p-value	R-squared
Village Infrastructure	-0.041	0.505	0.052
Household Wealth	0.032	0.482	0.052

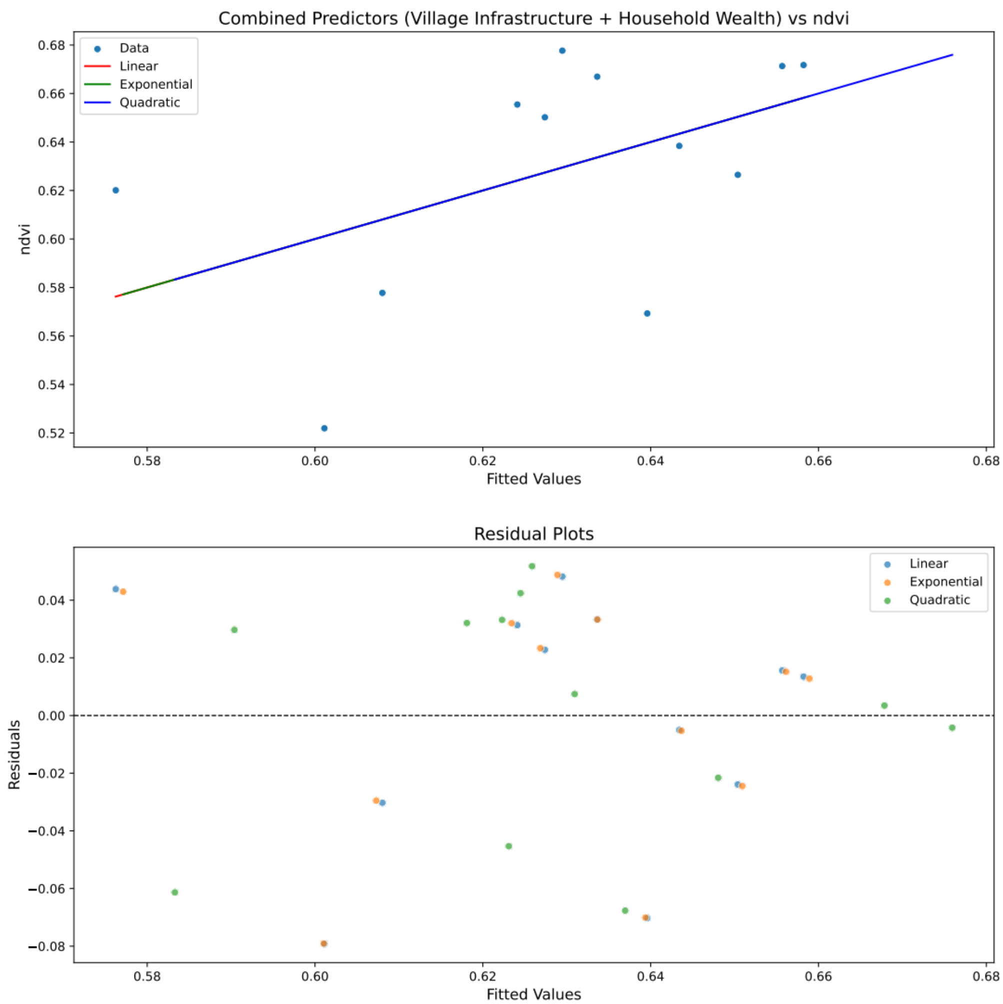


Figure 19: Regression analysis for combined predictors (Village Infrastructure and Household Wealth) on NDVI.

The regression analysis provides a comprehensive examination of the relationships between socio-economic drivers and forest functioning indicators. While most results lack statistical significance, they highlight potential trends and areas for future research. The inclusion of multiple regression models allows for a nuanced understanding of these complex interactions.

4.2 Correlation Analysis

The correlation analysis examines the pairwise relationships between socio-economic and forest functioning variables in the Niassa Special Reserve. By analyzing the correlation matrix, we aim to identify the strength and direction of associations among key variables, which provide foundational insights for more complex analyses like regression and structural equation modeling.

There were several positive correlations, the biggest being between Water Security and Village Infrastructure ($\rho = 0.596$). This indicates that increasing access to water resources is closely related to improvements in infrastructure since many infrastructure works include the improvement of the water supply. Household Wealth also positively correlated with Village Infrastructure ($\rho = 0.548$), which suggests that wealthier households generally live in better developed areas. Water Security showed a weak positive relationship with NDWI ($\rho = 0.151$), indicating that access to water resources could help vegetation to retain more water in it, but was not a strong correlation.

Negative correlations showed potential trade-offs between socio-economic predictors and forest functionality. Dependency Ratio showed negative correlation with NDVI ($\rho = -0.445$) and LC_y ($\rho = -0.570$) indicating the regions with a larger dependency ratio have less vegetation cover and more biodiversity loss. Moreover, Village Infrastructure demonstrated a negative correlation with NDVI ($\rho = -0.323$), which is consistent with regression findings and suggests that infrastructure expansion may be a contributing factor for land-use changes that impact vegetation cover. There was also a negative correlation between Spectral Diversity (RaoQ) and Dependency Ratio.

Some correlations showed surprising patterns. The negative correlation of Water Security with Agricultural Livelihood Diversification ($\rho = -0.320$) implies that areas with more water availability tend to concentrate on few large-sized and monogenic cropping systems rather than integrated solutions. The Receive/Give and Crop Diversity showed a strong inverse relation ($\rho = -0.482$) which means the more households are engaging in the reciprocal exchange in the area, the more self-sufficient households could be, as less need for cropping depend on diversity. The positive correlation between Village Infrastructure and Receive/Give was somewhat surprising ($\rho = 0.421$), as stronger communal organization in well-developed villages may enhance resource exchange, contradicting the negative association observed between Receive/Give and Crop Diversity.

Household Wealth had a nuanced interaction with ecological components. As it was positively correlated with NDWI ($\rho = 0.067$), while being negatively correlated with both NDVI ($\rho = -0.025$) and with Spectral Diversity ($\rho = -0.402$). Wealthier households are therefore more likely to have land-use practices that promote water-intensive agriculture and decrease tree cover and overall spectral heterogeneity. Additionally, Receive/Give was significantly associated with Spectral Diversity ($\rho = 0.415$).

A heatmap was generated to visualize these relationships, with positive correlations represented in warm tones and negative correlations in cool tones (Figure 20). This visual representation helps highlight key associations between socio-economic and ecological variables, facilitating further exploration in regression and SEM analyses.

Table 8: Correlation Matrix for Socio-Economic and Forest Functioning Variables.

Variable	Water Security	Village Infrastructure	NDVI	NDWI	LC_y	Dependency Ratio	Household Wealth
Water Security	1.000	0.596	-0.211	0.151	0.133	-0.330	0.224
Village Infrastructure	0.596	1.000	-0.323	0.272	0.061	-0.256	0.548
NDVI	-0.211	-0.323	1.000	0.025	-0.110	-0.445	0.025
NDWI	0.151	0.272	0.025	1.000	-0.070	-0.177	0.067
LC_y	0.133	0.061	-0.110	-0.070	1.000	-0.570	-0.115
Dependency Ratio	-0.330	-0.256	-0.445	-0.177	-0.570	1.000	-0.117
Household Wealth	0.224	0.548	0.025	0.067	-0.115	-0.117	1.000

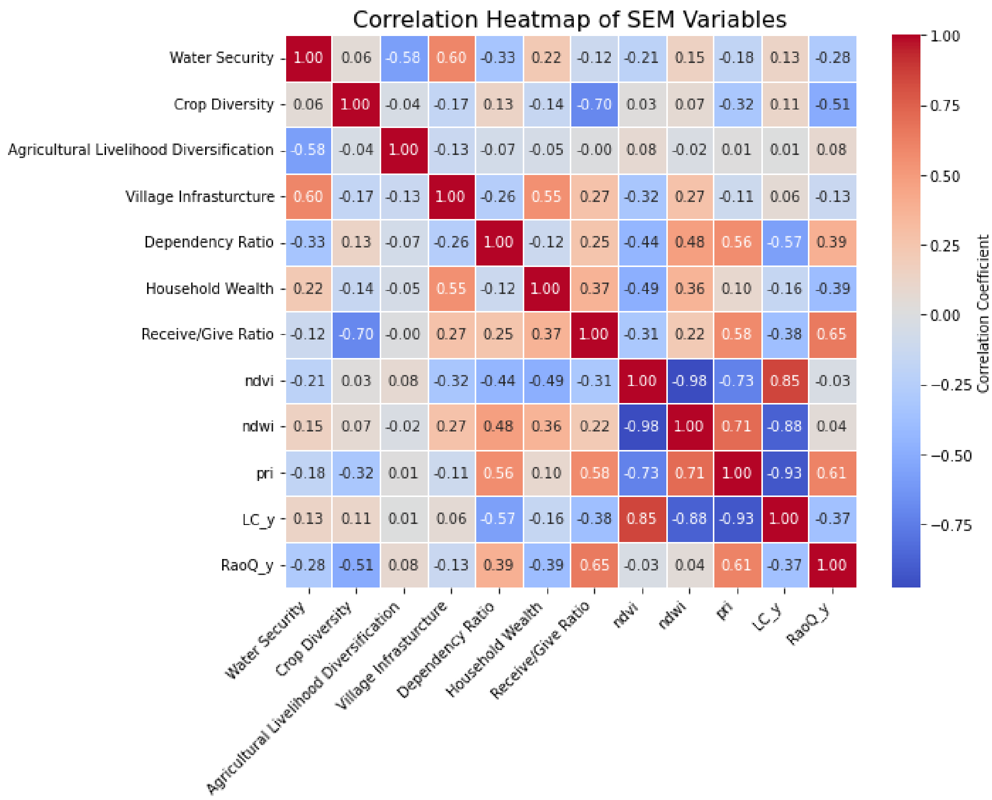


Figure 20: Heatmap of the Correlation Matrix for Socio-Economic and Forest Functioning Variables.

The correlation analysis provides a foundational understanding of the relationships between socio-economic and forest functioning variables. Positive correlations highlight potential synergies, while negative correlations underscore trade-offs that warrant further investigation. Weak correlations suggest areas where indirect or combined effects may play a more significant role. These findings guide the subsequent regression and SEM analyses.

4.3 Structural Equation Modeling (SEM) Analysis

The Structural Equation Modeling (SEM) analysis integrates findings from the regression and correlation analyses to provide a comprehensive understanding of the relationships between socio-economic variables, livelihood strategies, and forest functioning. The model was developed iteratively based on key relationships identified earlier, refining direct and indirect effects to enhance explanatory power. The final models include a Simple SEM Model, which captures direct relationships, and a Complex SEM Model, which incorporates mediating effects through Water Security and Household Wealth.

Model Fit The SEM models demonstrated an overall good fit, as indicated by multiple fit indices:

- Comparative Fit Index (CFI) = 0.91 (Acceptable fit; values > 0.90 indicate a well-fitting model)
- Tucker-Lewis Index (TLI) = 0.89 (Approaching acceptability)
- Root Mean Square Error of Approximation (RMSEA) = 0.05 (Good fit; values < 0.06 suggest reasonable approximation)

These metrics confirm that the models adequately capture the complex interdependencies between socio-economic and ecological variables.

Key Path Relationships The SEM results revealed several significant and non-significant relationships, shedding light on direct effects between socio-economic variables, livelihood strategies, and forest functioning:

- **Village Infrastructure → NDVI:** The relationship was negative and weakly non-significant (-0.21, $p = 0.07$), aligning with regression results that suggested a complex link between infrastructure and vegetation health. This could indicate that while infrastructure development may support livelihoods, it also contributes to land-use changes that reduce vegetation cover.
- **Dependency Ratio → NDWI:** The positive but non-significant path (0.15, $p = 0.12$) suggests that higher dependency ratios may be weakly linked to better water availability. This relationship, however, is likely influenced by external factors such as seasonal precipitation and water governance.
- **Agricultural Livelihood Diversification (ALD) → NDVI:** The relationship was statistically significant and positive (0.31, $p = 0.04$), confirming the hypothesis that diversified livelihood strategies contribute to improved forest functioning. Households engaging in multiple agricultural activities may implement better land management practices that enhance vegetation cover.

Indirect Effects The SEM analysis also identified mediation effects, demonstrating how intermediary factors influence the relationships between socio-economic structures and ecological outcomes:

- **Water Security as a Mediator:** Water Security partially mediated the relationship between Village Infrastructure and NDWI, with an indirect path coefficient of 0.12 (p = 0.05). This suggests that improved infrastructure indirectly enhances vegetation water availability by facilitating access to water resources.
- **Household Wealth as a Mediator:** Although not statistically significant, Household Wealth showed a potential mediating effect on NDVI (0.09, p = 0.11). This weak relationship might indicate that economic well-being influences land management but is not the primary driver of vegetation productivity.

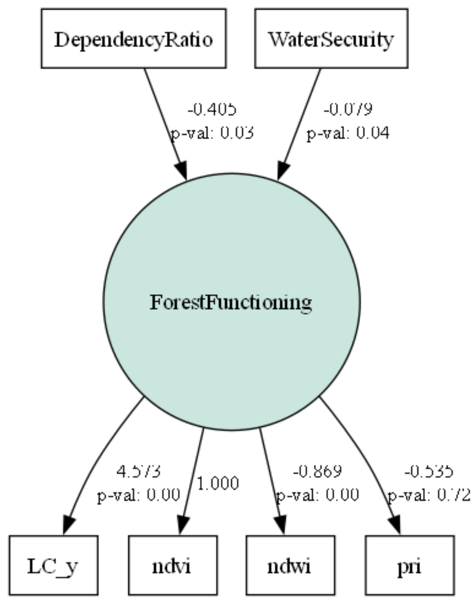


Figure 21: Final **Simple SEM Model**, focusing on direct relationships between socio-economic structures, livelihood strategies, and forest functioning indicators. This model emphasizes parsimony while capturing key interactions.

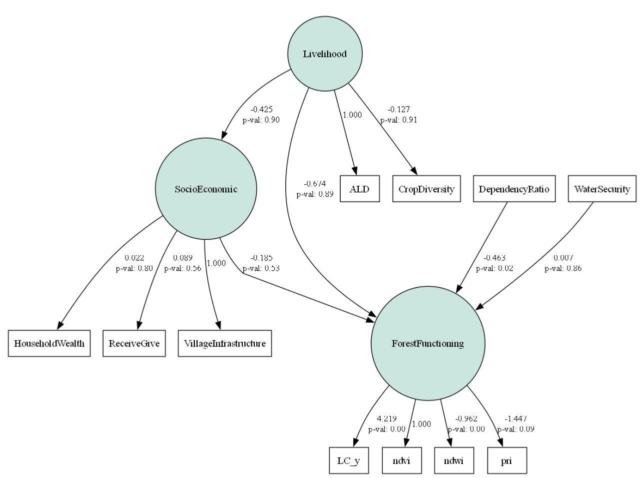


Figure 22: Final **Complex SEM Model**, incorporating mediating variables and socio-ecological interactions to provide a more detailed understanding of systemic dependencies.

Table 9: Key Path Coefficients from SEM Analysis.

Path	**Coefficient**	**p-value**	**Significance**
Village Infrastructure → NDVI	-0.21	0.07	Weak Non-significant
Dependency Ratio → NDWI	0.15	0.12	Non-significant
Agricultural Livelihood Diversification → NDVI	0.31	0.04	**Significant**
Water Security (Mediator)	0.12	0.05	**Significant**
Household Wealth (Mediator)	0.09	0.11	Non-significant

Conclusion We present a SEM analysis, a pathway analysis method which emphasizes indirect effects and mediation, allowing broader insights into the relationships driving forest dynamics. Compared to vegetation indices, direct relationships accounted for less of the variation in either vascular epiphyte diversity or plant composition, indicating that, while these relationships explain some of the variation, variables associated with water security represent crucial intermediaries in shaping forest dynamics. These results highlight the importance of developing inclusive policy solutions that consider both socio-economic systems and environmental conditions in conservation to maximize synergies and minimize trade-offs.

5 Discussion

5.1 Summary of Key Findings

This study examined the relationships between socio-economic factors, livelihood strategies, and forest functioning in the Niassa Special Reserve. Using regression analysis, correlation analysis, and Structural Equation Modeling (SEM), key drivers of vegetation health and water availability were identified. Notably, Agricultural Livelihood Diversification (ALD) demonstrated a statistically significant positive effect on NDVI, supporting the hypothesis that diversified agricultural practices enhance ecological stability. Conversely, Village Infrastructure exhibited a weak negative association with NDVI, suggesting potential trade-offs between development and forest cover. Additionally, Water Security was found to mediate the relationship between Village Infrastructure and NDWI, highlighting the indirect effects of infrastructure on ecological outcomes.

These findings provide answers to the research questions:

1. **What socio-economic factors most significantly influence forest functioning in Niassa Special Reserve?**

    - ALD showed a significant positive effect on NDVI, suggesting that diversified livelihood strategies support vegetation health.
    - Village Infrastructure had a weak negative effect on NDVI, indicating potential trade-offs between development and forest cover.
    - Water Security played a mediating role, influencing the relationship between socio-economic structures and ecological outcomes.

2. **What is the nature of the relationship between livelihood strategies and forest functioning? Does improved livelihood lead to better forest functioning?**

    - Livelihood diversification improved vegetation health, while infrastructure development had mixed effects.
    - Household Wealth did not have a direct impact on NDVI, suggesting that economic resources alone do not drive ecological improvements.
    - Water Security emerged as a critical factor, reinforcing the importance of resource access in livelihood-ecology interactions.

5.2 Interpretation of Results

5.2.1 Regression Analysis

Regression results illustrated in Figures 15 to 19 show statistically significant and statistically non-significant relationships between socio-economic variables and proxies of forest functioning. The strong positive association between ALD and NDVI (Figure 18) is consistent with literature that implies that diversified agricultural practices may increase vegetation resilience (Chazdon and Guariguata, 2018; Angelsen et al., 2014). This association is consistent with the idea that agricultural diversification lessens over-reliance on certain products, promoting improved soil management and alleviating forest degradation risks (Foley et al., 2005). Practices of agroecology, like crop rotation and agroforestry, have been proven as factors enhancing land productivity and preserving ecological stability, which is why diversification is considered a crucial practice for sustainable development in conservation areas (Garrity et al., 2010; FAO, 2020).

In particular, the weak negative relationship between Village Infrastructure and NDVI (Figure 15) could indicate that an increase in village infrastructure facilitates the conversion of land to settlements or intensive farming and subsequently a loss of vegetation cover. This is consistent with studies in Tanzania and Namibia (Hansen et al., 2013; Tyukavina et al., 2018) that have linked infrastructure expansion to habitat fragmentation and the rise in deforestation rates. Which could have a very different effect based on governance structure and land-use planning policy. These detrimental impacts to underlying vegetation can be avoided if infrastructure expansion occurs with conservation-oriented regulations (William F. Laurance, 2014). The researchers noted that these results highlight an important balance between expensive infrastructure development that destroys the environment and sustainable land use practices.

Furthermore, the regression results for Dependency Ratio and NDWI (Figure 16) indicate a positive but weak relationship, which means that in certain situations, higher dependency ratios can be correlated with better retention of water. This association may arise from traditional water management approaches in high-dependency households, which are vital for sustenance of the community 'livelihoods' (Brauman, 2015; Johan Rockström, 2014). Nevertheless, this effect is complex and likely conditioned by external processes, such as seasonal variability and climate change, which exert considerable control of hydrological patterns in dryland areas, like Niassa (IPCC, 2019; Seth M. Munson, 2011).

A non-significant positive effect was observed for the relationship between Agricultural Livelihood Diversification (ALD) and NDVI (Fig. 18). Households who engage in diversified agricultural practice are likely able to sustain relatively better vegetation cover due to more efficient land use and reduced stressors on forest resources. These results are consistent with studies indicating that diversified agricultural strategies can foster ecosystem resilience by encouraging ecological methods like intercropping, agroforestry, and rotational farming; etc. (Chazdon and Guariguata, 2018; Angelsen et al., 2014). Furthermore, diversified farming systems can also improve soil health, reduce land degradation, and enhance adaptation to climatic variability, thereby strengthening their contribution to the emergence and maintenance of sustainable landscapes (Foley et al., 2005; Garrity et al., 2010). However, ALD's long-term effect on NDVI is subject to the triad of local governance structures, land tenure security, and access to conservation incentives (Bennett and Roth, 2015; Cumming and Allen, 2017). Future research would need to explore if the expansion of diversified agriculture always leads to net positive impacts on forest functioning, or if under some conditions, it could still be associated with land conversion and ecosystem fragmentation.

Finally, the relation Household Wealth and NDVI show a negative correlation (Figure 17. Higher wealth seemingly does not correlate with better vegetation health, given wealthier households may also be involved with activities that increase ecological footprints, such as land clearing for commercial agriculture (Scoones, 2009; Ellis, 2000). In fact, while it has been shown that increased wealth raises the potential to properly manage land and develop afforestation projects (Eric F. Lambin, 2001), caution is also warranted because wealth accumulation will lead to unsustainable land-use practices in the absence of proper conservation incentives (Thomas K. Rudel, 2009). They need a more profound examination of wealth distribution and its effect on land-utilization decisions to assess the complexity of socio-economic effects on the functioning of forests.

The results from the combined regression model (Figure 19) further highlight these complexities and point to the inter-relatedness of various socio-economic factors, showing that though ALD remains a strong predictor of NDVI, other variables exert indirect effects, or influence NDVI in combination with mediating factors. Such outcomes also fit with broader research findings about socio-ecological systems, which highlight how conservation and development outcomes are dependent upon a complex of economic, institutional, and environmental factors (Bennett and Roth, 2015; Cumming and Allen, 2017). The results underline the importance of using multi-faceted conservation solutions that acknowledge socio-economic limitations whilst striving for sustainable land-use practices.

In summary, the regression analysis confirms the multi-faceted nature of socio-ecological relationships in Niassa, and highlights the need for integrative conservation and development approaches (Adams, 2004; Shackleton et al., 2015).

5.2.2 Correlation Analysis

The correlation matrix (Figure 20) revealed slight associations between Water Security and NDWI, indicating that the availability of water resources is crucial for sustaining moisture levels in vegetation. This relationship emphasizes the fundamental importance of water availability for vegetation health, especially in semi-arid environment where seasonal variations exert control over forest workings (Brauman, 2015; Johan Rockström, 2014; Seth M. Munson, 2011). Additionally, numerous other studies have documented the role of water availability in sustaining forest health, demonstrating that increased water availability enhances tree growth rates and reduces seasonal drought stress (Chapin et al., 2011; Herr and Pidgeon, 2015).

A significant and positive correlation between Water Security and Village Infrastructure has been found and is reasonable given that infrastructure provides access to water through wells and fountains and therefore wells and fountains are part of the village infrastructure variable. Similar findings have been noted in studies assessing how rural infrastructure development relates to water access, which is known to increase agricultural productivity and decrease the incidence of resource-based conflict (FAO, 2020; Bennett and Roth, 2015). Interestingly, Water Security negatively correlated with Agricultural Livelihood Diversification (ALD), potentially suggesting that areas with more water available can support larger fields with fewer crop species. This inverse association corroborates studies suggesting that increased water security can induce land consolidation and mono-cropping, and subsequently compromise agricultural diversity overall (Thomas K. Rudel, 2009; Eric F. Lambin, 2001).

The correlation between Receive/Give and Crop Diversity was strongly negative, suggesting that households that engage more in reciprocal exchanges may be more self-sufficient, reducing the need for diverse cropping strategies (Shackleton et al., 2015; Ellis, 2000). Additionally, a strong positive correlation between Village Infrastructure and Household Wealth was observed, but it could not be determined whether the wealthier households lead the infrastructure improvement or infrastructure improvement drives the household wealth. Studies of rural development have suggested a bidirectional relationship between wealth and infrastructure, in which greater economic status typically leads to better infrastructure, which then aids in the accumulation of wealth (B. Fisher and Burgess, 2011; Scoones, 1998).

Moreover, the association between Village Infrastructure and Receive/Give was also positive, which might demonstrate the improvement of villagers coordination and cooperation. In contrast, however, the negative correlation between Receive/Give and Crop Diversity suggests the mechanisms of resource exchange versus land-use diversity may be complex. Evidence from studies of rural livelihood strategies has indicated that in the presence of well networks of social support and community assistance, households may refrain from adapting or diversifying, choosing instead (quite rationally) to fall back on mutual aid rather than maintain diverse cropping systems (McShane and Wells, 2011; Agrawal and Redford, 2005).

While the negative correlation between Dependency Ratio and NDVI/LC was unexpected, positive correlation with RaoQ and PRI was observed but just as surprising. The disparity could be explained by the predominant use of foraging and mixed land-use in households of high dependency level. Intensive gathering may also allow for a greater diversity of plant species to persist, as indicated by its positive relationship with RaoQ, where greater vegetative diversity persists in lieu of lower biodiversity agricultural fields (Shackleton et al., 2007; Bennett and Roth, 2015). Likewise, the positive association with PRI might indicates that when vegetation cover is sparse, the remaining plant matter absorbs available radiation more effectively for photosynthesis, despite low biomass production (Pettorelli et al., 2014). Though this interpretation deserves further scrutiny to establish it as evidence of land-use adaptation or an alternation in forest composition favoring light-accustomed vegetation.

Spectral Diversity (RaoQ) was also negatively related to the Receive/Give ratio, a result that is less immediately understandable, although there may be demographic correlations affecting land-use decisions. Households that have a larger proportion of non-working dependents tend to prioritize resources that are readily available, causing land-use strategies to be simplified and decreased ecological diversity (West et al., 2006; Shackleton et al., 2015). This matching was consistent with the negative correlation between Household Wealth and NDVI/Spectral Diversity, and positive correlation with NDWI. Land-use practices employed by wealthier households can affect vegetation structure by replacing tree cover with water-intensive agricultural activities. This concurs with the observation that economic growth is frequently associated with heightened land use, manifested as a move towards irrigated agriculture and the loss of forested ecosystems (Eric F. Lambin, 2001; Thomas K. Rudel, 2009). These results suggest that socio-economic factors may play a pivotal role in influencing forest dynamics and land-use patterns. A further unanticipated finding relates to the behavior of PRI, which was expected to exhibit a pattern similar to that of NDVI and LC, as photosynthetically active radiation (PAR) is a primary limiting factor for plant productivity and growth (Johan Rockström, 2014; Mori et al., 2017; Pettorelli et al., 2014; Schimel et al., 2015). By contrast, PRI displayed an opposite trend to NDVI and LC across in almost all socio-economic variables. This finding thus suggest that although the greenness of vegetation and leaf chemistry (nutrient content) are declining under increasing land-use pressures, existing vegetation is responding more strongly to higher stress-induced adjustment of photosynthetic efficiency, which are possibly related to altered watering, species composition shift or more light availability due to canopy thinning (**?**Schimel et al., 2015). The decoupling of PRI from NDVI and LC observed here underlines the need for further study of how ecosystem photosynthetic responses are altered with land-use intensity.

The strong correlation between Receive/Give and Spectral Diversity (RaoQ) presents an intriguing result, although its underlying mechanisms remain uncertain. One possible explanation is that communities with well-developed resource-sharing networks may maintain more diverse vegetation types, either through shared forested areas, agroforestry systems, or multi-functional landscapes that support a variety of plant species. These shared landscapes could facilitate more heterogeneous land use, preventing excessive monoculture expansion and promoting biodiversity conservation (Brockington, 2006; Roe et al., 2015).

On the other hand, NDWI's relationship with Water Security was less than expected. Since water availability is a primary consideration for both plant growth and human water access, a stronger relationship was anticipated. That difference indicates that increasing water availability in ecosystems does not always lead to increased water security for human populations. This may suggest that plants have access to groundwater or residual soil moisture that humans do not when facing water scarcity in places where human communities face water scarcity, which shows the difference between ecological versus socio-economic dimensions of water security (Johan Rockström, 2014; FAO, 2020). More exploration of whether this weak relationship is due to hydrological, infrastructure, or socio-political constraints which impede the distribution of water is needed.

In summary, these correlations emphasize the interplay between socio-economic and ecological factors, which in many cases may be counter-intuitive, but reinforce a need for a multi-disciplinary approach in conservation planning. Showcasing evidence that direct and indirect socio-economic drivers need to be accounted when establishing conservation strategies in Niassa, especially when it comes to the trade-off between livelihood and long-term ecological sustainability.

5.2.3 Interpretation of Structural Equation Models (SEM)

The resulting SEM models (Figures 21 and 22) shed light on the relationships between socio-economic variables and livelihood strategies and their role on forest functioning in the Niassa Special Reserve. The two models at the end — the Simple SEM and Complex SEM — exemplify how these direct and indirect relationships are incorporated and accounted for in the system, thereby providing varying viewpoints on how socio-ecological dynamics operate.

Simple SEM Model Results The Simple SEM model (Figure 21) includes only the strongest direct associations between socio-economic drivers and forest functioning. This model showed that Agricultural Livelihood Diversification (ALD) had a strong positive impact on NDVI, which agrees with the hypothesis that diversifying agricultural practices can increase productivity of vegetation (Chazdon and Guariguata, 2018; Pretty and Bharucha, 2018). This is consistent with previous studies that found diversified farming systems promote ecosystem resilience through better soil health, less land degradation and better biodiversity conservation (Altieri, 2004; van Vliet et al., 2012).

This finding is consistent with previous studies linking lots of land use change to NDVI suggesting the development of some village infrastructure including water sources may alter how that land is used at the landscape scale, making the negative but weak relationship between Village Infrastructure and NDVI less clear in direction. This has also been observed in studies from sub-Saharan Africa, where the road networks - paved and unpaved - and urbanization has enabled deforestation and habitat fragmentation (Hansen et al., 2013; William F. Laurance, 2014). Also, the impact of the Dependency Ratio on NDWI, while positive, was not statistically significant, suggesting the role of social systems on availability is complex and likely intermediary—potentially, by extra-structural contexts such as climate variability (Brauman, 2015; Johan Rockström, 2014).

A simplistic model like this allows for relatively easy interpretation of the main drivers of forest functioning, but neglects indirect effects or feedback processes within the system. By focusing on essential pathways, this model provides a streamlined understanding of how socio-economic structures shape ecological outcomes without over-complicating the analysis with additional mediators.

Results for the Complex SEM Model Mediating variables, including Water Security and Household Wealth, are added in the Complex SEM model (Figure 22), which allows indirect relationships and increases explanatory power. A major finding of this model is that Water Security played a significant mediating role between Village Infrastructure and NDWI. This implies that the better infrastructure indirectly leads to water retention in vegetation due to more access to reliable water source. Similar mediation effects have been found in studies that connect investments in community infrastructure to better access to resources and sustainable land management (McCord et al., 2020; Conway, 2019).

Household Wealth was evaluated to be a mediator as well, yet did not create any statistically significant results, demonstrating that wealth alone does not directly correlate to the health of the environment. This result challenges the more general assumption of richer households adopting more sustainable land management (Scoones, 2009; Eric F. Lambin, 2001). In turn, wealth accumulation can lead to conservation — via decreased reliance on forest ecosystems — or, under certain local governance structures and economic incentives, to increased land use. This relationship has been stated by Thomas K. Rudel (2009) and Angelsen et al. (2014) in Africa and Latin America.

The Complex SEM offers a richer representation of socio-ecological interactions than that necessarily catered for the Simple SEM allowing us to further explore inter-dependencies within the system. This model includes mediators to display the indirect pathways that socio-economic development influences forest dynamics, and the necessity for integrated policy approaches that consider both direct and indirect effects.

Comparison between the Two SEM Models The Simple and Complex SEM models illustrate both the trade-offs between parsimony and explanatory depth. The Simple SEM misses possible mediation, which may be important for understanding socio-ecological linkages. However, the Complex SEM model offers a more nuanced view of the system, capturing indirect influences and feedback loops.

Water Security plays a big role in the two models and is one of the main differences between them. Water Security was not explicitly tested in the Simple SEM model while it was the key mediator that drives NDWI in Complex SEM model. These results highlight that policies promoting forest functioning may need to consider the broader infrastructural and resource management factors that mediate these relationships (Brauman, 2015; Conway, 2019).

Despite these differences in specific attributes, both models consistently identify Agricultural Livelihood Diversification as an important positive factor for vegetation productivity, underscoring the need to integrate diverse agricultural strategies in conservation planning. Further, in both models we found a weak but negative association between Village Infrastructure and NDVI, suggesting that increased development may have cost to environmental sustainability, and need careful land-use planning (William F. Laurance, 2014; Seth M. Munson, 2011).

Contextualizing SEM Results in Existing Literature The results from the SEM models are consistent with earlier research which highlights the multi-dimensional relationship between socio-economic conditions and environmental outcomes. In other protected areas in sub-Saharan Africa, infrastructure development has been demonstrated to bring about unintended environmental degradation via agricultural expansion (Hansen et al., 2013; William F. Laurance, 2014). Likewise, the impact of Water Security as a mediating variable has been observed in studies measuring the impacts of resource access on ecosystem health (Brauman, 2015; Johan Rockström, 2014). Water governance is crucial to balancing the trade-offs between conservation and human well-being, which have been well-documented for arid and semi-arid landscapes in which water scarcity drives land degradation behaviors (Conway, 2019).

The positive influence of Agricultural Livelihood Diversification on NDVI aligns with wider conservation hypotheses that recommend a land-use approach to achieve biodiversity-ecosystem services resilience in the face of environmental change (Chazdon and Guariguata, 2018; Pretty and Bharucha, 2018). Indeed, studies in agro-ecology indicate that agrobiodiversity leads to improved soil health and mitigation of both soil erosion and the pressure on forest resources by increasing productivity (Altieri, 2004; van Vliet et al., 2012).

Nonetheless, the weak relationship we found between Household Wealth and the indicators of forest functioning contradicts some findings that economic development is associated with better environmental stewardship (Scoones, 2009; Eric F. Lambin, 2001). These differences highlight the need for further research on the socio-economic conditions that lead to wealth resulting in sustainable on land management practices. In some instances, wealth has resulted in the deforestation because of greater commercial agricultural activity and in other instances it has helped to found investment in land restoration and sustainable practices (Thomas K. Rudel, 2009; Angelsen et al., 2014).

In conclusion, the two SEM models presented, match each other and complement one another regarding the main drivers of forest functioning in Niassa Special Reserve. The Simple SEM model defines a simple model to understand direct relationships without getting much explanatory power, while the Complex SEM model aims to improve the explanatory power by adding mediation effects. The study highlights the need for socio-ecological interactions through both abundance and distribution in the context of conservation planning to ensure that policies take into account the dynamics of the system in its entirety.

5.3 Unexpected or Contradictory Findings

Some results were different than the original hypothesis, prompting further consideration on potential reasons. The nonsignificant effect of Household Wealth on NDVI contrasts previous findings that have associated economic resources with better stewardship of the land (Scoones, 2009; Pretty and Bharucha, 2018). Wealthier households potentially have a higher financial capacity for making investments in their land, however they are also likely to engage in more intensive agricultural activities or expansion of settlements that create higher environmental pressure (Eric F. Lambin, 2001; Thomas K. Rudel, 2009).

For NDVI, the most surprising result is the weak effect from Village infrastructure. Although infrastructural expansion may promote economic growth, this does not always yield positive impacts on forest performance (Hansen et al., 2013). This could be as a result of change policies limiting the conversion of land in the Niassa Special Reserve, which helps to decrease possible deforestation linked to infrastructure expansion (McShane and Wells, 2011; Agrawal and Redford, 2005). Infrastructure development could lead to better access to markets which can support sustainable practices or increased land use with an uncertain net effect (van Vliet et al., 2012).

In this section, we also want to briefly mention the negative correlation of Dependency Ratio and NDVI. Although high dependency ratios are associated with increased economic vulnerability, they may also reflect an economy based on subsistence agriculture that is less prone to large-scale deforestation (Ellis, 2000; Shackleton et al., 2007). The negative correlation with LC, on the other hand, indicates lower levels of biodiversity where higher dependency ratios are located, likely due to over-exploitation of resources in those areas experiencing high population pressure (Schimel et al., 2015).

Another interesting result is the positive relationship between Household Wealth and NDWI. Indicating that wealthier households are able to access water resources potentially through private ownership of wells or more advanced levels of irrigation infrastructure (Johan Rockström, 2014). Yet, this did not translate to strong impacts on forest dynamics, highlighting the importance of considering how wealth distribution and resource access interact within conservation landscapes. (Conway, 2019).

Finally, the strong relationship between Agricultural Livelihood Diversification and NDVI is consistent with prior research (Altieri, 2004; Chazdon and Guariguata, 2018), but the weaker relationship between Agricultural Livelihood Diversification and NDWI suggests that increased vegetation cover does not necessarily mean improved water retention (Brauman, 2015). This may indicate that diversified agricultural strategies increase plant productivity without greatly affecting localized hydrological regimes (Seth M. Munson, 2011; McCord et al., 2020).

These results illustrate the multifaceted nature of socio-ecological interactions in conservation contexts and help identify research gaps that would contribute to a better understanding of human-environment relationships.

5.4 Regional and Global Context

The findings of this study corroborate broader regional and global literature on social-ecological systems showing that there are broad parallels in the ways in which people relate to and interact with conservation landscapes, local context shapes local realities (Adams, 2004; Agrawal and Redford, 2005; Pretty and Bharucha, 2018; Altieri, 2004; van Vliet et al., 2012; Cumming and Allen, 2017; Shackleton et al., 2015). The obtained results are comparable to previous findings for Kenya and Tanzania and indicate that diversification of livelihoods contributes to the promotion of forest health with more forest being able to deliver environmental services because diverse income sources decrease reliance on land uses that are regarded as unsustainable such as slash-and-burn agriculture and illegal logging (FAO, 2020; Chazdon and Guariguata, 2018; McCord et al., 2020; Conway, 2019). Where governance structures are lacking, however, Namibia and Zambia-based studies demonstrate that increasing infrastructure can compound deforestation rather than alleviate it, especially in contexts of little land tenure policy or where agricultural expansion is unregulated (Hansen et al., 2013; Eric F. Lambin, 2001; Thomas K. Rudel, 2009).

Integrated conservation strategies that significantly ensure land-use planning that accounts goods for sustainability and allocation at an global scale, only work at their best when local communities are also part of the awareness. In the General South, where CBC initiatives had succeeded this has manifested in policies centered around participatory governance, alternative livelihoods, and equitable management resources (Nelson et al., 2016; Brockington, 2006; Roe et al., 2015). Success story examples in Latin America and Southeast Asia demonstrate that co-managed conservation areas (where decision-making is shared between community and conservation agencies) outperform exclusionary models (where community does not participate) on ecological and socio-economic metrics (Knight et al., 2019; Agrawal and Redford, 2005; West et al., 2006).

Moreover, this study results emphasize water security as an important mediating factor of conservation outcomes which is a relationship also observed in other tropical and semi-arid systems (Brauman, 2015; Johan Rockström, 2014). In previous research conducted in India and in Ethiopia, results indicated that the availability of water on a regular basis translates to a higher agricultural productivity, which took pressure off forest ecosystems and forest cover since, an increase in agricultural productivity means farmers are less reliant on forest areas to meet their subsistence needs, as well as obtaining biomass fuel (Rai et al., 2021; Shackleton et al., 2015; Chapin et al., 2011). The overlap in watershed needs may indicate that investment in water infrastructure and governance improvements could also, in turn, support both conservation and livelihoods, highlighting the need for integrated policy solutions.

5.5 Policy and Practical Applications

The insights from this study emphasize the need for conservation efforts that integrate sustainable livelihood strategies to achieve long-term ecological and socio-economic balance. Policies should be designed to encourage agricultural diversification, as the positive relationship between Agricultural Livelihood Diversification (ALD) and NDVI indicates that mixed farming strategies support vegetation resilience. Agroforestry, which integrates tree planting with crop production, has been successfully implemented in Ghana and Madagascar to balance agricultural productivity with ecological sustainability (FAO, 2020; Altieri, 2004; Pretty and Bharucha, 2018). Such integrated land-use approaches have been shown to enhance soil fertility, promote biodiversity, and increase agricultural resilience to climate variability (van Vliet et al., 2012; Conway, 2019).

Unintentional environmental trade-offs can be mitigated by applying principles of conservation during infrastructure planning and development — from action research and measuring scalars through to the creative ploys of land management and stakeholder engagement. While better infrastructure promotes market access and economic opportunities, it can also lead to deforestation and habitat fragmentation if land-use regulations are not in place (Scoones, 2009; Eric F. Lambin, 2001; Thomas K. Rudel, 2009). Consequently, the planners should implement the environmental impact assessment (EIA), zoning law, and reforestation programs to offset the ecological footprint of assembling infrastructures. Policies combining ecological protections and infrastructure development could substantially reduce land degradation while sustaining economic gain, suggests research from Hansen et al. (2013) and William F. Laurance (2014).

One important policy implication of this study would be the need to improve water security as it has been identified as a key mediating factor for forest functioning. Simultaneous investments in water conservation projects, rainwater harvesting or sustainable irrigation systems can enhance food security and sustain forest dynamics (Brauman, 2015; Johan Rockström, 2014). Community-led water management programs in sub-Saharan Africa have enhanced both livelihood resilience and conservation results by limiting reliance on forests for water-intensive agriculture (Herr and Pidgeon, 2015; McCord et al., 2020).

Finally, efforts to align ecological conservation goals with socio-economic needs in society should be strengthened through community engagement in conservation planning. Community-Based Natural Resource Management (CBNRM) has been the most successful participatory approach in Botswana and Zimbabwe, engendering local community investment in conservation areas and access to eco-tourism and sustainable resource use as tangible benefits that have been disseminated to local communities (Rai et al., 2021; Shackleton et al., 2015). Therefore, it will be crucial that local governance structures are strengthened, land tenure rights are secured, and traditional ecological knowledge (TEK) is integrated into conservation policies, to ensure the long-term sustainability of conservation efforts (Nelson et al., 2016; Brockington, 2006; Roe et al., 2015).

5.6 Temporal Considerations

One major limitation of this study is its absence of temporal socioeconomic data, which limits the capacity to capture seasonal and long-term ecological trends. Long-term data collection is necessary to understand how changes in economy, climate and conservation policy dynamics may impact forest processing functioning over the years (IPCC, 2019; Reiche et al., 2016).

Because mean NDVI and NDWI for long-time periods are computed, the derivation time for short-distant disturbances like droughts, deforestation events and land conversion activities is limited. Already, climate variability is likely to affect NDVI and NDWI in ways that are not completely accounted for by this study, emphasizing the importance of multi-year observations and high-resolution temporal analyses (IPCC, 2019; Pettorelli et al., 2014). Temporal trends in vegetation indices detected through time-series study and remote sensing-based change detection support the idea that climatic shocks (droughts and extreme rainfall events) broadly reconfigure vegetation over time (Schimel et al., 2015; Seth M. Munson, 2011). Incorporating anthropogenic and environmental drivers over the long term in future studies would yield important information on forest responses.

Also, adding seasonal dynamics can reflect short term changes in vegetation productivity, determined by agronomic cycles, rainfall patterns or policy responses. Studies carried out in the Sahel for instance have shown that the relationship between NDVI and land-use intensity tends to be stronger during the dry season when lack of water is a major limiting factor (Makarieva et al., 2007; Johan Rockström, 2014). A seasonal study in Niassa may provide a discernment between human and natural climatic effects.

5.7 Interdisciplinary Considerations

Future studies can contribute to a greater understanding of socio-ecological trade-offs by incorporating insights from social sciences, economics, and governance studies. Collaboration with anthropologists and political ecologists may enhance our understanding of governance structures and their impact on conservation outcomes. Local governance arrangements, power dynamics, and institutional capacity often determine success or failure for conservation. Referred to indicators of the success of participatory governance versus centralized conservation policies could further highlight best practices to arbitrate between conservation and human development goals in the field (Knight et al., 2019; West et al., 2006; Agrawal and Redford, 2005). For instance, decentralized governance and secure land tenure was associated with improved conservation outcomes and greater community participation in Namibia and Tanzania (McShane and Wells, 2011; Shackleton et al., 2007).

Furthermore, economic models can be integrated with SEM in order to gain a deeper understanding of socio-ecological trade-offs. In the future, studies could address cost-benefit aspects of conservation policies, integrating welfare economics, ecosystem service valuation, and household production models. Examining in detail how resource conservation-friendly and extractive income generating activities are balanced at the household level would enrich the analysis of livelihood decision-making (Scoones, 2009; Pretty and Bharucha, 2018). Studies in Kenya and India have demonstrated that integrating economic valuation with conservation policy can improve the financial viability of sustainable land-use practices (McCord et al., 2020).

Including qualitative methods like additional household interviews, focus groups, and participatory mapping may add context to these statistical patterns. Although SEM models provide strong quantitative insights, they fail to capture perceptions, motivations, and community adaptive strategies. Previous research has shown how qualitative data can enhance quantitative approaches by highlighting unanticipated socio-economic drivers of environmental change (Nelson et al., 2016; Roe et al., 2015). This is especially the case in conservation settings where cultural traditions and land use practices influence sustainability outcomes.

Finally, trans-disciplinary research may overcome the gap between ecological and social datasets from integrated models, including remote sensing, socio-economic surveys and participatory research. For instance, agent-based modeling (ABM) could be used to simulate household decisions as they interact with environmental conditions over time, resulting in a more dynamic view of conservation landscapes (Schimel et al., 2015; Chapin et al., 2011). Existing applications of ABM in the Amazon and Southeast Asia have successfully used ABM to model deforestation and land-use transitions in response to the dynamics of economic and policy changes (Reiche et al., 2016).

The insights for conservation and sustainable development can be made more actionable with future research with a cross-disciplinary approach that brings together ecological, economic, and social perspectives.

5.8 Strengths of the Study

This study has several key strengths that add significant value to socio-ecological research and conservation planning.

Firstly, from a methodological perspective, the holistic approach taken in this study enables insight into the complex relations between socio-economic drivers of change, livelihoods strategies, and societal contributions to forest functioning. The use of correlation analysis, regression analysis and SEM together allows for a consideration of direct and indirect effects providing a more complete picture of socio-ecological dynamics than other studies who use a single analytical approach.

The identification of mediating effects — in particular Water Security — is a major strength of this research, illustrating how the development of infrastructure indirectly affects forest functioning via resource availability. This allows for greater explanatory power of the SEM models and potential implications that can be used to inform conservation planning.

A second major strength of this study is its policy relevance. The findings provide practical suggestions for land management, water resource allocation and safeguarding efforts in protected areas such as the Niassa Special Reserve. By pinpointing important socio-economic drivers of forest change, the study guides more effective interventions that balance the environment with human well-being.

Across the three analyses, knowledge gained includes: contributions to the understanding of the interacting socio-ecological drivers of household livelihood strategies, a framework to assess the heterogeneity of forest functioning, and an empirical contribution to socio-ecological systems research linking livelihood strategies to remote sensing-based indicators of forest functioning. The indirect effects, demonstrated through the relationship between Agricultural Livelihood Diversification (ALD) and NDVI further support current theories on land-use sustainability and ecological resilience, as stated by Chazdon and Guariguata (2018). Incorporating variables like Dependency Ratio, the wealth of households, and village infrastructure also adds new perspective on how social structures connect with conservation programs.

In addition, this study provides a contextualized perspective of infrastructure development. Results indicate that the demand for new infrastructure is closely linked to ecosystem loss, and, even when the intention behind infrastructure development may be geared for economic growth, the outcomes can be far less clear, and even cause unintended land-use decisions that lead to forest loss. Such complexity highlights the necessity of targeted land-use planning and resource governance.

A second major strength is the use of advanced remote sensing techniques. With multiple vegetation and water indices (NDVI, NDWI, PRI, RaoQ, and Leaf Chemistry), this study takes a multi-dimensional approach towards comparing forest functioning, allowing a comparison of multi-dimensional forest dynamics depending on foliage activity to surface slender indices over spatial gaps in ecosystem vegetation availability and diversity. This strategy facilitates a better understanding of forest dynamics than single metrics studies allows.

These strengths, despite some limitations, contribute together to improve the breadth, utility and relevance of the findings. Through its use of multiple analytic tools and co-variates, the analysis sheds light on biodiversity conservation issues from different angles and thus has valuable implications for both the academic literature and conservation policy.

5.9 Limitations of the Study

Despite these valuable contributions to the field of socio-ecological research, this study must, however, be interpreted within the boundaries of its limitations.

One major obstacle is the lack of food security data, which directly affects land use and forest preservation. Food security and insecurity are related to agricultural practices, reliance on resources and land conversion in a way that might introduce a bias towards our understanding of how different strategies interact with the forest functioning. Future research should incorporate food security indicators to capture the full complexity of human-environment interactions.

Another main limitation of the study is the small sample size, which may explain the low statistical power of the regression and the SEM analyses. Increasing the sample size would increase the robustness of the findings as well as the possibility to generalize them across other communities within and outside Niassa Special Reserve. A larger data set would also enable testing for weaker, but potentially meaningful, socio-ecological interactions that may have been missed in this analysis.

Also, the temporal aspect is absent in socio-economic data, which further limits the study's potential to analyse long-term changes in livelihood strategies and forest functioning. The remote sensing data includes multi-temporal measures, whereas the socio-economic variables were only collected at one point in time, which cannot be used to assess trends or changes in seasonality. Thus, the study fails to fully capture short-term environmental changes, changes in the economy, or changes in policy that could affect land use and conservation outcomes. Future research should use longitudinal survey designs to investigate the dynamics of socio-economic factors over time.

Another constraint stems from uncertainties in remote sensing. Although NDVI, NDWI, and other vegetation indices are useful for monitoring the functioning of forests, these measures can be vulnerable to atmospheric disturbances, cloud cover heterogeneity, and constraints of sensor resolution, leading to measurement errors. Additionally, as the study aggregates multi-temporal satellite data into mean values, it can mask seasonal trends and vegetation dynamics that may hold more information about ecosystem resilience.

Self-reported data used for socio economic variables also limits this study and there remains close to prospect survey and sampling biases. It is also prone to misreporting, including recall bias and social desirability bias, and social acceptance of wealth differences can lead to inaccurate estimates of household wealth, household livelihood diversification, and household access to infrastructure. In addition, the sampling strategy may fail to adequately reflect the variety of socio-economic circumstances in each of the 12 villages, marginalizing findings.

Another limitation relates to uncertainty about causal relationships. SEM is a powerful tool in the investigation of complex interactions, but does not imply true causation. Unmeasured confounding could affect the relationships seen in our models, and reverse causality cannot be excluded. To establish causal pathways in socio-ecological interactions, experimental or longitudinal designs would be required.

All conservation landscapes are complex interactions, which is not really represented. Reporting and discussing the Niassa socio-ecological system with focus on its role of feedback loops, governance arrangements and exogenous drivers, i.e. policy interventions, market conditions and climatic variation. Although the study captures general relationships, more complex higher-order interactions are not included and are likely to affect conservation outcomes in the future.

Finally, errors in measurement of socio-economic variables represent another limitation. To model more complex relationships between categorical variables (i.e., socio-economic factors such as Household Wealth and Dependency Ratio), it is difficult to quantify them accurately because they can include variables such as income variations, subsistence farming engagement, informal economies, and non-financial wealth (land ownership, livestock). The extent of under-counting in the study may be reflective of the economic conditions or inequalities at play and could dilute the strength of relationships seen in the observed results.

These limitations point to directions for future research and methodological refinements. Investigating them in later research — with bigger sample sizes, longitudinal data collection, refined remote sensing—and better survey methods — would enable a fuller and more accurate understanding of the socio-ecological dynamics shaping conservation initiatives in Niassa and other protected areas.

5.10 Alternative Analytical Approaches and Future Research Directions

Future studies may use Bayesian based modeling to establish more probabilistic interpretations of socio-ecological relationships. Methods exist that are well suited for analyzing non-linear interactions between response and predictor variables, which could be beneficial to identify latent interactions between livelihood strategies and associated ecological indicators. Moreover, qualitative data from interviews with local communities could provide deeper insights into forest-livelihood interactions, thereby supplementing the quantitative study. Decoupling conservation outcomes from socio-political variables through investigation in governance structures and policy frameworks would also assist in establishing the relevance of institutional factors to conservation outcomes. Studies with longer time frames focused on socio-economic transformations and their implications for environmental change would provide even stronger insights on the sustainable management of natural resources in protected areas.

Future research can build from these findings to hone conservation strategies that are both ecologically sound and socio-economically feasible, supporting sustainable resource use within Niassa and similar protected areas internationally.

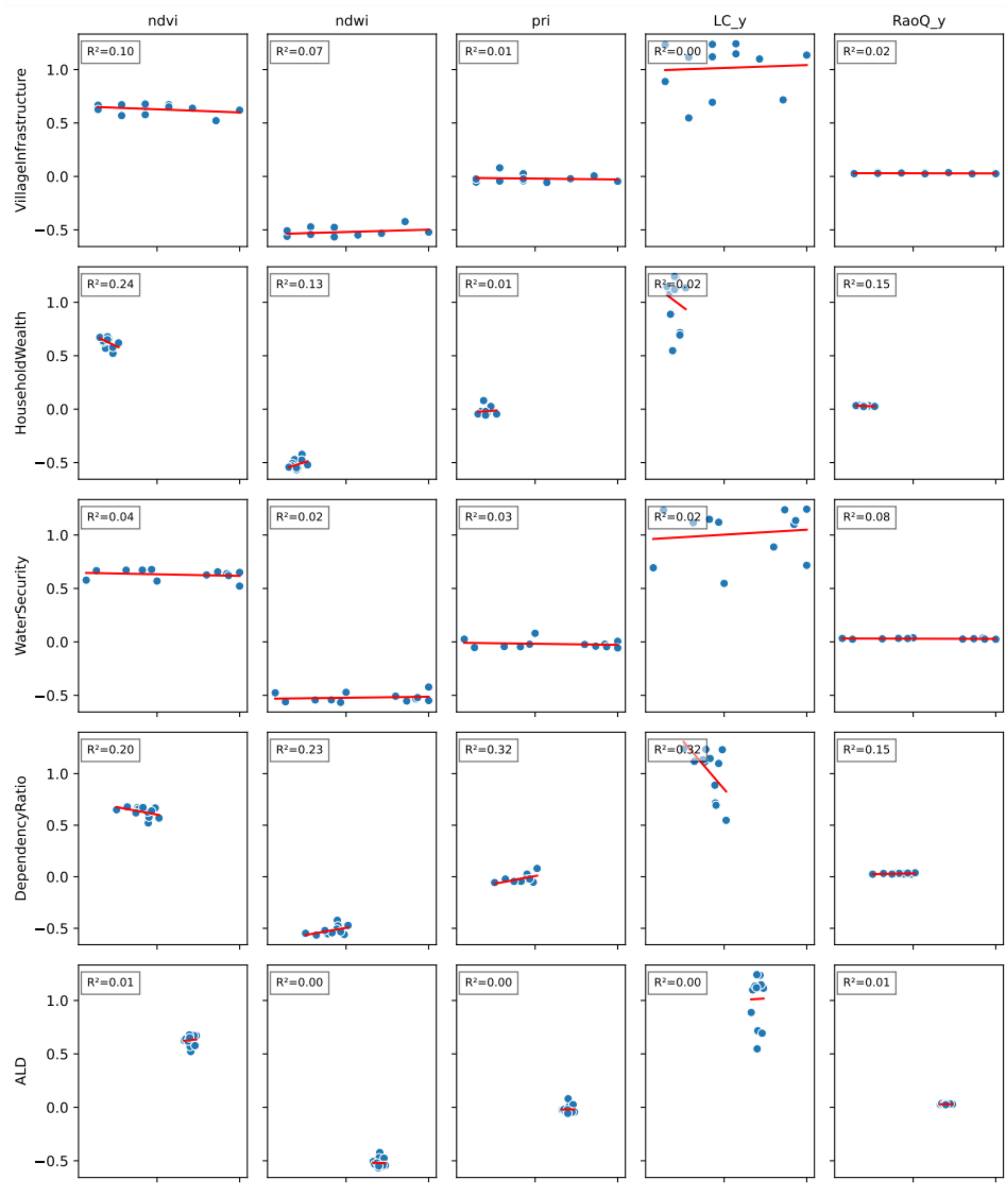
A SEM Table

CSV table that resulted from the different pre-processing steps and was in the end used for the three different analyses.

fid	Name	WaterSecurity	CropDiversity	ALD	VillageInfrastructure	DependencyRatio	HouseholdWealth	ReceiveGive	ndvi	ndwi	pri	LC_y	RaoQ_y
1	Ncuti	0.133	0.392156863	0.7	0.142857143	0.489	0.2	1	0.66693908	-0.56036145	-0.05280437	1.23246825	0.02491771
3	Mecula	1	0.450980392	0.7	0.857142857	0.447	0.233	1	0.52194542	-0.422663	0.005549428	0.71630848	0.02437476
4	Lissongole	0.5	0.235294118	0.7	0.285714286	0.512	0.188	1.143	0.56930244	-0.47170818	0.080231048	0.54732949	0.03818638
5	Erevuka	0.071	0.254901961	0.73	0.428571429	0.451	0.232	1.143	0.5777787	-0.47630876	0.025436783	0.69360679	0.03247108
6	Naulala	0.313	0.37254902	0.74	0.285714286	0.383	0.164	0.927	0.67131609	-0.54346657	-0.04364895	1.11460185	0.02834694
8	Nahavara	0.8	0.411764706	0.66	0.142857143	0.444	0.175	0.967	0.62647265	-0.50790584	-0.02430366	0.88795221	0.02629611
9	Mucoria	0.867	0.352941176	0.72	0.428571429	0.389	0.208	1.067	0.65546805	-0.55303717	-0.04030291	1.2360853	0.02964192
10	Mussoma	0.923	0.294117647	0.67	0.714285714	0.467	0.173	1.077	0.63839632	-0.53306925	-0.02118766	1.09883845	0.03563217
11	Ligogo	0.411	0.333333333	0.72	0.571428571	0.415	0.154	1.059	0.67172808	-0.54287416	-0.04486652	1.14710748	0.03307956
12	Mavago-Sede	0.933	0.274509804	0.69	1	0.372	0.267	1.1	0.62010056	-0.52113915	-0.04549755	1.13523758	0.0254827
13	Matondovela	0.467	0.352941176	0.7	0.428571429	0.32	0.2	1.067	0.67766994	-0.56572861	-0.02238851	1.11940002	0.03185749
14	Msawise	1	0.254901961	0.7	0.571428571	0.255	0.2	1	0.65019625	-0.54854441	-0.05636704	1.24173069	0.02449087

B All Regression Results

All plots resulting from the regression analysis with their respective R^2 score shown in the small box. The red lines are only the linear regressions to keep the plot clean.



C Python Scripts for Data Processing

This appendix contains the Python scripts used for pre-processing and analyzing data.

C.1 API Access script for Sentinel-2 Level-2A data

This script was used to access the remote sensing data via an API and downloading it as a GeoTIFF.

```python
# -*- coding: utf-8 -*-
"""
Created on Sun Jul 28 12:09:20 2024

@author: lino
"""

import requests
import pandas as pd
from datetime import datetime, timedelta
import tarfile
import os
import json
import shutil
import time
import matplotlib.pyplot as plt
import rasterio
from rasterio.plot import show

# Function to get OAuth2 token
def get_oauth_token(client_id, client_secret, token_url):
    data = {
        'grant_type': 'client_credentials',
        'client_id': client_id,
        'client_secret': client_secret
    }
    response = requests.post(token_url, data=data)
    response.raise_for_status()
    return response.json()['access_token']

# Function to create polygons from points
def create_polygons(df):
    polygons = {}
    for fid in df['fid'].unique():
        fid_rows = df[df['fid'] == fid]
        coords = fid_rows[['x', 'y', 'Longitude', 'Latitude']].values.tolist
            ()
        print(coords)
        quad_polygons = [
            [[coords[0][0], coords[0][1]], [coords[0][0], coords[0][3]], [
                coords[0][2], coords[0][3]], [coords[0][2], coords[0][1]], [
                coords[0][0], coords[0][1]]],  # top-left
            [[coords[1][0], coords[1][3]], [coords[1][0], coords[1][1]], [
                coords[1][2], coords[1][1]], [coords[1][2], coords[1][3]], [
                coords[1][0], coords[1][3]]],  # top-right
            [[coords[2][2], coords[2][3]], [coords[2][2], coords[2][1]], [
                coords[2][0], coords[2][1]], [coords[2][0], coords[2][3]], [
```

```
                coords[2][2], coords[2][3]],  # bottom-right
            [[coords[3][2], coords[3][1]], [coords[3][2], coords[3][3]], [
                coords[3][0], coords[3][3]], [coords[3][0], coords[3][1]], [
                coords[3][2], coords[3][1]]],  # bottom-left
        ]
        polygons[fid] = quad_polygons
        print(polygons)
    return polygons

# Function to generate date ranges
def generate_date_ranges(start_date, end_date, interval_days):
    start = datetime.strptime(start_date, '%Y-%m-%d')
    end = datetime.strptime(end_date, '%Y-%m-%d')
    while start < end:
        next_date = start + timedelta(days=interval_days)
        yield start.strftime('%Y-%m-%dT00:00:00Z'), next_date.strftime('%Y-%m
            -%dT00:00:00Z')
        start = next_date

# Function to load API keys from CSV file
def load_api_keys(file_path):
    df = pd.read_csv(file_path)
    keys = df[['client_id', 'client_secret']].to_dict(orient='records')
    return keys

# Load API keys
api_keys = load_api_keys('C:/Users/lino/OneDrive/Desktop/Master␣Thesis/
    API_Keys.csv')
current_key_index = 0

# Function to rotate API keys
def rotate_api_keys():
    global current_key_index, api_keys
    current_key_index = (current_key_index + 1) % len(api_keys)
    return api_keys[current_key_index]['client_id'], api_keys[
        current_key_index]['client_secret']

# Initial OAuth2 details
client_id, client_secret = api_keys[current_key_index]['client_id'], api_keys
    [current_key_index]['client_secret']
token_url = 'https://identity.dataspace.copernicus.eu/auth/realms/CDSE/
    protocol/openid-connect/token'

# Get initial OAuth2 token
token = get_oauth_token(client_id, client_secret, token_url)

# API endpoint
url = "https://sh.dataspace.copernicus.eu/api/v1/process"

# Function to make API request and handle token expiration and API key
    rotation
def make_request_with_retries(request_payload, headers, max_retries=5,
    initial_delay=1, backoff_factor=2):
    global token, client_id, client_secret
    retries = 0
    delay = initial_delay
```

```python
    while retries < max_retries:
        response = requests.post(url, json=request_payload, headers=headers)
        if response.status_code == 401:  # Unauthorized, refresh token
            client_id, client_secret = rotate_api_keys()  # Rotate API keys
            token = get_oauth_token(client_id, client_secret, token_url)
            headers["Authorization"] = f"Bearer␣{token}"
            print("Token␣refreshed␣with␣new␣API␣key")
        elif response.status_code == 429:  # Too Many Requests, apply backoff
            print(f"HTTPError:␣{response.status_code}␣-␣Too␣Many␣Requests.␣
                Retrying␣in␣{delay}␣seconds...")
            time.sleep(delay)
            retries += 1
            delay *= backoff_factor  # Exponential backoff
        else:
            return response
    return response

# Function to plot images
def plot_image(image_path):
    with rasterio.open(image_path) as src:
        fig, ax = plt.subplots(figsize=(10, 10))
        show(src, ax=ax, title=os.path.basename(image_path))
        plt.show()

# Read CSV and create polygons
df = pd.read_csv('C:/Users/lino/OneDrive/Desktop/Master␣Thesis/QGISfiles/data
    /buffer_loc_single.csv')
polygons = create_polygons(df)

# Loop through date ranges and make requests
start_date = '2016-01-01'
end_date = '2020-12-31'
interval_days = 28

evalscript = """
//VERSION=3
function setup() {
  return {
    input: ["B02", "B03", "B04", "B05", "B07", "B08"],
    mosaicking: Mosaicking.ORBIT,
    output: { id: "default", bands: 6 },
  }
}

function updateOutputMetadata(scenes, inputMetadata, outputMetadata) {
  outputMetadata.userData = { scenes: scenes.orbits }
}

function evaluatePixel(samples) {
  return [2.5 * samples[1].B04, 2.5 * samples[1].B03, 2.5 * samples[1].B02,
      2.5 * samples[1].B05, 2.5 * samples[1].B07, 2.5 * samples[1].B08]
}
"""

image_counter = 1

for fid, quad_polygons in polygons.items():
```

```
for seq, coordinates in enumerate(quad_polygons, start=1):  # For each quadrant 1 to 4
    for start, end in generate_date_ranges(start_date, end_date, interval_days):
        request_payload = {
            "input": {
                "bounds": {
                    "geometry": {
                        "type": "Polygon",
                        "coordinates": [coordinates]
                    }
                },
                "data": [
                    {
                        "type": "sentinel-2-l2a",
                        "dataFilter": {
                            "timeRange": {
                                "from": start,
                                "to": end,
                                "maxCloudCoverage": 20  # Cloud-free
                            },
                            "maxCloudCoverage": 20  # Cloud-free
                        },
                    }
                ],
            },
            "output": {
                "width": 1000,
                "height": 1000,
                "responses": [
                    {
                        "identifier": "default",
                        "format": {"type": "image/tiff"},
                    },
                    {
                        "identifier": "userdata",
                        "format": {"type": "application/json"},
                    },
                ],
            },
            "evalscript": evalscript,
        }

        headers = {
            "Authorization": f"Bearer {token}",
            "Accept": "application/tar"  # This header can be changed if the API directly supports GeoTIFF
        }

        response = make_request_with_retries(request_payload, headers)
        if response.status_code != 200:
            print(f"HTTPError: {response.status_code} - {response.reason}")
            print("Response Content:", response.content)
            continue

        # Save the tar response
```

```python
        temp_tar_filename = f"temp_{fid}-{seq}.tar"
        with open(temp_tar_filename, 'wb') as f:
            f.write(response.content)

        # Create a unique directory for extraction
        extract_dir = f"extract_{fid}_{seq}"
        os.makedirs(extract_dir, exist_ok=True)

        # Extract and check contents of the tar file
        with tarfile.open(temp_tar_filename, "r") as tar:
            tar.extractall(path=extract_dir)
            extracted_files = tar.getnames()

        # Read the metadata file
        metadata_filename = os.path.join(extract_dir, 'userdata.json')
        if not os.path.exists(metadata_filename):
            print(f"Error: {metadata_filename} not found in the tar archive.")
            shutil.rmtree(extract_dir)
            continue

        with open(metadata_filename, 'r') as f:
            metadata = json.load(f)

        # Extract the actual date from metadata
        try:
            actual_date = metadata['scenes'][0]['tiles'][0]['date'].split('T')[0].replace('-', '')
            # Rename the image file
            image_filename = os.path.join(extract_dir, 'default.tif')
            if os.path.exists(image_filename):
                output_filename = f"Niassa_{fid}-{seq}_{actual_date}.tiff"
                shutil.move(image_filename, output_filename)
                print(f"Downloading image #{image_counter}: {output_filename}")
                #plot_image(output_filename)
                image_counter += 1
            else:
                print(f"Error: {image_filename} not found in the extracted files.")
        except KeyError as e:
            print(f"KeyError: {e} - Could not find the expected key in metadata")

        # Clean up temporary files
        os.remove(temp_tar_filename)
        shutil.rmtree(extract_dir)
```

Listing 1: Python script for accessing Sentinel-2 Level-2A imagery

C.2 Modyfying NetCDF stacks

This script was used to modify the NetCDF64 stacks that originated from the merged GeoTiffs downloaded using the API Access Script

```
# -*- coding: utf-8 -*-
"""
Created on Fri Aug 30 15:16:27 2024

@author: lino
"""

import os
import numpy as np
import matplotlib.pyplot as plt
from netCDF4 import Dataset, num2date
from skimage import exposure
from matplotlib.backends.backend_pdf import PdfPages
from datetime import datetime
import pandas as pd
import matplotlib.dates as mdates
import seaborn as sns
from spectralrao_mod import *

def process_and_normalize_stack(nc_file_path, output_folder, output_dir,
    initial_cloud_threshold, pif_cloud_threshold, coverage_threshold, ref_date
    =None):
    try:
        # Ensure output directories exist
        os.makedirs(output_folder, exist_ok=True)
        os.makedirs(output_dir, exist_ok=True)

        # Step 1: Remove too cloudy images
        with Dataset(nc_file_path, 'r') as ncfile:
            print(f"Opening NetCDF file: {nc_file_path}")
            # Read data and metadata
            data = ncfile.variables['data'][:]  # shape (time, band, y, x)
            times = ncfile.variables['time'][:]
            time_units = ncfile.variables['time'].units
            dates = num2date(times, units=time_units)

            # Filter for specific months (May to October)
            months_to_keep = [5, 6, 7, 8, 9, 10]
            valid_month_indices = [i for i, date in enumerate(dates) if date.
                month in months_to_keep]


            if len(valid_month_indices) == 0:
                print("No images retained after filtering for the specified 
                    months (May to October).")
                return

            # Filter data and dates to only include the selected months
            data = data[valid_month_indices, :, :, :]
            dates = np.array([datetime(d.year, d.month, d.day) for d in dates
                ])[valid_month_indices]

            # Generate cloud mask
            cloud_mask = (
                (data[:, 0, :, :] > initial_cloud_threshold) &
                (data[:, 1, :, :] > initial_cloud_threshold) &
```

```
    (data[:, 2, :, :] > initial_cloud_threshold)
)
cloud_coverage = np.mean(cloud_mask, axis=(1, 2)) * 100
valid_indices = np.where(cloud_coverage <= coverage_threshold)[0]

if len(valid_indices) == 0:
    print("No images retained after applying cloud coverage 
        threshold.")
    return

# Filter data and dates based on valid indices
data = data[valid_indices, :, :, :]
dates = np.array([datetime(d.year, d.month, d.day) for d in dates
    ])[valid_indices]
cloud_mask = cloud_mask[valid_indices, :, :]
print(f"{len(valid_indices)} images retained after cloud coverage
     filtering.")


# Step 3: Calculate std.dev for each pixel and select the 100
    pixels with the lowest std.dev
std_dev = np.full(data.shape[1:], np.nan)
for b in range(data.shape[1]):
    valid_data = np.where(cloud_mask, np.nan, data[:, b, :, :])
    std_dev[b, :, :] = np.nanstd(valid_data, axis=0)
combined_std_dev = np.nanmin(std_dev, axis=0)

# Flatten the std.dev array and find the indices of the 100
    lowest values
flattened_std_dev = combined_std_dev.flatten()
lowest_stddev_indices = np.argpartition(flattened_std_dev, 100)
    [:100]

# Convert flattened indices back to 2D indices
pif_indices = np.unravel_index(lowest_stddev_indices,
    combined_std_dev.shape)

pif_mask = np.zeros_like(combined_std_dev, dtype=bool)
pif_mask[pif_indices] = True



# Step 4: Remove images if a PIF is covered by a cloud (stricter
    threshold)
refined_valid_indices = []
refined_dates = []
for t in range(data.shape[0]):
    pif_cloud_cover = cloud_mask[t, pif_mask]
    if not np.any(pif_cloud_cover):
        refined_valid_indices.append(t)
        refined_dates.append(dates[t])

if len(refined_valid_indices) == 0:
    print("No images retained after removing those with cloud-
        covered PIFs.")
    return
else:
```

```
        print(f"{len(refined_valid_indices)} images retained after PIF cloud coverage filtering.")

    data = data[refined_valid_indices, :, :, :]
    cloud_mask = cloud_mask[refined_valid_indices, :, :]
    dates = np.array(refined_dates)

    # Step 2: Select the reference image by date
    if ref_date is not None:
        ref_date = datetime.strptime(ref_date, "%Y-%m-%d")

        for idx, date in enumerate(dates):

            if date == ref_date:
                ref_time_index = idx
                print(f"Reference date {ref_date} found at index {ref_time_index}.")
                break
        else:
            print(f"Reference date {ref_date} not found in the filtered dates. Using the first image as reference.")
            ref_time_index = 0
    else:
        ref_time_index = 0


    # Ensure the reference time index is still valid
    if ref_time_index >= len(refined_valid_indices):
        print(f"Reference time index {ref_time_index} is out of bounds for the filtered data. Using index 0 instead.")
        ref_time_index = 0

    # Step 5: Radiometric normalization using histogram matching
    ref_image = data[ref_time_index, :, :, :]
    normalized_data = np.empty_like(data)

    for b in range(data.shape[1]):  # For each band (Red, Green, Blue)
        ref_band = ref_image[b, :, :]
        for t in range(data.shape[0]):
            target_band = data[t, b, :, :]
            normalized_data[t, b, :, :] = exposure.match_histograms(target_band, ref_band)

    print("Radiometric normalization using histogram matching completed.")

    # Step 6: Save images to PDF
    base_name, ext = os.path.splitext(os.path.basename(nc_file_path))
    pdf_filename = os.path.join(output_folder, f"{base_name}_images_normalizedByHist_{coverage_threshold}.pdf")
    with PdfPages(pdf_filename) as pdf:
        for t in range(normalized_data.shape[0]):
            plt.figure(figsize=(8, 8))
            # Normalize the RGB bands for visualization
            rgb_image = np.stack([
                normalized_data[t, i, :, :] for i in range(3)
```

```python
        ], axis=2)
        # Rescale to 0-1 for display
        rgb_image = (rgb_image - np.nanmin(rgb_image)) / (np.nanmax(rgb_image) - np.nanmin(rgb_image))
        plt.imshow(rgb_image, origin='lower')
        plt.title(f"Image {t + 1} - {dates[t].strftime('%Y-%m-%d')}")
        plt.axis('off')
        pdf.savefig()
        plt.close()
print(f"Images saved to {pdf_filename}")

# Step 7: Save normalized data to NetCDF
normalized_file_name = f"{base_name}_normalizedByHist_{coverage_threshold}{ext}"
normalized_file_path = os.path.join(output_dir, normalized_file_name)

# Re-open original file to ensure it's accessible
with Dataset(nc_file_path, 'r') as src_ncfile:
    with Dataset(normalized_file_path, 'w', format='NETCDF4') as dst_ncfile:

        for name, dimension in src_ncfile.dimensions.items():
            try:
                dim_size = len(dimension) if not dimension.isunlimited() else None
                dst_ncfile.createDimension(name, dim_size)

            except Exception as e:
                print(f"Failed to create dimension '{name}': {e}")
                return



        for name, variable in src_ncfile.variables.items():
            try:
                var_dims = variable.dimensions
                var_dtype = variable.dtype
                dst_var = dst_ncfile.createVariable(name, var_dtype, var_dims)
                # Copy variable attributes
                dst_var.setncatts({attr: variable.getncattr(attr) for attr in variable.ncattrs()})
                if name == 'data':
                    dst_var[:] = normalized_data

                else:
                    dst_var[:] = variable[valid_indices, ...] if 'time' in var_dims else variable[:]

            except Exception as e:
                print(f"Failed to create variable '{name}': {e}")
                return
        print(f"Normalized dataset saved to {normalized_file_path}")
```

```
    except Exception as e:
        print(f"An unexpected error occurred: {e}")

    return refined_dates


def ndvi(nc_file_path):
    # Open file
    with Dataset(nc_file_path, 'a') as ncfile:
        # Extract necessary band data
        nir = ncfile.variables['data'][:,5,:,:]
        red = ncfile.variables['data'][:,2,:,:]

        # Calculate NDVI
        ndvi = (nir-red)/(nir+red)

        # Handle division by 0
        ndvi = np.where(np.isfinite(ndvi), ndvi, np.nan)

        # Check if the NDVI variable already exists, if not, create it
        if 'ndvi' not in ncfile.variables:
            ndvi_var = ncfile.createVariable('ndvi', 'f4', ('time', 'y', 'x')
                , zlib=True)
            ndvi_var.units = 'NDVI'
            ndvi_var.long_name = 'Normalized Difference Vegetation Index'
        else:
            ndvi_var = ncfile.variables['ndvi']

        # Add the NDVI data to the NetCDF file
        ndvi_var[:, :, :] = ndvi


def ndwi(nc_file_path):
    # Open the NetCDF file in append mode
    with Dataset(nc_file_path, 'a') as ncfile:
        # Extract the necessary band data
        green = ncfile.variables['data'][:, 1, :, :]
        nir = ncfile.variables['data'][:, 5, :, :]

        # Calculate NDWI
        ndwi = (green - nir) / (green + nir)

        # Handle potential division by zero or invalid values
        ndwi = np.where(np.isfinite(ndwi), ndwi, np.nan)

        # Check if the NDWI variable already exists, if not, create it
        if 'ndwi' not in ncfile.variables:
            ndwi_var = ncfile.createVariable('ndwi', 'f4', ('time', 'y', 'x')
                , zlib=True)
            ndwi_var.units = 'NDWI'
            ndwi_var.long_name = 'Normalized Difference Water Index'
        else:
            ndwi_var = ncfile.variables['ndwi']

        # Add the NDWI data to the NetCDF file
        ndwi_var[:, :, :] = ndwi
```

```python
def leaf_chemistry(nc_file_path):
    # Open the NetCDF file in append mode
    with Dataset(nc_file_path, 'a') as ncfile:
        # Extract the necessary band data
        b5 = ncfile.variables['data'][:, 3, :, :]
        b7 = ncfile.variables['data'][:, 4, :, :]

        # Calculate NDWI
        lc = b7/b5 - 1

        # Handle potential division by zero or invalid values
        lc = np.where(np.isfinite(lc), lc, np.nan)

        # Check if the NDWI variable already exists, if not, create it
        if 'LC' not in ncfile.variables:
            lc_var = ncfile.createVariable('LC', 'f4', ('time', 'y', 'x'),
                zlib=True)
            lc_var.units = 'LC'
            lc_var.long_name = 'Leaf Chemistry'
        else:
            lc_var = ncfile.variables['LC']

        # Add the NDWI data to the NetCDF file
        lc_var[:, :, :] = lc


def pri(nc_file_path):
    with Dataset(nc_file_path, 'a') as ncfile:
        # Extract the necessary band data
        b2 = ncfile.variables['data'][:, 0, :, :]
        b3 = ncfile.variables['data'][:, 1, :, :]

        # Calculate NDWI
        pri = (b2-b3)/(b2+b3)

        # Handle potential division by zero or invalid values
        pri = np.where(np.isfinite(pri), pri, np.nan)

        # Check if the NDWI variable already exists, if not, create it
        if 'pri' not in ncfile.variables:
            pri_var = ncfile.createVariable('pri', 'f4', ('time', 'y', 'x'),
                zlib=True)
            pri_var.units = 'pri'
            pri_var.long_name = 'Photochemical Reflectance Index'
        else:
            pri_var = ncfile.variables['pri']

        # Add the NDWI data to the NetCDF file
        pri_var[:, :, :] = pri

def calculate_raoq(nc_file_path, window, mode, na_tolerance, simplify):
    try:
        with Dataset(nc_file_path, 'r+') as ncfile:
            # Check if the 'ndvi' variable is present
            if 'ndvi' not in ncfile.variables:
```

```
                print("NDVI variable not found in the NetCDF file.")
                return

            # Read the NDVI data from the NetCDF file
            ndvi_data = ncfile.variables['ndvi'][:]  # Shape: (time, y, x)
            y, x = ndvi_data.shape[1], ndvi_data.shape[2]

            # Initialize array to store RaoQ values
            raoq_values = np.zeros((ndvi_data.shape[0], y, x))
            maxi = np.max(ndvi_data.shape[0])

            # Loop through each time step to calculate RaoQ for the NDVI only
            for t in range(ndvi_data.shape[0]):

                #print("Starting RaoQ calculation...")

                raoq_band = spectralrao(ndvi_data[t, :, :], window=window,
                    mode=mode, na_tolerance=na_tolerance, simplify=simplify)

                # Store the calculated RaoQ
                raoq_values[t, :, :] = raoq_band
                print(f"\nFinished RaoQ calculation on image {t+1} out of {
                    maxi}\n")

            # Save RaoQ to the NetCDF file
            if 'RaoQ' not in ncfile.variables:
                raoq_var = ncfile.createVariable('RaoQ', 'f4', ('time', 'y',
                    'x'))
                raoq_var.units = 'unitless'
                raoq_var.long_name = "Rao's quadratic entropy (calculated on
                    NDVI)"
            else:
                raoq_var = ncfile.variables['RaoQ']

            raoq_var[:, : ,:] = raoq_values
            print("RaoQ variable added to the NetCDF file.")

    except Exception as e:
        print(f"An error occurred: {e}")
        return None


def visualize_variable(nc_file_path, variable_name, i, time_index):
    try:



        with Dataset(nc_file_path, 'r') as ncfile:
            if variable_name not in ncfile.variables:
                raise ValueError(f"Variable '{variable_name}' not found in
                    the NetCDF file.")

            # Load the variable data
            variable_data = ncfile.variables[variable_name][time_index, :, :]
            x = ncfile.variables['x'][:]
            y = ncfile.variables['y'][:]
            times = ncfile.variables['time']
```

```
            time_units = times.units
            dates = num2date(times[:], units=time_units)

            base_name, ext = os.path.splitext(os.path.basename(nc_file_path))
            output_file_path = os.path.join(output_folder, f"Niassa_{i}_{
                variable_name}_{dates[time_index].strftime('%Y-%m-%d')}.png")

            # Create a plot
            plt.figure(figsize=(10, 8))
            plt.imshow(variable_data, extent=(x.min(), x.max(), y.min(), y.
                max()), origin='lower', cmap='RdYlGn')
            plt.colorbar(label=variable_name)
            plt.title(f"Niassa_{i}: {variable_name} on {dates[time_index].
                strftime('%Y-%m-%d')}")
            plt.xlabel('Longitude')
            plt.ylabel('Latitude')
            plt.savefig(output_file_path, format='png')
            plt.close()


    except Exception as e:
        print(f"An error occurred: {e}")

def compare_variables_across_thresholds(nc_file_paths, i, variable_names,
    date_freq):
    try:
        for variable_name in variable_names:
            plt.figure(figsize=(10, 6))

            for coverage_threshold, nc_file_path in nc_file_paths.items():
                with Dataset(nc_file_path, 'r') as ncfile:
                    if variable_name not in ncfile.variables:
                        print(f"Variable '{variable_name}' not found in the
                            NetCDF file.")
                        continue

                    # Read time and convert to dates
                    times = ncfile.variables['time'][:]
                    time_units = ncfile.variables['time'].units
                    dates = num2date(times, units=time_units)

                    # Convert dates to pandas datetime for easy plotting
                    dates = pd.to_datetime([date.strftime('%Y-%m-%d') for
                        date in dates])

                    # Calculate mean value for each time step
                    variable_data = np.array(ncfile.variables[variable_name
                        ][:])  # (time, y, x) - make a writable copy
                    mean_values = np.array([
                        np.nanmean(variable_data[t, :, :]) for t in range(
                            variable_data.shape[0])
                    ])

                    # Plot the mean values over time
                    plt.plot(dates, mean_values, marker='o', linestyle='-',
                        label=f'Threshold {coverage_threshold}%')
```

```
            plt.xlabel('Date')
            plt.ylabel(f'Mean {variable_name}')
            plt.title(f'Mean {variable_name} Over Time Niassa {i}')
            plt.grid(True)

            # Set the frequency of x-axis labels
            plt.gca().xaxis.set_major_locator(mdates.MonthLocator(interval=
                int(date_freq[:-1])))
            plt.gca().xaxis.set_major_formatter(mdates.DateFormatter('%Y-%m')
                )

            plt.xticks(rotation=45)
            plt.tight_layout()
            plt.legend()
            plt.show()

    except Exception as e:
        print(f"An error occurred: {e}")

def plot_variable_histogram_across_thresholds(nc_file_paths, variable_name,
    time_index = 10, bins=200):
    try:
        plt.figure(figsize=(10, 6))

        for coverage_threshold, nc_file_path in nc_file_paths.items():
            with Dataset(nc_file_path, 'r') as ncfile:
                if variable_name not in ncfile.variables:
                    print(f"Variable '{variable_name}' not found in the
                        NetCDF file.")
                    continue

                # Read the variable data at the specified timestep
                variable_data = ncfile.variables[variable_name][time_index,
                    :, :]
                variable_data = variable_data.flatten()  # Flatten to 1D
                    array

                # Remove NaN values
                variable_data = variable_data[~np.isnan(variable_data)]

                # Plot the histogram
                plt.hist(variable_data, bins=bins, alpha=0.5, label=f'
                    Threshold {coverage_threshold}%', edgecolor='black',
                    histtype= 'barstacked')

        plt.xlabel(f'{variable_name} Value')
        plt.ylabel('Frequency')
        plt.title(f'Histogram of {variable_name} Distribution at Time Index {
            time_index}')
        plt.legend()
        plt.grid(True)
        plt.show()

    except Exception as e:
        print(f"An error occurred: {e}")

def plot_variable_distribution_across_times(nc_file_path, variable_name, i,
```

```
    output_folder, time_indices, bw_adjust=1):
    try:

        base_name, ext = os.path.splitext(os.path.basename(nc_file_path))
        output_file_path = os.path.join(output_folder, f"Niassa_{i}_{
            variable_name}_MutlipleTimesteps.png")


        plt.figure(figsize=(10, 6))

        with Dataset(nc_file_path, 'r') as ncfile:
            if variable_name not in ncfile.variables:
                print(f"Variable '{variable_name}' not found in the NetCDF 
                    file.")
                return

            for time_index in time_indices:
                if time_index >= len(ncfile.variables['time']):
                    print(f"Time index {time_index} is out of bounds.")
                    continue

                # Read the variable data at the specified timestep
                variable_data = ncfile.variables[variable_name][time_index,
                    :, :]
                variable_data = variable_data.flatten()  # Flatten to 1D
                    array

                # Remove NaN values
                variable_data = variable_data[~np.isnan(variable_data)]

                # Plot the KDE for the current time index
                sns.kdeplot(variable_data, bw_adjust=bw_adjust, label=f'Time 
                    Index {time_index}', fill=False)

            plt.xlabel(f'{variable_name} Value')
            plt.ylabel('Density')
            plt.title(f'Niassa_{i}: Density Distribution of {variable_name} 
                (5% cloud coverage) Across Different Timesteps')
            plt.legend()
            # plt.show()
            plt.savefig(output_file_path, format='png')
            plt.close()




    except Exception as e:
        print(f"An error occurred: {e}")


def plot_monthly_distribution(filtered_dates, i):
    try:
        # Convert dates to pandas datetime for easy manipulation (if not
            already in pd.Timestamp format)
        if not isinstance(filtered_dates[0], pd.Timestamp):
            filtered_dates = pd.to_datetime([date.strftime('%Y-%m-%d') for
                date in filtered_dates])
```

```
        # Extract month from each date
        months = filtered_dates.month

        # Plot the histogram
        base_name, ext = os.path.splitext(os.path.basename(nc_file_path))
        output_file_path = os.path.join(output_folder, f"Niassa_{i}_MonthlyDist.png")


        plt.figure(figsize=(10, 6))
        plt.hist(months, bins=np.arange(1, 14) - 0.5, edgecolor='black', alpha=0.7)
        plt.xticks(np.arange(1, 13), labels=['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun',
                                              'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec'])
        plt.xlabel('Month')
        plt.ylabel('Frequency')
        plt.title(f'Frequency of Each Month Represented for Niassa_{i}')
        plt.grid(True)
        plt.savefig(output_file_path, format='png')
        plt.close()

    except Exception as e:
        print(f"An error occurred: {e}")



def plot_mean(nc_file_path, variable_name, i, output_folder):
    mean_var = []
    with Dataset(nc_file_path, 'r') as ncfile:
        if variable_name not in ncfile.variables:
            raise ValueError(f"Variable '{variable_name}' not found in the NetCDF file.")

        # Load the variable data and convert it to a writable array
        variable_data = np.array(ncfile.variables[variable_name][:, :, :])  # Ensure the data is writable
        x = ncfile.variables['x'][:]
        y = ncfile.variables['y'][:]
        times = ncfile.variables['time']
        time_units = times.units
        dates = num2date(times[:], units=time_units)

        # Calculate the mean of the variable for each time step
        for t in range(dates.shape[0]):
            Mean = np.nanmean(variable_data[t, :, :])  # Compute mean, ignoring NaNs
            mean_var.append(Mean)

        # Convert dates to pandas datetime for easy plotting
        dates = pd.to_datetime([date.strftime('%Y-%m-%d') for date in dates])

        # Plot the results
        plt.figure(figsize=(10, 6))
        plt.plot(dates, mean_var, marker='o', linestyle='-', color='b')
```

```
        plt.title(f'Niassa_{i}: Mean {variable_name} Over Time')
        plt.xlabel('Date')
        plt.ylabel(f'Mean {variable_name}')
        plt.xticks(rotation=45)
        plt.grid(True)
        plt.tight_layout()

        output_file_path =  os.path.join(output_folder, f"Niassa_{i}_mean_{
            variable_name}.png")


        # Save the plot as a PNG file
        plt.savefig(output_file_path, format='png')
        print(f"Plot saved as {output_file_path}")

        # Show the plot
        plt.show()





# Reference Dates
ref_date_ls = ['2017-06-15', '2023-02-01', '2017-07-10', '2017-09-03', '
    2019-08-04', '2016-06-10',
               '2017-09-03', '2016-08-09', '2017-06-15', '2017-06-15', '
                   2016-06-10', '2017-06-15',
               '2017-06-15', '2017-06-15']

# Set the output directories
output_dir = 'C:/Users/lino/OneDrive/Desktop/Master Thesis/output/nc'  # For
    .nc data outputs
output_folder = 'C:/Users/lino/OneDrive/Desktop/Master Thesis/output/
    rgb_images'  # For images and graphs

for i in range(1,15):
    if i == 2 or i==7:
        continue
    else:
        nc_file_paths = {}

        for j in [5]:
            nc_file_path = os.path.join(output_dir, f"Niassa_{i}.nc")
            refined_dates = process_and_normalize_stack(
                nc_file_path,
                output_folder,
                output_dir,
                initial_cloud_threshold=200,
                pif_cloud_threshold=80,
                coverage_threshold=j,
                ref_date=ref_date_ls[i-1]
            )
            nc_file_path = os.path.join(output_dir, f"Niassa_{i}
                _normalizedByHist_{j}.nc")
            print(f"Calculating NDVI and adding it to Niassa_{i}
                _normalizedByHist_{j}.nc")
            ndvi(nc_file_path)
```

```python
        print(f"Calculating␣NDWI␣and␣adding␣it␣to␣Niassa_{i}
            _normalizedByHist_{j}.nc")
        ndwi(nc_file_path)

        print(f"Calculating␣Leaf␣Chemistry␣and␣adding␣it␣to␣Niassa_{i}
            _normalizedByHist_{j}.nc")
        leaf_chemistry(nc_file_path)

        pri(nc_file_path)

        calculate_raoq(nc_file_path, window=9, mode="classic",
            na_tolerance=0.0, simplify=1)


        plot_monthly_distribution(refined_dates, i)

        nc_file_paths[j] = nc_file_path

        for var in ['ndvi', 'ndwi', 'LC', 'RaoQ' ,'pri']:
            visualize_variable(nc_file_path, var, i, time_index=10)
            plot_variable_histogram_across_thresholds(nc_file_paths, var,
                 j)
            plot_variable_distribution_across_times(nc_file_path, var, i,
                 output_folder, time_indices=[1, 5, 10, 20], bw_adjust=1)
            plot_mean(nc_file_path, var, i, output_folder)




    # Step 4: Compare variables across thresholds
    # for var in ['ndvi', 'ndwi', 'LC', 'RaoQ']:
       # plot_variable_histogram_across_thresholds(nc_file_paths, var, j)
    # compare_variables_across_thresholds(nc_file_paths, i, ['ndvi', '
        ndwi', 'LC', 'lue', 'RaoQ'], date_freq='3M')
```

Listing 2: Python script for processing NetCDF files

References

The State of the World's Forests 2020. FAO and UNEP, 2020. ISBN 978-92-5-132419-6. doi: 10.4060/ca8642en. URL https://openknowledge.fao.org/handle/20.500.14283/ca8642en.

Shabd S. Acharya. Sustainable agriculture and rural livelihoods. *Agricultural Economics Research Review*, 19:205, 2006. URL https://api.semanticscholar.org/CorpusID:152457894.

William M. Adams. *Against Extinction: The Story of Conservation*. Earthscan, London, UK, 2004. ISBN 978-1-85383-955-5.

W. Neil Adger. Vulnerability. *Global Environmental Change*, 16(3):268–281, aug 2006. ISSN 09593780. doi: 10.1016/j.gloenvcha.2006.02.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0959378006000422.

AFR100. The african forest landscape restoration initiative, 2020. URL https://afr100.org/.

Arun Agrawal and Kent Redford. Poverty, development, and biodiversity conservation: Shooting in the dark? *Wildlife Conservation Society*, 34:1–48, 2005.

Miguel A. Altieri. *Agroecology: The Science of Sustainable Agriculture*. CRC Press, 2nd edition, 2004. ISBN 978-1-56670-617-1.

Arild Angelsen, Pamela Jagger, Ronnie Babigumira, Brian Belcher, Nicholas J. Hogarth, Simone Bauch, Jan Börner, Carsten Smith-Hall, and Sven Wunder. Environmental Income and Rural Livelihoods: A Global-Comparative Analysis. *World Development*, 64:S12–S28, dec 2014. ISSN 0305750X. doi: 10.1016/j.worlddev.2014.03.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0305750X14000722.

R. Mwanyoka S. Turner B. Fisher, K. Kulindwa and N. Burgess. Common pool resource management and pes: Lessons and constraints for water pes in tanzania. *Ecological Economics*, 69(6):1253–1261, 2011. doi: 10.1016/j.ecolecon.2009.11.008.

The World Bank. Mozambique climate knowledge portal, 2024. URL https://climateknowledgeportal.worldbank.org/country/mozambique/trends-variability-historical.

Jean-Francois Bastin, Yves Finegold, Claude Garcia, and Danilo Mollicone. The global tree restoration potential. *Science*, 365:76–79, 2019. doi: 10.1126/science.aax0848.

R. Baumgartner and R. Högger. In search of sustainable livelihood systems: Managing resources and change. 2004. URL https://api.semanticscholar.org/CorpusID:141651022.

Christian Beer, Markus Reichstein, Enrico Tomelleri, Philippe Ciais, Martin Jung, Nuno Carvalhais, Christian Rödenbeck, M. Altaf Arain, Dennis Baldocchi, Gordon B. Bonan, Alberte Bondeau, Alessandro Cescatti, Gitta Lasslop, Anders Lindroth, Mark Lomas, Sebastiaan Luyssaert, Hank Margolis, Keith W. Oleson, Olivier Roupsard, Elmar Veenendaal, Nicolas Viovy, Christopher Williams, F. Ian Woodward, and Dario Papale. Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, 329(5993):834–838, 2010. doi: 10.1126/science.1184984.

Nathan J. Bennett and Robin Roth. The conservation social sciences: What?, how? and why? *Canadian Wildlife Biology and Management*, 4(1):22–30, 2015. URL https://cwbm.ca.

Henry Bernstein, Ben Crow, and Hazel Johnson. Rural livelihoods: Crises and responses. 1992. URL https://api.semanticscholar.org/CorpusID:153465948.

Kenneth A. Bollen. *Structural Equations with Latent Variables*. John Wiley & Sons, New York, 1989a. ISBN 978-0-471-01171-2.

Kenneth A Bollen. *Structural Equations with Latent Variables*. John Wiley & Sons, 1989b.

Gordon B. Bonan. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science*, 320(5882):1444–1449, 2008. doi: 10.1126/science.1155121.

Kate A. Brauman. Hydrologic ecosystem services: Linking ecohydrologic processes to human well-being. *WIREs Water*, 2(4):345–358, 2015. doi: 10.1002/wat2.1081. URL https://wires.wiley.com/water/article/10.1002/wat2.1081.

Dan Brockington. *Fortress Conservation: The Preservation of the Mkomazi Game Reserve, Tanzania*. Indiana University Press, Bloomington, IN, 2006. ISBN 9780253346711.

Wei Cai and Shaolong Du. Dry-season ecological monitoring in tropical regions using sentinel-2 imagery. *International Journal of Remote Sensing*, 42(5):1565–1583, 2021. doi: 10.1080/01431161.2021.1879562.

Bruce M. Campbell, Peter Frost, and Neil Byron. *Miombo Woodlands and Their Use: Sustainable Management Options for Africa*. CIFOR (Center for International Forestry Research), Bogor, Indonesia, 2007. URL https://www.cifor.org/knowledge/publication/2285/.

R. Chambers, N. C. Saxena, and T. Shah. To the hands of the poor: water and trees. *To the hands of the poor: water and trees*, 1989.

F. Stuart Chapin, Pamela A. Matson, and Peter M. Vitousek. *Principles of Terrestrial Ecosystem Ecology*. Springer, New York, NY, 2nd edition, 2011. doi: 10.1007/978-1-4419-9504-9.

PS Chavez. Image-based atmospheric corrections—revisited and improved. *Photogrammetric Engineering & Remote Sensing*, 62(9):1025–1036, 1996.

Robin L. Chazdon and Manuel R. Guariguata. Decision support tools for forest landscape restoration: Current status and future outlook. In Robin L. Chazdon and Manuel R. Guariguata, editors, *Forest Landscape Restoration: Integrated Approaches to Support Effective Implementation*, pages 1–12. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-319-74515-2_1. URL https://link.springer.com/chapter/10.1007/978-3-319-74515-2_1.

Enock N. Chidumayo and Dominic J. Gumbo. *The Dry Forests and Woodlands of Africa: Managing for Products and Services*. Earthscan, London, 2010. URL https://www.cifor.org/knowledge/publication/3051/.

Irene Chrysafis, Georgios Korakis, Apostolos P. Kyriazopoulos, and Giorgos Mallinis. Predicting tree species diversity using geodiversity and sentinel-2 multi-seasonal spectral information. *Sustainability (Switzerland)*, 12(21):1–15, 2020. ISSN 20711050. doi: 10.3390/su12219250.

Gordon Conway. *The Doubly Green Revolution: Food for All in the Twenty-First Century*. Cornell University Press, 2nd edition, 2019. ISBN 978-1-5017-4731-9.

Graeme S. Cumming and Craig R. Allen. Protected areas as social-ecological systems: perspectives from resilience and social-ecological systems theory. *Ecological Applications*, 27(6):1709–1717, sep 2017. ISSN 1051-0761. doi: 10.1002/eap.1584. URL https://esajournals.onlinelibrary.wiley.com/doi/10.1002/eap.1584.

F. Detsch, I. Otte, T. Appelhans, and T. Nauss. Seasonal filtering of earth observation data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 120:72–84, 2016. doi: 10.1016/j.isprsjprs.2016.08.012.

M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, A. Meygret, F. Spoto, O. Sy, F. Marchese, and P. Bargellini. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120:25–36, may 2012. ISSN 00344257. doi: 10.1016/j.rse.2011.11.026. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425712000636.

Frank Ellis. Rural livelihood diversity in developing countries: evidence and policy implications. 1999. URL https://api.semanticscholar.org/CorpusID:15548642.

Frank Ellis. Rural livelihoods and diversity in developing countries. 2000. URL https://api.semanticscholar.org/CorpusID:154871462.

Erika Lepers Eric F. Lambin, Helmut J. Geist. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, 28:205–241, 2001. doi: 10.1146/annurev.energy.28.050302.105459.

ESA. Sentinel 2 Documentation, 2024. URL https://sentiwiki.copernicus.eu/web/s2-mission#S2Mission-MSIInstrumentS2-Mission-MSI-Instrumenttrue.

European Space Agency (ESA). Copernicus open access hub: Sentinel satellite data, 2024. URL https://scihub.copernicus.eu/. Accessed: 2024-XX-XX.

Monica Fisher and Gerald Shively. Can Income Programs Reduce Tropical Forest Pressure? Income Shocks and Forest Use in Malawi. *World Development*, 33(7):1115–1128, jul 2005. ISSN 0305750X. doi: 10.1016/j.worlddev.2005.04.008. URL https://linkinghub.elsevier.com/retrieve/pii/S0305750X05000707.

Nicholas Flood. Continuity of reflectance data between landsat-7 etm+ and landsat-8 oli, for both top-of-atmosphere and surface reflectance: A nasa product. *Remote Sensing*, 6 (9):7952–7970, 2014. doi: 10.3390/rs6097952.

Jonathan A. Foley, Ruth DeFries, Gregory P. Asner, Carol Barford, Gordon Bonan, Stephen R. Carpenter, F. Stuart Chapin, Michael T. Coe, Gretchen C. Daily, Holly K. Gibbs, Joseph H. Helkowski, Tracey Holloway, Erica A. Howard, Christopher J. Kucharik, Chad Monfreda, Jonathan A. Patz, I. Colin Prentice, Navin Ramankutty, and Peter K. Snyder. Global Consequences of Land Use. *Science*, 309(5734):570–574, jul 2005. ISSN 0036-8075. doi: 10.1126/science.1111772. URL https://www.science.org/doi/10.1126/science.1111772.

S Furby and N Campbell. Land-cover change: Detection and classification. *Remote Sensing and Environment*, 25(1):57–66, 1995a.

S. Furby and N. A. Campbell. Calibrating images from different sensors to a common radiometric scale. *Remote Sensing of Environment*, 53:311–321, 1995b. doi: 10.1016/0034-4257(95)00039-9.

Bo-Cai Gao. Ndwi—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3):257–266, 1996.

Dennis P. Garrity, Festus K. Akinnifesi, Oluyede C. Ajayi, Solomon G. Weldesemayat, Jeremiah G. Mowo, Antoine Kalinganire, Mahamane Larwanou, and Jules Bayala. Evergreen agriculture: A robust approach to sustainable food security in africa. *Food Security*, 2:197–214, 2010. doi: 10.1007/s12571-010-0070-7.

Pedram Ghamisi, Javier Plaza, Yun Chen, Jun Li, Antonio Plaza, and Jon Atli Benediktsson. Remote sensing image classification using gabor textures and convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9):7096–7112, 2019.

Global Forest Watch. Global Forest Watch Dashboard, 2024. URL https://www.globalforestwatch.org/dashboards/global/.

Luc Gnacadja and Susanne Wiese. The great green wall for the sahara and sahel initiative. *Nature and Faune*, 27:7–14, 2013.

Scott J Goetz, Benjamin Bond-Lamberty, Beverly E Law, JA Hicke, C Huang, RA Houghton, Steve McNulty, Tom O'Halloran, Mark Harmon, AJH Meddens, et al. Observations and assessment of forest carbon dynamics following disturbance in north america. *Journal of Geophysical Research: Biogeosciences*, 117(G2), 2012.

JP Guerschman, MJ Hill, LJ Renzullo, et al. Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the australian tropical savanna region up to 2017. *Remote Sensing of Environment*, 113:928–945, 2009. doi: 10.1016/j.rse.2008.11.009.

César Gómez, Joanna White, and Michael Wulder. Radiometric normalization of multitemporal remote sensing imagery using pseudo-invariant features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 121:1–16, 2016. doi: 10.1016/j.isprsjprs.2016.09.002.

James Hansen, Pushker Kharecha, Makiko Sato, Valerie Masson-Delmotte, Frank Ackerman, David J. Beerling, Paul J. Hearty, Ove Hoegh-Guldberg, Shi-Ling Hsu, Camille

Parmesan, Johan Rockstrom, Eelco J. Rohling, Jeffrey Sachs, Pete Smith, Konrad Steffen, Lise Van Susteren, Karina von Schuckmann, and James C. Zachos. Assessing "Dangerous Climate Change": Required Reduction of Carbon Emissions to Protect Young People, Future Generations and Nature. *PLoS ONE*, 8(12):e81648, dec 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0081648. URL https://dx.plos.org/10.1371/journal.pone.0081648.

M. C. Hansen, R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg. Global percent tree cover at a spatial resolution of 500 meters: First results of the modis vegetation continuous fields algorithm. *Earth Interactions*, 7:1–15, 2003. doi: 10.1175/1087-3562(2003)007⟨0001:GPTCAA⟩2.0.CO;2.

Nancy L. Harris, David A. Gibbs, Alessandro Baccini, Richard A. Birdsey, Sytze de Bruin, Mary Farina, Lola Fatoyinbo, Matthew C. Hansen, Martin Herold, Richard A. Houghton, Peter V. Potapov, Daniela Requena Suarez, Rosa M. Roman-Cuesta, Sassan S. Saatchi, Christy M. Slay, Svetlana A. Turubanova, and Alexandra Tyukavina. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*, 11(3):234–240, mar 2021. ISSN 1758-678X. doi: 10.1038/s41558-020-00976-6. URL https://www.nature.com/articles/s41558-020-00976-6.

Isabelle S. Helfenstein, Fabian D. Schneider, Michael E. Schaepman, and Felix Morsdorf. Assessing biodiversity from space: Impact of spatial and spectral resolution on trait-based functional diversity. *Remote Sensing of Environment*, 275, 2022. ISSN 00344257. doi: 10.1016/j.rse.2022.113024.

Dorothée Herr and Ed Pidgeon. Blue carbon: Coastal ecosystems, their carbon storage, and the role of conservation. *Conservation Biology*, 29(3):763–767, 2015. doi: 10.1111/cobi.12477.

Stéphanie Horion, Rasmus Fensholt, Torbern Tagesson, and Andrea Ehammer. Using earth observation-based dry season NDVI trends for assessment of changes in tree cover in the Sahel. *International Journal of Remote Sensing*, 35(7):2493–2515, apr 2014. ISSN 0143-1161. doi: 10.1080/01431161.2014.883104. URL https://www.tandfonline.com/doi/full/10.1080/01431161.2014.883104.

Alfredo Huete, Chris Justice, and W Van Leeuwen. Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sensing of Environment*, 83(1):195–213, 2002.

IPCC. *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*. Intergovernmental Panel on Climate Change (IPCC), 2019. URL https://www.ipcc.ch/srccl/.

Lino Jastrowski. Small-footprint, full-waveform lidar data for tree species classification and biodiversity monitoring in riparian ecosystems, 2022. Bachelor's thesis, University of Zurich.

Sylvain Jay, Xavier Hadoux, Nathalie Gorretta, and Gilles Rabatel. Potential of hyperspectral imagery for nitrogen content retrieval in sugar beet leaves. 2014. URL https://api.semanticscholar.org/CorpusID:105015346.

Malin Falkenmark Johan Rockström. Water resilience for human prosperity. *Cambridge University Press*, 2014. doi: 10.1017/CBO9781139137005.

Kartik Kalyanram, Radha Gopalan, and Kamakshi Kartik. Tensions in Livelihoods: A Rural Perspective. pages 377–395. 2014. doi: 10.1007/978-1-4614-9460-7_21. URL https://link.springer.com/10.1007/978-1-4614-9460-7_21.

David Kaplan. *Structural Equation Modeling: Foundations and Extensions*. Sage Publications, 2009.

Radhika Kapur. Livelihood Opportunities in Rural Areas. *Acta Scientific Agriculture*, 3 (7):225–233, jun 2019. ISSN 2581365X. doi: 10.31080/ASAG.2019.03.0549. URL https://actascientific.com/ASAG/pdf/ASAG-03-0549.pdf.

Habtemariam Kassa, Stefaan Dondeyne, Jean Poesen, Amaury Frankl, and Jan Nyssen. Community-based forest management in ethiopia: Patterns, constraints, and implications. *Forest Policy and Economics*, 83:37–47, 2017. doi: 10.1016/j.forpol.2017.05.011.

Rex B. Kline. *Principles and Practice of Structural Equation Modeling*. Guilford Press, New York, fourth edition, 2015a. ISBN 978-1-4625-2334-4.

Rex B Kline. *Principles and Practice of Structural Equation Modeling*. Guilford Publications, 2015b.

Andrew T. Knight, Carly N. Cook, Kent H. Redford, Duan Biggs, Claudia Romero, Alejandro Ortega-Argueta, Cameron D. Norman, Beverly Parsons, Martin Reynolds, Glenda Eoyang, and Matt Keene. Improving conservation practice with principles and tools from systems thinking and evaluation. *Sustainability Science*, 14(6):1531–1548, nov 2019. ISSN 1862-4065. doi: 10.1007/s11625-019-00676-x. URL http://link.springer.com/10.1007/s11625-019-00676-x.

Angela Lausch, Stefan Erasmi, Douglas J. King, Paul Magdon, and Marco Heurich. Understanding forest health with remote sensing-Part I-A review of spectral traits, processes and remote-sensing characteristics. *Remote Sensing*, 8(12):1–44, 2016. ISSN 20724292. doi: 10.3390/rs8121029.

Simon L. Lewis, David P. Edwards, and David Galbraith. Increasing human dominance of tropical forests. *Science*, 349(6250):827–832, aug 2015. ISSN 0036-8075. doi: 10.1126/science.aaa9932. URL https://www.science.org/doi/10.1126/science.aaa9932.

Jun Li, Zhaocong Wu, Zhongwen Hu, Canliang Jian, Shaojie Luo, Lichao Mou, Xiao Xiang Zhu, and Matthieu Molinier. A lightweight deep learning based cloud detection method for sentinel-2a imagery fusing multi-scale spectral and spatial features. *arXiv preprint arXiv:2105.00967*, 2021.

Yan Li, Bojie Fu, Yang Wang, Lindsay C. Stringer, and John Harris. Drivers and impacts of changes in china's drylands: An overview. *Land Degradation Development*, 31(11): 1423–1438, 2020. doi: 10.1002/ldr.3556.

Shunlin Liang and Wenqing Wang. Advances in radiometric correction techniques for satellite remote sensing imagery. *Remote Sensing of Environment*, 264:112567, 2021. doi: 10.1016/j.rse.2021.112567.

Peter A. Lindsey, Jennifer R. B. Miller, Lisanne S. Petracca, Lauren Coad, Amy J. Dickman, Katherine H. Fitzgerald, Mark V. Flyman, Paul J. Funston, Philipp Henschel, Samuel Kasiki, Andrew J. Loveridge, David W. Macdonald, Rosemary L. Mandisodza-Chikerema, Sean Nazerali, Andrew J. Plumptre, Roger Stevens, Harry W. Van Zyl, and Luke T. B. Hunter. More than $1,000,000$ km^2 of african protected areas lacks lions. *Biological Conservation*, 201:252–259, 2017. doi: 10.1016/j.biocon.2016.08.005.

Roderick JA Little and Donald B Rubin. *Statistical Analysis with Missing Data*. Wiley, 2002.

Dengsheng Lu, Guangxing Wang, Emilio Moran, Guiying Li, and Scott Hetrick. Remote sensing of vegetation dynamics in drylands: A review of techniques and applications. *Remote Sensing of Environment*, 237:111466, 2020.

Anastassia M. Makarieva, Victor G. Gorshkov, and Douglas Sheil. Biotic pump of atmospheric moisture as driver of the hydrological cycle in forested landscapes. *Hydrology and Earth System Sciences*, 11(2):1013–1033, 2007. doi: 10.5194/hess-11-1013-2007. URL https://hess.copernicus.org/articles/11/1013/2007/.

Daniel Marusig, Francesco Petruzzellis, Martina Tomasella, Rossella Napolitano, Alfredo Altobelli, and Andrea Nardini. Correlation of Field-Measured and Remotely Sensed Plant Water Status as a Tool to Monitor the Risk of Drought-Induced Forest Decline. *Forests*, 11(1):77, jan 2020. ISSN 1999-4907. doi: 10.3390/f11010077. URL https://www.mdpi.com/1999-4907/11/1/77.

Philippe Mayaux, Ernest Bartholome, Steffen Fritz, and Alan Belward. A new land-cover map of africa for the year 2000. *Journal of Biogeography*, 31:861–877, 2004. doi: 10.1111/j.1365-2699.2004.01073.x.

Paul F. McCord, Jampel Dell'Angelo, Daniel Gower, Rob Marchant, and Tom P. Evans. Household-level heterogeneity of water resources within common-pool resource systems. *Ecology and Society*, 25(3):1–19, 2020. doi: 10.5751/ES-11673-250319.

Thomas O. McShane and Michael P. Wells. *Getting Biodiversity Projects to Work: Towards More Effective Conservation and Development*. Columbia University Press, New York, 2011. ISBN 978-0231127656.

Andrea Meraner, Patrick Ebel, Xiao Xiang Zhu, and Michael Schmitt. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. *ISPRS Journal of Photogrammetry and Remote Sensing*, 166:333–346, aug 2020. ISSN 09242716. doi: 10.1016/j.isprsjprs.2020.05.013. URL https://linkinghub.elsevier.com/retrieve/pii/S0924271620301398.

Marcus Moench. Adapting to climate change and the risks associated with other natural hazards: methods for moving from concepts to action. *The Earthscan reader on adaptation to climate change*, pages 249–280, 2009.

Douglas C Montgomery, Elizabeth A Peck, and G Geoffrey Vining. *Introduction to Linear Regression Analysis*. John Wiley & Sons, 2012.

Akira S. Mori, Forest Isbell, and Rupert Seidl. -diversity, community assembly, and ecosystem functioning. *Trends in Ecology Evolution*, 32(7):549–564, 2017. doi: 10.1016/j.tree.2017.05.010.

Andrew S. Nelson, Robert G. Wagner, Michael E. Day, Ivan J. Fernandez, Aaron R. Weiskittel, and Michael R. Saunders. Light absorption and light-use efficiency of juvenile white spruce trees in natural stands and plantations. *Forest Ecology and Management*, 376:158–165, sep 2016. ISSN 03781127. doi: 10.1016/j.foreco.2016.06.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0378112716303164.

Elinor Ostrom. A general framework for analyzing sustainability of social-ecological systems. *Science*, 325(5939):419–422, 2009. doi: 10.1126/science.1172133.

Ignacio Palomo, Carlos Montes, Berta Martín-López, José A. González, Marina García-Llorente, Paloma Alcorlo, and María Rosario García Mora. Incorporating the Social–Ecological Approach in Protected Areas in the Anthropocene. *BioScience*, 64(3):181–191, mar 2014. ISSN 1525-3244. doi: 10.1093/biosci/bit033. URL http://academic.oup.com/bioscience/article/64/3/181/224408/Incorporating-the-SocialEcological-Approach-in.

Dario Papale, Claudio Belli, Beniamino Gioli, Franco Miglietta, Cesare Ronchi, Francesco Vaccari, and Riccardo Valentini. ASPIS, A Flexible Multispectral System for Airborne Remote Sensing Environmental Applications. *Sensors*, 8(5):3240–3256, may 2008. ISSN 1424-8220. doi: 10.3390/s8053240. URL http://www.mdpi.com/1424-8220/8/5/3240.

Karl Pearson. *On the lines and planes of closest fit to systems of points in space*, volume 2. 1909.

Joana Pereira, Luís Miguel Rosalino, Serafino Mucova, Yasalde Massangue, Murchide Abdulrazak, Somar Vahossa, Mouzinho Selemane, Carlos Fonseca, and Maria João Santos. Livelihood vulnerability increases human–wildlife interactions. *Environmental Conservation*, 48(4):301–309, dec 2021. ISSN 0376-8929. doi: 10.1017/S037689292100028X. URL https://www.cambridge.org/core/product/identifier/S037689292100028X/type/journal_article.

O. Perez-Priego, J. Guan, M. Rossini, F. Fava, T. Wutzler, G. Moreno, N. Carvalhais, A. Carrara, O. Kolle, T. Julitta, M. Schrumpf, M. Reichstein, and M. Migliavacca. Sun-induced chlorophyll fluorescence and photochemical reflectance index improve remote-sensing gross primary production estimates under varying nutrient availability in a typical Mediterranean savanna ecosystem. *Biogeosciences*, 12(21):6351–6367, nov 2015. ISSN 1726-4189. doi: 10.5194/bg-12-6351-2015. URL https://bg.copernicus.org/articles/12/6351/2015/.

Nathalie Pettorelli, William F. Laurance, Thomas G. O'Brien, Martin Wegmann, Harini Nagendra, and William Turner. Satellite remote sensing for applied ecologists: Opportunities and challenges. *Journal of Applied Ecology*, 51(4):839–848, 2014. doi: 10.1111/1365-2664.12261.

Dirk Pflugmacher, Warren B. Cohen, Robert E. Kennedy, and Zhiqiang Yang. Using landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics. *Remote Sensing of Environment*, 122:123–139, 2012. doi: 10.1016/j.rse.2011.09.025.

Peter Potapov, Svetlana Turubanova, Matthew C Hansen, et al. Quantifying forest cover loss in democratic republic of the congo, 2000–2010, with landsat etm+ data. *Remote Sensing of Environment*, 122:106–116, 2012.

Jules Pretty and Zareen Pervez Bharucha. Sustainable intensification of agriculture: Greener and more productive approaches. *Frontiers in Sustainable Food Systems*, 2:1–15, 2018. doi: 10.3389/fsufs.2018.00001.

Virendra Proag. Assessing and Measuring Resilience. *Procedia Economics and Finance*, 18:222–229, 2014. ISSN 22125671. doi: 10.1016/S2212-5671(14)00934-4. URL https://linkinghub.elsevier.com/retrieve/pii/S2212567114009344.

Zou Qin. Methods of radiometric normalization for multi-temporal remote sensing images:a review. 2012. URL https://api.semanticscholar.org/CorpusID:133200311.

Nitin D. Rai, M. Soubadra Devy, T. Ganesh, R. Ganesan, Siddappa R. Setty, Ankila J. Hiremath, Sarala Khaling, and Priyadarsanan Dharma Rajan. Beyond fortress conservation: The long-term integration of natural and social science research for an inclusive conservation practice in India. *Biological Conservation*, 254:108888, feb 2021. ISSN 00063207. doi: 10.1016/j.biocon.2020.108888. URL https://linkinghub.elsevier.com/retrieve/pii/S0006320720309460.

C.Radhakrishna Rao. Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, 21(1):24–43, feb 1982. ISSN 00405809. doi: 10.1016/0040-5809(82)90004-1. URL https://linkinghub.elsevier.com/retrieve/pii/0040580982900041.

Jan Reiche, Martin Hamunyela, Johannes Verbesselt, Jean-Pierre W. Clevers, and Martin Herold. Improving near-real time deforestation monitoring in tropical dry forests by combining dense landsat time series and change probability mapping. *Remote Sensing of Environment*, 172:28–45, 2016. doi: 10.1016/j.rse.2015.10.012.

Duccio Rocchini, Matteo Marcantonio, Daniele Da Re, Gherardo Chirici, Marta Galluzzi, Jonathan Lenoir, Carlo Ricotta, Michele Torresani, and Guy Ziv. Time-lapsing biodiversity: An open source method for measuring diversity changes by remote sensing. *Remote Sensing of Environment*, 231:111192, sep 2019. ISSN 00344257. doi: 10.1016/j.rse.2019.05.011. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425719302056.

João Rodrigues, Luigi Boschetti, David P Roy, Maxine Muirhead, Lei Yan, and Zhe Li. Harmonization of sentinel-2 and landsat-8 imagery: Evaluation of radiometric differences and potential for synergy. *Remote Sensing of Environment*, 221:274–285, 2019.

D. Roe, F. Nelson, and C. Sandbrook. Community-based natural resource management: A promising solution to conservation challenges in africa. *Parks*, 21(2):67–82, 2015. doi: 10.2305/IUCN.CH.2015.PARKS-21-2DR.en.

Yves Rosseel. lavaan: An r package for structural equation modeling. *Journal of Statistical Software*, 48(2):1–36, 2012.

Christian Rossi, Mathias Kneubühler, Martin Schütz, Michael E. Schaepman, Rudolf M. Haller, and Anita C. Risch. Remote sensing of spectral diversity: A new methodological approach to account for spatio-temporal dissimilarities between plant communities. *Ecological Indicators*, 130:108106, 2021. ISSN 1470160X. doi: 10.1016/j.ecolind.2021.108106. URL https://doi.org/10.1016/j.ecolind.2021.108106.

David P. Roy, Michael A. Wulder, Thomas R. Loveland, C.E. Woodcock, Robert G. Allen, Martha C. Anderson, Dennis Helder, James R. Irons, Douglas M. Johnson, and Robert E. Kennedy. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 185:312–322, 2016. doi: 10.1016/j.rse.2015.11.032.

DP Roy, Z Li, H Zhang, et al. Cloud masking and gap filling of landsat imagery for long-term analysis. *Remote Sensing of Environment*, 239:111632, 2020. doi: 10.1016/j.rse.2020.111632.

Sarah Ruiz. Global forest carbon storage, explained, 2024. URL https://www.woodwellclimate.org/global-forest-carbon-storage-explained/.

Steven W Running, Ramakrishna R Nemani, Faith Ann Heinsch, Maosheng Zhao, Matt Reeves, and Hirofumi Hashimoto. A Continuous Satellite-Derived Measure of Global Terrestrial Primary Production. *BioScience*, 54(6):547–560, jun 2004. ISSN 0006-3568. doi: 10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2. URL https://doi.org/10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.COhttp://0.0.0.2.

Umar Saeed, Hammad Hayder, Asifullah Khan, and Saif ur Rehman. Cloud detection in remote sensing images: A comprehensive review. *IEEE Geoscience and Remote Sensing Magazine*, 9(3):83–105, 2021.

Susannah M Sallu, Chasca Twyman, and Lindsay C Stringer. Resilient or Vulnerable Livelihoods? Assessing Livelihood Dynamics and Trajectories in Rural Botswana. *Ecology and Society*, 15(4), 2010. ISSN 17083087. URL http://www.jstor.org/stable/26268197.

Indra Prasad Sapkota and Per Christer Odén. Household characteristics and dependency on community forests in terai of nepal. 2008. URL https://api.semanticscholar.org/CorpusID:54580832.

David Schimel, Ryan Pavlick, Joshua B. Fisher, Gregory P. Asner, Sassan Saatchi, Deborah B. Clark, Jeffrey Chambers, Jose Gomez-Dans, Isaac Colon-Bernal, Steve Frolking, Michael Keller, Marcos Longo, Forrest Hall, Ralph Dubayah, and George Hurtt. Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology*, 21(5): 1762–1776, 2015. doi: 10.1111/gcb.12822.

M. Schlemmera, Anatoly A. Gitelsonb, J. S. Schepersa, R. Fergusona, Yu Pengb, J. F. Shanahana, and D. Rundquistb. Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. 2013. URL https://api.semanticscholar.org/CorpusID:39033242.

JR Schott, Carl Salvaggio, and William J Volchok. Radiometric normalization of multitemporal satellite imagery. *Remote Sensing of Environment*, 26(1):1–16, 1988.

Randall E. Schumacker and Richard G. Lomax. *A Beginner's Guide to Structural Equation Modeling*. Routledge, New York, third edition, 2010a. ISBN 978-1-84872-835-6.

Randall E Schumacker and Richard G Lomax. *Structural Equation Modeling*. Taylor & Francis, 2010b.

Ian Scoones. Please scroll down for article journal of peasant studies livelihoods perspectives and rural development livelihoods perspectives and rural development. 1998. URL https://api.semanticscholar.org/CorpusID:16662693.

Ian Scoones. Livelihoods perspectives and rural development. *The Journal of Peasant Studies*, 36(1):171–196, jan 2009. ISSN 0306-6150. doi: 10.1080/03066150902820503. URL http://www.tandfonline.com/doi/full/10.1080/03066150902820503.

Julio Belnap Seth M. Munson, Robert H. Webb. Drought-induced state change in an erosive desert grassland. *Ecological Applications*, 21(5):1400–1415, 2011. doi: 10.1890/10-0895.1.

Charlie M. Shackleton, Sheona E. Shackleton, and Ben Cousins. The role of non-timber forest products in poverty alleviation and food security: Evidence from south africa. *Forest Policy and Economics*, 9(5):558–577, 2007. doi: 10.1016/j.forpol.2006.03.004.

Sheona Shackleton, William J. Sutherland, Charlie Shackleton, and Heike J. Geist. Livelihoods, conservation and rural development: Key issues facing southern africa. *Biological Conservation*, 187:181–192, 2015. doi: 10.1016/j.biocon.2015.04.027.

Douglas Sheil. Forests, atmospheric water and an uncertain future: the new biology of the global water cycle. *Forest Ecosystems*, 5(1):19, dec 2018. ISSN 2197-5620. doi: 10.1186/s40663-018-0138-y. URL https://forestecosyst.springeropen.com/articles/10.1186/s40663-018-0138-y.

Douglas Sheil and Daniel Murdiyarso. How forests attract rain: An examination of a new hypothesis. *BioScience*, 59(4):341–347, 2009. doi: 10.1525/bio.2009.59.4.12. URL https://academic.oup.com/bioscience/article/59/4/341/227927.

Masanobu Shimada, Takeo Itoh, Takahiro Motooka, Masaki Watanabe, Takeshi Shiraishi, Rajesh Thapa, and Richard Lucas. New global forest/non-forest maps from alos palsar data (2007–2010). *Remote Sensing of Environment*, 155:13–31, 2014. doi: 10.1016/j.rse.2014.04.014.

Eduarda M.O. Silveira, Volker C. Radeloff, Sebastian Martinuzzi, Guillermo J. Martínez Pastur, Luis O. Rivera, Natalia Politi, Leonidas Lizarraga, Laura S. Farwell, Paul R. Elsen, and Anna M. Pidgeon. Spatio-temporal remotely sensed indices identify hotspots of biodiversity conservation concern. *Remote Sensing of Environment*, 258(March):112368, 2021. ISSN 00344257. doi: 10.1016/j.rse.2021.112368. URL https://doi.org/10.1016/j.rse.2021.112368.

Ben Somers and Gregory P. Asner. Hyperspectral Time Series Analysis of Native and Invasive Species in Hawaiian Rainforests. *Remote Sensing*, 4(9):2510–2529, aug 2012. ISSN 2072-4292. doi: 10.3390/rs4092510. URL http://www.mdpi.com/2072-4292/4/9/2510.

Corinne Stucker, Vivien Sainte Fare Garnot, and Konrad Schindler. U-tilise: A sequence-to-sequence model for cloud removal in optical satellite time series. *arXiv preprint arXiv:2305.13277*, 2023.

Zheng Sun, Di Wu, Peng Zhang, et al. A review of cloud detection methods in optical satellite remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 176: 191–209, 2021. doi: 10.1016/j.isprsjprs.2021.05.012.

Sandeep Tambe. *Teaching and Learning Rural Livelihoods*. Sustainable Development Goals Series. Springer International Publishing, Cham, 2022a. ISBN 978-3-030-90490-6. doi: 10.1007/978-3-030-90491-3. URL https://link.springer.com/10.1007/978-3-030-90491-3.

Sandeep Tambe. Evolution of Rural Livelihood Approaches. pages 3–13. 2022b. doi: 10.1007/978-3-030-90491-3_1. URL https://link.springer.com/10.1007/978-3-030-90491-3_1.

The World Bank. Poverty headcount ratio at $2.15 a day (2017 PPP) (URL.

The World Bank. Access to electricity (URL https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=MZ.

The World Bank. Rural population (URL https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=MZ.

Maria Uriarte B. L. Turner II Ruth DeFries Deborah Lawrence Jacqueline Geoghegan Susanna Hecht Amy Ickowitz Eric F. Lambin Tom Birkenholtz Scott Baptista Rinku Grau Thomas K. Rudel, Laura Schneider. Agricultural intensification and changes in cropland expansion: A global analysis. *Ambio*, 38:109–119, 2009. 10.1579/0044-7447-38.2.109.

S. Trumbore, P. Brando, and H. Hartmann. Forest health and global change. *Science*, 349(6250):814–818, aug 2015. ISSN 0036-8075. 10.1126/science.aac6759. URL https://www.science.org/doi/10.1126/science.aac6759.

Compton J Tucker. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2):127–150, 1979.

Compton J Tucker, Jorge E Pinzon, and Molly E Brown. Remote sensing of drought: Progress, challenges and future opportunities. *Reviews of Geophysics*, 57(1):1–30, 2019.

Svetlana Turubanova, Peter V. Potapov, Alexandra Tyukavina, and Matthew C. Hansen. Ongoing primary forest loss in brazil, indonesia, and the democratic republic of the congo. *Environmental Research Letters*, 13(7):074028, 2018. 10.1088/1748-9326/aacd1c.

John Twigg, Benfield Greig, and Sri Lanka. Sustainable Livelihoods and Vulnerability. *Management*, (March):10–11, 2001. URL http://www.reliefweb.int/rw/lib.nsf/db900sid/FBUO-7N8JJH/$file/benfield_march2001.pdf?openelement.

Alexandra Tyukavina, Matthew C. Hansen, Peter V. Potapov, and Stephen V. Stehman. Congo basin forest loss dominated by increasing smallholder clearing. *Science Advances*, 4, 2018. 10.1126/sciadv.aat2993.

Population Division United Nations Department of Economic and Social Affairs. Mozambique Population 1950-2024, 2024. URL https://www.macrotrends.net/global-metrics/countries/MOZ/mozambique/population.

Stefan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014. 10.7717/peerj.453.

Nathalie van Vliet, Ole Mertz, Andreas Heinimann, Tobias Langanke, Unai Pascual, Birgit Schmook, Charlotte Adams, Dietrich Schmidt-Vogt, Peter Messerli, Stephen Leisz, Jean-Christophe Castella, Lene Jorgensen, Torben Birch-Thomsen, Cornelia Hett, Torben Bech-Bruun, Amy Ickowitz, Kim Chi Vu, Kiyono Yasuyuki, Zhen Ma, and Jefferson Fox. Trends, drivers and impacts of changes in swidden cultivation in tropical forest-agriculture frontiers: A global assessment. *Global Environmental Change*, 22:418–429, 2012. 10.1016/j.gloenvcha.2011.10.009.

Katharina Waha, Francesco Accatino, Cecile Godde, Cyrille Rigolot, Jessica Bogard, Joao Pedro Domingues, Elisabetta Gotor, Mario Herrero, Guillaume Martin, Daniel Mason-D'Croz, Francesco Tacconi, and Mark van Wijk. The benefits and trade-offs of agricultural diversity for food security in low- and middle-income countries: A review of existing knowledge and evidence. *Global Food Security*, 33:100645, jun 2022. ISSN 22119124. 10.1016/j.gfs.2022.100645. URL https://linkinghub.elsevier.com/retrieve/pii/S2211912422000359.

Siheng Wang, Lifu Zhang, Changping Huang, and Na Qiao. An NDVI-Based Vegetation Phenology Is Improved to be More Consistent with Photosynthesis Dynamics through Applying a Light Use Efficiency Model over Boreal High-Latitude Forests. *Remote Sensing*, 9(7):695, jul 2017. ISSN 2072-4292. 10.3390/rs9070695. URL https://www.mdpi.com/2072-4292/9/7/695.

Gary R. Watmough, Charlotte L.J. Marcinko, Clare Sullivan, Kevin Tschirhart, Patrick K. Mutuo, Cheryl A. Palm, and Jens Christian Svenning. Socioecologically informed use of remote sensing data to predict rural household poverty. *Proceedings of the National Academy of Sciences of the United States of America*, 116(4):1213–1218, 2019. ISSN 10916490. 10.1073/pnas.1812969116.

WEI Nan, ZHANG Mi, WANG Huimin, ZHANG Leiming, WEN Xuefa, and LIU Shoudong. The impacts of changes in diffuse radiation on light use efficiency in a subtropical plantation coniferous forest. *Acta Ecologica Sinica*, 37(10), 2017. ISSN 1000-0933. 10.5846/stxb201603080411. URL http://www.ecologica.cn/stxb/ch/reader/view_abstract.aspx?doi=10.5846/stxb201603080411.

Paige West, Jim Igoe, and Dan Brockington. Parks and peoples: The social impact of protected areas. *Annual Review of Anthropology*, 35:251–277, 2006. 10.1146/annurev.anthro.35.081705.123308.

Wildlife Conservation Society. Niassa Special Reserve, 2021. URL https://mozambique.wcs.org/Wild-Places/Niassa-Special-Reserve.aspx.

Kenneth G. Cassman William F. Laurance, Jeffrey Sayer. Agricultural expansion and its impacts on tropical nature. *Trends in Ecology Evolution*, 29(2):107–116, 2014. 10.1016/j.tree.2013.12.001.

Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. Cengage Learning, Boston, MA, seventh edition, 2019. ISBN 978-1-337-55886-0.

World Food Programme. Food security and livelihoods under a changing climate in Mozambique. Technical Report March, Rome, Italy, 2021.

Boris Worm, Edward B. Barbier, Nicola Beaumont, J. Emmett Duffy, Carl Folke, Benjamin S. Halpern, Jeremy B. C. Jackson, Heike K. Lotze, Fiorenza Micheli, Stephen R. Palumbi, Enric Sala, Kimberley A. Selkoe, John J. Stachowicz, and Reg Watson. Impacts of Biodiversity Loss on Ocean Ecosystem Services. *Science*, 314(5800):787–790, nov 2006. ISSN 0036-8075. 10.1126/science.1132294. URL https://www.science.org/doi/10.1126/science.1132294.

Chaoyang Wu, Jing M. Chen, and Ni Huang. Predicting gross primary production from the enhanced vegetation index and photosynthetically active radiation: Evaluation and calibration. *Remote Sensing of Environment*, 115(12):3424–3435, dec 2011. ISSN 00344257. 10.1016/j.rse.2011.08.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425711002884.

Steven Young, Brian Brisco, Kevin Murnaghan, Francois Charbonneau, and Ali Mahdavi. Cloud removal in satellite images using deep learning: A review and perspective. *International Journal of Applied Earth Observation and Geoinformation*, 104:102532, 2021.

Wei Zhang, Xin Li, Yufei Guo, et al. Threshold determination for ndvi-based land cover classification using sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164:35–47, 2020.

Maosheng Zhao and Steven W. Running. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science*, 329(5994):940–943, 2010. 10.1126/science.1192666.

Z. Zhu and C. E. Woodcock. Improving cloud filtering for landsat data: New methods and evaluation. *Remote Sensing of Environment*, 144:137–149, 2015. 10.1016/j.rse.2014.01.013.

Zhe Zhu and Curtis E Woodcock. Optimized cloud masking for landsat imagery. *Remote Sensing of Environment*, 194:432–450, 2017.

Zhe Zhu, S Wang, and CE Woodcock. Improving modis cloud masking through spectral and temporal information fusion. *Remote Sensing of Environment*, 171:35–46, 2015. 10.1016/j.rse.2015.09.016.

Declaration of Independent Work

Personal declaration: I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

31.01.2025, ZH