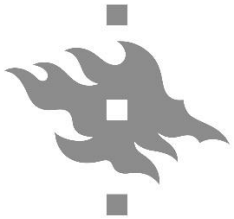

Master's thesis in Geography

Geoinformatics

Modelling dry season grass biomass in savanna rangelands in Kenya using imaging spectroscopy

Linda Pesonen

2024

Supervisors:
Ilja Vuorinne
Janne Heiskanen
Petri Pellikka

Master's Programme in Geography

Faculty of Science

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of Science	Department of Geosciences and Geography

Author

Linda Pesonen

Title of thesis

Modelling dry season grass biomass in savanna rangelands in Kenya using imaging spectroscopy

Programme and study track

Master's programme in geography, Geoinformatics

Level of the thesis	Date	Number of pages
Master's thesis, 30 credits	January 2024	70 pages

Abstract

Grass biomass has many important and diverse roles for ecosystems functioning, the carbon cycle, rangeland productivity and local livelihoods. Quantifying and understanding grass biomass in dynamic savanna ecosystems during dry season is important for sustainable land management and monitoring grazing pressures, especially amidst climate change. Traditional ground-based methods to assess vegetation are subjective and time consuming, while remote sensing provides efficiency in monitoring grass biomass at large scales. Grass biomass assessments using remote sensing data have been extensively conducted worldwide, but such research in African savannas remains rare.

This study aimed to study connections between dry season grass biomass measured in savanna rangelands and airborne hyperspectral imagery data obtained simultaneously in LUMO Conservancy area of South-Eastern Kenya. Two modelling techniques were compared: averaged plot values (n=24) and individual sample values (n=96). Three vegetation indices (RSI, NDSI, RDSI) were computed and Generalised Additive Models (GAM) were applied to portray the relationship between measured grass biomass and VIs. The highest explanatory power for both modelling techniques was found with RSI and NDSI indices with averaged plot level values having the highest performance ($D^2 = 0.79$, RMSE = 40.15 g/m^2), with the band combination of B78 and B43 (908 nm / 667 nm). The best performing vegetation index (RSI) was used to predict grass biomass in the study area, which indicated a biomass range of 0 to 2894 g/m^2.

The study highlights the potential of using hyperspectral imagery to assess grass biomass in the savanna environments. However, challenges and limitations were faced related to the heterogeneous nature of savannas, varying weather conditions affected by rainfall, the temporal limits of the study, and disturbances in spectral information caused by heavily grazed areas, dead material, and preprocessing techniques. It is suggested that future research considers these factors by incorporating a broader set of variables, extending the duration of the study, exploring various preprocessing techniques, increasing the sample size, and employing additional data sources, such as active sensors and hyperspectral satellite imagery, to enhance model performance and improve accuracy.

Keywords

grass biomass, remote sensing, vegetation indices, GAM, linear regression, hyperspectral data, imaging spectroscopy, savanna, Kenya, rangelands, agropastoral

Where deposited

University of Helsinki electronic theses library E-thesis/HELDA

Additional information

Contents

Abbreviations

CART	Classification and Regression Trees
CB	Cubist
CCCI	Canopy Chlorophyll Content index
CHM	Canopy Height Model
DSM	Digital Surface Model
EVI	Enhanced Vegetation Index
GAM	Generalised additive model
LAI	Leaf area index
LiDAR	Light detecting and ranging
LOOCV	Leave-one-out cross-validation
MAE	Mean absolute error
NDVI	Normalized difference vegetation index
NIR	Near infrared
NRDVI	Near Infrared Reciprocal Vegetation Index
PLS	Partial Least Squares regression
RDSI	Reciprocal difference index
RSI	Ratio based index
SAVI	Soil Adjusted Vegetation
SR	Simple ratio
SVM	Support Vector Machine
SWIR	Short Wave Infrared
VI	Vegetation index

1. Introduction

Grass biomass has many important and diverse roles for ecosystem functioning, the carbon cycle, rangeland productivity and local livelihoods (Ali et al., 2017; Vertès et al., 2007). Quantifying and understanding grass biomass during dry season in dynamic savanna ecosystems is important for sustainable land management and monitoring grazing pressures, especially amidst climate change (Ramoelo et al., 2012). Over one-fifth of the world's population reside in or around savannas and many of these local communities rely on pastoralism or subsistence agriculture (Bouvet et al., 2018). Global change can impact the structure of these sensitive ecosystems and communities through alterations in rainfall and productivity of crops and pasture (Mistry, Jayalaxshmi and Beradi, Andrea, 2014; W. Zhang et al., 2019). Savannas are far less studied compared to other ecosystems like rainforests, despite the significance and fragility of savannas (Bouvet et al., 2018).

Traditional ground-based methods to assess vegetation are subjective and time consuming, while remote sensing provides efficiency in monitoring grass biomass at large scales (Xu et al., 2008). Accurate biomass estimates and modern techniques utilizing remote sensing methods enhance the science-based management of grassland ecological resources, monitoring of desertification, and ensuring sustainable long-term ecosystem management (Psomas et al., 2011; X. Zhang et al., 2019). Grass biomass can be assessed through various remote sensing techniques and methods but applying vegetation indices (VIs) is one of the most known processing methods (Tucker and Sellers, 1986). A widely popular approach for estimating grass biomass involves integrating remote sensed data with VIs and *in-situ* measurements (Anderson et al., 1993) to create regression models (Dusseux et al., 2015; Huang et al., 2013; Jin et al., 2014; Song et al., 2014). Assessing vegetation biomass with hyperspectral imagery and vegetation indices has been conducted in various savannas and grasslands globally including Alpine grasslands in Tibet, Brazilian Cerrados and North-West Mexico (Jacon et al., 2021; Kong et al., 2018; Raya-Sereno et al., 2021). Multispectral data can lose its sensitivity to variations in biomass values when it reaches a certain limit, resulting in a phenomena called saturation (Li et al., 2021). However, employing hyperspectral imagery and e.g. narrow bands for biomass estimation can provide means to overcome the saturation issue (Chen et al., 2009; O. Mutanga and Skidmore, 2004). Additionally, hyperspectral data offers a wide range of bands and wavelengths not available

in multispectral data (Galvão et al 2020). Currently, there's no literature in modelling grass biomass with hyperspectral imagery in savanna rangelands during dry season in Southeast Kenya.

The aim of this thesis was to explore the connections between airborne hyperspectral data and dry season grass biomass in savanna rangelands. This study focused on creating a modelling workflow with Generalised Additive Models (GAM) to quantify savanna grass biomass, evaluate the most important spectral features to explain grass biomass, and to compute optimal spectral band combinations for VIs (RSI, NDSI, RDSI) with two techniques to extract spectral values from the imagery. Finally, the objective was to map dry season grass biomass patterns and variation in the intensively grazed study area utilizing the best performing model and VI.

The following research questions were addressed:

1) What are the relationships between grass biomass and vegetation indices calculated from hyperspectral data?
2) How feasible is the use of airborne hyperspectral data in visible to range for modelling grass biomass patterns in savanna rangelands of Kenya during dry season?
3) What are the most important spectral features to explain grass biomass?
4) Based on best model predictions, how is grass biomass geographically distributed in the study area?

2. Background

2.1. The savanna ecosystem

Savannas are ecologically dynamic, heterogenous and affected by a combination of human-induced economic, social, and environmental elements in a world of globalization and global change (Hill et al., 2011). Accordingly, it is crucial to monitor key structural and functional elements of these complex ecosystems, such as biomass, spatial extent and biodiversity – especially as human need and land use changes are inevitable (Hill et al., 2011). Savannas are an important component of the terrestrial biosphere located close to the equator spreading out over broad plains of three continents (Pennington et al., 2018). They cover approximately 20 % of Earth's surface with most of these ecosystems located in Africa, but also spread widely across America, Asia, India and Australia (Fig. 1) (Pennington et al., 2018). They play an important role in the global carbon cycle, and annually sequester about 36 to 42 tonnes of carbon per hectare (Huston and Wolverton, 2009). Although the carbon density of African savannas and woodlands is lower compared to those of closed forests, these areas are greatly larger and cover about 50% of the continent and represent a significant carbon stock (Bartholomé and Belward, 2005). They represent roughly a third of the net primary production of all the terrestrial vegetation, making savannas a crucial regulating component of the land carbon sink (Ahlstrom et al., 2015; Moeckel et al., 2017; Poulter et al., 2014). A notable portion of this carbon is stored as aboveground biomass (AGB), which contains the carbon stored through photosynthesis within vegetation, making it both an essential climate variable as well as an important ecosystem service in savanna environments (Bojinski et al., 2014).

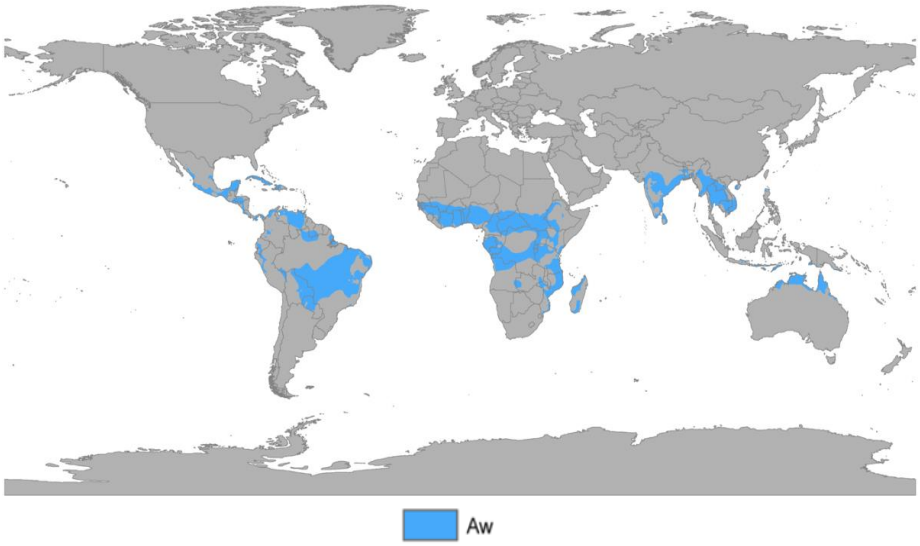

Fig. 1. Global distribution of savannas and dry forests. Savannas are spread widely across Africa, America, Asia and Australia. The blue areas of "Aw" category indicate a tropical savanna climate with a dry winter season (Figure from 'Köppen climate classification map' by Peel, M. C., Finlayson, B. L., and McMahon, T. A. (University of Melbourne), under CC BY-SA 3.0. Available at: https://openverse.org/image/6ff19ae0-4bdd-4ba1-a632-d83415ebd2bb?q=savanna%20map).

Savanna is a grassy plain characterized by the coexistence of sparse tree and shrub cover on a bed of grass, forming mixed woodland-grassland ecosystems as stated by Pennington et al. (2018) and Shorrocks and Bates (2014). Both woody and herbaceous plants are a result of long-term adaptation and have successfully overcome two major environmental challenges: seasonal drought and/or periodic phenomenon of burning and flood. These open tree canopies have a grassy ground layer, which provides forage for mammalian grazers such as rhinos, zebras, buffalos, and antelopes (Fig. 2.). Other common animals in the savannas are elephants, giraffes, lions, leopards, cheetahs, hyenas, and wildebeests, and some of these species migrate long distances between regions due to rainfall and forage availability (Shorrocks and Bates, 2014). Annual rainfall affects the quantity of plant biomass in an area, thus significantly impacting the feed of herbivores and carnivores (Pennington et al., 2018). Savannas are typically found in climatic regions between deserts and tropical forests (Shorrocks and Bates, 2014). Many factors contribute to the unique spatial structuring and heterogenous dynamic of these ecosystems (Shorrocks and Bates, 2014). The distribution and abundance of trees, shrub and grass are controlled by large herbivores, reoccurring fires, and a limited annual rainfall (Abdi et al., 2022; Shorrocks and Bates, 2014; W. Zhang et al., 2019).

Fig. 2. Commonly seen animals in the African savannas, such as zebras, buffalos, and giraffes. Grazing livestock can be seen browsing the landscape frequently (bottom right). Pictures taken in the study area of LUMO, January 2022.

The climate of savannas includes two distinct seasons; a growing season during which most of the photosynthesis takes place and a dry season lasting from 2 to 9 months, in which most of the herbaceous vegetation senesces and dries out (Eamus et al., 2016; Kahiu and Hanan, 2018; Mistry, Jayalaxshmi and Beradi, Andrea, 2014). The mean annual rainfall of savannas ranges

from 300 to 2500 mm and temperatures vary between 14 and 30°C (Harris et al., 2020, p. 4; Pennington et al., 2018). During the dry season, the grass layer dries which accumulates and forms as a fuel for future fires (Scogings and Sankaran, 2019). Savannas impact majorly to the seasonal fluctuations of global greenhouse gas emissions from fires, as this biome contributes more than 80 percent of the 4 million km^2 of land that burns worldwide every year (Scogings and Sankaran, 2019).

Bouvet et al (2018) note that savannas and grasslands are influenced by a range of ecological and anthropogenic factors in addition to climatic conditions like precipitation, although vegetation types usually align with rainfall patterns in Africa (Fig. 3). Mean Annual Precipitation (MAP) defines the maximum attainable woody cover in African savannas, however, other influencing factors include soil properties, topography, and human activities such as clearing and logging (Bouvet et al., 2018) Additionally, humans suppressing fires has led to increase in woody cover and shrubs (Archer et al., 2017). Many areas in Africa have the potential to turn into savannas or forests but are eventually formed based on the local circumstances (Bouvet et al., 2018).

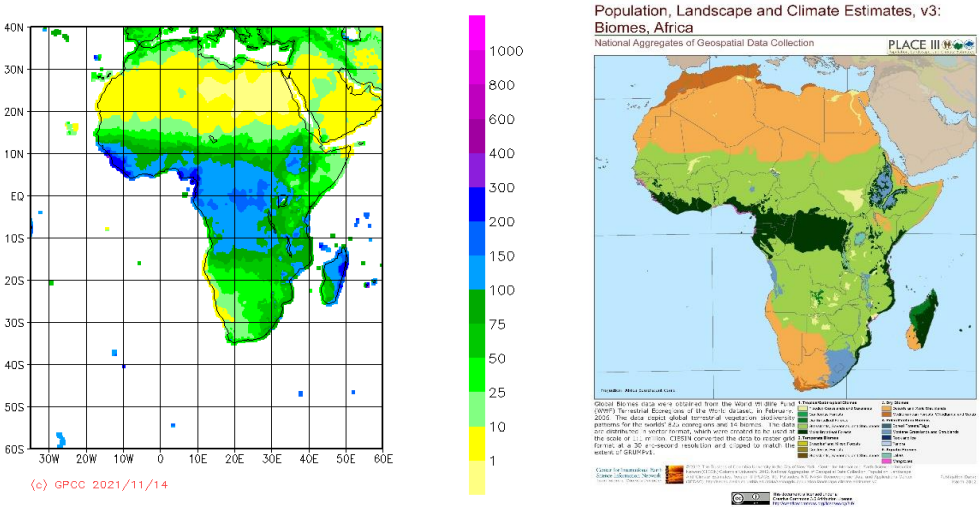

Fig. 3. Demonstration of precipitation patterns in Africa according to GPCC climatology on the left and biome map of Africa on the right where light green represents grasslands, savannas and shrublands (Left figure from "Contributions to the Improvement of Climate Data Availability and Quality for Sub-Saharan Africa (2022)" by Frank Kaspar, created using the GPCC Visualizer (https://kunden.dwd.de/GPCC/Visualizer), under CC BY-SA 3.0, available at: https://openverse.org/image/00b9625cb59d41c7ab05526142dc56a3?q=precipitation%20map%20africa. Right picture from SEDACMaps under a CC BY 2.0 license, available at: https://www.flickr.com/photos/54545503@N04/7242976044).

An analysis by Doi et al. (2022) reported of the severe drought in East Africa during 2021. They highlighted the critical need for advanced climate modelling as the impacts due to natural climate variability are having more serious effects under the ongoing global warming. These extreme dry conditions during short rains seasons led to severe food insecurity in the region and typically result in other critical issues too, such as unsafe drinking water and potential appearance of infectious diseases (Azage et al., 2017; Doi et al., 2022, 2020). Additionally, extremes weather conditions, such as drought, can lead to increase in conflicts (Michelini et al., 2023).

Over one-fifth of the world's population reside in or around savannas and many of these communities rely on pastoralism or subsistence agriculture through activities like subsistence farming, producing timber, livestock grazing and fuelwood harvesting (Bouvet et al., 2018; Chidumayo and Gumbo, 2010). Pellikka et al. (2013) highlighted the impacts and consequences of agricultural expansions in the Taita Hills of Kenya, adjacent to the study area of LUMO. The expansions can result in reduced soil water-holding capacity and increased risks of hydrological droughts during dry seasons can impact local economies, agriculture, and ecosystems (Alcántara-Ayala et al., 2006). Agricultural expansion might lead to increased food production for the growing population, but the level of productivity of these lands will remain uncertain, with various negative consequences not fully understood (Pellikka et al., 2013).

Global change in rainfall climatology may affect the productivity of crops and pasture and the structure of tropical savanna ecosystems by favouring woody plants (Mistry, Jayalaxshmi and Beradi, Andrea, 2014; W. Zhang et al., 2019). Increase in woody plants (woody encroachment) contributes to both carbon losses and gains (Bodart et al., 2013). Multiple factors lead to woody encroachment such as more intense rainfall, changing fire regimes, CO_2 fertilization, land use changes, grazing and increased aridity (Berdugo et al., 2020; Bond and Midgley, 2012; Devine et al., 2017; Hill and Hanan, 2010; Kulmatiski and Beard, 2013; O'Connor et al., 2014; Rosan et al., 2019; Stevens et al., 2017). While woody encroachment can result in carbon gains, it can also negatively impact the local communities by reducing grazing areas (Eldridge et al., 2011).

Moreover, climate change affects to the carbon balance of Africa, mainly through the carbon absorption and release by land ecosystems (Valentini et al., 2014). Savannas and woodlands contribute significantly to Africa's carbon emission due to the periodic fires (Valentini et al., 2014). Africa has a key role globally contributing to the interannual variations in the atmospheric CO2 levels also through deforestation (Brink and Eva, 2009; Ciais et al., 2011;

Williams et al., 2007). However, the exact impacts of these human activities, such as highly variable deforestation rates and changes in croplands are poorly estimated (Houghton and Hackler, 2006).

Savannas and tropical dry forests are in need of sustainable management and monitoring, as they are noticeably under-represented in protected areas (Pennington et al., 2018). Large-scale monitoring, including AGB estimation with remote sensing, is crucial for a better understanding and managing of the impacts of climate change and local activities on these ecosystems (Bouvet et al., 2018; Hill et al., 2011).

2.2. Role and importance of grass biomass in the savanna

The role and importance of grass biomass in African savannas is multidimensional and relates to wildlife and livestock feed. Pastoralism (livestock grazing) has a long history in Africa, as communities have been accustomed seasonal movement in search of greener pastures for their livestock as nomads (Shackleton et al., 2002). Livestock production is strongly dependent on the quality and quantity of grass and pastures and remains a major source of income in the rural economy (Shackleton et al., 2002). This is evident in areas like savanna lowlands of southern Kenya, where grazing is the primary land use (Sheila Wachiye et al., 2022). Pasture quality can be determined by leaf nutrient content and the quantity is estimated by biomass (mass per unit area) (McNaughton, 1988). Both quality and quantity are essential for understanding the distribution, population dynamics and feeding patterns of livestock and wildlife (Ramoelo et al., 2012). This geographic information of grass quality and quantity is important for effective planning and management of savanna rangelands. (Ramoelo et al., 2012).

Local farmers manage production systems based on spatial and temporal variations of the forage. The variability in forage provision is increasing due to climate change, which is further complicated by the heterogeneity of pastures varying in structure, composition, and phenology, as seen in Fig. 4 (Ferner et al., 2021; Moeckel et al., 2017). The effects of changes in plant species composition and the increasing pressure of population growth have resulted in a decrease in grass biomass productivity (Ferner et al., 2021). While grassland assessment studies using remote sensing data have been extensively conducted worldwide, such research in African savannas is rare (Zumo et al., 2022).

Fig. 4. Savannas can have great variation in quantity and quality of vegetation even within a small area. Pictures taken in LUMO Conservancy in January 2022.

Grass biomass plays many important and diverse roles for ecological aspects as well. Grasslands occupy at least two-thirds of global agricultural land and cover approximately 25 % of the Earth's terrestrial surface (Ferner et al., 2021; Wood et al., 2000). The amount of grass biomass directly affects to the carrying capacity of pastures and is crucial for providing forage for grazing livestock and wildlife (Ali et al., 2017). Grasslands are often transformed or maintained as agricultural lands due to their high productivity (Psomas et al., 2011). They also contribute to the diversity, cultural history and ecological goods and services of rural and agricultural landscapes (Psomas et al., 2011). Sustainable use of grasslands will be necessary to provide food for the estimated nine billion people inhabiting Earth by 2050 (Zhang et al., 2022). Additionally, grass biomass impacts ecosystem functioning, biodiversity, the carbon cycle, air and soil quality, and the regulation of water, nitrogen and pollutant flows (Batáry et al., 2007; Lobell and Field, 2007; Soussana and Lüscher, 2007; Vertès et al., 2007).

The assessment of biomass in savannas during dry season is important for several reasons, mostly related to grazing and livestock production. Ramoelo et al. (2012) note that during the dry season, the quantity of grass limits the grazers more than its quality, thus quantifying the variability of biomass during this time is crucial for decision makers in planning and management of the grazing systems. In addition, this information can be used for assessing fire risks since biomass acts as fuel load for fires. Consequently, having detailed information on the spatial distribution of grass biomass provides sustainability for livestock production, rangeland productivity and preservation of these ecosystems (Kumar and Mutanga, 2017; Ramoelo and Cho, 2014). Utilizing remote sensing methods for accurate biomass estimates is essential to enhancing the science-based management of grassland ecological resources, monitoring desertification, and ensuring long term ecosystem management (Psomas et al., 2011; X. Zhang et al., 2019).

2.3. Remote sensing of grass biomass

2.3.1. Principles of remote sensing

Remote sensing is the science and process of deriving physical information of the Earth's surface by measuring its reflected and emitted electromagnetic radiation from a distance, typically with imagery acquired by satellites or aircraft (Campbell and Wynne, 2011). Remote sensing methods enable quick and cost-effective assessment of biomass over large areas, each

having their own benefits depending on the application and context of the research. Assessing vegetation with traditional ground-based methods is subjective, time consuming and suitable only for monitoring at a limited scale (Xu et al., 2008). The literature of biomass assessments has highlighted various remote sensing techniques, such as Light Detection and Ranging (LiDAR), optical satellite and aerial imagery, synthetic aperture radar (SAR), unmanned aerial vehicles (UAV), thermal infrared sensors, and ground-based remote sensing with multispectral and hyperspectral cameras and spectrometers (Ahamed et al., 2011; Amiri et al., 2010; Kumar et al., 2015). Typically, remote sensing in this field uses two types of optical images: multispectral, which are the most used, and hyperspectral imagery. The coarse spatial and spectral resolution can limit the accuracy and robustness of research results (Wang et al., 2012). In the meantime, the utilizing of hyperspectral imagery especially in grassland monitoring have gained popularity in recent years (Balzarolo et al., 2015; Ling et al., 2019; Lyu et al., 2020).

Campbell and Wynne (2011) emphasize the importance of having a sound understanding of the concept of electromagnetic radiation to interpret remote sensing imagery correctly. Classically electromagnetic radiation is described as waves with a corresponding wavelength. Wavelength is the distance between two consecutive peaks of a wave, which can be used to characterize the magnitude of radiation, for example in nanometres (nm). The wavelength of radiation determines its colour and properties within the spectrum, which provides information that can be used to interpret vegetation characteristics such as biomass. Visible light is the most known form of electromagnetic radiation but large segments outside this range are of interest when studying vegetation (Fig. 5) (Campbell and Wynne, 2011). The segment important for remote sensing of vegetation is near infrared (NIR, 720-1500 nm), which is very wide relative to the visible region (Campbell and Wynne, 2011).

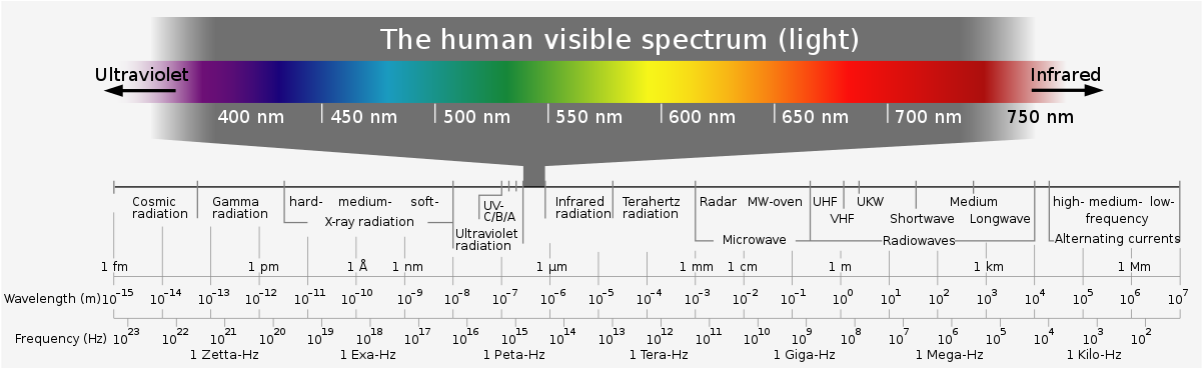

Fig. 5. The electromagnetic spectrum that shows the wavelength ranges of electromagnetic radiation (Figure from Horst Frank (2006), licensed under CC BY-SA 4.0. Available at: https://commons.wikimedia.org/w/index.php?curid=76714066).

Hyperspectral remote sensing, also known as imaging spectroscopy, is a powerful remote sensing technique for monitoring the Earth's atmosphere and surface (Guanter et al., 2019). Hyperspectral remote sensing has multiple advantages in capturing vegetation as it has the benefit of representing fine spectral signatures (Lyu et al., 2020; Räsänen et al., 2020; Shen et al., 2008). Hyperspectral sensors can collect vast amounts of information with their large number of narrow and contiguous spectral bands, enabling the detection of finer variations in reflectance and absorption patterns, which leads to higher mapping accuracies and novel ideas for mapping grassland biomass (Campbell and Wynne, 2011; Kong et al., 2018; Wang et al., 2012). Hyperspectral instruments provide a larger number of features for above ground biomass assessments, compared to multispectral sensors (Galvão et al., 2020). The large number of fine resolution bands enable the calculation of multiple narrowband indices (VIs) and absorption band parameters (Jacon et al., 2021). These features can be used as input variables for AGB estimates using field inventory and machine learning models (Jacon et al., 2021; Toniol et al., 2017).

Hyperspectral instruments differ from traditional spectroscopy by gathering spectra of areas on Earth's surface, not only point targets (Campbell and Wynne, 2011). Some hyperspectral sensors can acquire data across 200 or more bands at 10–12-bit radiometric resolution. (Campbell and Wynne, 2011). The fine spectral bands of hyperspectral sensors enables them to identify subtle spectral features and patterns that otherwise might be hidden in the broader bands of typical multispectral sensors (O. Mutanga and Skidmore, 2004). In this regard, various studies have assessed the efficacy of using hyperspectral data to overcome the traditional issues encountered with the wider bands of multispectral data, for example, the well-known saturation issue in biomass estimation, especially for grassland/vegetation or agricultural imagery (Adam and Mutanga, 2012; Compton J. Tucker, 1977; Imran et al., 2020; O. Mutanga and Skidmore, 2004).

However, employing hyperspectral data in applications has its own challenges, including high cost, limited availability, heavy processing, and high dimensionality which leads to costly processing (Adam and Mutanga, 2012). The cost of acquiring these data often ends up being a significant obstacle for conservation organizations, managers, and researchers (Adam and Mutanga, 2012; Tsalyuk, 2014). Atmosphere can substantially affect the data, especially with sensors carried by satellites (Campbell and Wynne, 2011). Atmospheric particles like water, smoke, clouds, haze as well as several physical processes, including scattering, absorption and refraction are impactful on the spectral information quality (Campbell and Wynne, 2011). Okin

et al. (2001) highlight some practical limits on hyperspectral vegetation differentiation in arid and semi-arid environments. Hyperspectral remote sensing faces challenges in retrieving vegetation types in areas with low vegetation cover, especially when the cover is below 30%. In many circumstances vegetation can still be estimated reliably, though overestimations of vegetation are likely. Dry plants do not often have strong spectral features, like distinguished red edge and deep chlorophyll absorptions, resulting in spectral variability and nonlinear mixing. Additionally, some plants in arid and semi-arid environments display evolutionary adaptations that make them able to survive high heat and drought, therefore affecting their spectral properties (Okin et al., 2001).

Regardless of the limits, new technologies and techniques are being developed continuously. Various applications exist in hyperspectral remote sensing beyond estimating biomass and detecting plant stress, such as monitoring of sea surface temperatures, coral reefs, mangrove forests and ocean colour (Atazadeh et al., 2021). Hyperspectral remote sensing is relevant in applications like land cover and land use change (Abera et al., 2022; Ahmad et al., 2023), environmental monitoring (Zhao et al., 2021) and conservation (Wang et al., 2010). Furthermore, its applications are gaining importance and relevance in various fields including, agriculture, forestry, geography, geology, archaeology, ecology and others (Campbell and Wynne, 2011).

2.3.2. Remote sensing methods for assessing grass biomass

Quantifying vegetation biomass with hyperspectral imagery and vegetation indices has been conducted in savannas and grasslands around the world, such as Alpine grasslands in Tibet, (Kong et al., 2018), Brazilian *Cerrados* (Jacon et al., 2021) and North-West Mexico (Raya-Sereno et al., 2021). Grass biomass can be assessed through various remote sensing techniques and methods, but applying vegetation indices is one of the most recognized processing methods (Gao et al., 2012). For over 40 decades, various methods have been developed to assess grassland biomass based on remote sensing data (Gao et al., 2012). The efficient assessment of aboveground net primary production (ANPP) is actively researched and can provide important insights about the productivity and ecosystem service value of grasslands, also serving as an indicator for monitoring vegetation degradation and productivity (Gao et al., 2012). Remote sensing can provide effective ways to capture vegetation conditions and useful approaches to monitor grassland biomass. Vegetation indices have been used widely in different applications

of remote sensing since the introduction of NDVI by Tucker (1977) which uses the red and NIR bands.

The most popular and well-studied approach for estimating grassland biomass involves combining remotely sensed data and *in-situ* measurements (Anderson et al., 1993) to develop regression models (Boschetti et al., 2007; Dusseux et al., 2015; Huang et al., 2013; Jin et al., 2014; Shen et al., 2008; Song et al., 2014; Ullah et al., 2012). Field measurements of biomass are unquestionably crucial for validating results and calibrating models in remote sensing methodologies (Adam et al., 2010). Several studies suggest that very high accuracies have been reached with a VI based regression models for biomass estimation (Bradford et al., 2005; Loris and Damiano, 2006; Olmanson et al., 2013; Vescovo and Gianelle, 2008; Wylie et al., 2002; Xu et al., 2007). However, a significant limitation of these models is that they are site specific and unable to learn the complex and non-linear patterns in the data (Ali et al., 2017).

Understanding the spectral characteristics of vegetation is crucial to interpret and process remote sensing data. The reflectance properties of vegetation are defined by various factors, such as leaf pigment, structure of plant cells and water content (Eamus et al., 2016; Knipling, 1970). Grass and other healthy vegetation have distinct reflectance patterns in the visible (400–700 nm) and NIR (700–1300 nm) regions of the electromagnetic spectrum (Fig. 6) (Collins, 1978). This is due to the absorption properties of biologically active pigments, such as chlorophylls, carotenoids, xanthophylls and anthocyanins, which result in low reflectance in the blue (450 nm) and red (680 nm) region and higher reflectance in NIR regions (Eamus et al., 2016). This gives the distinct characteristic green appearance to human eyes as green light (500–600 nm) is reflected more than absorbed. The reflectance is relatively high in the NIR region due to internal structure of the leaves and light scattering, and less reflectance is expected beyond 1300 nm in the infrared because light its strongly absorbed by water in the region (Eamus et al., 2016). The steep curve in between the red (680 nm) and NIR (780 nm) is referred to as the “red edge” area, which is defined by the abrupt rise in reflectance between the chlorophyll absorption feature in the red wavelengths and NIR region (Collins, 1978; Horler et al., 1983). However, when vegetation is under a physiological stress, for example during dry season, the most profound changes in the spectral characteristics are seen in the visible spectral region rather than the infrared, because chlorophyll is sensitive to physiological changes (Eamus et al., 2016).

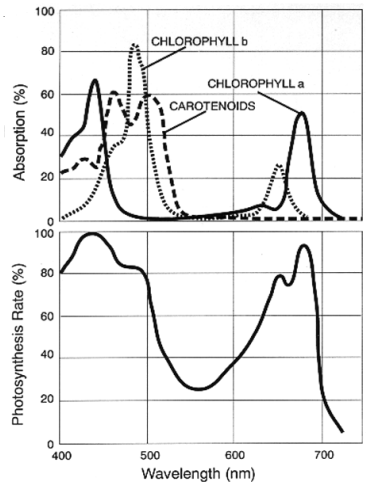

Fig. 6. The absorption spectra of typical photosynthetically active radiation, chlorophyll-A, chlorophyll-B and carotenoids (Figure from 'Concepts in Photobiology: Photosynthesis and Photomorphogenesis', edited by GS Singhal, G Renger, SK Sopory, K-D Irrgang, and Govindjee, Narosa Publishers/New Delhi; and Kluwer Academic/Dordrecht, pp. 11-51, by John Whitmarsh and Govindjee, licensed under CC BY-SA 2.0. Source: http://www.life.uiuc.edu/govindjee/paper/gov.html. From unpublished data available at: https://en.wikipedia.org/wiki/Normalized_difference_vegetation_index#/media/File:Par_action_spectrum.gif).

Various studies have reported of the importance of selecting appropriate wavelengths in grass biomass assessments through remote sensing techniques. A study by Mutanga and Skidmore (2004) states that combination of narrow bands (MNDVI) in the shorter wavelengths of the red edge (700–750 nm) and longer wavelengths of the red edge (750–780 nm) resulted in higher correlation with biomass, outperforming NDVI which yielded lowest correlations (O. Mutanga and Skidmore, 2004). Similarly, Kong et al. (2016) reported that good correlation with biomass was achieved with spectral bands of 550, 680, 860, and 900 nm as well as their combination forms.

Highlighted by Jackson and Huete (1991), vegetation indices (VIs) are metrics developed in the field of remote sensing to characterize various aspects of vegetation, such as type, amount, and condition by combining spectral data from different electromagnetic wavelengths. Simple measurements of light reflection from a surface are insufficient to provide a consistent picture of the surface due to differences in atmospheric conditions over time. Consequently, this problem can be partially addressed by combining data from two or more spectral bands to a vegetation index. They can be computed by calculating differences, ratios, combining ratios and differences with sums and forming linear combinations of spectral data. Jackson and Huete (1991) state that vegetation index is designed to minimize the solar irradiance variations and

soil background effects, while enhancing the vegetation signal. VIs are widely used and have established methodologies in numerous disciplines focused on biomass assessment, plant health, plant stress, crop production and water use. Despite developed for extracting spectral characteristics of vegetation, VIs are also influenced by various factors, such as moisture condition, soil background, atmospheric condition and the angles of the solar and the viewer (Jackson and Huete, 1991).

Due to the unique reflection pattern to vegetation, popular method to study the quantity of vegetation existing in a pixel of an imagery has been to compare the reflectance between red and NIR wavelengths (Eamus et al., 2016). Therefore, a higher ratio of NIR to red reflectance presumably indicates that there is an abundance of green vegetation. This ratio is the basis of various vegetation indices, such as Simple Ratio (SR), which was first reported by Jordan (1969). Later, research by Colwell (1974) found that the ratio of NIR and red light is a feasible method to assess grass biomass. The Simple Ratio is represented by the formulation:

$$\text{SR} = \frac{NIR}{red}$$

Additionally, plenty of other VIs have been developed throughout the years with the most popular being the normalized difference vegetation index (NDVI) by Deering (1978). It is formed by the difference of NIR, and red reflectance divided by their sum:

$$\text{NDVI} = \frac{NIR - red}{NIR + red}$$

Thus, the range of NDVI varies from -1 to +1 where negative values imply non-vegetated to sparse vegetation cover, and high positive values indicate productive green vegetation (Deering, 1979). SR and NDVI are mathematically equivalent thus providing the same information (Jackson and Huete, 1991). Another vegetation index, used for example in quantifying desert shrub communities (Mao et al., 2022) and studying optical characteristics of plant leaves (Gitelson et al., 2001), is referred to as reciprocal difference spectral index:

$$\text{RDSI} = \frac{1}{Band\ A} - \frac{1}{Band\ B}$$

Vegetation indices are known to be sensitive to background factors, especially in the case of high-resolution imagery. VIs can be affected by the background soil brightness stemming from bare ground, leading to uncertainties and overestimations in grass biomass assessment (Huete, 1988; Ren et al., 2018; Shi et al., 2021; Tang et al., 2019; Xue and Su, 2017). The ground surface spectral reflectance can also be affected by various elements in the landscape, such as soil colour (Gao et al., 2000), distribution of plant communities (Boelman et al., 2005), topography (Kawamura et al., 2005), and hydrology (Todd and Hoffer, 1998). Many studies have highlighted the saturation problem of NDVI at high canopy density biomass after a certain biomass density or leaf area index (LAI) which yields poor NDVI estimates in more densely vegetated areas or during peak growing season (Compton J. Tucker, 1977; Gao et al., 2000; Sellers, 1985; Thenkabail et al., 2000; Todd et al., 1998). NDVI computed from broad band sensors, typically multispectral, reach a saturation level after aboveground biomass is about 0.3 g/cm^{-1} or when vegetation cover is 100% (Hurcom and Harrison, 1998). NDVI is more effective when detecting sparse vegetation (Jackson and Huete, 1991). SR is highly sensitive to changes in dense vegetation but less effective in sparse vegetation (Jackson and Huete, 1991). Using narrow band vegetation indices formed from hyperspectral imagery for high canopy density biomass estimation has proved to overcome the saturation issue (Chen et al., 2009; O. Mutanga and Skidmore, 2004). Studies have shown that high accuracy in estimating biomass at full canopy cover can be reached when modifying vegetation indices to combine red edge and NIR wavelengths (O. Mutanga and Skidmore, 2004).

The improved performance of red edge based indices compared to the traditional NDVI could be explained by their sensitiveness to vegetation properties like chlorophyll content and canopy biomass versus other wavelengths of the electromagnetic spectrum (O. Mutanga and Skidmore, 2004; Onisimo Mutanga and Skidmore, 2004). Todd et al. (1998) reported that red edge displayed higher sensitivity to biomass variations in the green vegetation as compared to senescing vegetation. Also, Lucas et al. (2000) demonstrated a strong relationship between leaf area index (LAI) and the red edge position using airborne imagery. Minor changes in vegetation properties can lead to a shift in the red edge position, and additionally indices incorporating red edge and NIR can minimize the influence of the atmospheric absorption and soil background effects (Kokaly and Clark, 1999).

Modelling aboveground biomass can be divided into parametric and non-parametric models (Wang et al., 2022). Generally, parametric models consist of linear, logarithmic, exponential and other statistical regression methods (Chu, 2020; Grüner et al., 2019; Van Der Merwe et al., 2020; Zhang et al., 2018). Non-parametric models are primarily machine learning methods, such as random forest (RF) (Naidoo et al., 2019), support vector machine (SVM) (Meng et al., 2020) and artificial neural networks (ANN) (Yang et al., 2018). While there is an increasing interest in using machine learning approaches, studies in grassland monitoring have mostly used empirical models and statistical regression models (Wang et al., 2022).

Generalised additive model (GAM) is a semi-parametric model that integrates both parametric and non-parametric components, serving as an extension of generalised linear models (GLMs). GAMs were first introduced by Hastie and Tibshirani (1986), and they have been used in various applications, such as in ecology (Yee and Mitchell, 1991) and biomass modelling (Fassnacht et al., 2021; Vuorinne et al., 2021). Hastie and Tibshirani (1986) note that GAMs provide flexibility, especially when portraying complex relationships, by using smoothers to form curves that display underlying trends between response, grass biomass measurements in this study, and predictor variables. A crucial parameter in GAM is its dimension k, which adjusts the smoothing function. (Li and Mao, 2020; Yee and Mitchell, 1991) They are data-driven so that data itself determines the shape of the response curves, rather than being restricted by fixed shapes. This enables the individual interpretation of variables and their contribution to the model. (Yee and Mitchell, 1991) The independent and dependent variables are not required to have any predetermined distributions (e.g. normality) or hypothetical relationships, unlike in other linear statistical models. Additionally, GAMs are known for their robust generalization abilities (Li and Mao, 2020).

However, despite modern advancements there are certain limitations and challenges in all the aforementioned methods. KarakoÇ et al. (2019) demonstrated that in the case of high biomass densities and degraded accuracy of biomass estimations, selection of appropriate indices depending on the spectral characteristics of grasslands was more critical than the choice of regression models. Wang et al. (2022) argued that instead of emphasising simple statistical regression models, most applications should involve advanced estimation methods, such as deep learning and fitting non-linear relationships.

A common issue in studying heterogenous vegetation and landscapes is the interference in spectral signatures, where soil can be mixed with pixels of vegetation and vice versa, therefore

impacting the accuracy of estimating vegetation properties (Du et al., 2023). The mixed pixel issue is particularly relevant with coarser resolution but does not disappear completely even when the resolution increases (Du et al., 2023). Additionally, factors such as herbaceous ground layer, sparse vegetation cover and GPS positional errors can lead to backscattering, distortions in data, and inaccurate reflectance values (Braun et al., 2018; Sarrazin et al., 2011).

Further, Sarrazin et al., (2011) note that the grazing effects of wildlife and agriculture in the landscapes can lead to considerable biomass differences within small areas. Additionally, the presence of soil background signals, varying vegetation types (grass, shrubs and trees) contribute to the difficulties in differentiating vegetation signals in savannas (Gessner et al., 2013; Huete, 1988; Okin et al., 2001).

The quick responses to rain, especially grasses, is both an ecological advantage and challenge for remote sensing in semi-arid ecosystems (Williams et al., 1998; Xu et al., 2015). This is crucial for the ecosystems success across different seasons but leads to challenges in remote sensing, since merely satellite based methods using VIs like NDVI may not fully represent temporal patterns of primary production, regardless of extensive spatial coverage of the data (de Jonge et al., 2022).

Additionally, highly seasonal rainfall occurrence poses challenges for remote sensing applications in savannas due to the high reflectance in red and NIR of dry vegetation (Okin et al., 2001; Xu et al., 2014). The low contrast NIR to red ratio of dry vegetation complicates biomass estimation with methods used in other ecosystems making the use of remote sensing to quantify dry biomass unreliable (Eisfelder et al., 2012; Xu et al., 2014). Xu et al. (2014) showed that NDVI may have negative or no correlation with vegetation biomass when dead plant cover exceeds 20%. Further, there is a challenge in transferring models used for remote sensing of vegetation measurements in semi-arid environments from one place to another (Eisfelder et al., 2012).

To address some of the identified challenges and the scarcity of similar studies, this research aims to estimate grass biomass in an East African savanna environment using hyperspectral airborne imagery and identifying the most effective wavelengths by applying vegetation indices and GAM (Generalised Additive Model) models.

3. Methodology

3.1. Study area

The study area – LUMO Community Wildlife Conservancy area – is located in the lowlands of Taita-Taveta County in southeastern part of Kenya (3°29'13.2"S, 38°12'36.0"E, Lion Rock) about 850 – 1050 m above sea level, covering approx. 460 m2 (Fig. 7) (Sheila Wachiye et al., 2022). LUMO Conservancy is an amalgamation made up of three community ranches: Lualenyi, Mramba and Oza, hence the acronym name LUMO. Lualenyi contributes 28,000 acres, Mramba and Oza each adding 10,000 acres which comprises a total of 48,000 acres (Sheila Wachiye et al., 2022). Situated along Tsavo East and Tsavo West National Parks, LUMO provides an important wildlife habitat for animals moving between these two larger protected areas but also enables the migration of wildlife populations across Mkomazi National Park in Tanzania (LUMO Conservancy, 2023).

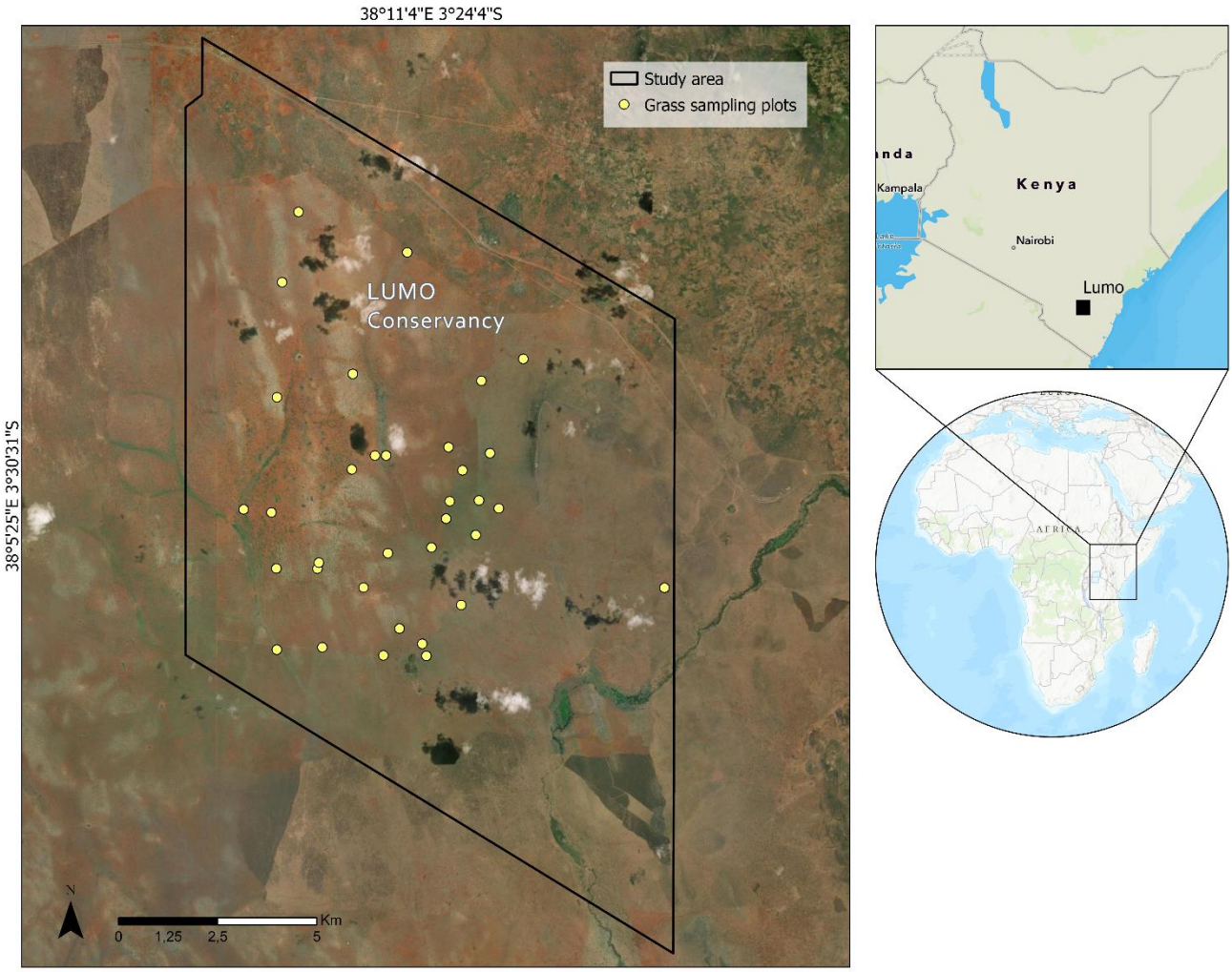

Fig. 7. Study area map in LUMO Conservancy (Kenya) including all the grass sampling plots of the study. All the basemaps are sourced from ArcGIS Pro basemap-gallery, provided by Esri Inc (ArcGIS Pro 3.1.0, 2023, Esri Inc).

The semi-arid conditions of LUMO and the lowlands of Taita-Taveta County are distinctively defining feature of its tropical savanna environment (Abera et al., 2022). The region has two rainy seasons annually and the rainfall is between 400 and 600mm. The long rains occur from March to May while short rains are expected from November to December (Pellikka et al., 2018). January and February experience usually the peak of dryness and heat, which is when the field work was conducted. The dry season taking place from June to October is cooler. At the end of the dry season in October, the maxima evapotranspiration in the lowlands is >7mm/day (Erdogan et al., 2011; Pellikka et al., 2018).

LUMO Conservancy represents the common environmental characteristics of a tropical savanna. The lowlands of Taita-Taveta County, including LUMO Conservancy, are mainly

composed of grasslands and shrublands (Abera et al., 2022). The wildlife conservation areas accommodate megafauna, for example elephants, giraffes, antelopes, rhinos, and buffalos (Abera et al., 2022). The prevalent soil in the lowlands is Ferralsol, which is acidic, dark red, sandy soil due to accumulation of metal oxides (Amara et al., 2020; Kögel-Knabner and Amelung, 2014). Further, these areas are characterized by *Acacia-Commiphora* type dry thickets, croplands by dryland agriculture, cattle grazing, sisal farming and other notable wildlife conservation efforts in addition to LUMO, such as Tsavo National Park and Taita Wildlife Conservancy (Pellikka et al., 2018).

The growing population in Kenya and Taita-Taveta County presents further challenges with increasing need for land resources, water supplies and energy production, accentuating the pressure on communities like LUMO (Pellikka et al., 2013). Since 1962 to 2019, the population of the whole Taita Taveta District has grown from 90,000 people to 340,671 (Platts et al., 2011; Kenya National Bureau of Statistics, 2019). The main source of income for 78% of people in the County is agriculture with most of them being small-scale farmers and there are 28 ranches designated for livestock production (CIDP, 2013-2017).

LUMO report (2019) states that rangeland management in the conservancy is facing notable challenges, predominantly due to overgrazing. The report also mentions that the issue is being aggravated by elephant destruction and the impacts of climate change. The western part of LUMO (West Mramba) is preserved for livestock management, while the eastern part (East Mramba) serves as habitat for wildlife but is very often invaded by large cattle herds (Amara et al., 2020). There is a large inflow of herds within LUMO by local herders, neighbouring communities, and illegal herders during dry seasons. The density of wildlife populations may differ significantly in the area during the wet and dry seasons. This leads to significant overgrazing challenges since the land is already under pressure during the peaks of heat and drought. (Amara et al., 2020) Herders move their cattle to LUMO in search of better forage and water supplies when their lands are depleted of forage and have most of their lands used for crop production (Sheila Wachiye et al., 2022). LUMO Conservancy is not fenced, which enables the free movement of livestock and wildlife, but meantime also contributes to the problem of overgrazing (CIDP, 2013-2017). Understanding the complexity of land use and management challenges, especially combined with local agricultural practices is essential when tackling with diverse and ongoing issues of overgrazing in dynamic savanna conservancies such as LUMO. The University of Helsinki have conducted studies in this area with advanced remote sensing techniques to enhance the understanding and mapping of ecological characteristics,

such as biomass distribution, in relation to land use management and other human activities (Abera et al., 2022; Amara et al., 2020; Sheila Wachiye et al., 2022).

3.2. Data collection

3.2.1 Grass sampling

In-situ measurements for the grass biomass modelling were collected in LUMO to represent the variety of grass biomass. The field work was carried out from 15th of January until 1st of February 2022 at 35 pre-selected study plot sites. The location of the plots was chosen to capture the environmental variability and heterogeneity in the study area. The variability in soil and grass were accounted for by relying on the expertise of team members, their previous knowledge of the area and recent Pleiades Neo satellite imagery. Some of the plots were selected among the plots studied by Wachiye et al. (2022) (Sheila Wachiye et al., 2022).

In each study plot of 20 m × 20 m, four subplots were chosen randomly, resulting in a total 140 subplots of 1 m × 1 m in size. First, the location of the study plot and the subplots were measured using RTK-GNSS (Trimble GeoXH 600 and NetR9) resulting in average accuracy of > 30 cm. Following this, all the measurements, notes and pictures were taken before sampling the grass. The biomass was harvested for each subplot at a ground surface level (Fig. 8) followed by categorization of samples into grass and shrubs. After sampling, the total grass and shrub fresh weight was measured in the field. Lastly, the grass was dried in the oven for 24-48 hours at 70°C and measured for dry weight in the laboratory of Taita Research Station (Fig. 8). These dry weight measurements served as reference values in the regression analysis framework.

Fig. 8. (A) Sampling grass biomass of a subplot on the field on the left and (B) conducting fresh grass measurements in the laboratory of Taita Research station on the right. Pictures taken in January 2022.

3.2.2 Hyperspectral data collection

The collection of hyperspectral airborne imagery was acquired in February 2022 during dry season using AisaKestrel 10 sensor by Specim mounted on Cessna C206. Table 1 presents overview of the data collection details. There was a two-week interval between the field and the airborne data collection campaign. The imagery of the study area in LUMO was collected on three distinct days: 14., 17. and 19.02.2022 with a flying height of 850 m above the ground. Hyperspectral AisaKestrel 10 sensor was used in four times spectral binning mode, reducing the number of bands from 192 to 92 wavelength bands with a wavelength range of 381 – 1003 nm and spectral resolution of 6.7–6.9 nm. The binning technique reduces data volume and noise. The imagery was collected between 08:30 – 11:05 under bright clear-sky conditions to keep the sun angle effects consistent and to avoid afternoon cloud build up. Keeping the flight lines approximately perpendicular to the sunlight direction assists to minimize bidirectional reflectance effects (Skidmore et al., 2010). The sensor was used in two times spatial binning mode, leading to a fine spatial resolution of 0.7 meters. Flight lines were flown on each flight day covering the whole study area. Visibility was 80 km and average water vapour 2.4 cm. AisaKestrel 10 sensor is a pushbroom-type sensor that scans the surface perpendicular to the

flight direction. These types of sensors image one line at a time, using a line of detectors that measure all the pixels in the line simultaneously (Sousa et al., 2022).

Table 1. Characteristics of the airborne imagery campaign taking place above the study area in 2022 and the hyperspectral sensor Aisa Kestrel 10 by Specim (Spectral Imaging Ltd. Oulu, Finland).

Sensor	AisaKestrel 10
Spectral range	381 - 1004 nm
Spectral resolution	6.7–6.9 nm
Spatial resolution	0.7 m
Frame rate	Up to 170 or 100Hz
Number of bands	92
Field of view (FOV)	39°
Platform	Cessna C206
Time of data collection	14., 17., & 19.02.2022
Total system weight	4.75 kg

3.3. Pre-processing of hyperspectral data

The raw data underwent through several pre-processing steps prior analysing: georectification, radiometric and atmospheric correction with DROACOR 2.0 (ReSe Applications Schläpfer, Wil, Switzerland) and Specim's CaliGeo Pro 2.2 (Spectral Imaging Ltd. Oulu, Finland).

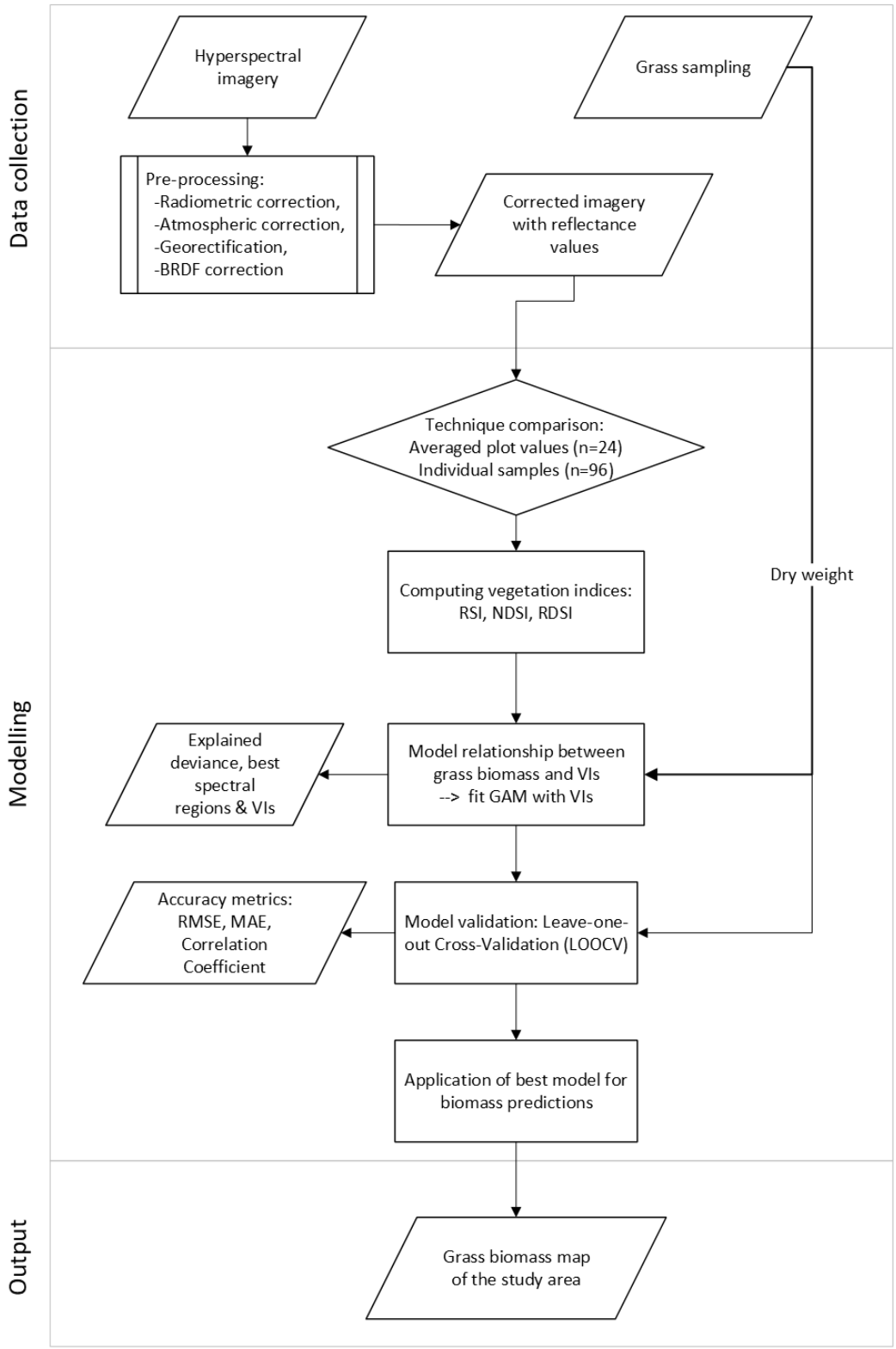

Fig. 9. Flowchart visualizing the workflow of data collection, preprocessing and modelling.

The geometric correction of the hyperspectral imagery includes various methods to ensure that the geometry and the location of the data is accurate. Geometric correction, also known as georectification is a process where the collected hyperspectral imagery is aligned with accurate geographic coordinates to correct any inconsistencies between the location coordinated of the raw data, and the true location coordinates on the ground (P.Dave et al., 2015). It is vital to ensure the geometric accuracy of each pixel in the image to avoid distortions and secure ground for accurate analysis. Georectification was performed using sensor location and orientation information. This information is acquired during the image acquisition with PPP-GNSS (location), and inertial measurement unit (orientation). Georectification was performed with a Digital Surface Model (DSM) for image alignment to consider the height variations of ground object, which impact their perceived locations. The information from DSM, GNSS for location and an IMU (Inertial Measurement Unit) information (roll, pitch, azimuth) was employed for position and imagery alignment. In addition, information collected from boresight flights were applied to the process. Boresight flights are used to calculate boresight angles which is the attitude difference between GNSS/INS and Aisa sensor. Boresight angles can be assessed through at least three flight lines flown in various directions and ground control points. Ultimately, optimal georeferencing results for airborne hyperspectral data is achieved when the effect of attitude difference is eliminated. (SPECIM, CaliGeoPRO User Guide)

Radiometric correction ensures to optimize the quality of the measured spectral radiance data using radiometric calibration data that was received from SPECIM to use for sensor calibration before the flights (Spectral Imaging Ltd. Oulu, Finland). The raw data digital number (DN) format of the sensor measurements is converted into spectral radiance. Secondly, the radiance values detected by the sensor is transformed into ground surface reflectance value using conversion factors from the calibration files. These reflectance values provide information of how much light is being reflected off the objects on the ground (Chen et al., 2005).

In addition to correcting scattered radiance, further steps were taken to mitigate atmospheric effects. Atmospheric correction was done with DROACOR software version 2.0 (ReSe Applications Schläpfer, Wil, Switzerland) to reduce the impacts of scattering and absorption, generating true reflectance values. Some of the flight lines had visible cumulus clouds which were removed with a cloud shadow removal mask. This noise in the data was corrected and filled by interpolating some of these absorption bands and applying a “smoothing” algorithm called Savitkzy-Golay filter which is related to bottom of atmosphere reflectance retrieval.

Ultimately, all the interpolated bands were excluded from the modelling workflow to get the original reflectance values.

As noted by Collings et al. (2010), the differences in the solar angle and the viewing angle of sensor can cause variations in reflectance between the flight lines and consequently inconsistencies in data. These differences, called BRDF effect, may lead to disparities in imagery captured at different times or angles, even if the characteristics of the area remain similar. To address this, BRDF correction was used following atmospheric corrections to correct these variations in reflectance measurements. This process considers different factors, such as vegetation to make necessary adjustments (Collings et al., 2010). The effects from bidirectional reflectance distribution function (BDRF) were corrected using DROACOR software version 2.0 (ReSe Applications Schläpfer, Wil, Switzerland). The method in DROACOR software uses the Ross-Thick-Li-Sparse reciprocal BRDF model kernels to correct the imagery (DROACOR User Manual, 2022). Finally, a tree mask was applied to the mosaicked and corrected images with a canopy height model (CHM) in RStudio version 12.0.353 (RStudio, 2022) to exclude all objects above two meters.

3.4. Data analysis and modelling

This study aimed to gain deeper understanding of the relationship between grass biomass measurements and vegetation indices (VIs) calculated from hyperspectral data. Three main strategies were adopted in this methodology: (i) calculating Generalised Additive Models (GAMs) for the VIs, (ii) conducting accuracy assessments and (iii) producing a grass biomass prediction map using the best performed VI.

The analytical process of calculating VIs, fitting GAMs, and performing accuracy assessment was executed on RStudio version 12.0.353 (RStudio, 2022). The VIs were derived from reflectance values acquired from study plots located within the study area. All potential two-band combinations were computed for three widely used standard VI formulations: 1) ratio based spectral index (RSI = $\frac{Band\ 1}{Band\ 2}$), 2) normalized difference spectral index (NDSI = $\frac{(Band\ 1 - Band\ 2)}{(Band\ 1 + Band\ 2)}$) and 3) reciprocal difference vegetation index (RDSI = $\frac{1}{Band\ 1} - \frac{1}{Band\ 2}$). The optimal band combinations for each vegetation index were identified by fitting GAMs and

examining the highest explained deviances (D^2). GAMs were implemented in RStudio with mgcv-package to explore the relationship between biomass and VIs.

The equation of GAM is as follows:

$$y = \beta_o + \sum_{j=1}^{p} f_j(\mathrm{X}_j) + \varepsilon$$

where y is the response vector, β_0 is the model intercept, $f_j\,(X)$ is a smooth function of predictor *X,* and ε is the residual (Wood, 2017). A link function is used in a GAM model, which connects the mean value of the dependent variable (grass biomass) to its linear predictor (Crawley, 2012). In this study, a smoothing function with $k = 3$ was chosen to prevent overfitting (Li and Mao, 2020) and a Gaussian-error structure was used as the identity link function. The degrees of freedom (flexibility) in the model are determined by the *k* parameter, which defines the upper limit on the degrees of freedom connected to the smoothing function (Wood, 2023).

The reflectance values were extracted from the hyperspectral mosaic using subplot centre points with a bilinear method, which interpolates the values of the four nearest raster cells. Next, an exploratory data analysis was conducted with NDVI to explore any outliers in the dataset.

Once the GAMs were fitted with all the VIs, the models and best spectral regions for biomass were assessed based on explained deviance (D^2), which is formulated as follows (Guisan and Zimmermann, 2000):

$$D^2 = \frac{(Null\ deviance - Residual\ deviance)}{Null\ deviance}.$$

Furthermore, the performance and reliability of models was evaluated with leave-one-out cross-validation (LOOCV) that resulted further information with values of Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and correlation coefficient. The LOOCV is a type of *k*-fold cross-validation method used as statistical estimation technique where the initial dataset is divided into *k* subsets of approximately equal size (Geisser, 1975; Stone, 1974). The model is trained *k* times, each round using a different subset for testing and the remaining for training. This process is repeated *k* times where each of the subsets is used once as a test set. The results from the tests are obtained by averaging the accuracy values computed on each subset. LOOCV

is widely used in different fields in addition to remote sensing, such as bioinformatics (Simon et al., 2003) and machine learning (Elisseeff and Pontil, 2002).

Two distinct modelling techniques were applied in the study, with the purpose of comparing different techniques and assess which one would produce better results. The plot level technique could potentially mitigate position accuracy errors by using average values from field measurements, thus providing a more stable and representative modelling results. The benefits of sample level technique could be related to bigger sample size which could result in increased reliability, reduced errors, and robustness.

4. Results

4.1. Descriptive statistics of grass biomass data

Summary of dry and wet weight grass biomass is shown in Table 2. The dry weight of sampled grass biomass varied between 0-348 g/m^2. The histogram of plot level measurements displays values being right skewed whilst slightly asymmetric, indicating dry weight grass biomass values are concentrated mostly in the lower part with some higher peaks across the scale (Fig. 10).

Table 2. Descriptive statistics of dry and wet weight grass biomass sample measurements collected from LUMO.

Grass biomass (g/m^2)	Min	1st Quantile	Median	Mean	3rd Quantile	Max
Dry weight	0.0	35.9	89.8	111.4	155.2	348.0
Wet weight	0.0	64.75	175.50	224.10	329.25	1132.50

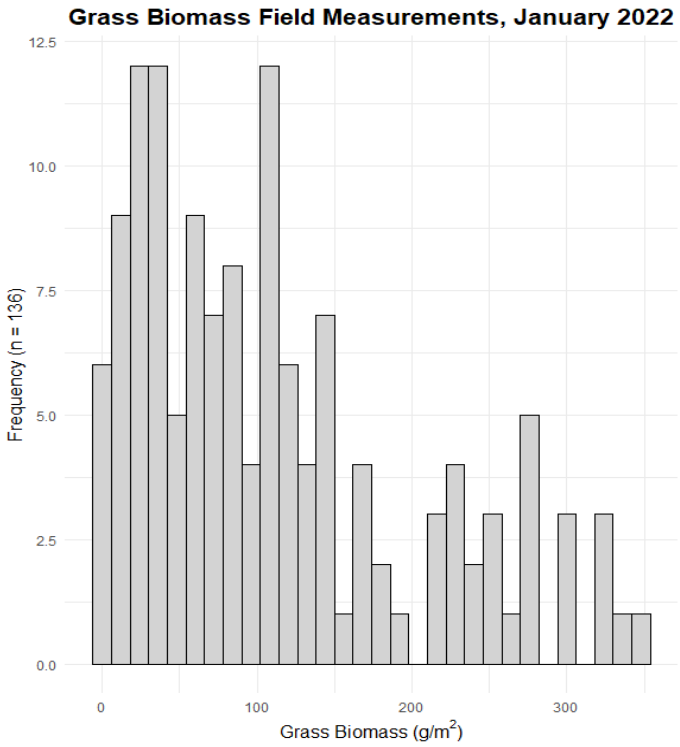

Fig 10. A histogram of dry weight grass biomass measurements.

4.2 Exploratory data analysis

To identify possible outliers in the dataset, exploratory data analysis of the relationship of NDVI and dry weight grass biomass was done prior to executing the GAM modelling workflow. The relationship between NDVI and dry weight grass biomass is shown in Fig. 11. All the figures show a positive relationship with NDVI and biomass. Figure of the complete dataset including all the image acquisition dates (14., 17., & 19.02.2022) shows a greater variability in values and a weaker relationship between the variables compared to the reduced data set of plot and sample levels (14. & 17.2.).

An exclusion of data was related to differing environmental conditions between image acquisition dates. Rainfall events took place in the study area, as shown in Fig 12., which caused differences in reflectance values between the first two flights (14.02, 17.02) and the last one (19.02.) (Fig. 13). Thus, it was necessary to exclude the images acquired 19.02. and all the plots located within this area to maintain the integrity and comparability between data. This demonstrates the importance of the right timing in field work measurements and acquiring airborne imagery at the same time in this kind of semi-arid savanna environment that is quickly reacting to rainfall. Altogether, the reduced dataset included in the final modelling workflow

consists of 24 plots and 94 subplots that were within the image acquisition dates of 14.02 and 17.02.2022.

Based on visual inspection of the dataset, one plot (P14) was identified as an outlier and consequently it was excluded from the final analysis of GAM modelling. The reflectance values of this plot were distinctly different from the rest of the dataset and did not represent typical conditions and patterns of grass biomass in the study area. The vegetation in this plot was quite scarce and located in an atypical place. The location was in a seasonal stream bed, a dent where water is stored. Consequently, despite having low biomass, the plot still shows as green and lush in the imagery, resulting in higher reflectance values.

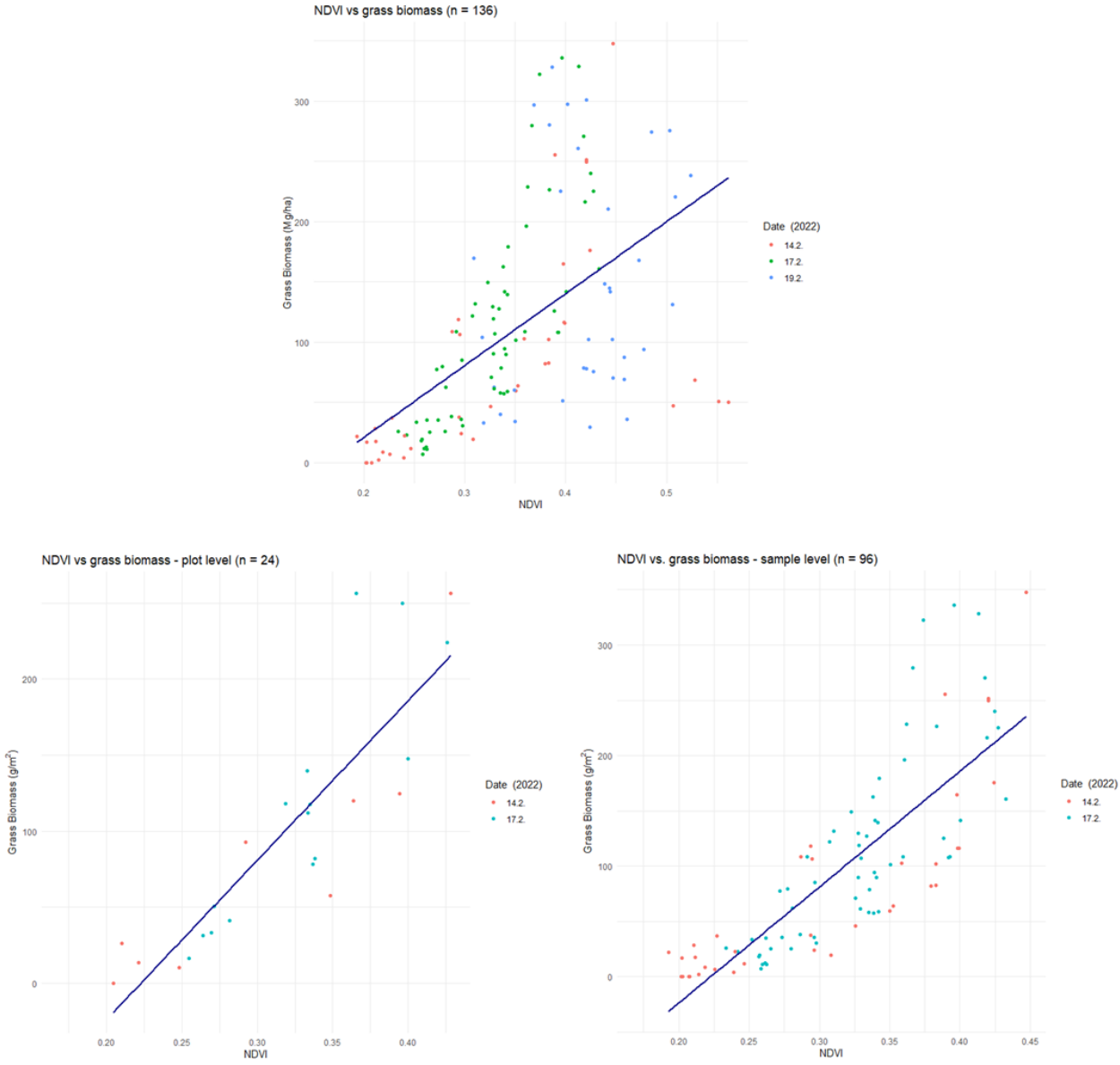

Fig 11. The relationship between NDVI and dry weight grass biomass samples in complete and reduced datasets.

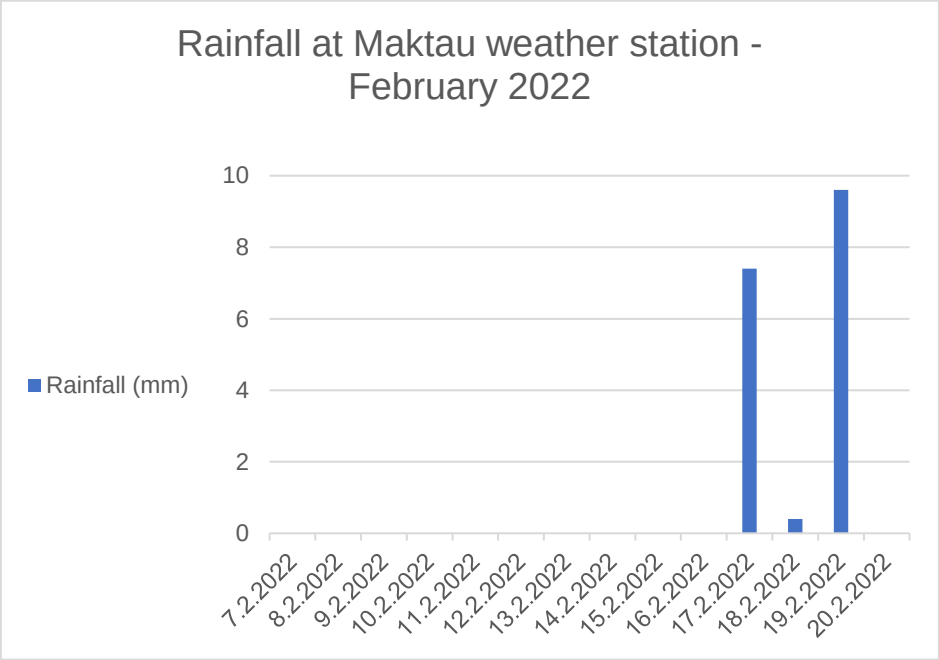

Fig 12. Histogram of a nearby a weather station in the study area during hyperspectral image collection campaign (weather station operated by University of Helsinki).

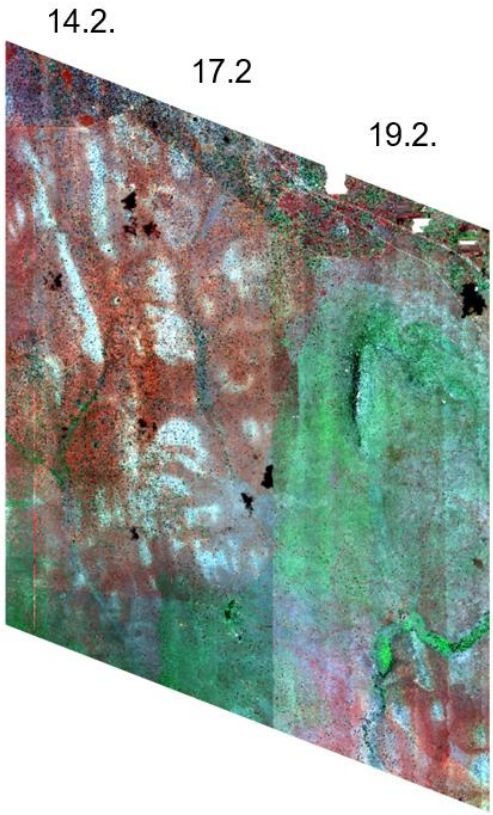

Fig. 13. Mosaic of acquired hyperspectral images showing the distinct differences between dates 14., 17., and 19.02.2022.

4.3 Best vegetation indices and band combinations

All the possible combinations of two spectral bands were computed on three individual vegetation indices: ratio based spectral index (Band 1 / Band 2), normalized difference spectral index (Band 1 – Band 2) / (Band 1 – Band 2) and reciprocal difference index (1 / Band 1 – 1 / Band 2). Explained deviances (D^2) between all band combinations were determined and are shown in Fig. 14, which illustrates how well each band combination in these vegetation indices predict data variability. Results of the analysis are displayed in the form of D^2 matrices for each Band 1 (401–997 nm) and Band 2 (401–997 nm) pair. The D^2 value is represented by colour intensity, where the most intensive yellow hue suggests the highest D^2 value and therefore being the most effective at explaining variations in the data.

As shown in Fig. 18., there are visible trends with the optimal band combinations and their explained deviances, despite of D^2 value variations across plot and sample level GAMs and different vegetation indices. The values for these optimal band combinations range from 0.61 to 0.79. Table 3 reveals a consistent pattern across both modelling levels: the same band combinations appear as the best for both RSI and NDSI (B78 / B43) and for RDSI (B46 / B47). Fig. 18. graphically highlights that an overwhelming proportion of grass biomass information is concentrated between approximately 680 nm and 997 nm, combined with bands around 600–700 nm. The highest explanatory power for both plot and sample level modelling was discovered in RSI and NDSI indices, with D^2 value being 0.79 at the plot-level and 0.68 at the sample-level. The most effective band combination for these indices was B78 and B43 (908 nm / 667 nm), where the near-infrared (NIR) portion around 908 nm coincides with red range wavelengths around 667 nm. Although RDSI performed weaker overall in both levels, its highest explained deviance was achieved with the combination of 688 nm and 695 nm. The highest D^2 values for RDSI at the plot level was 0.71 and 0.61 at the sample level. Poor explained deviance values were obtained approximately beyond 590 nm.

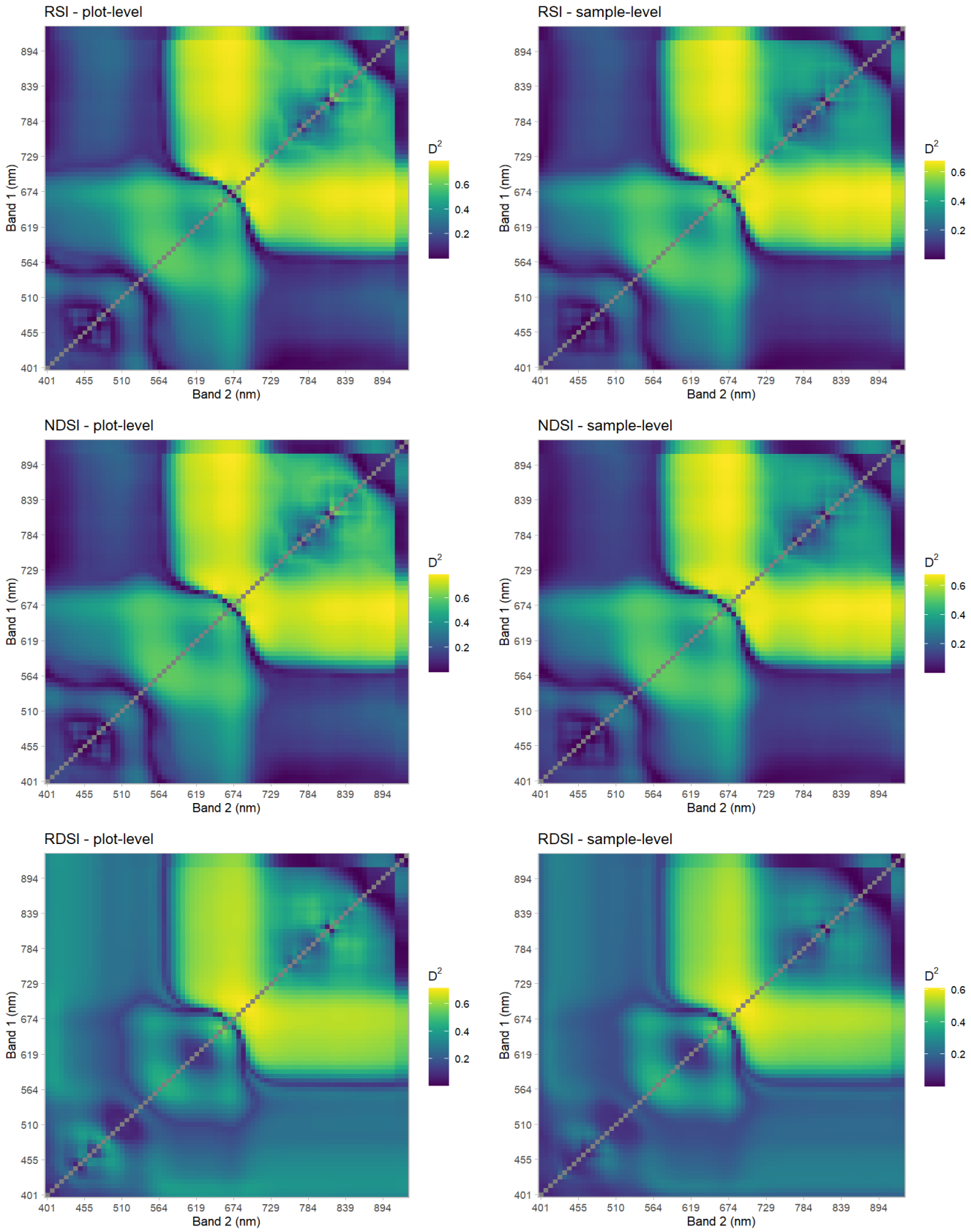

Figure 14. Matrixes of plot and sample level GAMs showing the highest and lowest explained deviance values (D^2) between dry weight grass biomass and vegetation indices (RSI, NDSI, RDSI) and their best band combinations.

Table 3. Summary of explained deviances (D^2) and the best band combinations for RSI, NDSI and RDSI vegetation indices by fitted Generalize additive models (GAMs).

D^2 values and band combinations for vegetation indices - plot level (n = 24)

Index	D^2	Best band combination	Wavelength (nm)
RSI	0.79	B78 / B43	908 / 667
NDSI	0.79	B78 / B43	908 / 667
RDSI	0.71	B46 / B47	688 / 695

D^2 values and band combinations for vegetation indices - sample level (n = 96)

Index	D^2	Best band combination	Wavelength (nm)
RSI	0.68	B78 / B43	908 / 667
NDSI	0.68	B78 / B43	908 / 667
RDSI	0.61	B46 / B47	688 / 695

4.4 Assessment of the best biomass models

4.4.2. Plot and sample level models

LOOCV was applied for the best performing RSI, NDSI and RDSI indices to assess the predictive accuracy of the GAMs predictions in both modelling levels. Table 4 presents an overview of the results of the validation assessment with RMSE (g/m^2), MAE (g/m^2) and correlation values. Based on the error metrics, RSI at plot level outperformed other indices both at plot and sample levels. This is demonstrated by previously mentioned D^2 values in Table 4 (0.79), RMSE of 40.15 (g/m^2), MAE of 30.64 (g/m^2) and a correlation coefficient of 0.74. Further, a correlation coefficient suggests a strong positive linear relationship between grass biomass and the RSI index, as Fig. 19. also highlights. However, NDSI at plot level showed

similar metrics but scored slightly higher RMSE (40.96) and MAE (31.49), and a lower correlation coefficient (0.73).

Table 4. Root mean squared error (RMSE), mean absolute error (MAE) and coefficient correlation metrics of leave-one-out cross validation (LOOCV) accuracy assessment of the best RSI, NDSI and RDSI indices at plot and sample levels.

LOOC-validation results

Plot-level (n = 24)

Index	RMSE (g/m^2)	MAE (g/m^2)	Correlation
RSI	40.15	30.64	0.74
NDSI	40.96	31.49	0.73
RDSI	49.04	38.37	0.63

Sample-level (n = 96)

Index	RMSE (g/m^2)	MAE (g/m^2)	Correlation
RSI	51.67	39.53	0.65
NDSI	51.16	38.83	0.66
RDSI	51.67	43.15	0.58

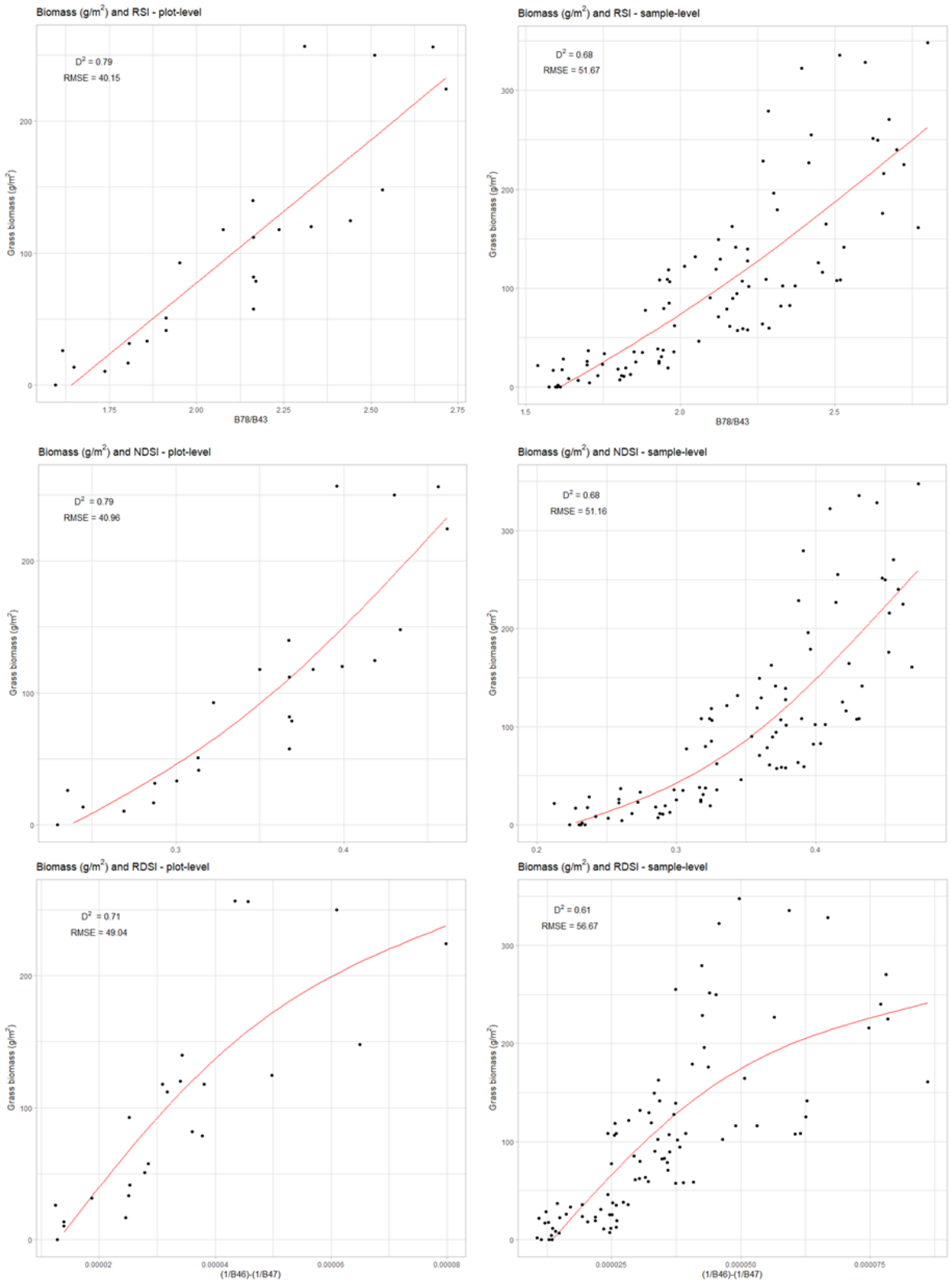

Fig. 15. Generalize additive model (GAM) regression response curves presenting the relationship between observed and predicted values of grass biomass and vegetation indices (RSI, NDSI, RDSI) at plot and sample level. D^2 indicates explained deviance values and RMSE the root mean squared error values (g/m^2).

4.5. Grass biomass map

Grass biomass was estimated in the study area using the best performing RSI vegetation index (Fig. 16). The estimated grass biomass values range between 0 – 2894 g/m^2, as seen in Table 5. The map displays notable variations in grass biomass within the area, especially in the southern region. The median values (101.36 g/m^2) are overall close to the field measurements, but there are very high estimates reaching up to 2894 g/m^2. These high values can be seen in water channels, especially in the western part of the map. Additionally, some bushes and trees have seemed to have affected to these particularly high biomass values. There are also big differences between the two imagery dates, occurring especially in the south.

Table 5. Descriptive summary of estimated grass biomass in the study area of LUMO Conservancy (2022).

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
Grass biomass (g/m^2)	0	47.98	101.36	154.4	226.5	2893.55

Fig. 16. Grass biomass map derived from the best index of RSI to demonstrate the spatial variability of grass biomass in the study area (January 2021). The map was done with ArcGIS Pro (ArcGIS Pro 3.1.0, 2023, Esri Inc) and a standard deviation stretch was applied for visualisation purposes.

5. Discussion

5.1. Performance of hyperspectral data in grass biomass modelling

One of the objectives was to explore the relationship between three key elements: the selected vegetation indices (RSI, NDSI, RDSI) and grass biomass measurements. This was carried out through a comparative approach, where averaged plot level (n=24) measurements were compared with individual sample level (n=96) measurements. Semi-arid environments like savanna are highly heterogeneous and dynamic and behold a lot of variation in both grass biomass and reflectance values of acquired hyperspectral imagery. The results of this study indicate the strongest relationship between grass biomass and the vegetation indices RSI and NDSI at plot level.

The results showed that RSI at plot level had the highest performance in predicting dry weight grass biomass among fitting GAMs and LOOCV accuracy assessment techniques (RMSE = 40.15, MAE = 30.64, Correlation = 0.74). The optimal band combinations for plot and sample levels showed consistency with all the indices achieving the same best band combinations. RSI and NDSI demonstrated the strongest connection to grass biomass at plot level with the most effective band combinations found in the 908 nm and 667 nm regions, resulting in an explained deviance of 0.79. These findings suggest a strong relationship between grass biomass, RSI and NDSI. The results align with those obtained by Thenkabail et al. (2002) who identified the best narrowband indices to be found with the combination of a band centered at 675 nm with a bandwidth of 15 nm for the red and another band centered around 906 nm or 920 nm for NIR.

RSI and NDSI formulas were found superior to RDSI at both plot and sample level estimation in grass biomass. The results differ to those of Mao et al. (2022) who estimated aboveground biomass in desert environments and highlighted the potential of RDSI to generate significantly better VIs compared to other formulas. For instance, a formula based on RDSI called NRDVI (Near Infrared Reciprocal Vegetation Index) showed strong explanatory power for all aboveground biomass satellite modelling. The study suggested that the reciprocal conversion method would imitate the absorption spectrum successfully, therefore making it effective in capturing the variations in spectral absorbance of desert shrubs. However, additional research is required as the RDSI is not as established in estimating biomass compared to other more commonly known formulas like RSI and NDSI.

Another objective of the study was to identify most important spectral features to explain grass biomass. Previous studies, such as those by Kong et al. (2018) and O. Mutanga and Skidmore (2004), have demonstrated that the use of red bands around 620–780 nm and NIR bands around 700–1300 nm, as well as their combinations, yield good correlation with biomass. It aligns with this study's results, where high explained deviance values were also found throughout the range of around 680–997 nm coinciding with bands from 620 to 720 nm, in addition to the most effective band combination of 908 nm and 667 nm. These identified spectral regions accord with the studies of Thenkabail et al. (2002), Cho and Skidmore (2006) and Horler et al. (1983).

The most effective spectral regions observed in this study can explained by the reflectance properties of vegetation. Thenkabail et al. (2002) highlighted the significance of the red spectral regions (660–675 nm) relating to chlorophyll absorption with two main points; pre-maxima at 660 nm, which is characterized by varying absorption due to factors like biomass and soil background, and maxima at 675 nm, which has the greatest crop-soil contrast and strong correlations with chlorophyll-a and chlorophyll-b. Similarly, Cho and Skidmore (2006) and Horler et al. (1983) reported that shifts in the red edge (680–760 nm) indicate changes in plant health and biomass, which is caused by its sensitivity to chlorophyll content of the vegetation. Mutanga and Skidmore (2004) showed that a significant proportion of biomass information is in the red edge portion (680–780 nm).

Using the combination of sensitive red and NIR in VIs (RSI, NDSI) proved effective in quantifying vegetation characteristics like biomass, as highlighted in the results of this study. Thenkabail et al. (2002) demonstrated the effectiveness of this approach in studying vegetation with hyperspectral imagery. The NIR region (750–1300 nm) shows strong reflectance from healthy vegetation due to internal structure of the leaves, thus corresponding to biomass and LAI well. They also stated that absorption increases in the NIR "shoulder" (940–1010 nm) as biomass and moisture increases, further emphasizing the importance of NIR in biomass assessments.

Xu et al. (2022) and Mutanga and Skidmore (2004) emphasized the effectiveness of using green region in addition to red and NIR for grassland vegetation studies. Even though this study aligns most strongly with the stated efficacy of red and NIR regions, a mild correlation was observed between biomass and VIs within the 550–680 nm coinciding with 530–590 nm. This subtly emphasizes the green region. However, it's important to note that the strength of these relationships was less prominent compared to the most significant spectral regions identified

(667–997 nm). Further, this study was focused on savanna rangelands, instead of controlled laboratory conditions or pure grasslands as studied by Mutanga and Skidmore (2004) and Xu et al. (2022). The observed superiority of red and NIR over green in this study could be related to numerous factors including processing techniques, sensor differences, temporality, and differences in study context, and would need further investigation.

In this study, the lowest D^2 values were obtained approximately in areas where 700–997 nm (red and NIR) is combined with the region of 401–580 nm (blue and green). Vegetation reflects more light in the green region (500–600 nm) due to lower chlorophyll absorption, conversely to blue (400–500 nm) and red, which are strongly absorbed for photosynthesis (Huete, 2004). Thenkabail et al. (2002) further stated that blue range is subject to atmospheric effects and provides only a small contrast in reflectance of vegetation and soil, making it questionable to use in biomass assessments. Consequently, this would explain the poor observed relationships of in this study, where combining the visible (401–580 nm) with regions of red and NIR (700–997 nm) showed less sensitivity to biomass compared to the most effective regions (680–997 nm).

The findings of this study showed that certain VIs have a stronger relationship with biomass, as RSI and NDSI (908 nm / 667 nm) demonstrated the most explanatory power. The performance of RSI and NDSI was quite similar, aligning with their mathematical equivalence as previously mentioned in the literature review. Jackson and Huete (1991) reported that RSI utilizing red and NIR was found highly effective in areas of dense green vegetation. RSI is highly sensitive to changes in vegetation particularly during growth peaks, and it increases substantially when red reflectance approaches zero. In contrast, NDVI is more sensitive to sparse vegetation densities but less effective in high vegetation due to the well-known saturation problems. According to these factors, NDVI would be expected to outperform RSI in sparse environments (Jackson and Huete, 1991) like savannas, which contrasts slightly the findings of this study. Our findings indicated that the models were effective in biomass estimations even at the highest values, therefore overcoming the saturation problem typically related to broadband data. However, grass biomass is generally lower during dry season, suggesting that the saturation problem might be more prominent at other times of the year. The results of this study showed no indications to saturation problems with the use of narrowband indices. Thus, at other times of the year, the best performing models could've resulted substantially different results, or the performance difference between RSI and NDVI could have been more prominent than the observed results. Additionally, Huete et al. (2002) stated that the performance of certain

vegetation indices depends on the geographical context, methods, modelling techniques, vegetation type and environmental conditions.

Several recent studies have highlighted the effectiveness of NDSI and RSI in semi-arid environments (Schucknecht et al., 2017; Vuorinne et al., 2021) and planted forests (Peng et al., 2019) utilizing additional variables in addition to VIs. They acknowledged the possibility of using more variables for model enhancement, such as topography, climatic factors, and other external biophysical variables. In relation to this study, including variables such as grass height, soil temperature and moisture might be beneficial and could enhance the model. However, Hawkings (2004) noted that employing more complex models in remote sensing data analysis including numerous variables to explain variance can result in overfitting and unreliable predictions. This emphasizes the advantage of this study's approach, which used a simpler modelling workflow with fewer variables. Nevertheless, future research could benefit exploring the impacts of additional variables in the modelling, as well as exploring the performance of more vegetation indices to gain further understanding. Such indices could be Soil Adjusted Vegetation index (SAVI), Enhanced Vegetation Index (EVI) and/or Canopy Chlorophyll Content index (CCCI), similarly to the study by Vuorinne et al. (2021).

Addressing the feasibility of utilizing airborne hyperspectral data was one of the objectives of this study. There are several benefits to using hyperspectral imagery in biomass modelling compared to broadband multispectral data, which is known to have limitations related to acquiring sufficient information on biophysical features in terrestrial ecosystems. Thenkabail et al. (2002) highlighted how hyperspectral imagery provides more detailed information sensitive to plant pigmentation, canopy structure and soil background effects. The study found that the use of narrowband models enhanced the explanation power for estimating various crop variables. The model accuracy improved biomass estimates by up to 24% and 27%, compared to the best performing broadband indices. Narrowband indices also offer robustness in accommodating variations across a wide range of conditions including growth stages, soil background effects, levels of pigmentation (Thenkabail et al., 2002). In this study, various vegetation covers and phenological phases were represented, such as senescent, overgrazed, green with high canopy, and sparse to no vegetation. Regardless of great variations, hyperspectral imaging enabled to model different forms of grass and vegetation covers in accordance with the study of Thenkabail et al. (2002).

The results of this study accord with Thenkabail et al (2002) and Mutanga and Skidmore (2004), who emphasize the importance of optimal band extraction to ensure mode performance. This approach mitigates the issues associated with the high dimensionality and large volumes of hyperspectral data and optimizes the use of rich information captured from vegetation (Thenkabail et al., 2002). Thenkabail et al. (2002) also reported that such flexibility allows to target particular applications and leads to more efficient data analysis. In this study, the use of narrowband indices was considered beneficial, as utilizing hyperspectral imagery allowed to extract specific narrowband indices. This led to be effective both in results and the computation of models. This study's findings indicate that the benefits of hyperspectral imaging are likely an advantage for developing detailed models representing dynamic savanna environments. The added value of narrowband indices requires further investigation.

Remote sensing research in heavily grazed grasslands, such as those in northern Tibet (Duan et al., 2011) has been known to face challenges. Decreased green cover and increased soil exposure in these environments have been observed to shift the red edge position (Xu et al., 2022). Additionally, Xu et al. (2022) noted that the presence of dead materials, a large component of native grasslands, or bare soil can cause a slight shift in the red valley position (a dip in reflectance between red and NIR), moving it away from 670 nm. A study by Wei et al. (2019) suggests how underlying surface in arid grasslands affects the red edge parameters. Our results indicated that the relationship between red edge and grass biomass weren't particularly notable, unlike the study of Mutanga and Skidmore (2004) demonstrated. This could be related to the influence of dead materials and bare soil that impact the acquired spectral information. There is notable lack of research on how dead materials influence red edge parameters, especially in the context of semi-arid savannas (Xu et al., 2022), and a better understanding of this could be beneficial for further investigation.

Schucknecht et al. (2017) proposes biomass estimations to be conducted early in the senescent phase, right after maximum vegetation development, which aligns the temporal timing of this study. This could help decrease errors in biomass measurements related to natural leaf decay or grazing taking place during senescence phases. It has been reported that late biomass measurements tend to underestimate the total biomass and not align with remote sensing data, making data during the vegetation peaks generally the most accurate (Schucknecht et al., 2017). This study was timed at the beginning of the dry season, which proved to be beneficial according to the proposals of Schucknecht et al. (2017), as a few months after the study the vegetation in the study area had already suffered greater levels of drought and senescence.

Furthermore, this study raises unanswered questions about how the best performing indices (RSI and NDSI) and spectral bands of this study (908 nm / 667 nm) would perform at the end of the dry season compared to the beginning, which could provide interesting insights for further investigation.

An interesting finding of this study demonstrated that using averaged values (n=24) compared to individual samples (n=96) yielded better performance in modelling results. Several factors could explain this observation. As noted in the literature review, fine spatial resolution provides more detail but can cause variability and noise in data that is not relevant. Rocchini et al. (2016) discussed that the use of very high spatial resolution imagery might lead to higher spatial heterogeneity in the spectra and would lead to noise instead of enhancing the spectral information. This occurs especially in heterogenic nature of semi-arid rangeland environments like savannas and can be seen as significant variations in hyperspectral remote sensing measurements, depending on whether the pixel captures shrubs, grass or bare soil (Braun et al., 2018; Gessner et al., 2013; Sarrazin et al., 2011). Additionally, the environmental conditions of savannas themselves are characterized by variations in moisture, grazing intensity, and vegetation cover (Huete, 1988; Okin et al., 2001). Consequently, the use of averaged values might have reduced these challenges by mitigating the noise and variability related to hyperspectral data and resulting in better performing models in heterogenous savanna environments, as was demonstrated in this study.

Lastly, one of the objectives of this thesis was to map grass biomass within the study area. The predictions were based on the best performing RSI index of B78/B43 bands (908 / 667 nm). The biomass map showed great variations within the area and between the two images, especially in the southern part of the map. There are uncertainties related to weather conditions, image acquisition and preprocessing that could explain this. Daytime heating in the savannas typically result in forming of cumulus clouds (Ruppert and Johnson, 2016). These atmospheric disturbances can directly affect reflectance values and occur in the modelling results as well as the estimated biomass map. Furthermore, rainfall events between the image acquisition dates can lead to inconsistencies between the images and affect the mapping. The grass biomass map showed significant variations (0-2893.55 g/m^2) in overall biomass and particular higher biomass seen in the southern part of the map. However, the median value (101.36 g/m^2) indicates that the estimations are overall realistic. It is suspected that the overestimations could be related to rainfall event taking place before image acquisition in the southern part of the

study area. Additionally, some water channels in the western part of the map and bushes/trees might have affected the overestimations.

There were also differences in flight times, as the border of the two mosaics had a difference of about 1h 30 mins in their acquisition times. BRDF correction and other preprocessing techniques can cause such discrepancies in ending results due to variability in viewing angles and atmospheric conditions. Jia et al. (2020) discuss the challenges associated with employing BRDF correction, such as terrains with high variation in topography. The variation can have impacts on the reliability of soil-vegetation BRDF retrieval and most current BRDF techniques for airborne imagery are created for flat surfaces (Jia et al., 2020). Finally, a conservative tree mask was applied to identify objects taller than two meters, therefore some artefacts are visible on the map which could impact the biomass estimates overoptimistically. Future work could benefit from exploring BRDF correction techniques to consider topographic factors. Additionally, timing the imagery acquisition strategically to avoid rainfall, and applying a more rigorous tree mask would probably improve the models.

5.2. Challenges and future prospects

It is important to highlight the limitations of this biomass estimation study in a semi-arid savanna. Temporal extent was a significant limitation of the study, as data observations were conducted during only one month, during which conditions varied significantly. The two-week gap between grass sampling and the acquisition of hyperspectral imagery was further complicated by the rainfall event that occurred during this period. An exploratory analysis of the data showed that the imagery from one day was not comparable with the other two days in the study area due to the rain, which lead to the exclusion of one acquisition day from the dataset. This demonstrates the sensitivity of environmental conditions and the challenges this presents to remote sensing research. As mentioned in the literature review, vegetation in savanna uses available water as quickly as possible (Xu et al., 2015), providing ecological competitive advantage in unpredictable environments. A sudden rainfall can significantly alter the reflectance values of the vegetation and can lead to quick greenup, as happened in this study. As mentioned in the previous section by Thenkabail et al. (2002), the absorption increases in the NIR (940–1010 nm) as biomass and moisture increase. It is important to consider the appropriate timing in field work measurements and the acquisition of airborne imagery

simultaneously. At the same time, perfect synchronization is challenging as the timing of the flights are heavily dependent on weather conditions.

Additionally, one plot located in a seasonal riverbed proved to be an anomaly in the dataset. Despite appearing lush and green in the imagery, field measurements showed low biomass, thus resulting in its exclusion from the dataset. This plot did not represent the general conditions and patterns of grass biomass in the study area as the biomass, as it was characterized by scarce biomass and atypical location. Despite its low biomass samples, the plot showed high values on the map and in the model which led to inconsistencies in the modelling due to an anomaly. Exclusion of data to gain better model performance raises questions of appropriate data cleaning and shows necessity of making compromises in modelling. This study aimed to represent the area while considering generalizability and excluding some extreme variations in the dataset. Additional research is needed to better understand plots with unique characteristics such as this plot.

Similar to the suggestions of Braun et al. (2018), this study also confirms the proposition for long term monitoring, potentially using automated active sensors operating at different wavelengths. The use of hyperspectral satellite imagery could have benefits in improving modelling workflows in savannas by increasing spatial coverage, ensuring continuous monitoring, and by employing bands that were not accessible in this study, such as SWIR (short-wave infrared). Future studies could consider utilizing SWIR bands to gain more insights into identifying the optimal spectral regions in assessing grass biomass. Shaik et al. (2023) discuss the benefits of new hyperspectral satellites with high spectral resolution, such as PRISMA, in various applications, including monitoring vegetation. PRISMA has global coverage which is an efficient advantage compared to airborne sensors that have a very limited coverage. It can also be fused with other remote sensing data, such as LiDAR, radar or multispectral, which would provide interesting insights in grass biomass assessments and are suggested for further investigation. The challenges of hyperspectral satellites are similar to those of hyperspectral airborne data that is also a subject to atmospheric interferences and requires high computational time for processing large datasets (Shaik et al., 2023). Moreover, further work is needed to identify the most efficient modelling workflows for the best understanding of fluctuations in biomass over time in dynamic savanna landscapes, which was a significant limitation of this study.

In the context of high spatial heterogeneity and complex grass species, traditional parametric models don't usually result in good performance, because they rely on assumptions of the data. In contrast, non-parametric models require the selection of variables and can freely learn from the data without any restricting pre-assumed functional relationships (Wang et al., 2022). Grazing has been reported to affect modelling performances by the complexity and variability of biomass, especially under intensive grazing practices (Ali et al., 2017). It was noted that the biomass curves of intensely grazed areas were more complex with significant interannual variation compared to less grazed areas in the study by Ali et al (2017). However, it has been emphasised that focusing on the most effective spectral regions of grasslands is more crucial than the proper choice of regression models (Karakoç and Karabulut, 2019). GAM proved to be a convenient model choice in the light of complex nature of savanna environment since it is in between parametric and non-parametric models and does not require specifying polynomial terms or predictor transformations to improve the model performance (Hastie and Tibshirani, 1987). GAM has succeeded well in biomass estimations, including a study by Vuorinne et al. (2020). The results showed strong relationship between biomass estimates and vegetation indices, and the best performance was achieved using RSI and NDSI, that are in agreement with the results of this study. However, the regression response curves of this study demonstrate that the regression lines were not particularly complex and suggest that employing simpler parametric models could've also resulted in effective results. Thus, extracting the most effective spectral regions likely had a bigger impact on the success of the modelling rather than choice of regression model itself, as stated by Karakoç and Karabulut (2019).

In general, combining ground truth data with remote sensing data can lead to highly effective results in grassland monitoring models. Lyu et al. (2021) demonstrated this by NDVI and EVI with meteorological, soil variables in an Artificial Neural Network (ANN) model and achieved a high model performance (R^2 = 0.91). Similarly, Meng et al. (2020) found that various regression and machine learning models, such as Random Forest with combined ground and satellite data resulted in R^2 = 0.78. Zhou et al (2021) also combined vegetation indices and ground data with Random Forest, leading to an effective performance in R^2 of 0.85. The employment of field samples in this study proved to be a necessity, as modelling results can vary greatly between ground truth data, which was the case with the beforementioned anomalous plot. Besides model optimization, study by Herrman et al. (2013) detecting weeds with fine resolution data discussed about accuracy techniques. They stated that improved classification outcomes could have been obtained if the cross-validation was employed from

one date, one plot or even one image. Their results raise questions on which accuracy method in remote sensing grass biomass is the best for remote sensing data, especially with a rather small dataset like in this study. Despite the effective results gained in this study, accuracy could've been improved and utilizing some other accuracy techniques would require further investigation.

Additionally, further research is required to assess the generalizability of GAM hyperspectral remote sensing of biomass, as models commonly tend to be site specific. Another suggestion incorporates testing the same dataset with different models, such as the previously mentioned well-performing Random Forest (RF). Jacon et al. (2021) states how other machine learning models can also provide more accurate estimates of biomass compared to regression models, including models that can handle a greater number of hyperspectral metrics. These include Classification and Regression Trees (CART), Cubist (CB), Partial Least Squares regression (PLS) and Support Vector Machine (SVM).

Only a limited number of studies exist on studying grass biomass using remote sensed data in the rangeland savannas of Kenya during dry season. Biomass modelling employing various techniques in other contexts, such as crops (Ji et al., 2007; Panda et al., 2010; Uno et al., 2005) and forests (Cutler et al., 2012) is well established in numerous locations globally. This study showed that assessing grass biomass in the semi-arid savanna in Kenya using hyperspectral data in the visible to NIR spectral range indicated promising results. These findings can be beneficial for applying similar workflows in other semi-arid environments, showing the efficacy of using hyperspectral imaging in biomass assessments in the beginning of dry season. Further, it was discovered that utilizing averaged values can lead to better model performance with hyperspectral data by smoothing variations and noise. Additionally, this study contributes to expand the database coverage of savanna environments, especially in Eastern Africa. It also addresses the existing research gap on the topic and supports sustainable science-based monitoring techniques for savannas amidst climate change. However, further research is needed to improve the models and consider limitations of this study, including limited temporality and spatial coverage, varying weather conditions, anomalous study plots and effects of certain pre-processing techniques.

Grasslands serve as a vital food source for grazing livestock. The quantity of biomass has a crucial role in determining the pastures carrying capacity, which indicates the number of animals that can graze a pasture sustainably. The most affordable forage for livestock is grazed

grass, which highlights the need for sustainable management of grasslands to ensure the availability of low-cost high-quality grass (Ali et al., 2017). Consequently, local livelihoods, food security and wellbeing are closely connected to the condition of rangelands, emphasizing the importance of studying key variables such as grass biomass and monitor sensitive, dynamic environments like savannas, particularly in the context of global environmental changes.

Acknowledgments:

I am grateful for the opportunity to be a part of the ESSA project, which plays a significant role in advancing sustainable pastoralism in East Africa. My sincere thanks also go to the staff and facilities at the Taita Research Station in Wundanyi, Kenya, whose hospitality made my stay feel like home and allowed me the use of their laboratory. I extend my gratitude to the University of Helsinki and the Department of Geosciences and Geography for funding my field trip to the beautiful country of Kenya.

Special thanks to all those who assisted in the fieldwork, including Ilja Vuorinne, Dr. Janne Heiskanen, Darius Kimuzi, Rechard Mwasi, Ian Ocholla, Mwadime Mjomba, Professor Enrico di Minin and Dr. Sheila Wachiye. Your contributions were vital for this study.

I owe immense thanks to my supervisors. My gratitude extends to Ilja Vuorinne for your astonishing expertise and patience in closely guiding me through the preprocessing and analysis, as well as leading the operations for the hyperspectral imagery campaign and preprocessing. Dr. Janne Heiskanen, I am deeply appreciative of your valuable support, wisdom and incredible knowledge that guided me throughout this process, and for proposing this topic to me. Professor Petri Pellikka, your decades of extensive and global work in geoinformatics, efforts in operating the Taita Research Station and organizing this thesis have been inspirational and appreciated. I am profoundly grateful for all your help.

Lastly, I express my heartfelt regards to my mother, the most supportive and inspiring figure in my life, who initially sparked my passion for science, especially in environmental studies. You are the greatest blessing in my life.

References:

Abdi, A.M., Brandt, M., Abel, C., Fensholt, R., 2022. Satellite Remote Sensing of Savannas: Current Status and Emerging Opportunities. Journal of Remote Sensing 2022. https://doi.org/10.34133/2022/9835284

Abera, T.A., Vuorinne, I., Munyao, M., Pellikka, P.K.E., Heiskanen, J., 2022. Land Cover Map for Multifunctional Landscapes of Taita Taveta County, Kenya, Based on Sentinel-1 Radar, Sentinel-2 Optical, and Topoclimatic Data. Data 7, 36. https://doi.org/10.3390/data7030036

Adam, E., Mutanga, O., 2012. Estimation of high density wetland biomass: Combining regression model with vegetation index developed from Worldview-2 imagery. Presented at the Proceedings of SPIE - The International Society for Optical Engineering, p. 85310V. https://doi.org/10.1117/12.970469

Adam, E., Mutanga, O., Rugege, D., 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. Wetlands Ecol Manage 18, 281–296. https://doi.org/10.1007/s11273-009-9169-z

Ahamed, T., Tian, L., Zhang, Y., Ting, K.C., 2011. A review of remote sensing methods for biomass feedstock production. Biomass and Bioenergy 35, 2455–2469. https://doi.org/10.1016/j.biombioe.2011.02.028

Ahlstrom, A., Raupach, M., Schurgers, G., Smith, B.B.N., Arneth, A., Jung, M., Reichstein, M., Canadell, J., Friedlingstein, P., Jain, A., Kato, E., Poulter, B., Sitch, S., Stocker, B., Viovy, N., Wang, Y., Wiltshire, A., Zaehle, S., Zeng, N., 2015. The dominant role of semi-arid ecosystems in the trend and variability of the land CO2 sink. Science 348, 895–899. https://doi.org/10.1126/science.aaa1668

Ahmad, M.N., Shao, Z., Javed, A., 2023. Modelling land use/land cover (LULC) change dynamics, future prospects, and its environmental impacts based on geospatial data models and remote sensing data. Environ Sci Pollut Res 30, 32985–33001. https://doi.org/10.1007/s11356-022-24442-2

Alcántara-Ayala, I., Esteban-Chávez, O., Parrot, J.F., 2006. Landsliding related to land-cover change: A diachronic analysis of hillslope instability distribution in the Sierra Norte, Puebla, Mexico. CATENA, Geomorphology and Land Degradation 65, 152–165. https://doi.org/10.1016/j.catena.2005.11.006

Ali, I., Cawkwell, F., Dwyer, E., Green, S., 2017. Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data—A Machine Learning Approach. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 10, 3254–3264. https://doi.org/10.1109/JSTARS.2016.2561618

Amara, E., Adhikari, H., Heiskanen, J., Siljander, M., Munyao, M., Omondi, P., Pellikka, P., 2020. Aboveground Biomass Distribution in a Multi-Use Savannah Landscape in Southeastern Kenya: Impact of Land Use and Fences. Land 9, 381. https://doi.org/10.3390/land9100381

Amiri, F., shariff, rashid, Shariff, B., 2010. Using Remote Sensing Data for Vegetation Cover Assessment in Semi-Arid Rangeland of Center Province of Iran. World Applied Sciences Journal 11, 1537–1546.

Anderson, G.L., Hanson, J.D., Haas, R.H., 1993. Evaluating landsat thematic mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands. Remote Sensing of Environment 45, 165–175. https://doi.org/10.1016/0034-4257(93)90040-5

ArcGIS Pro 3.1.0 Software. 2023. Esri Inc.

Archer, S.R., Andersen, E.M., Predick, K.I., Schwinning, S., Steidl, R.J., Woods, S.R., 2017. Woody Plant Encroachment: Causes and Consequences, in: Briske, D.D. (Ed.), Rangeland Systems: Processes, Management and Challenges, Springer Series on Environmental Management. Springer International Publishing, Cham, pp. 25–84. https://doi.org/10.1007/978-3-319-46709-2_2

Atazadeh, E., Mahdavifard, M., Atazadeh, E., Mahdavifard, M., 2021. Application of Remote Sensing in Natural Sciences, in: Remote Sensing. IntechOpen. https://doi.org/10.5772/intechopen.94468

Azage, M., Kumie, A., Worku, A., Bagtzoglou, A.C., Anagnostou, E., 2017. Effect of climatic variability on childhood diarrhea and its high risk periods in northwestern parts of Ethiopia. PLOS ONE 12, e0186933. https://doi.org/10.1371/journal.pone.0186933

Balzarolo, M., Vescovo, L., Hammerle, A., Gianelle, D., Papale, D., Tomelleri, E., Wohlfahrt, G., 2015. On the relationship between ecosystem-scale hyperspectral reflectance and CO_2 exchange in European mountain grasslands. Biogeosciences 12, 3089–3108. https://doi.org/10.5194/bg-12-3089-2015

Bartholomé, E., Belward, A.S., 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing 26, 1959–1977. https://doi.org/10.1080/01431160412331291297

Batáry, P., Báldi, A., Erdős, S., 2007. Grassland versus non-grassland bird abundance and diversity in managed grasslands: local, landscape and regional scale effects. Biodivers Conserv 16, 871–881. https://doi.org/10.1007/s10531-006-9135-5

Berdugo, M., Delgado-Baquerizo, M., Soliveres, S., Hernández-Clemente, R., Zhao, Y., Gaitán, J.J., Gross, N., Saiz, H., Maire, V., Lehmann, A., Rillig, M.C., Solé, R.V., Maestre, F.T., 2020. Global ecosystem thresholds driven by aridity. Science 367, 787–790. https://doi.org/10.1126/science.aay5958

Bodart, C., Brink, A.B., Donnay, F., Lupi, A., Mayaux, P., Achard, F., 2013. Continental estimates of forest cover and forest cover changes in the dry ecosystems of Africa between 1990 and 2000. Journal of Biogeography 40, 1036–1047. https://doi.org/10.1111/jbi.12084

Boelman, N.T., Stieglitz, M., Griffin, K.L., Shaver, G.R., 2005. Inter-annual variability of NDVI in response to long-term warming and fertilization in wet sedge and tussock tundra. Oecologia 143, 588–597. https://doi.org/10.1007/s00442-005-0012-9

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A., Zemp, M., 2014. The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. Bulletin of the American Meteorological Society 95, 1431–1443. https://doi.org/10.1175/BAMS-D-13-00047.1

Bond, W.J., Midgley, G.F., 2012. Carbon dioxide and the uneasy interactions of trees and savannah grasses. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 601–612. https://doi.org/10.1098/rstb.2011.0182

Boschetti, M., Bocchi, S., Brivio, P.A., 2007. Assessment of pasture production in the Italian Alps using spectrometric and remote sensing information. Agriculture, Ecosystems & Environment 118, 267–272. https://doi.org/10.1016/j.agee.2006.05.024

Bouvet, A., Mermoz, S., Le Toan, T., Villard, L., Mathieu, R., Naidoo, L., Asner, G.P., 2018. An above-ground biomass map of African savannahs and woodlands at 25m resolution derived from ALOS PALSAR. Remote Sensing of Environment 206, 156–173. https://doi.org/10.1016/j.rse.2017.12.030

Bradford, J.B., Hicke, J.A., Lauenroth, W.K., 2005. The relative importance of light-use efficiency modifications from environmental conditions and cultivation for estimation of large-scale net primary productivity. Remote Sensing of Environment 96, 246–255. https://doi.org/10.1016/j.rse.2005.02.013

Braun, A., Wagner, J., Hochschild, V., 2018. Above-ground biomass estimates based on active and passive microwave sensor imagery in low-biomass savanna ecosystems. JARS 12, 046027. https://doi.org/10.1117/1.JRS.12.046027

Brink, A.B., Eva, H.D., 2009. Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach. Applied Geography 29, 501–512. https://doi.org/10.1016/j.apgeog.2008.10.004

Campbell, J.B., Wynne, R.H., 2011. Introduction to Remote Sensing, Fifth Edition. Guilford Press.

Chen, J., Gu, S., Shen, M., Tang, Y., Matsushita, B., 2009. Estimating aboveground biomass of grassland having a high canopy cover: an exploratory analysis of in situ hyperspectral data. International Journal of Remote Sensing 30, 6497–6517. https://doi.org/10.1080/01431160902882496

Chen, X., Vierling, L., Deering, D., 2005. A simple and effective radiometric correction method to improve landscape change detection across sensors and across time. Remote Sensing of Environment 98, 63–79. https://doi.org/10.1016/j.rse.2005.05.021

Chidumayo, E.N., Gumbo, D.J., 2010. The Dry Forests and Woodlands of Africa: Managing for Products and Services. Earthscan.

Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. Remote Sensing of Environment 101, 181–193. https://doi.org/10.1016/j.rse.2005.12.011

Chu, D., 2020. ABOVEGROUND BIOMASS ESTIMATES OF GRASSLAND IN THE NORTH TIBET USING MODIS REMOTE SENSING APPROACHES. Appl. Ecol. Env. Res. 18, 7655–7672. https://doi.org/10.15666/aeer/1806_76557672

Ciais, P., Bombelli, A., Williams, M., Piao, S.L., Chave, J., Ryan, C.M., Henry, M., Brender, P., Valentini, R., 2011. The carbon balance of Africa: synthesis of recent research studies. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 369, 2038–2057. https://doi.org/10.1098/rsta.2010.0328

Collings, S., Caccetta, P., Campbell, N., Wu, X., 2010. Techniques for BRDF Correction of Hyperspectral Mosaics. IEEE Trans. Geosci. Remote Sensing 48, 3733–3746. https://doi.org/10.1109/TGRS.2010.2048574

Collins, W., 1978. Remote sensing of crop type and maturity. Photogrammetric Engineering and Remote Sensing 44.

Colwell, J.E., 1974. Vegetation canopy reflectance. Remote Sensing of Environment 3, 175–183. https://doi.org/10.1016/0034-4257(74)90003-0

Compton J. Tucker, 1977. Asymptotic nature of grass canopy spectral reflectance. URL https://opg.optica.org/view_article.cfm?pdfKey=4ca08562-1c23-495c-ab7615cb9c65bcc9_21018

Compton J. Tucker, 1977. Asymptotic nature of grass canopy spectral reflectance [WWW Document]. URL https://opg.optica.org/view_article.cfm?pdfKey=4ca08562-1c23-495c-ab7615cb9c65bcc9_21018 (accessed 4.12.23).

Crawley, M.J., 2012. The R Book. John Wiley & Sons.

Cutler, M.E.J., Boyd, D.S., Foody, G.M., Vetrivel, A., 2012. Estimating tropical forest biomass with a combination of SAR image texture and Landsat TM data: An assessment of predictions between regions. ISPRS Journal of Photogrammetry and Remote Sensing 70, 66–77. https://doi.org/10.1016/j.isprsjprs.2012.03.011

de Jonge, I.K., Veldhuis, M.P., Vrieling, A., Olff, H., 2022. Camera traps enable the estimation of herbaceous aboveground net primary production (ANPP) in an African savanna at high temporal resolution. Remote Sensing in Ecology and Conservation 8, 583–600. https://doi.org/10.1002/rse2.263

Deering, D., 1979. Rangeland reflectance characteristics measured by aircraft and spacecraft sensors.

Devine, A.P., McDonald, R.A., Quaife, T., Maclean, I.M.D., 2017. Determinants of woody encroachment and cover in African savannas. Oecologia 183, 939–951. https://doi.org/10.1007/s00442-017-3807-6

Doi, T., Behera, S.K., Yamagata, T., 2022. On the Predictability of the Extreme Drought in East Africa During the Short Rains Season. Geophysical Research Letters 49, e2022GL100905. https://doi.org/10.1029/2022GL100905

Doi, T., Nonaka, M., Behera, S., 2020. Skill Assessment of Seasonal-to-Interannual Prediction of Sea Level Anomaly in the North Pacific Based on the SINTEX-F Climate Model. Frontiers in Marine Science 7.

DROACOR software 2.0. ReSe Applications Schläpfer, Wil, Switzerland.

DROACOR. 2022. User Manual, Version 2.0, September 2022

Du, M., Li, M., Noguchi, N., Ji, J., Ye, M. (George), 2023. Retrieval of Fractional Vegetation Cover from Remote Sensing Image of Unmanned Aerial Vehicle Based on Mixed Pixel Decomposition Method. Drones 7, 43. https://doi.org/10.3390/drones7010043

Duan, M., Gao, Q., Wan, Y., Li, Y., Guo, Y., Ganzhu, Z., Liu, Y., Qin, X., 2011. Biomass estimation of alpine grasslands under different grazing intensities using spectral vegetation indices. Canadian Journal of Remote Sensing 37, 413–421. https://doi.org/10.5589/m11-050

Dusseux, P., Hubert-Moy, L., Corpetti, T., Vertès, F., 2015. Evaluation of SPOT imagery for the estimation of grassland biomass. International Journal of Applied Earth Observation and Geoinformation 38, 72–77. https://doi.org/10.1016/j.jag.2014.12.003

Eamus, D., Huete, A., Yu, Q., 2016. Vegetation Dynamics: A Synthesis of Plant Ecophysiology, Remote Sensing and Modelling. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781107286221

Eisfelder, C., Kuenzer, C., Dech, S., 2012. Derivation of biomass information for semi-arid areas using remote-sensing data. International Journal of Remote Sensing 33, 2937–2984.

Eldridge, D.J., Bowker, M.A., Maestre, F.T., Roger, E., Reynolds, J.F., Whitford, W.G., 2011. Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis: Synthesizing shrub encroachment effects. Ecology Letters 14, 709–722. https://doi.org/10.1111/j.1461-0248.2011.01630.x

Elisseeff, A., Pontil, M., 2002. Leave-one-out error and stability of learning algorithms with applications.

Erdogan, H.E., Pellikka, P.K.E., Clark, B., 2011. Modelling the impact of land-cover change on potential soil loss in the Taita Hills, Kenya, between 1987 and 2003 using remote-

sensing and geospatial data. International Journal of Remote Sensing 32, 5919–5945. https://doi.org/10.1080/01431161.2010.499379

Fassnacht, F.E., Poblete-Olivares, J., Rivero, L., Lopatin, J., Ceballos-Comisso, A., Galleguillos, M., 2021. Using Sentinel-2 and canopy height models to derive a landscape-level biomass map covering multiple vegetation types. International Journal of Applied Earth Observation and Geoinformation 94, 102236. https://doi.org/10.1016/j.jag.2020.102236

Ferner, J., Linstädter, A., Rogaß, C., Südekum, K.-H., Schmidtlein, S., 2021. Towards Forage Resource Monitoring in subtropical Savanna Grasslands: going multispectral or hyperspectral? European Journal of Remote Sensing 54, 364–384. https://doi.org/10.1080/22797254.2021.1934556

Frank, H. .2006. URL: https://commons.wikimedia.org/w/index.php?curid=76714066

Galvão, L.S., Almeida de Souza, A., Breunig, F.M., 2020. A hyperspectral experiment over tropical forests based on the EO-1 orbit change and PROSAIL simulation. GIScience & Remote Sensing 57, 74–90. https://doi.org/10.1080/15481603.2019.1668595

Gao, J.-X., Chen, Y.-M., Lü, S.-H., Feng, C.-Y., Chang, X.-L., Ye, S.-X., Liu, J.-D., 2012. A ground spectral model for estimating biomass at the peak of the growing season in Hulunbeier grassland, Inner Mongolia, China. International Journal of Remote Sensing 33, 4029–4043. https://doi.org/10.1080/01431161.2011.639401

Gao, X., Huete, A.R., Ni, W., Miura, T., 2000. Optical–Biophysical Relationships of Vegetation Spectra without Background Contamination. Remote Sensing of Environment 74, 609–620. https://doi.org/10.1016/S0034-4257(00)00150-4

Geisser, S., 1975. The Predictive Sample Reuse Method with Applications. Journal of the American Statistical Association 70, 320–328. https://doi.org/10.2307/2285815

Gessner, U., Machwitz, M., Conrad, C., Dech, S., 2013. Estimating the fractional cover of growth forms and bare surface in savannas. A multi-resolution approach based on regression tree ensembles. Remote Sensing of Environment 129, 90–102. https://doi.org/10.1016/j.rse.2012.10.026

Gitelson, A., Merzlyak, M., Chivkunova, O., 2001. Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves. Photochemistry and photobiology 74, 38–45. https://doi.org/10.1562/0031-8655(2001)074<0038:OPANEO>2.0.CO;2

Grüner, E., Astor, T., Wachendorf, M., 2019. Biomass Prediction of Heterogeneous Temperate Grasslands Using an SfM Approach Based on UAV Imaging. https://doi.org/10.17170/kobra-20190204158

Guanter, L., Brell, M., Chan, J.C.-W., Giardino, C., Gomez-Dans, J., Mielke, C., Morsdorf, F., Segl, K., Yokoya, N., 2019. Synergies of Spaceborne Imaging Spectroscopy with Other Remote Sensing Approaches. Surv Geophys 40, 657–687. https://doi.org/10.1007/s10712-018-9485-z

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci Data 7, 109. https://doi.org/10.1038/s41597-020-0453-3

Hastie, T., Tibshirani, R., 1987. Generalized Additive Models: Some Applications. Journal of the American Statistical Association 82, 371–386. https://doi.org/10.1080/01621459.1987.10478440

Hawkins, D.M., 2004. The Problem of Overfitting. J. Chem. Inf. Comput. Sci. 44, 1–12. https://doi.org/10.1021/ci0342472

Herrmann, I., Shapira, U., Kinast, S., Karnieli, A., Bonfil, D.J., 2013. Ground-level hyperspectral imagery for detecting weeds in wheat fields. Precision Agric 14, 637–659. https://doi.org/10.1007/s11119-013-9321-x

Hill, M.J., Hanan, N.P., 2010. Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales. CRC Press.

Hill, M.J., Hanan, N.P., Hoffmann, W., Scholes, R., Prince, S., Ferwerda, J., Lucas, R.M., Baker, I., Arneth, A., Higgins, I., Barrett, D.J., Disney, M., Hutley, L., 2011. Remote Sensing and Modeling of Savannas: The State of the Dis-Union.

Horler, D.N.H., Dockray, M., Barber, J., 1983. The red edge of plant leaf reflectance. International Journal of Remote Sensing. https://doi.org/10.1080/01431168308948546

Houghton, R.A., Hackler, J.L., 2006. Emissions of carbon from land use change in sub-Saharan Africa. Journal of Geophysical Research: Biogeosciences 111. https://doi.org/10.1029/2005JG000076

Huang, N., He, J.-S., Niu, Z., 2013. Estimating the spatial pattern of soil respiration in Tibetan alpine grasslands using Landsat TM images and MODIS data. Ecological Indicators 26, 117–125. https://doi.org/10.1016/j.ecolind.2012.10.027

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

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

Huete, A.R., 2004. 11 - REMOTE SENSING FOR ENVIRONMENTAL MONITORING, in: Artiola, J.F., Pepper, I.L., Brusseau, M.L. (Eds.), Environmental Monitoring and Characterization. Academic Press, Burlington, pp. 183–206. https://doi.org/10.1016/B978-012064477-3/50013-8

Hurcom, S.J., Harrison, A.R., 1998. The NDVI and spectral decomposition for semi-arid vegetation abundance estimation. International Journal of Remote Sensing 19, 3109–3125. https://doi.org/10.1080/014311698214217

Huston, M.A., Wolverton, S., 2009. The global distribution of net primary production: resolving the paradox. Ecological Monographs 79, 343–377. https://doi.org/10.1890/08-0588.1

Imran, A.B., Khan, K., Ali, N., Ahmad, N., Ali, A., Shah, K., 2020. Narrow band based and broadband derived vegetation indices using Sentinel-2 Imagery to estimate vegetation biomass. Global J. Environ. Sci. Manage. 6. https://doi.org/10.22034/GJESM.2020.01.08

Jackson, R.D., Huete, A.R., 1991. Interpreting vegetation indices. Preventive Veterinary Medicine 11, 185–200. https://doi.org/10.1016/S0167-5877(05)80004-2

Jacon, A.D., Galvão, L.S., Dalagnol, R., dos Santos, J.R., 2021. Aboveground biomass estimates over Brazilian savannas using hyperspectral metrics and machine learning models: experiences with Hyperion/EO-1. GIScience & Remote Sensing 58, 1112–1129. https://doi.org/10.1080/15481603.2021.1969630

Ji, B., Sun, Y., Yang, S., Wan, J., 2007. Artificial neural networks for rice yield prediction in mountainous regions. The Journal of Agricultural Science 145, 249–261. https://doi.org/10.1017/S0021859606006691

Jia, W., Pang, Y., Tortini, R., Schläpfer, D., Li, Z., Roujean, J.-L., 2020. A Kernel-Driven BRDF Approach to Correct Airborne Hyperspectral Imagery over Forested Areas with Rugged Topography. Remote Sensing 12, 432. https://doi.org/10.3390/rs12030432

Jin, Y., Yang, X., Qiu, J., Li, J., Gao, T., Wu, Q., Zhao, F., Ma, H., Yu, H., Xu, B., 2014. Remote Sensing-Based Biomass Estimation and Its Spatio-Temporal Variations in Temperate Grassland, Northern China. Remote Sensing 6, 1496–1513. https://doi.org/10.3390/rs6021496

Jordan, C.F., 1969. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. Ecology 50, 663–666. https://doi.org/10.2307/1936256

Kahiu, M.N., Hanan, N.P., 2018. Fire in sub-Saharan Africa: The fuel, cure and connectivity hypothesis. Global Ecology and Biogeography 27, 946–957. https://doi.org/10.1111/geb.12753

Karakoç, A., Karabulut, M., 2019. Ratio-based vegetation indices for biomass estimation depending on grassland characteristics. Turkish Journal of Botany 43, 619–633. https://doi.org/10.3906/bot-1902-50

Kaspar, F. .2022. 'Contributions to the Improvement of Climate Data Availability and Quality for Sub-Saharan Africa', GPCC Visualizer. URL: https://openverse.org/image/00b9625cb59d41c7ab05526142dc56a3?q=precipitation%20map%20africa

Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, S., 2005. Quantifying grazing intensities using geographic information systems and satellite remote sensing in the Xilingol steppe region, Inner Mongolia, China. Agriculture, Ecosystems & Environment 107, 83–93. https://doi.org/10.1016/j.agee.2004.09.008

Kenya National Bureau of Statistics. 2019. Kenya Population and Housing Census Results.

Knipling, E.B., 1970. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. Remote Sensing of Environment 1, 155–159. https://doi.org/10.1016/S0034-4257(70)80021-9

Kögel-Knabner, I., Amelung, W., 2014. 12.7 - Dynamics, Chemistry, and Preservation of Organic Matter in Soils, in: Holland, H.D., Turekian, K.K. (Eds.), Treatise on Geochemistry (Second Edition). Elsevier, Oxford, pp. 157–215. https://doi.org/10.1016/B978-0-08-095975-7.01012-3

Kokaly, R.F., Clark, R.N., 1999. Spectroscopic Determination of Leaf Biochemistry Using Band-Depth Analysis of Absorption Features and Stepwise Multiple Linear Regression. Remote Sensing of Environment 67, 267–287. https://doi.org/10.1016/S0034-4257(98)00084-4

Kong, B., Yu, H., Du, R., Wang, Q., 2018. Quantitative Estimation of Biomass of Alpine Grasslands Using Hyperspectral Remote Sensing. Rangeland Ecology & Management 72. https://doi.org/10.1016/j.rama.2018.10.005

Kong, B., Yu, H., Du, R., Wang, Q., 2018. Quantitative Estimation of Biomass of Alpine Grasslands Using Hyperspectral Remote Sensing. Rangeland Ecology & Management 72. https://doi.org/10.1016/j.rama.2018.10.005

Kulmatiski, A., Beard, K.H., 2013. Woody plant encroachment facilitated by increased precipitation intensity. Nature Clim Change 3, 833–837. https://doi.org/10.1038/nclimate1904

Kumar, L., Mutanga, O., 2017. Remote Sensing of Above-Ground Biomass. Remote Sensing 9, 935. https://doi.org/10.3390/rs9090935

Kumar, L., Sinha, P., Taylor, S., Alqurashi, A.F., 2015. Review of the use of remote sensing for biomass estimation to support renewable energy generation. JARS 9, 097696. https://doi.org/10.1117/1.JRS.9.097696

Li, J., Mao, X., 2020. Comparison of Canopy Closure Estimation of Plantations Using Parametric, Semi-Parametric, and Non-Parametric Models Based on GF-1 Remote Sensing Images. Forests 11, 597. https://doi.org/10.3390/f11050597

Li, X., Zhang, M., Long, J., Lin, H., 2021. A Novel Method for Estimating Spatial Distribution of Forest Above-Ground Biomass Based on Multispectral Fusion Data and Ensemble Learning Algorithm. Remote Sensing 13, 3910. https://doi.org/10.3390/rs13193910

Ling, B., Raynor, E.J., Goodin, D.G., Joern, A., 2019. Effects of Fire and Large Herbivores on Canopy Nitrogen in a Tallgrass Prairie. Remote Sensing 11, 1364. https://doi.org/10.3390/rs11111364

Lobell, D.B., Field, C.B., 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environ. Res. Lett. 2, 014002. https://doi.org/10.1088/1748-9326/2/1/014002

Loris, V., Damiano, G., 2006. Mapping the green herbage ratio of grasslands using both aerial and satellite-derived spectral reflectance. Agriculture, Ecosystems & Environment 115, 141–149. https://doi.org/10.1016/j.agee.2005.12.018

Lucas, N.S., Curran, P.J., Plummer, S.E., Danson, F.M., 2000. Estimating the stem carbon production of a coniferous forest using an ecosystem simulation model driven by the remotely sensed red edge. International Journal of Remote Sensing 21, 619–631. https://doi.org/10.1080/014311600210461

LUMO Conservancy, 2019. LUMO annual report. URL: https://lumoconservancy.com/wp-content/uploads/2023/12/LUMO_Report-_2019-1.pdf.

LUMO Conservancy. 2023. Who We Are. URL: https://lumoconservancy.com/who-we-are/.

Lyu, X., Li, X., Dang, D., Dou, H., Xuan, X., Liu, S., Li, M., Gong, J., 2020. A new method for grassland degradation monitoring by vegetation species composition using hyperspectral remote sensing. Ecological Indicators 114, 106310. https://doi.org/10.1016/j.ecolind.2020.106310

Lyu, X., Li, X., Gong, J., Li, S., Dou, H., Dang, D., Xuan, X., Wang, H., 2021. Remote-sensing inversion method for aboveground biomass of typical steppe in Inner Mongolia, China. Ecological Indicators 120, 106883. https://doi.org/10.1016/j.ecolind.2020.106883

Mao, P., Ding, J., Jiang, B., Qin, L., Qiu, G.Y., 2022. How can UAV bridge the gap between ground and satellite observations for quantifying the biomass of desert shrub community?

ISPRS Journal of Photogrammetry and Remote Sensing 192, 361–376. https://doi.org/10.1016/j.isprsjprs.2022.08.021

McNaughton, S.J., 1988. Mineral nutrition and spatial concentrations of African ungulates. Nature 334, 343–345. https://doi.org/10.1038/334343a0

Meng, B., Liang, T., Yi, S., Yin, J., Cui, X., Ge, J., Hou, M., Lv, Y., Sun, Y., 2020. Modeling Alpine Grassland Above Ground Biomass Based on Remote Sensing Data and Machine Learning Algorithm: A Case Study in East of the Tibetan Plateau, China. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 13, 2986–2995. https://doi.org/10.1109/JSTARS.2020.2999348

Michelini, S., Šedová, B., Schewe, J., Frieler, K., 2023. Extreme weather impacts do not improve conflict predictions in Africa. Humanit Soc Sci Commun 10, 1–10. https://doi.org/10.1057/s41599-023-01996-1

Mistry, Jayalaxshmi, Beradi, Andrea, 2014. World of savannas: Ecology and Human Use, in: World of Savannas: Ecology and Human Use. Routledge.

Moeckel, T., Safari, H., Reddersen, B., Fricke, T., Wachendorf, M., 2017. Fusion of Ultrasonic and Spectral Sensor Data for Improving the Estimation of Biomass in Grasslands with Heterogeneous Sward Structure. Remote Sensing 9, 98. https://doi.org/10.3390/rs9010098

Mutanga, O., Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. International Journal of Remote Sensing 25, 3999–4014. https://doi.org/10.1080/01431160310001654923

Mutanga, Onisimo, Skidmore, A.K., 2004. Hyperspectral band depth analysis for a better estimation of grass biomass (Cenchrus ciliaris) measured under controlled laboratory conditions. International Journal of Applied Earth Observation and Geoinformation 5, 87–96. https://doi.org/10.1016/j.jag.2004.01.001

Naidoo, L., Van Deventer, H., Ramoelo, A., Mathieu, R., Nondlazi, B., Gangat, R., 2019. Estimating above ground biomass as an indicator of carbon storage in vegetated wetlands of the grassland biome of South Africa. International Journal of Applied Earth Observation and Geoinformation 78, 118–129. https://doi.org/10.1016/j.jag.2019.01.021

O'Connor, T.G., Puttick, J.R., Hoffman, M.T., 2014. Bush encroachment in southern Africa: changes and causes. African Journal of Range & Forage Science 31, 67–88. https://doi.org/10.2989/10220119.2014.939996

Okin, G.S., Roberts, D.A., Murray, B., Okin, W.J., 2001. Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. Remote Sensing of Environment 77, 212–225. https://doi.org/10.1016/S0034-4257(01)00207-3

Olmanson, L.G., Brezonik, P.L., Bauer, M.E., 2013. Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. Remote Sensing of Environment 130, 254–265. https://doi.org/10.1016/j.rse.2012.11.023

P.Dave, C., Joshi, R., Srivastava, S., 2015. A Survey on Geometric Correction of Satellite Imagery. International Journal of Computer Applications 116, 24–27. https://doi.org/10.5120/20389-2655

Panda, S.S., Ames, D.P., Panigrahi, S., 2010. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. Remote Sensing 2, 673–696. https://doi.org/10.3390/rs2030673

Peel, M.C., Finlayson, B.L. and McMahon, T.A. n.d. 'Köppen climate classification map'. University of Melbourne. URL: https://openverse.org/image/6ff19ae0-4bdd-4ba1-a632-d83415ebd2bb?q=savanna%20map

Pellikka, P.K.E., Clark, B.J.F., Gosa, A.G., Himberg, N., Hurskainen, P., Maeda, E., Mwang'ombe, J., Omoro, L.M.A., Siljander, M., 2013. Agricultural Expansion and Its Consequences in the Taita Hills, Kenya, in: Developments in Earth Surface Processes. Elsevier, pp. 165–179. https://doi.org/10.1016/B978-0-444-59559-1.00013-X

Pellikka, P.K.E., Heikinheimo, V., Hietanen, J., Schäfer, E., Siljander, M., Heiskanen, J., 2018. Impact of land cover change on aboveground carbon stocks in Afromontane landscape in Kenya. Applied Geography 94, 178–189. https://doi.org/10.1016/j.apgeog.2018.03.017

Peng, D., Zhang, H., Liu, L., Huang, W., Huete, A.R., Zhang, X., Wang, F., Yu, L., Xie, Q., Wang, C., Luo, S., Li, C., Zhang, B., 2019. Estimating the Aboveground Biomass for Planted Forests Based on Stand Age and Environmental Variables. Remote Sensing 11, 2270. https://doi.org/10.3390/rs11192270

Pennington, R.T., Lehmann, C.E.R., Rowland, L.M., 2018. Tropical savannas and dry forests. Current Biology 28, R541–R545. https://doi.org/10.1016/j.cub.2018.03.014

Platts, P., Burgess, N., Gereau, R., Lovett, J., Marshall, A., Mcclean, C., Pellikka, P., Swetnam, R., Marchant, R., 2011. Delimiting tropical mountain ecoregions for conservation. Environmental Conservation 38, 312–324. https://doi.org/10.1017/S0376892911000191

Poulter, B., Frank, D., Ciais, P., Myneni, R.B., Andela, N., Bi, J., Broquet, G., Canadell, J.G., Chevallier, F., Liu, Y.Y., Running, S.W., Sitch, S., Van Der Werf, G.R., 2014. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. Nature 509, 600–603. https://doi.org/10.1038/nature13376

Psomas, A., Kneubühler, M., Huber, S., Itten, K., Zimmermann, N.E., 2011. Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats. International Journal of Remote Sensing 32, 9007–9031. https://doi.org/10.1080/01431161.2010.532172

Ramoelo, A., Cho, M.A., 2014. Dry season biomass estimation as an indicator of rangeland quantity using multi-scale remote sensing data.

Ramoelo, A., Cho, M.A., Mathieu, R.S., Skidmore, A.K., Schlerf, M., Heitkönig, I.M.A., 2012. Estimating grass nutrients and biomass as an indicator of rangeland (forage) quality and quantity using remote sensing in Savanna ecosystems.

Räsänen, A., Juutinen, S., Kalacska, M., Aurela, M., Heikkinen, P., Mäenpää, K., Rimali, A., Virtanen, T., 2020. Peatland leaf-area index and biomass estimation with ultra-high resolution remote sensing. GIScience & Remote Sensing 57, 943–964. https://doi.org/10.1080/15481603.2020.1829377

Raya-Sereno, M.D., Ortiz-Monasterio, J.I., Alonso-Ayuso, M., Rodrigues, F.A., Rodríguez, A.A., González-Perez, L., Quemada, M., 2021. High-Resolution Airborne Hyperspectral Imagery for Assessing Yield, Biomass, Grain N Concentration, and N Output in Spring Wheat. Remote Sensing 13, 1373. https://doi.org/10.3390/rs13071373

Ren, H., Zhou, G., Zhang, F., 2018. Using negative soil adjustment factor in soil-adjusted vegetation index (SAVI) for aboveground living biomass estimation in arid grasslands. Remote Sensing of Environment 209, 439–445. https://doi.org/10.1016/j.rse.2018.02.068

Rocchini, D., Boyd, D.S., Féret, J.-B., Foody, G.M., He, K.S., Lausch, A., Nagendra, H., Wegmann, M., Pettorelli, N., 2016. Satellite remote sensing to monitor species diversity: potential

and pitfalls. Remote Sensing in Ecology and Conservation 2, 25–36. https://doi.org/10.1002/rse2.9

Rosan, T.M., Aragão, L., Oliveras, I., Phillips, O., Malhi, Y., Gloor, M., Wagner, F., 2019. Extensive twenty-first century woody encroachment in South America's Savanna. Geophysical Research Letters. https://doi.org/10.1029/2019GL082327

RStudio version 12.0.353. RStudio, 2022.

Ruppert, J.H., Johnson, R.H., 2016. On the cumulus diurnal cycle over the tropical warm pool. Journal of Advances in Modeling Earth Systems 8, 669–690. https://doi.org/10.1002/2015MS000610

Sarrazin, M.J.D., Van Aardt, J.A.N., Asner, G.P., McGlinchy, J., Messinger, D.W., Wu, J., 2011. Fusing small-footprint waveform LiDAR and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. Canadian Journal of Remote Sensing 37, 653–665. https://doi.org/10.5589/m12-007

Schucknecht, A., Meroni, M., Kayitakire, F., Boureima, A., 2017. Phenology-Based Biomass Estimation to Support Rangeland Management in Semi-Arid Environments. Remote Sensing 9, 463. https://doi.org/10.3390/rs9050463

Scogings, P.F., Sankaran, M., 2019. Savanna Woody Plants and Large Herbivores. John Wiley & Sons.

SEDACMaps. n.d. URL: https://www.flickr.com/photos/54545503@N04/7242976044

Sellers, P.J., 1985. Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing 6, 1335–1372. https://doi.org/10.1080/01431168508948283

Shackleton, S.E., Shackleton, C.M., Netshiluvhi, T.R., Geach, B.S., Ballance, A., Fairbanks, D.H.K., 2002. Use patterns and value of Savanna resources in three Rural villages in South Africa. Econ Bot 56, 130–146. https://doi.org/10.1663/0013-0001(2002)056[0130:UPAVOS]2.0.CO;2

Shaik, R.U., Periasamy, S., Zeng, W., 2023. Potential Assessment of PRISMA Hyperspectral Imagery for Remote Sensing Applications. Remote Sensing 15, 1378. https://doi.org/10.3390/rs15051378

Sheila Wachiye, Petri Pellikka, Janne Rinne, Janne Heiskanen, Sheil Abwanda, Lutz Merbold, 2022. Effects of livestock and wildlife grazing intensity on soil carbon dioxide flux in the savanna grassland of Kenya | Elsevier Enhanced Reader [WWW Document]. https://doi.org/10.1016/j.agee.2021.107713

Shen, M., Tang, Y., Klein, J., Zhang, P., Gu, S., Shimono, A., Chen, J., 2008. Estimation of aboveground biomass using in situ hyperspectral measurements in five major grassland ecosystems on the Tibetan Plateau. Journal of Plant Ecology 1, 247–257. https://doi.org/10.1093/jpe/rtn025

Shi, Y., Gao, J., Li, X., Li, J., dela Torre, D.M.G., Brierley, G.J., 2021. Improved Estimation of Aboveground Biomass of Disturbed Grassland through Including Bare Ground and Grazing Intensity. Remote Sensing 13, 2105. https://doi.org/10.3390/rs13112105

Shorrocks, B., Bates, W., 2014. The Biology of African Savannahs. OUP Oxford.

Simon, R., Radmacher, M.D., Dobbin, K., McShane, L.M., 2003. Pitfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification. JNCI: Journal of the National Cancer Institute 95, 14–18. https://doi.org/10.1093/jnci/95.1.14

Skidmore, A.K., Ferwerda, J.G., Mutanga, O., Van Wieren, S.E., Peel, M., Grant, R.C., Prins, H.H.T., Balcik, F.B., Venus, V., 2010. Forage quality of savannas — Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. Remote Sensing of Environment 114, 64–72. https://doi.org/10.1016/j.rse.2009.08.010

Song, W., Jia, H., Liu, S., Liang, S., Wang, Z., Hao, L., Chai, S., 2014. A Remote Sensing Based Forage Biomass Yield Inversion Model of Alpine-cold Meadow during Grass-withering Period in Sanjiangyuan Area. IOP Conf. Ser.: Earth Environ. Sci. 17, 012042. https://doi.org/10.1088/1755-1315/17/1/012042

Sousa, J.J., Toscano, P., Matese, A., Di Gennaro, S.F., Berton, A., Gatti, M., Poni, S., Pádua, L., Hruška, J., Morais, R., Peres, E., 2022. UAV-Based Hyperspectral Monitoring Using Push-Broom and Snapshot Sensors: A Multisite Assessment for Precision Viticulture Applications. Sensors 22, 6574. https://doi.org/10.3390/s22176574

Soussana, J.-F., Lüscher, A., 2007. Temperate grasslands and global atmospheric change: a review. Grass and Forage Science 62, 127–134. https://doi.org/10.1111/j.1365-2494.2007.00577.x

Specim CaliGeo Pro 2.2. Spectral Imaging Ltd. Oulu, Finland

SPECIM CaliGeoPRO 2.2 User Guide. Edition 1.0.

Stevens, N., Lehmann, C.E.R., Murphy, B.P., Durigan, G., 2017. Savanna woody encroachment is widespread across three continents. Global Change Biology 23, 235–244. https://doi.org/10.1111/gcb.13409

Stone, M., 1974. Cross-Validatory Choice and Assessment of Statistical Predictions. Journal of the Royal Statistical Society: Series B (Methodological) 36, 111–133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Tang, Z., Zhang, Y., Cong, N., Wimberly, M., Wang, L., Huang, K., Li, J., Zu, J., Zhu, Y., Chen, N., 2019. Spatial pattern of pika holes and their effects on vegetation coverage on the Tibetan Plateau: An analysis using unmanned aerial vehicle imagery. Ecological Indicators 107, 105551. https://doi.org/10.1016/j.ecolind.2019.105551

Thenkabail, P.S., Smith, R.B., De Pauw, E., 2000. Hyperspectral Vegetation Indices and Their Relationships with Agricultural Crop Characteristics. Remote Sensing of Environment 71, 158–182. https://doi.org/10.1016/S0034-4257(99)00067-X

Thenkabail, P.S., Smith, R.B., De Pauw, E., 2002. Evaluation of narrowband and broadband vegetation indices for determining optimal hyperspectral wavebands for agricultural crop characterization. Photogramm. eng. remote sensing 68, 607–621.

Todd, S.W., Hoffer, R.M., 1998. Responses of Spectral Indices to Variations in Vegetation Cover and Soil Background.

Todd, S.W., Hoffer, R.M., Milchunas, D.G., 1998. Biomass estimation on grazed and ungrazed rangelands using spectral indices. International Journal of Remote Sensing 19, 427–438. https://doi.org/10.1080/014311698216071

Toniol, A.C., Galvão, L.S., Ponzoni, F.J., Sano, E.E., de Jesus Amore, D., 2017. Potential of hyperspectral metrics and classifiers for mapping Brazilian savannas in the rainy and dry seasons. Remote Sensing Applications: Society and Environment 8, 20–29. https://doi.org/10.1016/j.rsase.2017.07.004

Tsalyuk, M., 2014. Vegetation - Herbivory Dynamics in Rangeland Ecosystems: Geospatial Modeling for Savanna and Wildlife Conservation in California and Namibia. UC Berkeley.

Tucker, C.J., Sellers, P.J., 1986. Satellite remote sensing of primary production. International Journal of Remote Sensing 7, 1395–1416. https://doi.org/10.1080/01431168608948944

Ullah, S., Si, Y., Schlerf, M., Skidmore, A.K., Shafique, M., Iqbal, I.A., 2012. Estimation of grassland biomass and nitrogen using MERIS data. International Journal of Applied Earth Observation and Geoinformation 19, 196–204. https://doi.org/10.1016/j.jag.2012.05.008

Uno, Y., Prasher, S.O., Lacroix, R., Goel, P.K., Karimi, Y., Viau, A., Patel, R.M., 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. Computers and Electronics in Agriculture 47, 149–161. https://doi.org/10.1016/j.compag.2004.11.014

Valentini, R., Arneth, A., Bombelli, A., Castaldi, S., Cazzolla Gatti, R., Chevallier, F., Ciais, P., Grieco, E., Hartmann, J., Henry, M., Houghton, R.A., Jung, M., Kutsch, W.L., Malhi, Y., Mayorga, E., Merbold, L., Murray-Tortarolo, G., Papale, D., Peylin, P., Poulter, B., Raymond, P.A., Santini, M., Sitch, S., Vaglio Laurin, G., Van Der Werf, G.R., Williams, C.A., Scholes, R.J., 2014. A full greenhouse gases budget of Africa: synthesis, uncertainties, and vulnerabilities. Biogeosciences 11, 381–407. https://doi.org/10.5194/bg-11-381-2014

Van Der Merwe, D., Baldwin, C.E., Boyer, W., 2020. An efficient method for estimating dormant season grass biomass in tallgrass prairie from ultra-high spatial resolution aerial imaging produced with small unmanned aircraft systems. Int. J. Wildland Fire 29, 696. https://doi.org/10.1071/WF19026

Vertès, F., Hatch, D., Velthof, G., Taube, F., Laurent, F., Loiseau, P., Recous, S., 2007. Short-term and cumulative effects of grassland cultivation on nitrogen and carbon cycling in ley-arable rotations. Permanent and temporary grassland: plant, environment and economy. Proceedings of the 14th Symposium of the European Grassland Federation, Ghent, Belgium, 3-5 September 2007 227–246.

Vescovo, L., Gianelle, D., 2008. Using the MIR bands in vegetation indices for the estimation of grassland biophysical parameters from satellite remote sensing in the Alps region of Trentino (Italy). Advances in Space Research 41, 1764–1772. https://doi.org/10.1016/j.asr.2007.07.043

Vuorinne, I., Heiskanen, J., Pellikka, P.K.E., 2021. Assessing Leaf Biomass of Agave sisalana Using Sentinel-2 Vegetation Indices. Remote Sensing 13, 233. https://doi.org/10.3390/rs13020233

Wang, J., Sun, L., Liu, H., 2012. Monitoring biomass of water hyacinth by using hyperspectral remote sensing, in: 2012 First International Conference on Agro- Geoinformatics (Agro-Geoinformatics). Presented at the 2012 First International Conference on Agro-Geoinformatics (Agro-Geoinformatics), pp. 1–4. https://doi.org/10.1109/Agro-Geoinformatics.2012.6311673

Wang, K., Franklin, S.E., Guo, X., Cattet, M., 2010. Remote Sensing of Ecology, Biodiversity and Conservation: A Review from the Perspective of Remote Sensing Specialists. Sensors 10, 9647–9667. https://doi.org/10.3390/s101109647

Wang, Z., Ma, Y., Zhang, Y., Shang, J., 2022. Review of Remote Sensing Applications in Grassland Monitoring. Remote Sensing 14, 2903. https://doi.org/10.3390/rs14122903

Wei, H., Yang, X., Zhang, B., Ding, F., Zhang, W., Liu, S., Chen, F., 2019. Hyper-spectral characteristics of rolled-leaf desert vegetation in the Hexi Corridor, China. J. Arid Land 11, 332–344. https://doi.org/10.1007/s40333-019-0013-x

Whitmarsh, J. and Govindjee. n.d. 'Concepts in Photobiology: Photosynthesis and Photomorphogenesis', in Singhal, G.S., Renger, G., Sopory, S.K., Irrgang, K.-D. and Govindjee (eds.) Narosa Publishers/New Delhi; and Kluwer Academic/Dordrecht, pp. 11-51. URL: http://www.life.uiuc.edu/govindjee/paper/gov.html

Williams, C.A., Hanan, N.P., Neff, J.C., Scholes, R.J., Berry, J.A., Denning, A.S., Baker, D.F., 2007. Africa and the global carbon cycle. Carbon Balance and Management 2, 3. https://doi.org/10.1186/1750-0680-2-3

Williams, K., Wilsey, B., Mcnaughton, S., Banyikwa, F., 1998. Temporally Variable Rainfall Does Not Limit Yields of Serengeti Grasses. Oikos 81, 463. https://doi.org/10.2307/3546768

Wood, S., 2023. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation.

Wood, S., Sebastian, K., Scherr, S.J., 2000. Pilot analysis of global ecosystems: agroecosystems. Pilot analysis of global ecosystems: agroecosystems.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, Second Edition. CRC Press.

Wylie, B.K., Meyer, D.J., Tieszen, L.L., Mannel, S., 2002. Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands: A case study. Remote Sensing of Environment, Recent Advances in Remote Sensing of Biophysical Variables 79, 266–278. https://doi.org/10.1016/S0034-4257(01)00278-4

Xu, B., Yang, X., Tao, W., Qin, Z., Liu, H., Miao, J., 2007. Remote sensing monitoring upon the grass production in China. Acta Ecologica Sinica 27, 405–413. https://doi.org/10.1016/S1872-2032(07)60012-2

Xu, B., Yang, X.C., Tao, W.G., Qin, Z.H., Liu, H.Q., Miao, J.M., Bi, Y.Y., 2008. MODIS-based remote sensing monitoring of grass production in China. International Journal of Remote Sensing 29, 5313–5327. https://doi.org/10.1080/01431160802036276

Xu, D., Guo, X., Li, Z., Yang, X., Yin, H., 2014. Measuring the dead component of mixed grassland with Landsat imagery. Remote Sensing of Environment 142, 33–43. https://doi.org/10.1016/j.rse.2013.11.017

Xu, D., Liu, Y., Xu, W., Guo, X., 2022. The Impact of NPV on the Spectral Parameters in the Yellow-Edge, Red-Edge and NIR Shoulder Wavelength Regions in Grasslands. Remote Sensing 14, 3031. https://doi.org/10.3390/rs14133031

Xu, X., Medvigy, D., Rodriguez-Iturbe, I., 2015. Relation between rainfall intensity and savanna tree abundance explained by water use strategies. Proceedings of the National Academy of Sciences 112, 12992–12996. https://doi.org/10.1073/pnas.1517382112

Xue, J., Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. Journal of Sensors 2017, 1–17. https://doi.org/10.1155/2017/1353691

Yang, S., Feng, Q., Liang, T., Liu, B., Zhang, W., Xie, H., 2018. Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. Remote Sensing of Environment 204, 448–455. https://doi.org/10.1016/j.rse.2017.10.011

Yee, T.W., Mitchell, N.D., 1991. Generalized additive models in plant ecology. J Vegetation Science 2, 587–602. https://doi.org/10.2307/3236170

Zhang, H., Sun, Y., Chang, L., Qin, Yu, Chen, J., Qin, Yan, Du, J., Yi, S., Wang, Y., 2018. Estimation of Grassland Canopy Height and Aboveground Biomass at the Quadrat Scale Using Unmanned Aerial Vehicle. Remote Sensing 10, 851. https://doi.org/10.3390/rs10060851

Zhang, W., Brandt, M., Penuelas, J., Guichard, F., Tong, X., Tian, F., Fensholt, R., 2019. Ecosystem structural changes controlled by altered rainfall climatology in tropical savannas. Nat Commun 10, 671. https://doi.org/10.1038/s41467-019-08602-6

Zhang, X., Tian, M., Chen, X., Fan, Y., Ma, J., Xing, D., 2019. An evaluation model for aboveground biomass based on Hyperspectral Data from field and TM8 in Khorchin grassland, China. https://doi.org/10.1101/792424

Zhang, Y., Huang, J., Huang, H., Li, X., Jin, Y., Guo, H., Feng, Q., Zhao, Y., 2022. Grassland Aboveground Biomass Estimation through Assimilating Remote Sensing Data into a Grass Simulation Model. Remote Sensing 14, 3194. https://doi.org/10.3390/rs14133194

Zhao, G., Sanchez-Azofeifa, A., Laakso, K., Sun, C., Fei, L., 2021. Hyperspectral and Full-Waveform LiDAR Improve Mapping of Tropical Dry Forest's Successional Stages. Remote Sensing 13, 3830. https://doi.org/10.3390/rs13193830

Zhou, W., Li, H., Xie, L., Nie, X., Wang, Z., Du, Z., Yue, T., 2021. Remote sensing inversion of grassland aboveground biomass based on high accuracy surface modeling. Ecological Indicators 121, 107215. https://doi.org/10.1016/j.ecolind.2020.107215

Zumo, I.M., Hashim, M., Hassan, N., 2022. Mapping grass above-ground biomass of grazing-lands using satellite remote sensing. Geocarto International 37, 4843–4856. https://doi.org/10.1080/10106049.2021.1899309