Retrieval of aboveground grass carbon stocks in savannah ecosystems using SAR data products

By

Reneilwe Maake
215081068

Submitted in partial fulfilment of the academic requirements of the Degree of Doctor of Philosophy in Environmental Science (Remote sensing)

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

June 2025

PREFACE

The research was completed by Ms R. Maake while based in the Discipline of Geography, School of Agricultural, Earth and Environmental Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg Campus, South Africa. The research was financially supported by the Agricultural Research Council and National Research Foundation under the SARChI in Land-use Planning and Management.

The work reported in this study has not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.

[redacted]

Signed: Reneilwe Maake

Date: June 2025

[redacted]

Signed: Prof Onisimo Mutanga

Date: June 2025

[redacted]

Signed: Prof George Chirima

Date: June 2025

DECLARATION 1: PLAGIARISM

I, Reneilwe Maake, declare that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii) this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii) this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v) where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi) this dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii) this dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

[redacted]

Signed: Reneilwe Maake

Date: 30 June 2025

DECLARATION 2: PUBLICATIONS

My work as the main author in each paper and presentation is indicated. The * indicates corresponding author.

Chapter 2

1. Reneilwe Maake*, Onisimo Mutanga, George Chirima, and Mbulisi Sibanda. "Quantifying aboveground grass biomass using space-borne sensors: A meta-analysis and systematic review." Geomatics 3, no. 4 (2023): 478-500. DOI: 10.3390/geomatics3040026

This paper presents an analysis of studies collected from peer-reviewed literature. I designed the experiment, collected and analysed the data and ultimately wrote the paper.

Chapter 3

2. Reneilwe Maake*, Onisimo Mutanga, Johannes George Chirima, and Mahlatse Kganyago. "Determining Optimal SAR Parameters for Quantifying Above-Ground Grass Carbon Stock in Savannah Ecosystems Using a Tree-Based Algorithm." Remote Sensing in Earth Systems Sciences (2024): 1-13. DOI: 10.1007/s41976-024-00170-8

This paper is an analysis of data collected across the wet season of 2020 at the Kruger National Park in Skukuza. I designed the experiment, collected and analysed the data, and ultimately wrote the paper.

The paper, based on Chapter 3, was also orally presented at the South African Society of Geographers (SSAG) Conference held at the University of the Free State, QwaQwa campus. South Africa on 2 – 4 October 2023. Paper presented by R. Maake.

Chapter 4

1. Reneilwe Maake*, Onisimo Mutanga, Johannes George Chirima, and Mahlatse Kganyago. "Intra-seasonal evaluation of above-ground grass carbon stocks in savannah ecosystems using remote sensing. In preparation.

This paper is an analysis of data collected across the wet and dry seasons of 2020 at the Kruger National Park in Skukuza. I designed the experiment, collected and analysed the data, and ultimately wrote the paper.

The paper, based on Chapter 4, was also orally presented at the Grassland Society of Southern Africa conference held at the ANEW Hotel Hilton in KwaZulu-Natal, South Africa on 21 to 25 July 2025. Paper presented by R. Maake.

Chapter 5

2. Reneilwe Maake*, Onisimo Mutanga, Johannes George Chirima, and Mahlatse Kganyago. "Inter-seasonal quantification of above-ground grass carbon stock in savannah ecosystems using SAR data." In preparation

This paper is an analysis of data collected across the wet and dry seasons of 2020 and 2021 respectively at the Kruger National Park in Skukuza. I designed the experiment, collected and analysed the data, and finally wrote the paper.

[redacted]

Signed: Reneilwe Maake

Date: 30 June 2025

DEDICATION

I dedicate this thesis to my family for their unwavering support and belief in my endeavours, particularly to my mother (Moyagabo Maake), my son (Tshepang Maake), my fiancé (Nchaka Dikotla), my late grandmother (Elina Sewapa), late aunt (Sandra Maake) and little brother (Karabo Maake).

ABSTRACT

Savannah grassland is one of the world's most extensive biomes, covering approximately 25% of the Earth's surface. The proliferation of natural and human-induced activities, such as grass sward removal, fire occurrence and climate change, deprives the ecosystems' ability to render services such as climate regulation. Degraded savannah grasslands tend to sequester and store carbon at a slower rate than usual. Consequently, high levels of greenhouse gases such as carbon dioxide (CO_2) accumulate in the atmosphere at a faster rate than normal. Savannah grasslands sequester and store carbon in their aboveground plant matter, aboveground litter as well as in belowground matter. Fifty percent (50%) of the aboveground grass biomass (AGGB) comprises carbon. As such, concerns about global warming have sparked studies on ways to limit the accumulation of CO_2 in the atmosphere. The decline in capacity to store carbon contributes to elevated atmospheric CO_2 levels, thereby accelerating global warming. However, limited understanding exists regarding how specific degradation factors affect the carbon storage potential of AGGB. This gap hinders the development of effective mitigation strategies aimed at enhancing the carbon sequestration function of savannah grasslands.

Therefore, the development and implementation of aboveground grass carbon stock (AGGCS) quantification methods and ultimately monitoring operations in savannahs are necessary to determine if the ecosystem is acting as a carbon source or sink. As such, remote sensing tools coupled with scalable machine learning and artificial intelligence offer potential for effective quantification and monitoring of AGGCS.

The purpose of this study was to determine if the Synthetic Aperture Radar (SAR) data products can quantify aboveground grass carbon stocks in savannah ecosystems with considerable predictive accuracy. The scope of this study was limited to (1) reviewing the progress and challenges in the retrieval of AGGCS in Savannah ecosystems using remote sensing, (2) determining the optimal SAR parameters suitable for AGGCS in Savannah ecosystems and further determining if incorporating ancillary variables can boost the predictive performance of SAR-based models in quantifying AGGCS in Savannah ecosystems, (4) evaluating intra-seasonal variations of AGGCS in savannah ecosystems using SAR data (5) assessing the impacts of seasons on the predictive performance of SAR data in quantifying the seasonal variation of AGGCS in Savannah ecosystems.

The Kruger National Park (KNP) in the Mpumalanga Province of South Africa served as an experimental site for extracting grass samples required to achieve the above-mentioned objectives. Ground observation and grass biomass measurements were taken across the wet and dry seasons of 2020 and 2021. To achieve this, Sentinel-1 was used to derive radar indices and texture matrices to quantify AGGCS. Next topographic variables were extracted from the SRTM (Shuttle Radar Topography Mission) Digital Elevation Model (DEM) and added into the model. The XGBoost algorithm was utilised to predict the aboveground grass carbon stocks and computed the regression coefficient (R^2) and error matrices such as MAE (Mean Absolute Error), RMSE (Root Mean Square Error) and RMSE% (Percent Root Mean Square Error) to assess the model performance.

The study found that SAR-derived parameters alone do not fully capture the variability in aboveground grass carbon stock, particularly in the complexly configured savannah ecosystems. The study highlighted the Radar Vegetation Index (RVI) as the most sensitive and consistent parameter for quantifying AGGCS within and across seasons. The study also found that the performance of SAR-based models can be improved with the incorporation of texture and topographic variables. However, when testing the models within seasons and across the years, the study found that the contribution of SAR, texture and topographic parameters varied between the wet and dry seasons in both years. Notably, the predictive capabilities of the models also differed within the seasons and across the years. Overall, RVI, Slope, Terrain Ruggedness Index (TRI), Aspect, Vertical Transmit/Horizontal Receive (VH_{cor}), Topographic Position Index (TPI), Topographic Wetness Index (TWI), Lat, Vertical Transmit/Vertical Receive (VV_{cor}) and Hill shade optimally predicted AGGCS. The XGBoost models that integrated all the variables (topographic, texture and SAR indices) showed the highest predictive power in both the wet and dry seasons across both years. Notably, the model developed for the dry season exhibited superior predictive power (R^2 = 0.86, RMSE% = 13.30 and MAE = 2.81) compared to that of the wet season. The findings underscore the dynamic and variable characteristics of ecosystem processes within savannah environments and their impact on the predictive power of remote sensing-based models. Models that are calibrated for a single season may effectively capture specific conditions, which could hinder their transferability across other temporal scales with different environmental conditions.

Overall, the findings obtained in this study underscore the importance of SAR sensors such as Sentinel-1 as proxies for quantifying aboveground grass carbon stocks in complex savannah ecosystems. These findings have implications for informing ecosystem conservation strategies and climate change modelling, and therefore, support Goal 13 of the Sustainable Development Goals (SDG), which calls for immediate action to combat climate change and mitigate its impacts.

Keywords: Savanna ecosystems, Aboveground grass carbon stocks, Synthetic Aperture Radar, Sentinel-1, Texture analysis, Topographic variables, Machine learning, Seasonal variation.

ACKNOWLEDGMENTS

1 Thessalonians 5:18: "*Give thanks in all circumstances; for this is God's will for you in Christ Jesus*." I am thankful for the lessons, wisdom and strength necessary for the successful completion of this study. It has been the longest of all my academic journeys and contributions to this journey stemmed from different individuals and institutions:

I am particularly grateful to my fathers in academia, Professors Onisimo Mutanga and George Chirima, for holding my hand and never letting go even in circumstances where progress seemed far-fetched. Your level of patience in this journey was beyond exceptional! I have without a doubt learned from the best, that I am now a critical thinker and an independent researcher.

Special thanks to the Agricultural Research Council-NRE for the funding, office space and ICT resource, the National Research Foundation (NRF) Research Chair in Land Use Planning and Management (Grant number: 84157) for funding some of my fieldwork, the University of KwaZulu-Natal for considering my application to pursue my studies, the Kruger National Park, Skukuza, for granting access during field work. All thanks to Samantha Mabuza and Tercia Strydom, Patricia Khoza for facilitating the logistics and documentation and all the game guards who protected us during different occasions of the fieldwork.

I acknowledge the valuable data collection assistance from (Kgaogelo Mogano, Sabelo Mazibuko, Eugene Ratshiedana, Eric Economon, Makgethwa Masemola, Sibongiseni Mgolozeli and Elvis Molobane). Here are the fruits of your labour.

Furthermore, I acknowledge the valuable contributions of Dr. Mahlatse Kganyago and Mbulisi Sibanda regarding the manuscripts resulting from this research. I am grateful for their unwavering willingness to provide technical advice and for their assistance in proofreading my work. Lastly, I would like to thank my Geoinformation and Agromet work family; you have all touched my academic journey in many ways than one, particularly (Zinhle Mashaba-Munghemezulu, Theresa Mazarire, Khaled Abutaleb, Richard Tswai, Siphokazi Gcayi, Lwandile Nduku, Moumakwe Peter, Vuwani Makuya, Solomon Newete, Idani Gundula). Special Thanks to Dr Thomas Fyfield for editing all my published manuscripts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: GENERAL INTRODUCTION

1.1 Background

As terrestrial ecosystems, savannahs represent one of the largest biomes, covering a quarter of the Earth's landscape (Scurlock and Hall, 1998). Tropical savannas cover 15 million km^2 globally, an aerial extent approximated to that of tropical forests (Parton *et al.*, 1993; D'Onofrio *et al.*, 2020). These ecosystems are mostly located within the tropics (Figure 1.1) (Mishra and Young, 2020) across Africa, South America, Central America, Australia, India, and Southeast Asia. Savannahs serve as the primary habitats for the majority of the world's domestic and wildlife animals (Mishra and Young, 2020). These pasturelands, along with designated wildlife conservation areas, underpin various socio-economic activities, including pastoralism (Schmidt and Skidmore, 2001), tourism (Schmidt *et al.*, 2016), and conservation efforts (de Lima *et al.*, 2018). In Africa, for example, rural communities practice grazing, fuel wood and medicinal plants collection on these ecosystems to sustain their livelihoods (Olsson and Ouattara, 2013).

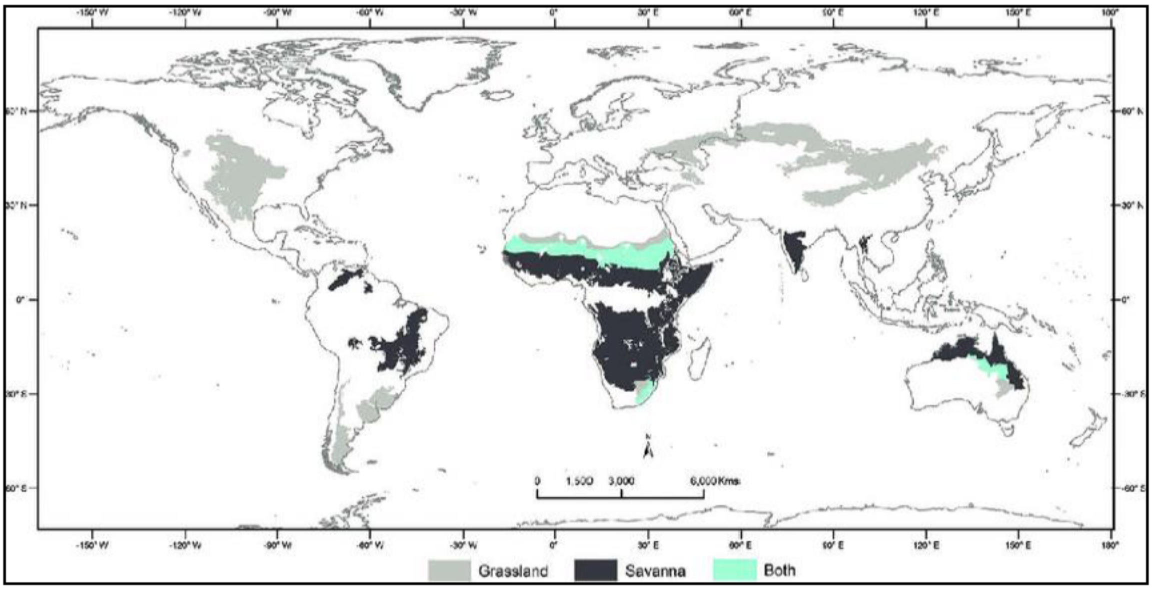

Figure 1. 1 Global geographical extent of natural savannahs and grassland, adopted from Mishra and Young (2020)

Through the climate change lens, savannahs are either a net sink or source for carbon. Nevertheless, owing to their extensive footprint, savannahs sequester carbon through gross primary productivity (GPP), which pockets over 20% of the global GPP annually (Grace *et al.*, 2006). This process plays a crucial role in influencing the global land-atmosphere energy balance and carbon cycles (Mistry, 2000) and, therefore, positioning this ecosystem as a natural sink. This GPP is an interaction between the extensive grassy vegetation layer that coexists with the patched woody vegetation layer (Moore *et al.*, 2016), which both form a characteristic of the savannah landscape (Scholes and Archer, 1997; Mordelet and Le Roux, 2006). The grass layer contributes approximately 10% of the savannah's GPP (Ren and Zhou, 2012) and can account for an annual sink of about 0.5 Petagram (Pg) Carbon (Scurlock and Hall, 1998). In addition, the grass layer encompasses C4 grass species, which possess vital traits to sequester carbon, thereby contributing to mitigating global warming (Ren and Zhou, 2012).

As a terrestrial carbon stock, the grass layer stores carbon in its live and dead organic matter (Lu, 2006), known as biomass, and is further divided into below (BGGB), including the roots and above-ground grass biomass (AGGB), including stems and leaves. The AGGB refers to all organic matter that exists above the soil surface (FAO, 2010), and roughly 50% of this matter contains carbon (Houghton *et al.*, 2009). Therefore, insights on the distribution of AGGB across various spatial and temporal windows is essential, particularly to expedite the full implementation of international Greenhouse gas (GHG) inventory initiatives (Myeong *et al.*, 2006). Furthermore, ecologists and environmentalists utilise this information in evaluating productivity rates, ensuring sustainable bionomics practices. Consequently, it is imperative to evaluate and monitor the biomass of grasses to maintain their ability to sequester carbon, as well as to sustain the various ecosystem services they provide, particularly in the context of climate change.

However, the effectiveness of the grass layer to sink carbon is highly dynamic and shaped by both natural processes (e.g. fire, season) and human activities (e.g. degradation) (Parr *et al.*, 2014). For example, the duration and intensity of the dry and wet seasons significantly affect the distribution of grasses across various regions and among different species (Xu and Baldocchi, 2004). However, human-induced processes, in particular, land use change, increase the rate of the physiological capacity of the grass layer to respire carbon (Kemp and Michalk, 2007). For example, the transformation of grasslands into agricultural land respires roughly 59% of the carbon stored in the soil (Bai and Cotrufo, 2022). The relationship between increased rates of respiration and elevated concentrations of greenhouse gases, particularly CO_2 in the atmosphere has been documented (White *et al.*, 2000). The chemistry between increased levels of CO_2 and global warming have also been documented (Laurin *et al.*, 2014). However, while the potential of the woody layer in the carbon budget is well documented (McKinley *et al.*, 2011), little is known about the potential of the grass layer, particularly in the savannah ecosystem characterised by complex vegetation structure. This is evidenced by global carbon and biodiversity initiatives, such as Reducing Emissions from Deforestation and Forest Degradation (REDD+), which are tailored for forests and often neglect the potential of the grass layer (Pillar and Overbeck, 2025). Consequently, overlooking the contribution of grasses in the global carbon budget will result in an imprecise quantification of carbon sinks and sources, leading to potential underestimations in climate change models.

The side effects of increased rates of carbon respiration are evidenced by high CO_2 concentrations in the atmosphere, which increased from 278 parts per million (ppm) to 420.54 ppm as of April 2023 (Lan *et al.*, 2023). The launch of international programs such as the Intergovernmental Panel on Climate Change (IPCC) and the Kyoto Protocol of the United Nations Framework Convention on Climate Change (UNFCCC) for monitoring and mitigating greenhouse gas emissions and global warming is further proof (IPCC, 2006). These programs underscore the need for reliable and precise information on current and future projections on grassland's status as a carbon sink (UNFCCC, 2007). Thus, quantifying terrestrial carbon stocks and its changes is essential for shedding light on the global carbon budget. However, the foremost setback of carbon stock estimation is the cost tied with the acquisition of site-specific data on a frequent basis (Butterfield and Malmström, 2009). Specifically, the distribution of over 70% world's savannahs is concentrated in the global south, where resource constraints impede the ability to conduct long-term monitoring and validation of carbon stocks (FAO, 2010). Hence, knowledge on the contribution of the grass layer to global carbon cycles remains poor (Dass *et al.*, 2018).

Edaphic conditions vary across different savannah landscapes. This influences the spatial distribution of AGGB carbon stocks which is likely to be dichotomised along the topographic gradient due to changes in environmental aspects such as topography, rainfall, temperature, and species type (Gao *et al.*, 2013; Jiao *et al.*, 2017). For example, Jiao *et al.* (2017) observed that the spatial pattern of AGGB varies across the horizontal and vertical zonality. In addition, Paruelo *et al.* (2010) analysed carbon stocks and fluxes across the environmental gradients of the most extensive temperate rangeland areas. They reported that plant biomass contribution to total carbon averaged 13 % and varied from 9.5% to 27% among sites. However, such studies have not assessed the processes controlling carbon storage in savannah ecosystems. In particular, an understanding of how specific degradation drivers (e.g., fire frequency, grazing intensity, vegetation structural changes) influence AGGB is still limited. As such, it is paramount to derive knowledge on the changes of AGGB carbon stocks in such ecosystems to understand regional carbon cycles and the sustainable use of grassland resources considering the changing climate (Yang *et al.*, 2010).

The habitual practices of retrieving carbon stocks of AGGB relied on plot-based surveys (i.e. visual assessments) and measurements such as clipping and weighing ('t Mannetje, 2000; Flombaum and Sala, 2007; Yang *et al.*, 2010; Redjadj *et al.*, 2012). However, not only are plot-based measurements restricted by extreme costs, but are also limited to individual point measurements, which are unfeasible for remote and inaccessible locations (Lu, 2006). Hence, they are unable to provide immediate and frequently updated estimates of AGGB carbon stocks over a larger footprint (Adjorlolo *et al.*, 2012). This prompts the need to seek immediate and economical techniques that can provide up-to-date, extensive datasets.

Remote sensing (RS) presents an alternative approach for the acquisition of extensive data to quantify AGGB carbon stocks (Kawamura *et al.*, 2005; Piao *et al.*, 2007; Ma *et al.*, 2010; Jin *et al.*, 2011). The quantification of AGGB carbon stocks using RS has for several years relied on multispectral optical sensors. These sensors ranged from course resolution sensors such as NOAA-AVHRR (Piao *et al.*, 2007; Ma *et al.*, 2010), MODIS (Kawamura *et al.*, 2005; Jin *et al.*, 2011), medium resolution such as Landsat (Numata *et al.*, 2007; Xie *et al.*, 2009; Barrachina *et al.*, 2015; Schmidt *et al.*, 2016), to high spatial resolutions (Sentinel-2 multispectral instrument) (Sibanda *et al.*, 2016; Rapiya *et al.*, 2023; Netsianda and Mhangara, 2025). For example, Netsianda and Mhangara (2025) used Sentinel-2 based model, which was able to predict AGGB with an R^2 value of 0.63 $m^2\ m^{-2}$ and an RMSE of 0.15 $m^2\ m^{-2}$. In another study, Sibanda *et al.* (2015) compared the performance of Sentinel-2 and Landsat-8 data for quantifying AGGB treated with different fertilisers. Their findings indicate that Sentinel-2 performed better than Landsat-8. These studies demonstrate the usefulness of remote sensing technology in carbon stock estimations.

However, the mixed woody-grass layers configuration of savannahs poses challenges on the discriminative capability of optical sensors when retrieving AGGCS. For example, optical sensors struggle to distinguish woody and grass layers (Abdi *et al.*, 2022). This is evidenced by for example, Basuki *et al.* (2012) who reported poor estimate of AGGB in heterogeneous landscapes from optical data due to the manifestation of mixed pixels. In addition, optical sensors mostly detect the reflectance of the top of grass canopies (Goh *et al.*, 2013) and therefore omitting vital structural parameters such as grass height, which contribute to the retrieval of carbon stocks, Moreover, these sensors are sensitive to atmospheric and topographic interferences. The acquisition of optical images is interrupted by cloud cover limiting their application in seasonal and long-term monitoring studies (Hong and Wdowinski, 2014). The inability of optical sensors to capture images in all-weather and night conditions (Sinha *et al.*, 2015b) results in the omission of useful data for a wall-to-wall analysis of changes in AGGB carbon stocks. The omission of useful data is more pronounced in the tropical regions such as the Sub-Saharan Africa which are characterized by repeated cloud cover and smoke haze (Pohl and Van Genderen, 1998). As such, testing the capabilities of alternative remote sensing products operating outside the optical region of the electromagnetic spectrum might overcome such challenges.

Sensors such as the SAR and Light Detection and Ranging (LIDAR) operating within the microwave region of the electro-magnetic spectrum can provide an alternative method of overcoming the challenges of quantifying AGGB carbon stocks in tropical savannahs using optical Remote Sensing (Ghasemi *et al.*, 2011; Sinha *et al.*, 2015b). Unlike, optical sensors, these sensors acquire images in all weather conditions and are, therefore, suitable for assessing the spatial and seasonal dynamics of AGGB carbon stocks; especially in the Sub-Saharan Africa with repeated cloud cover conditions (Mitchard *et al.*, 2009; Malenovský *et al.*, 2012; Schmidt *et al.*, 2016; Ali *et al.*, 2017). Radar sensors boost about cloud penetration, therefore, extending the monitoring of carbon stocks to the wet season when optical sensors are obscured by clouds. Moreover, SAR sensors have the ability to detect within-canopy traits of the vegetation, which are almost impossible to detect when using optical sensors (Sinha *et al.*, 2015b).

Realising this potential, attempts have been made to quantify AGGB carbon stocks using SAR data (Tanase *et al.*, 2013; Dusseux *et al.*, 2014; Bouvet *et al.*, 2018; Sinha *et al.*, 2019; Madundo *et al.*, 2023). For instance, Dusseux *et al.* (2014) compared the performance of variables extracted from SPOT-5 and RADARSAT-2 satellite images to monitor grassland biomass. Their results indicated that RADARSAT-2 (96 %) performed better than SPOT-5 (88%) in terms of predictive capability. In addition, Schmidt *et al.* (2016) tested the utility of TerraSAR-X (TSX) radar in relation to Landsat Enhanced Thematic Mapper Plus (ETM+) and Operational Land Imager (OLI) for pasture biomass estimation in open savannah woodland. They concluded that estimates from to Landsat ETM+ and Operational Land Imager (OLI) alone tend to saturate in the cover-to-mass relationship and fail to differentiate standing dry matter from litter. In a synthesis study, Maake *et al.* (2023) assessed performance Radar-based AGGB models against optical-based models. Their results show that Radar-based AGGB retrieval algorithms achieved the highest median coefficient of determination (R^2) compared to optical data.

Nevertheless, timely and reliable Radar-based carbon stock estimates have mostly been achieved using commercial sensors such as RADARSAT-2, TerrarSAR-X, COSMO-SkyMed (Dusseux *et al.*, 2014; Schmidt *et al.*, 2016; Stelmaszczuk-Górska *et al.*, 2018). However, the foremost setback of using such data is their association with extreme costs. As such, they remain operationally underutilised (Sinha *et al.*, 2015a). Hence, the lack of studies in the global south compared to the global north (e.g. Europe, China and North America). The paucity of commercial radar is more pronounced in the African continent. Furthermore, the temporal windows (i.e. 24, 25, 35 and 46 days) of RADARSAT-2, RISAT-1, ENVISAT ASAR and ALOS PALSAR, respectively are inadequate for timelessly characterising seasonal carbon stocks (Ouchi, 2013). Hence, the need to explore other robust, yet affordable microwave sensors that would overcome the above-mentioned shortfalls.

The radar C-SAR sensor on board of Sentinel-1 (S-1) is paving the way to overcome the challenge associated with the costs of acquiring regular data for carbon stocks estimation. S-1 offers valuable data for a wide range of land applications at no cost (Torres *et al.*, 2012; Bargiel, 2017). The mission continues the C-band SAR Earth Observation (EO) of ESA's ERS-1, ERS-2 and ENVISAT, and Canada's RADARSAT-1 and RADARSAT-2 satellites. The mission (incorporating Sentinel-1A and Sentinel-1B satellites) is a polar orbiting day and night SAR imaging mission at an altitude of 700 km. S-1A was launched in April 2014 and S-1B in April 25th 2016 (Torres *et al.*, 2012). The satellite works on the C-band, capable of penetrating through leaves of most grasslands (Ghasemi *et al.*, 2011). In addition, it has improved spatial windows. It consists of four modes: interferometric wide-swath (IW) with a swath width of 250

Kilometres (km) and 5 x 20 m^2 pixel resolution, wave-mode (WV) at 20 x 20 km^2 and 5 x 5 m^2 pixel resolution, strip map (SM) mode at 80 km swath width and 5 x 5 m^2 pixel resolution, and extra wide-swath (EW) at 400 km swath width and 20 x 40 m^2 pixel resolution. Compared to its predecessors, S-1's (IW) mode has much improved spatial (5 x 20 m^2) and temporal (12 days) resolutions suitable for studying the temporal dynamics of AGGCS. S-1 provides selectable polarisation capabilities in a single (HH or VV) or dual polarization (VV+VH or HH+HV) suitable for detecting grassland parameters (Nickmilder *et al.*, 2021; Li *et al.*, 2024).

Attempts have been made to characterise the seasonal variations of AGGB with a single SAR pass in the peak growing season with poor results (Kuplich *et al.*, 2005; Shang *et al.*, 2009). Literature suggests that a series of imagery acquired throughout the growing season provides ample of information for characterizing seasonal dynamics in grasslands that might not be detected with a single pass (O'Connor *et al.*, 2012; Dusseux *et al.*, 2014). Huang and Geiger (2008) observed that inclusion of grass phonological stages increased the accuracy of mapping grass cover. Studies have tested the capability of multi-temporal SAR data for estimating AGB and carbon stocks in the field of forestry (Santos *et al.*, 2003; Mitchard *et al.*, 2009; Kumar *et al.*, 2015; Omar *et al.*, 2017). Meanwhile, studies that attempted to characterize temporal dynamics of AGGB using SAR data were conducted outside sub-Saharan Africa with different environmental settings (Schmidt *et al.*, 2016; Ali *et al.*, 2017; Rapiya *et al.*, 2023). Except for Rapiya *et al.* (2023), these studies have not attempted to quantify temporal dynamics of grass stocks across inter-annual wet and dry seasons. Furthermore, even (Rapiya *et al.*, 2023)'s study was conducted along a climate, soil and topographic gradient that significantly differs from the conditions of the savannah examined in this study. As such, the capability of Sentinel-1 for quantifying the seasonal variability of carbon stocks across the wet and dry seasons in savannah ecosystems is yet to be fully explored.

Similarly, studies also demonstrated signal saturation in SAR-based AGGB estimations (Mutanga *et al.*, 2023). To deal with this, studies opted for the multi-parameter (Kuplich *et al.*, 2005; Cutler *et al.*, 2012; Liu *et al.*, 2019; Yao and Ren, 2024; Maake *et al.*, 2025), others have opted for multi-sensor fusion (Holmström and Fransson, 2003; Dusseux *et al.*, 2014; Hong and Wdowinski, 2014; Shao and Zhang, 2016; Rapiya *et al.*, 2023; Khan *et al.*, 2025). For the multi-parameter approach, where for example SAR and texture parameters are utilized, studies have noted improved prediction powers (Maake *et al.*, 2025) and highlighted the ability of SAR to characterize the structural attributes of grasses and the ability of texture matrices to detect brightness that stem from the canopy structure of vegetation complimentarily improve the model performance. For the multi-sensor approach, Zhu *et al.* (2012) observed that optical instruments capture top of canopy grass attributes (e.g. surface reflectance and emissivity), while SAR instruments capture within grass canopy attributes (e.g. physical constituents, structure and dielectric properties. For example, when comparing the predictive power from Sentinel-1 and Sentinel-2 images separately and in combination, Holmström and Fransson (2003) found better accuracies when the two datasets are combined compared to when they are used separately. de Araujo *et al.* (1999) used a combination of Landsat TM and JERS-1 data to develop a multiple regression model to estimate forest and savannah biomass in Brazil. While most of the abovementioned studies were conducted in a forest environment (Holmström and Fransson, 2003; Goh *et al.*, 2013; Dusseux *et al.*, 2014), few studies have been conducted in a grassland setting (Hong and Wdowinski, 2014), but not for AGGCS estimation in savannah ecosystems. Consequently, testing the potential of combining SAR data with other parameters such as texture and topographic variables for AGGCS estimation in savannah ecosystems is of equal importance.

The development of a reliable model for biomass estimation also depends on the selection of optimal regression algorithms. Several statistical analysis techniques have been utilised for retrieving carbon stocks from remotely sensed data. The most commonly used are simple linear regressions, including stepwise multiple linear regressions (SMLR) (Zhao *et al.*, 2007). Although these models have demonstrated satisfactory predictive capabilities, they also exhibited limitations in their capacity to extract more intricate non-linear relationships that are often associated with heterogeneous landscapes. For example, linear regressions assume additive, monotonic relationships between variables (e.g., RVI and AGGCS). However, ecological processes often follow thresholds, saturations, or interactive effects (e.g., soil moisture and temperature) that linear models fail to represent (Xu *et al.*, 2020). Moreover, SMLR suffer from multicollinearity and over-fitting (Lorber and Kowalski, 1988). Alternatively, machine-learning techniques such as XGBoost offer an inherent capacity to deal with the over-fitting problem of full-spectrum calibration (Chen *et al.*, 2009). Studies show XGBoost consistently outperforms linear regression, Random Forest (RF), and Support Vector Machines (SVM) in biomass estimation accuracy. For example, Li *et al.* (2020), XGBoost achieved higher R^2 values (up to 0.82) and lower RMSE than RF and SVM when using Sentinel-

2 and Landsat-9 data for grassland and forest biomass mapping. It is hypothesised that advanced machine learning algorithms are capable of improving the estimation of AGGCS.

The leveraging of the freely available SAR data and the regression capabilities of machine learning models paves the way for the immediate, continuous, and accurate estimation of AGGCS in the heterogeneously configured savannah ecosystem. These advancements are instrumental in the formulation of credible national climate policies, guiding sustainable management practices in savannah ecosystems, and aiding in the achievement of both national and international carbon accounting objectives, as delineated by the IPCC and the Kyoto Protocol (IPCC, 2006). Consequently, this study seeks to develop a framework for mapping and monitoring the spatial distribution of AGGCS within South African savannah landscapes, illustrating the effective integration of remote sensing, particularly SAR data and machine learning methodologies into carbon stock assessment and management strategies.

1.2 Aims of the study

Based on the gaps identified in the literature, the main aim of the study is to test the utility of SAR data in quantifying seasonal variation of aboveground carbon stocks in savannah ecosystems

1.3 Objectives of the study

Based on the above-mentioned aim was achieved through the following set of objectives:

1. To review the progress and challenges in the retrieval of AGGCS in Savannah ecosystems using remote sensing
2. To determine the optimal SAR parameters suitable for quantifying AGGCS in Savannah ecosystems
3. To evaluate the sensitivity of SAR-derived indices coupled with topo-edaphic and texture parameters in quantifying within-season variation of above-ground grass carbon stocks in savannah ecosystems
4. To evaluate the utility of SAR data in quantifying the inter-annual variation of AGGCS in Savannah ecosystems

1.4 Hypothesis of the study

The study sought to validate the following research hypotheses:

1. The retrieval of aboveground biomass carbon stocks in tropical savannas using remote sensing is influenced by specific challenges and progress in the methodology.
2. The combination of specific SAR-derived and texture parameters improves the accuracy of quantifying AGGCS in Savanna ecosystems than others
3. The within-seasonal variations in above-ground biomass in savanna ecosystems can be accurately estimated with the combined use of SAR-derived indices, texture and topographic variables
4. The predictive capability of SAR-derived indices coupled with topography and texture measures in quantifying AGGCS within Savanna ecosystems is affected by seasonal variation

1.5 Scope of the study

The study sought to assess how SAR-based models perform in quantifying AGGCS in a complex Savannah ecosystem in South Africa. The study was based on the freely available of Sentinel-1 images for extracting predictor variables used to train the models to quantify AGGCS. The predictor variables from the aforementioned remote sensing platform, coupled with topographic data, were used as the foundation for building and assessing the models for quantifying AGGCS. Located within a Savannah ecosystem, the KNP, Skukuza, Mpumalanga Province, dominated by the grass layer co-existing with the woody layer, served as the study area.

1.6 General study site description

Kruger National Park (KNP) is located in the north-eastern South Africa, within the savannah ecosystem (Figure. 1.2). The park, particularly the southern part (24°97' to 25°45' S and 31°00' to 32°01' E), which is bounded by Sabi Park in the north, Marloth Park in the south, and Bushbuckridge in the west, served as an experimental site for this research. Established in 1926, the KNP is one of Africa's most renowned and iconic wildlife reserves and is approximately 20,000 square kilometers (Km^2) in extent (MacFadyen et al., 2019). The KNP has a subtropical climate with warm summers and moderate winters. The summer months, typically from October to March, bring higher temperatures and rainfall to the region. The average annual rainfall in the park ranges from 500 to 700 millimetres, mostly occurring during the summer months (Zambatis *et al.*, 2006). The winters in this area are brief and dry (Kruger *et al.*, 2002) with an average annual minimum temperature of 21.9 °C and maximum temperature of 34 °C. The vegetation comprises dense woodlands and grassy plains (Higgins *et al.*, 2007), the latter being dominated by C4-type grass species (Low, 1998). The vegetation types present in the area include clay thorn bush, mixed bushveld, and sweet and sour lowveld bushveld (Mucina *et al.*, 2006). The grass layer comprises various species such as *Themeda triandra, Panicum maximum Cymbopogon spp and Chlis gayana* (Gertenbach, 1983; Theron *et al.*, 2020). The geological substrate consists of granite and gabbro (Janecke, 2020). The choice for the KNP as an experimental site was informed by its status as one of the most ecologically diverse savanna ecosystems in Southern Africa. The park consists of grassland-dominated areas that extend across diverse gradients of rainfall, soil types, geology, fire regimes, and grazing intensities. These environmental heterogeneities provide an ideal context for investigating the effects of various biophysical factors on AGGCS as well as testing the sensitivity of SAR data in predicting AGGCS.

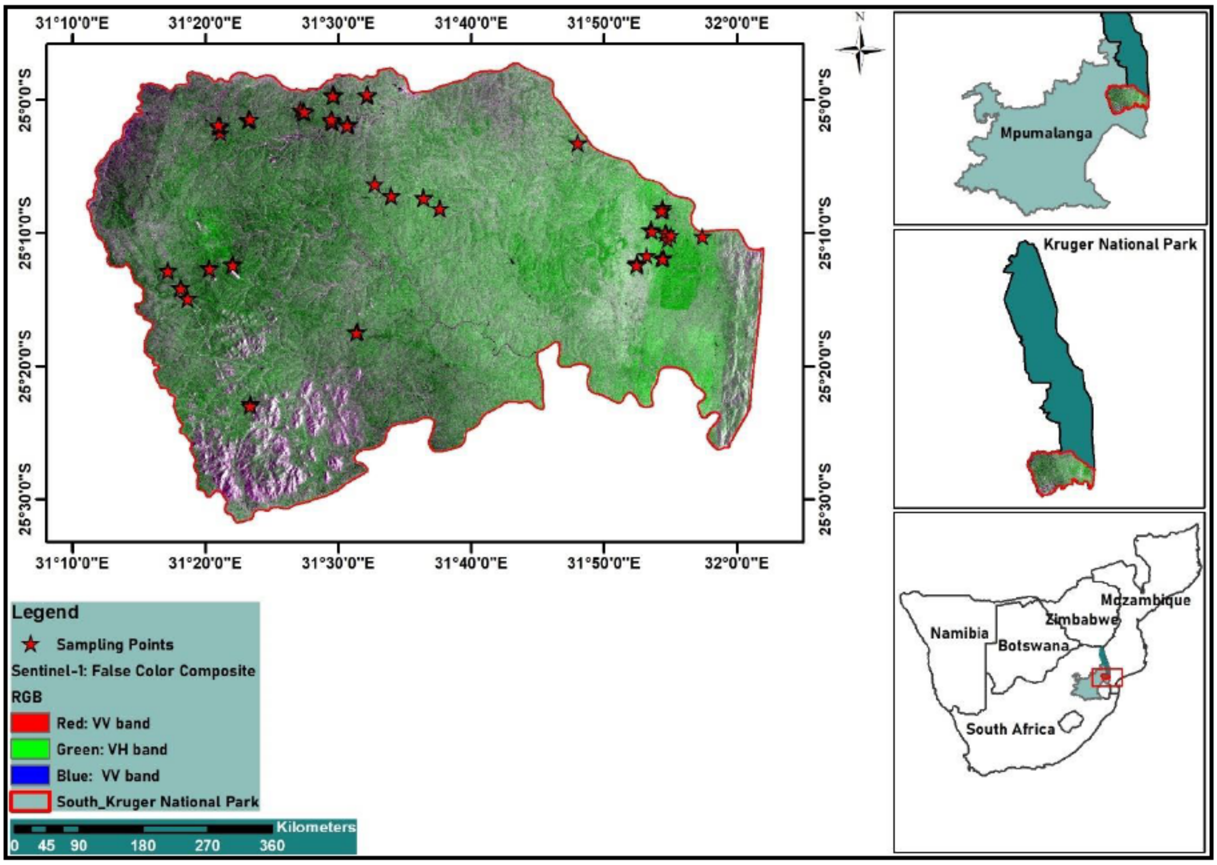

Figure 1. 2 Map showing the location of Kruger National Park and the sampling points. An inset of a Sentinel-1 satellite image is also shown.

1.7 Site and plot demarcation

With the KNP's veld condition assessment (VCA) map (Trollope *et al.*, 1989) as a reference, 25 sites and clustered 5 plots were randomly selected each of these sites. The adopted sampling approach in this study was not only guided by considerations of spatial coverage, logistical practicality, and configuration with remote sensing data, but also by the necessity to capture variability in grass compositions. The VCA sites are approximately 3-12 km apart from each other (Adjorlolo and Mutanga, 2013). As such, the clustering technique was adopted to reduce these significant distances. Significantly, the clusters were situated in areas where the grass layer demonstrated a consistent structural uniformity. With a Garmin Montana 650 Global Positioning System (GPS), we located and marked 10 m x 10 m (0.01 ha) plots at the chosen sites. This plot size was selected to maintain alignment with the spatial resolution of the Sentinel-1 image. A quadrat measuring 50 cm x 50 cm were randomly tossed twice within the designated 100 m^2 plot. The grasses were then cut to ground level within each 50 cm x 50 cm quadrat and weighed the measured fresh biomass with a scale calibrated to a 0.5 g margin of error. The samples were subsequently delivered to a lab, where they were subjected to oven-drying at 65 °C for 48 hours (Maake *et al.*, 2023) until a consistent dry biomass was achieved and then weighed to obtain the dry weight (Figure 1.3). Then, the average of the two quadrats was calculated to indicate the AGGB of that specific plot. The field data collection spanned from 2020 to 2021, covering the wet and dry seasons across the years.

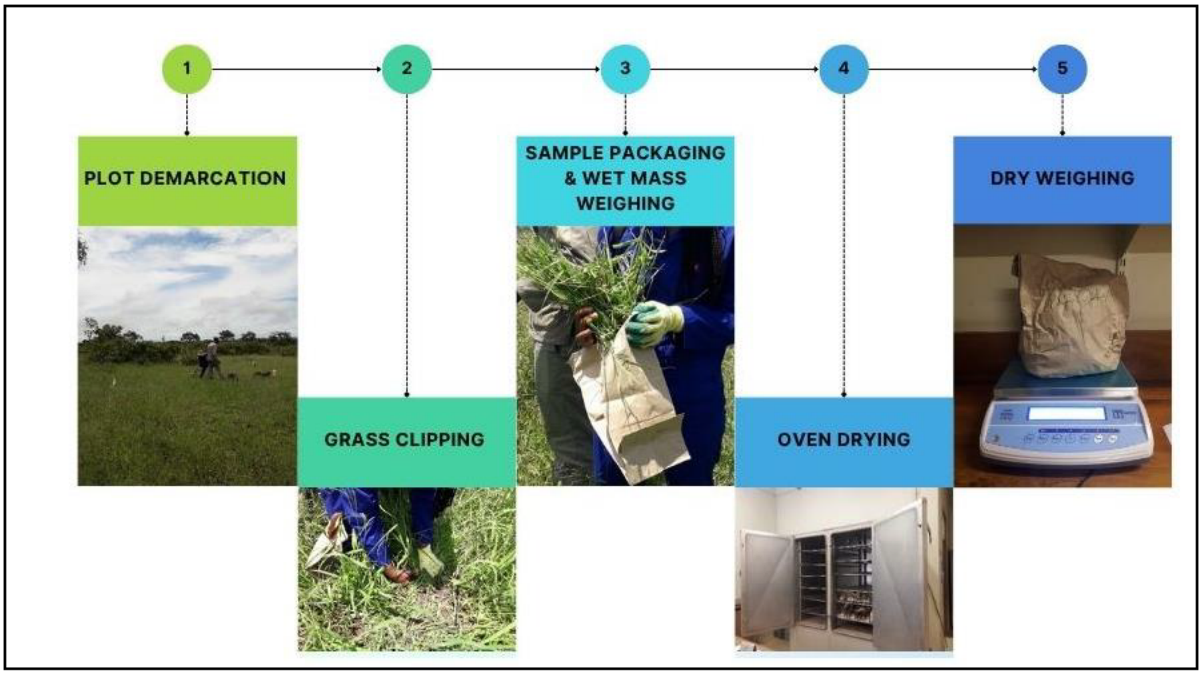

Figure 1. 3 Field data collection and processing workflow, adopted from Maake *et al.* (2025)

1.8 General satellite data acquisition and pre-processing

The Sentinel mission consists of a series of Earth observation satellites launched and operated by the European Space Agency (ESA) as part of the Copernicus program. For this study, satellite images from Sentinel-1C (S1) mission were utilised. The two satellites consist of two constellations (S-1A/B and S-2A/B) respectively, which provide a global median average revisit period of 5 days (Sola *et al.*, 2018). The Sentinel-1 captures images in the Vertical transit/Vertical receive (VH) mode as well as the Vertical transmit/Horizontal receive (VH) channels in an ascending mode at a 10 m spatial resolution (Sentinel-1 Team, 2013). Sentinel-1 was chosen due to its ability to penetrate cloud cover, facilitating consistent year-round monitoring of AGGB in areas such as KNP with frequent cloud cover in the wet season, which often restricts optical data availability. Furthermore, its sensitivity to grass structural attributes including height and density renders it particularly effective for detecting biomass variability across savannah ecosystems. In addition, the sensor’s high temporal resolution (i.e. 6 to 12 days) allows for intra and inter-seasonal analyses, while the 10-meter spatial resolution relatively corresponds with the 10 m × 10 m field plots, thereby minimizing spatial scale mismatch. Sentinel-1 was pre-processed using the Sentinel Application Platform (SNAP) software. Firstly, a subset of the image was created to speed up the subsequent analysis processes. The orbit file was then applied to determine the satellite's position in space during the data acquisition. A radiometric calibration was additionally conducted to transform the image's raw digital numbers into backscatter coefficients known as sigma nought (σ0) values. Afterwards, the image was adjusted for surface features using the SRTM DEM at a resolution of 1-second grid (30 meters); to account for terrain elevation affects the backscatter values. Speckle filtering was further applied to decrease image noise and improve image quality for visualization purposes. Finally, the backscatter coefficients were transformed from linear (sigma nought (σ0) values) to decibels (dB) (Figure 1.4). To complement the SAR data, SRTM DEM was also sourced from NASA (https://www.earthdata.nasa.gov/) and used to derived terrain variables (i.e. slope, aspect, TWI, TRI, TPI and hill hade) to incorporate into the model to account for topographic influences on grass biomass distribution.

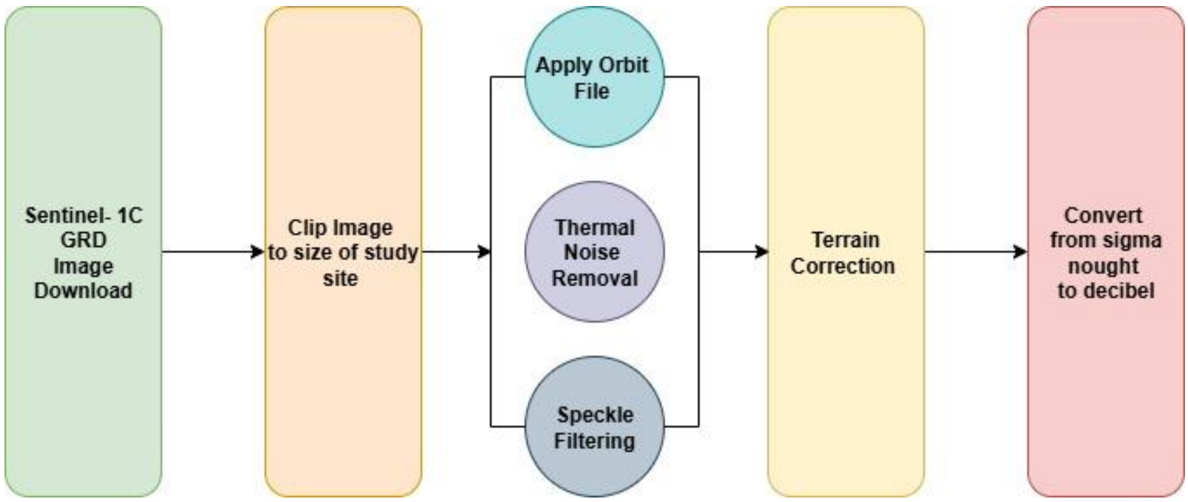

Figure 1. 4 A depiction of Sentinel-1 processing workflows

1.9 Outline of the thesis structure

The dissertation consists of six chapters, with chapter one providing the general introduction of the overall thesis and chapter seven providing the synthesis of all the chapters in the thesis. Five manuscripts were prepared to answer each research question as outlined in subsection 1.4. The literature review and methodology are embedded in each of the mentioned chapters.

1.9.1 Chapter One: *"General Introduction"*

Chapter 1 serves as the introductory section of the study, detailing the relationships among grass carbon stock quantification in savannah ecosystems observed in the literature. Quantifying and monitoring essential vegetation traits, such as grass carbon stock, improves our comprehension of the global carbon cycle. The chapter also explores different techniques for assessing carbon stocks in savannahs, along with their possible constraints and difficulties. It finishes with an in-depth explanation of the research issue, detailing the study's aims and objectives, and highlighting the significance of the grass layer in savannah ecosystems in sequestrating carbon. The chapter established the framework for investigating grass carbon stocks within a savannah ecosystem and informed the research aims and objectives of this study.

1.9.2 Chapter Two: *"Reviewing the progress and challenges in the retrieval of AGGCS in Savanna ecosystems using remote sensing"*

Chapter 2 provides a systematic review of the literature in quantifying aboveground grass carbon stock using space-borne sensors. It provides a comprehensive synthesis of the milestones achieved by studies in pursuit of optimal models for quantifying AGGCS in savannah ecosystems and further provides an overview of how these models perform and the challenges faced. This chapter aims to critically assess how various remote sensing-based models perform in predicting grass carbon stocks and to identify key limitations that constrain their accuracy and operational applicability. The findings indicate that of the publications reviewed (n = 108), only 19% of the studies developed remote sensing-based models for estimating AGGCS in savannah ecosystems, indicating a notable underrepresentation of this ecological system. The underrepresentation becomes more pronounced when focusing on the specific sensors employed, where a limited number of studies developed SAR-based models for estimating AGGCS in savannah ecosystems. Addressing this gab is crucial for developing robust approaches for monitoring AGGCS across savannah landscapes and supporting global carbon accounting, ecosystem management, and climate mitigation strategies.

1.9.3 Chapter three: *"Determine the optimal SAR parameters suitable for quantifying AGGCS in Savanna ecosystems"*

Chapter 3 determines optimal SAR parameters for quantifying AGGCS in Savannah ecosystems. It explores several SAR band and indices configurations that provide the most accurate and reliable estimates of AGGCS in Savannah ecosystems. The chapter examines how various combinations of SAR parameters, such as polarisation, ratios and indices influence the ability to assess carbon stocks in grasslands, offering insights into the most effective combinations of SAR parameters for predicting and monitoring carbon stocks. The findings indicate that parameters derived from a single-date Sentinel-1 image alone were inadequate for quantifying AGGCS in a savannah ecosystem. However, when incorporating texture matrices with the SAR parameters, improved predictive power was observed.

1.9.4 Chapter four: *"Evaluating the sensitivity of SAR-derived indices coupled with topo-edaphic and texture parameters in quantifying within-season variation of above-ground grass carbon stocks in savannah ecosystems"*

Chapter 4 focuses on enhancing the prediction of above-ground grass carbon stocks using a multi-parameter coalescing approach, leveraging data from Sentinel-1, DEM and texture parameters. The chapter explores how integrating various parameters from both radar (SAR) and DEM (terrain) improves the accuracy and reliability of grass carbon stock estimations. By combining the strengths of Sentinel-1's radar and DEM, and texture, this approach aims to refine carbon stock predictions, offering a more comprehensive understanding of grass biomass and its role in carbon sequestration within savannah ecosystems. The findings of this chapter indicate that, as stand-alone predictors, SAR indices, texture and topographic variables produce weak correlations, suggesting that this approach struggles to effectively explain the variability in AGGCS in the complex savannas. Improved predictive powers of the XGBoost model were observed with the coalescing SAR, texture, and topographic across the wet and the dry of one season cycle.

1.9.5 Chapter five: "Evaluating the utility of SAR data in quantifying the inter-annual variation of AGGCS in Savanna ecosystems"

Chapter 5 focuses on the seasonal quantification of aboveground grass carbon stocks in savannah ecosystems using SAR. This chapter investigates how SAR data, collected during different seasons, can be used to monitor and estimate variations in grass biomass and carbon content. By analysing seasonal changes in vegetation cover and structure, the chapter aims to enhance the understanding of grass carbon stocks throughout the year, considering factors such as rainfall, temperature, and growth cycles in savannah environments. The findings of this study show that different combinations of predictors (i.e. radar indices, texture, and topography) contributed differently across the wet and dry seasons of 2020 and 2021, respectively. Nevertheless, among the variables utilised in the study, RVI and TRI consistently contributed the most in modelling AGGCS across the wet and dry seasons of 2020 and the dry season of 2021, respectively. However, the XGBoost models based on the combination of all the variables outperformed those that employed a single-variable model. Overall, the dry season-based AGGCS models consistently outperformed the wet-season-based models within and across the years under investigation. The findings contribute to a more accurate and dynamic carbon stock assessments in these ecosystems.

1.9.6 Chapter six: *"Synthesis"*

Chapter 6 synthesises the findings presented in chapter two to five and offers a thorough conclusion that encapsulates the principal contributions of this research. It emphasises the progress achieved in the comprehension and quantification of aboveground grass carbon stocks within savannah ecosystems, particularly through the application of remote sensing technologies, including SAR and optical sensors. Furthermore, the chapter examines the broader implications of the study's findings in relation to carbon sequestration and climate change mitigation efforts. It also delineates potential future research trajectories, focusing on opportunities to refine biomass estimation methodologies, broaden seasonal analyses, and improve the integration of multi-sensor data to enhance the accuracy of carbon stock predictions. Overall, this study contributes to our understanding of remote sensing-based methodologies for quantifying AGGCS, particularly the complexly configured savannahs. It demonstrates the potential of SAR coalesced with texture and topographic variables in improving the predictive capabilities of AGGCS models.

1.10 Linking of Chapters

Chapter 1 serves as an introduction to the study, offering a comprehensive background on the interrelationship between carbon stock estimation, climate change, and the significance of vegetation parameters, particularly biomass. This chapter establishes the framework for investigating grass carbon stocks within a savannah ecosystem and delineates the research aims and objectives.

Chapter 2 builds on the foundational knowledge outlined in Chapter 1 and expands the scope by conducting a thorough quantitative review of the existing literature regarding the utilisation of space-borne sensors for the quantification of aboveground grass biomass. This chapter underscores the importance of Synthetic Aperture Radar (SAR) and optical satellite imagery in monitoring carbon stocks, thereby contextualising the subsequent application of Sentinel-1 in the subsequent chapters. The findings indicated an underrepresentation of SAR-based models for predicting AGGCS in savannah ecosystems. The observed underrepresentation of SAR-based models tailored towards quantifying AGGCS in savannah ecosystems in the global south directed this study towards addressing this ecological and geographical imbalance. As such, subsequent chapters were diverted towards developing algorithms for quantifying AGGCS in the savannah ecosystem using SAR data products.

Chapter 3 explores the application of SAR, particularly Sentinel-1 derived parameters, including backscatter coefficients, intensity ratios, normalised radar backscatter, arithmetic computations, and the XGBoost tree-based algorithm, to predict the AGGCS. It builds on the findings presented in chapter 2, demonstrating the capabilities of Active remote sensing technologies in estimating carbon stocks in savannah ecosystems, thereby enhancing our comprehension of carbon dynamics within savannah ecosystems. The findings indicate that parameters derived from a single-date Sentinel-1 image alone were inadequate for quantifying AGGCS in a savannah ecosystem. However, incorporating texture matrices with SAR parameters improved predictive power of the models.

Chapter 4 builds upon the findings of Chapter 3 by acknowledging the weakness of SAR data and proposing a multi-parameter coalescing approach that employs both Sentinel-1, DEM and texture matrices to leverage their strengths to enhance the accuracy of grass carbon stock predictions. This chapter extends the SAR analysis presented in Chapter 3 by incorporating ancillary data to refine and improve the estimations of aboveground carbon stocks. It further assesses performance the optimal parameters presented in chapter across the wet and dry seasons of single year. Overall, the integration of Sentinel-1-derived matrices coupled with texture metrics topographic variables and the XGBoost algorithm showed potential in predicting AGGCS in savannah ecosystems.

Chapter 5 expands the scope of chapter 4 and focuses on the assessing the impact of season on the predictive power of the optimal coalesced model presented in chapter 4 across multiple years. It demonstrates that remote sensing technologies are essential for monitoring biomass fluctuations across different seasons, thereby enhancing our comprehension of carbon dynamics within savannah ecosystems. The findings of this study show that different combinations of predictors (i.e. radar indices, texture, and topography) contributed differently across the wet and dry seasons of 2020 and 2021, respectively. Nevertheless, among the variables utilised in the study, RVI and TRI consistently contributed the most in modelling AGGCS across the wet and dry seasons of 2020 and the dry season of 2021, respectively. However, the XGBoost models based on the combination of all the variables outperformed those that employed a single-variable model. Overall, the dry season-based AGGCS models consistently outperformed the wet-season-based models within and across the years under investigation.

Chapter 6 synthesises the key findings of the thesis. It ties up all the individual chapters into coherent whole research. The chapter puts the findings into context by relating them to the broader literature and theoretical framework. It shows how the results align with or differ from existing knowledge. Most importantly, it responds directly to the research Aim. This chapter explains the significance of the findings and what they mean in practice. It also identifies gaps and suggests directions for future research. Overall, it completes the thesis by showing how everything fits together.

CHAPTER 2: QUANTIFYING ABOVEGROUND GRASS BIOMASS USING SPACE-BORNE SENSORS: A META-ANALYSIS AND SYSTEMATIC REVIEW

This chapter is based on:

geomatics

MDPI

Review

Quantifying Aboveground Grass Biomass Using Space-Borne Sensors: A Meta-Analysis and Systematic Review

Reneilwe Maake [1,2,*], Onisimo Mutanga [1], George Chirima [2,3] and Mbulisi Sibanda [1,4]

Abstract

Recently, the move from cost-tied to open-access data has led to the mushrooming of research in pursuit of algorithms for estimating the aboveground grass biomass (AGGB). Nevertheless, a comprehensive synthesis or direction on the milestones achieved or an overview of how these models perform is lacking. This study synthesises the research from decades of experiments in order to point researchers in the direction of what was achieved, the challenges faced, as well as how the models perform. A pool of findings from 108 remote sensing-based AGGB studies published from 1972 to 2020 show that about 19% of the remote sensing-based algorithms were tested in the savannah grasslands. An uneven annual publication yield was observed with approximately 36% of the research output from Asia, whereas countries in the global south yielded few publications (<10%). Optical sensors, particularly MODIS, remain a major source of satellite data for AGGB studies, whilst studies in the global south rarely use active sensors such as Sentinel-1. Optical data tend to produce low regression accuracies that are highly inconsistent across the studies compared to radar. The vegetation indices, particularly the Normalised Difference Vegetation Index (NDVI), remain as the most frequently used predictor variable. The predictor variables such as the sward height, red edge position and backscatter coefficients produced consistent accuracies. Deciding on the optimal algorithm for estimating the AGGB is daunting due to the lack of overlap in the grassland type, location, sensor types, and predictor variables, signalling the need for standardised remote sensing techniques, including data collection methods to ensure the transferability of remote sensing-based AGGB models across multiple locations.

Keywords: meta-analysis; grass biomass; savannah ecosystems; remote sensing

2.1 Introduction

The Savannah grassland is one of the world's most extensive biomes, covering approximately 25% of the Earth's surface (Scurlock and Hall, 1998). These grasslands house C4 grass species (Osborne *et al.*, 2018), which are capable of storing large amounts of carbon, contributing roughly 10% of the global terrestrial carbon stock. In addition, this biological configuration renders the Savannah grassland's essential role players in regulating the world's carbon cycle. In Africa, the savannah grasslands serve as grazing and browsing grounds for domesticated and wild animals and further consist of a distinctive biodiversity with scenic views and wildlife that attracts tourism (Osborne *et al.*, 2018).

Nevertheless, the proliferation of natural and human-induced degradation such as grass sward removal and climate change jeopardises the services that the savannah grasslands provide (Parr *et al.*, 2014). Degraded savannah grasslands tend to sequester and store carbon at a slower rate than normal (Ghosh and Mahanta, 2014), a process which is being intensified by climate change. Consequently, high levels of greenhouse gases such as carbon dioxide (CO2) accumulate in the atmosphere at a faster rate than normal. For example, the literature indicates that global CO2 levels have accrued from 278 ppm six decades ago to 420.54 ppm by April 2023 (Lan, 2023).

Savannah grasslands sequester and store carbon mostly in their above-ground plant matter as well as in above-ground litter and to a limited extent in below-ground matter (Chen *et al.*, 2003). Fifty percent (50%) of the aboveground grass biomass (AGGB) comprises carbon. Therefore, the development and implementation of grass biomass quantification methods and ultimately monitoring operations in savannahs are needed to preserve and sustain the ecosystem's capacity to sequester carbon, and other services it renders, thus contributing to slowing down the current rise in greenhouse gases and subsequently, the resultant global warming effects.

Estimates of the AGGB have for decades been obtained from plot-based surveys (i.e. visual assessments) (Fiala, 2010) involving clipping and weighing (Adjorlolo *et al.*, 2012; Shoko *et al.*, 2016; Sibanda *et al.*, 2016; Meng *et al.*, 2020). Nevertheless, studies have indicated that walking the fields and demarcating baseline plots for making estimates is unfeasible in terms of money, time, and manpower over extensive, remote, and inaccessible geographical areas (Lu, 2006; Eisfelder *et al.*, 2012). However, the immediate need to update AGGB estimates timeously across larger geographical footprints has triggered the need for scientists to seek alternative methods to achieve this goal in an economical and repeatable manner at various spatial scales. Remote sensing in combination with machine learning and artificial intelligence holds great potential as an alternative method and is gaining traction in many biomass quantification studies.

A plethora of studies were undertaken in the pursuit of a suitable algorithm for deriving the AGGB from satellite-based data (Svoray and Shoshany, 2002; Mutanga and Skidmore, 2004; Eisfelder *et al.*, 2012; Shoko *et al.*, 2018b; Debastiani *et al.*, 2019). As indicated by Joshi *et al.* (2016) the development of an appropriate method for quantifying carbon stocks remains a continuing field of exploration. As such, systematically reviewing the volume of published literature will not only provide a rigorous assessment of how remote sensing techniques best estimate AGGB carbon stocks but will also serve as a guideline for future researchers to promptly identify the research needs and direction in this area of study.

2.2 Materials and Methods

Electronic databases such as Scopus and Web of Science are essential sources of scientific publications for peer-reviewed journals suitable for systematic reviews. For this review, we utilised Web of Science, IEEE Explorer Scopus, and Google Scholar to search for relevant peer-reviewed publications. A broad mixture of systematically generated search words was used, including “remote sensing” and “grassland biomass”, to yield a widespread list of publications. A search with these keywords returned a number of peer-reviewed publications with 466 from Web of Science, 78 from IEEE Explorer, 537 from Scopus and 1000 from Google Scholar (Figure 2.1). We then applied a search filter to select the studies that achieved the following criteria: (1) attempted to develop remote sensing-based AGGB models; (2) were published in English from 1972 to 2020, following the launch of the first space-borne satellite; and (3) were peer-reviewed and published in scientific journals. This systematic review primarily excluded relevant studies in other major languages due to constraints related to time, translation resources, and the scope of the study. Performing a multilingual search would have necessitated specialized translation skills that exceeded the scope of this study. Figure 2.1 portrays the PRISMA flow of the publication retrieval procedure. Based on the titles and abstracts, we performed a high-level screening of the articles to check if they met the stated filtering condition. We then downloaded the articles that met the filtering criteria, were available in a portable document format (pdf), were accessible in full length for further reading and analysis and excluded those that were outside the scope of the above-mentioned filtering criteria. The publication that met the selection criteria summed to 108 and was subjected to data extraction and analysis.

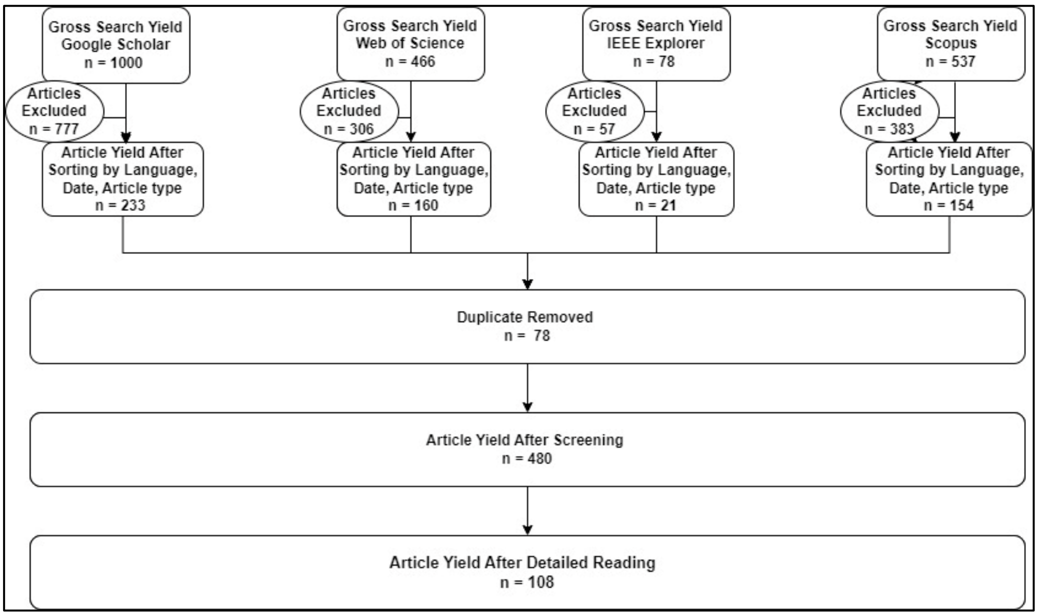

Figure 2. 1 The PRISMA workflow of literature search and selection used for the review

We used the built-in export function in the Web of Science, Scopus, IEEE Explorer, and Google Scholar databases to export the metadata of the searched publications to a Microsoft Excel spreadsheet (Microsoft Corporation, Redmond, WA, USA). Basic publication attributes such as the author, title, journal name, and year were extracted and exported. We then manually extracted additional publication attributes including the study site, geographic coordinates, sampling procedure, platform and sensor type, and spatial resolution, as well as predictor variables and methods through reading to expand and design a comprehensive database to achieve the intended goal of our review. As in Zolkos *et al.* (2013), we also extracted the coefficient of determination statistic (R^2) per data type, sensor as well as grassland types for the best remote sensing-based AGGB algorithm. For the purpose of this review, we only considered the coefficient of determination of the models developed using the training data.

To track the milestones of the remote sensing-based AGGB estimations, statistical frequencies were computed using MS Excel and SPSS. This assisted in observing and identifying trends in terms of the grassland type, geo-location and publications, platforms and sensors, predictor variables and algorithms, as well as the sampling procedure used in the literature. For the geographical trends, we imported the extracted geographical coordinates of where the studies were conducted into a Geographical Information System (GIS) software (Version 10.10) and created a proportional symbol map to visualise the frequency per country. As in in Ma *et al.* (2010), we also computed the mean measure and standard deviation of the coefficient of determination for sensor type and grassland type. We did this to compare the accuracies of the models across data types, sensor and grassland types using the coefficient of determination (R^2) statistic as a measure. To do this, we grouped the multiple R statistics data per sensor and grassland types and plotted box plots. Furthermore, based on the range of spatial resolution found, we categorised the spatial resolution into six subcategories to be able to detect the trends in the use of spatial resolution in the images utilised. The studies reported a mix of findings and necessitated the need to omit some studies that did not report the required attributes. As a result, the number of collected statistics depends on whether the reviewed studies reported it or not. To identify the gaps and challenges, we manually extracted the text through reading and grouped it into themes.

2.3 Results

2.3.1 Milestones of Remote Sensing-Based AGGB Retrieval

2.3.1.1 Sampling and Analysis Protocol for Training and Validation Purposes

Grasslands store carbon in the form of living and dead biomass (Lu, 2006). Grass biomass is the organic material from both below and above the ground (either living or dead). Above-ground grass biomass encompasses all organic matter living above the ground (FAO, 2010). While direct measurements of AGGB are obtainable through unfeasible and costly ground surveys, remotely sensed methods serve as an indirect means of retrieving AGGB (Chave *et al.*, 2019). With this approach, the grass biomass is indirectly inferred from satellite imagery where the relationships between satellite-derived indices and field biomass samples are established (Butterfield and Malmström, 2009; Eisfelder *et al.*, 2012). According to Chave *et* *al.* (2019) "Remote sensing missions are dependent on accurate and representative in situ datasets for the training of their algorithms and product validation".

Ali *et al.* (2016) listed several methods used to collect AGGB samples for training and validating remote sensing data including visual, cut and dry, rising plate meter, and field spectroscopy. Within the reviewed literature, the harvesting (cut and dry) method remains the most heavily utilised destructive method to collect in situ AGGB samples for training and validating remote sensing algorithms (Figure 2.2A). A small number of studies utilised the rising plate meter, followed by the visual technique. For studies that adopted the clipping approach, the next step was to dry and weigh the clipped grass biomass samples in the laboratory at specific timeframes and temperatures. As shown in Figure 2.2B, 65 degrees Celsius was the most frequently used temperature to dry the grass samples, followed by 80 and 70 °C. A few studies dried the grass biomass samples at 95 °C. Figure 2.2C shows that most studies dried the clipped grass samples over 48 h, followed by 72 and 24 h. Figure 2.2D shows the sizes of the subplots (n = 74) used when clipping the grass biomass samples. Of the 74 studies, 52% used $1m^2$ subplots, followed by 0.25 m^2.

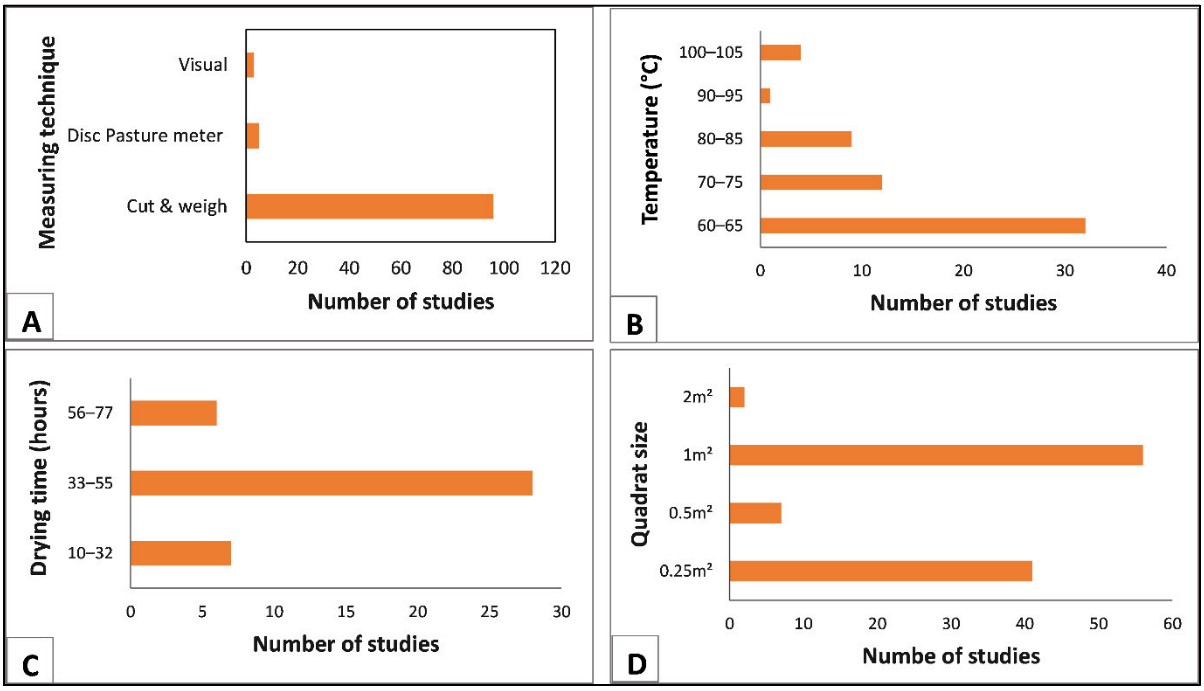

Figure 2. 2 A depiction of (A) field AGGB sample collection techniques; (B) drying temperature expressed in degrees Celsius adopted in the studies; (C) drying time expressed in hours; and (D) the subplot size utilised in the literature

2.3.1.2 Grassland Types Covered in the Reviewed Studies

In order to identify the trends in grassland types, we grouped them based on their geographical location. The trends from the surveyed literature show that four types of grasslands were studied (Figure 2.3A). Of the publications reviewed (n = 108), 62% of the remote sensing-based algorithms were developed and tested in the steppe grasslands, followed by the savannah grasslands (19%) and the prairie grasslands (13%), while the Pampas grasslands were rarely investigated (6%).

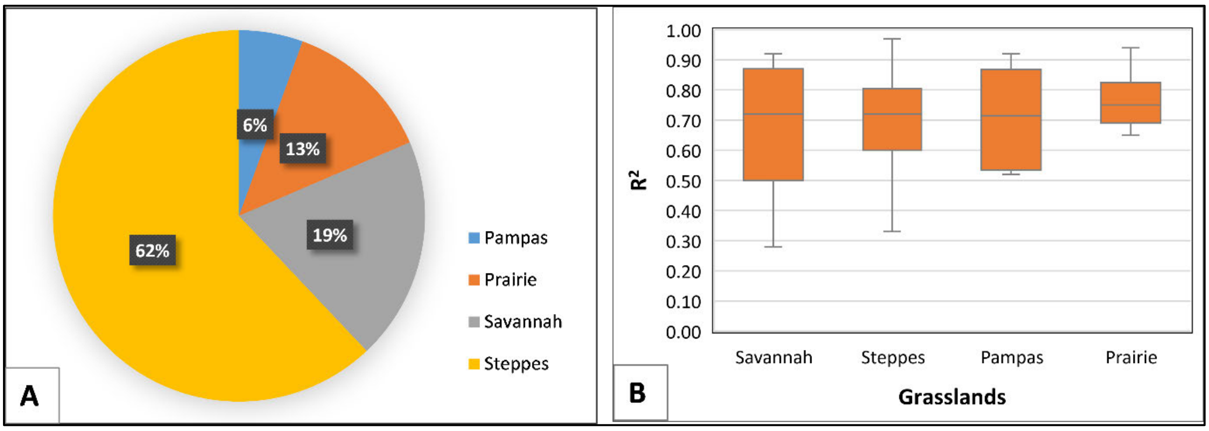

Figure 2. 3 A depiction of (A) grassland types investigated in the reviewed studies, and (B) box plots showing median R^2 for the distribution of grassland type investigated in the selected literature

Figure 2.3B shows a box plot of the coefficient of determination of the remote sensing-based AGGB retrieval algorithm categorised by the grassland type. The grassland retrieval algorithms from the selected studies have medians that are in close range across the grassland type. Specifically, the savannah grasslands show more variability of R^2 values (0.28–0.94), followed by the Pampas (0.33–0.97) and the steppe, while the prairie grasslands show the least variability. The prairie grasslands had a marginally higher mean (0.75) R^2 compared to the savannah grasslands. The models developed in the savannah grasslands show high variability compared to the prairie grasslands.

2.3.1.3 Geographical and Temporal Gradients Covered

The results from a geographical trend analysis show that remote sensing-based algorithms for estimating the AGGB were tested in 24 countries. However, an alarming gap was observed in terms of the regions studied (Figure 2.4A). Specifically, the top five countries in which the algorithms were tested the most were China, South Africa, Germany, the United States of America, and Italy. China yielded the highest proportion of studies (36%) that developed remote sensing-based algorithms for grass biomass retrieval. South Africa had the second largest number of studies (12%) while the USA and Germany contributed small percentages (8 and 6%, respectively). The remaining countries such as Australia, Ireland, and Japan contributed the least (1%)

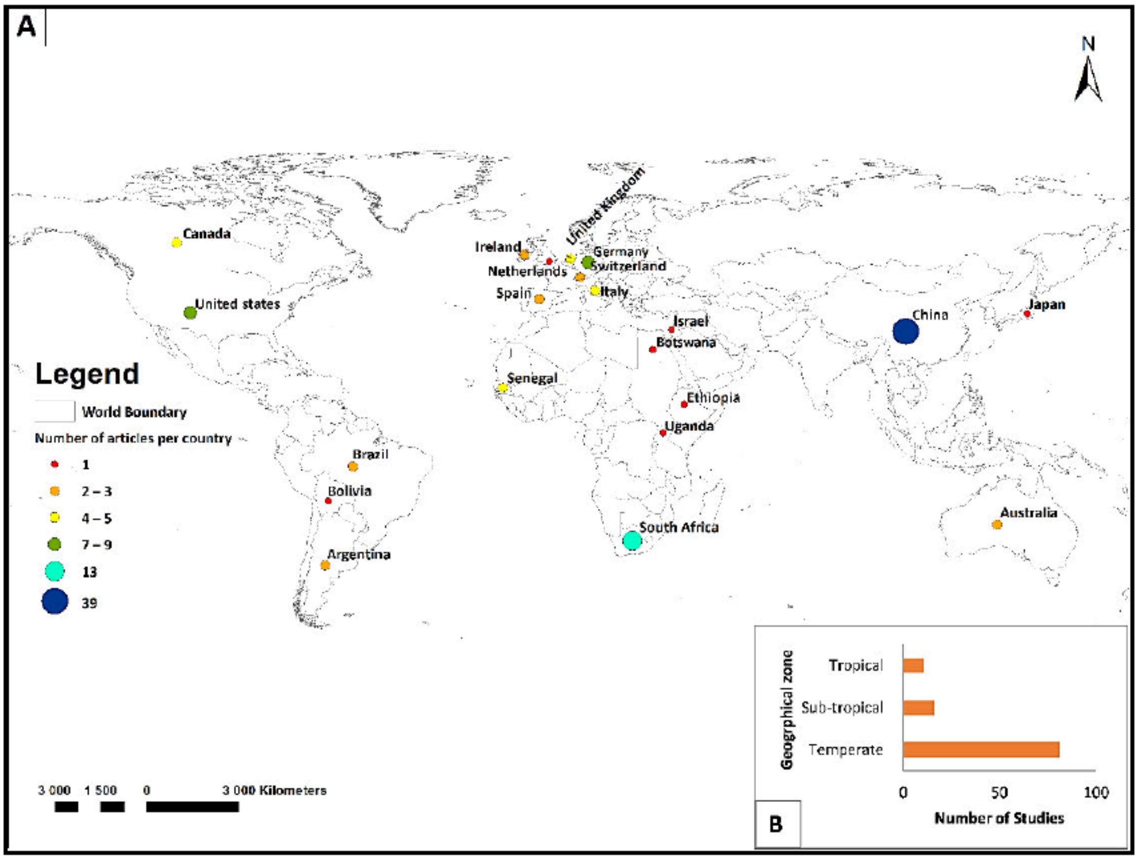

Figure 2. 4 A depiction of (A) the geographical distribution of extracted literature by country, and (B) the geographical climate zones observed in the selected literature

There are four geographical climate zones worldwide, viz. the frigid, temperate, sub-tropical, and tropical zones (Mavridou *et al.*, 2018). The results from a further assessment of the geographical zones in which the studies were undertaken shows that the biomass retrieval algorithms were developed mostly in the temperate zone. Fewer studies (<20) were conducted in the tropical and subtropical zones while no studies were conducted in the frigid zone (Figure 2.4B). Our findings show that Prince and Tucker made the first attempt to develop remote sensing-based algorithms to retrieve the AGGB in 1986. As shown in Figure 2.5A, based on the surveyed literature, two studies on AGGB estimation were conducted between 1986 and 2005, followed by a gradual increase between 2006 and 2009. Subsequently, a decrease in the number of publications occurred between 2010 and 2013. The number of remote sensing-based AGGB studies gradually increased from 2014, with more than six publications per year. The largest number of studies were carried out in 2019 (13 studies).

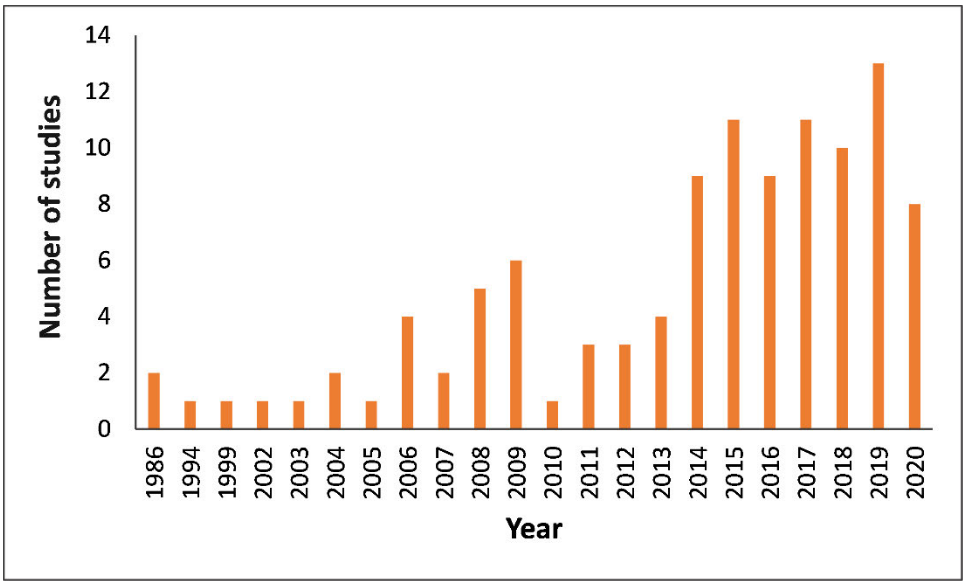

Figure 2. 5 A depiction of the number of publications on remote sensing-based AGGB retrieval by year

An inconsistent trend in the number of publications was observed across the years of the survey (Fig 2.5). In particular, the years 1986–2005 accumulated the lowest yield in the number of publications, followed by a steady increase in the period 2006–2009. Another drop in the number of publications was observed in the period 2010–2013 and a gradual increase in the period 2014–2020.

2.3.1.4 Platform and Sensor Configurations

The findings from the investigated literature (n = 108) show that a variety of platforms and sensors were used. From a platform perspective, the majority of the studies used sensors from space-borne platforms (69%), followed by sensors from ground-based platforms (20%), and lastly air-borne (11%) (Figure 2.6A). The images from optical sensors are the most heavily utilised amongst those selected for developing grass biomass retrieval models, followed by those from radar sensors. Light detecting and ranging (lidar) sensors were rarely utilised (Figure 2.6B).

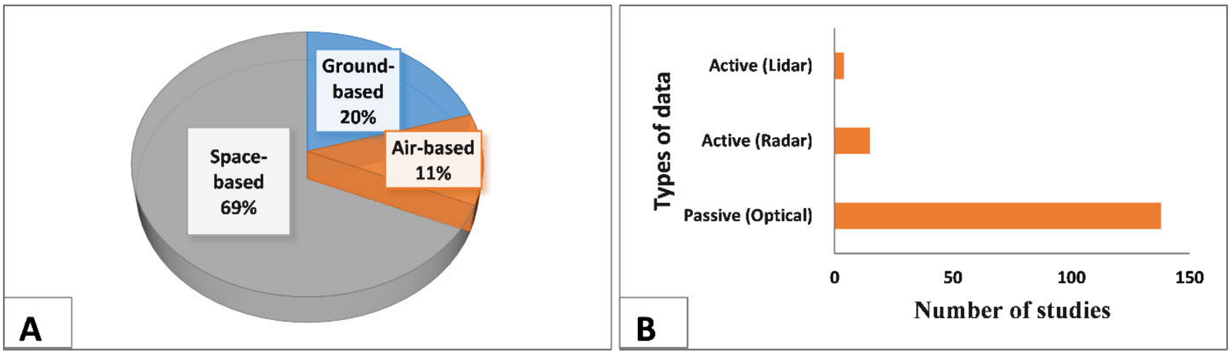

Figure 2. 6 depiction of (A) remote sensing platform, and (B) type of data used in the selected literature categories of remote sensing data used for AGGB retrieval performance by sensor type.

We further assessed if the remote sensing data type utilised had an impact on the performance of the AGGB algorithm by plotting boxplots using the R^2 as presented in Figure 2.7. The median overall coefficient of determination for all the remote sensing data types was above 0.6. Radar-based AGGB retrieval algorithms achieved the highest median R^2 compared to optical data. As shown in Figure 2.7, the accuracies reported for the optical data in the surveyed literature show a large variability compared to the accuracies of radar. The graph also indicates that there are no significant differences between data types although there is much variability in the accuracies of the algorithm derived using passive data.

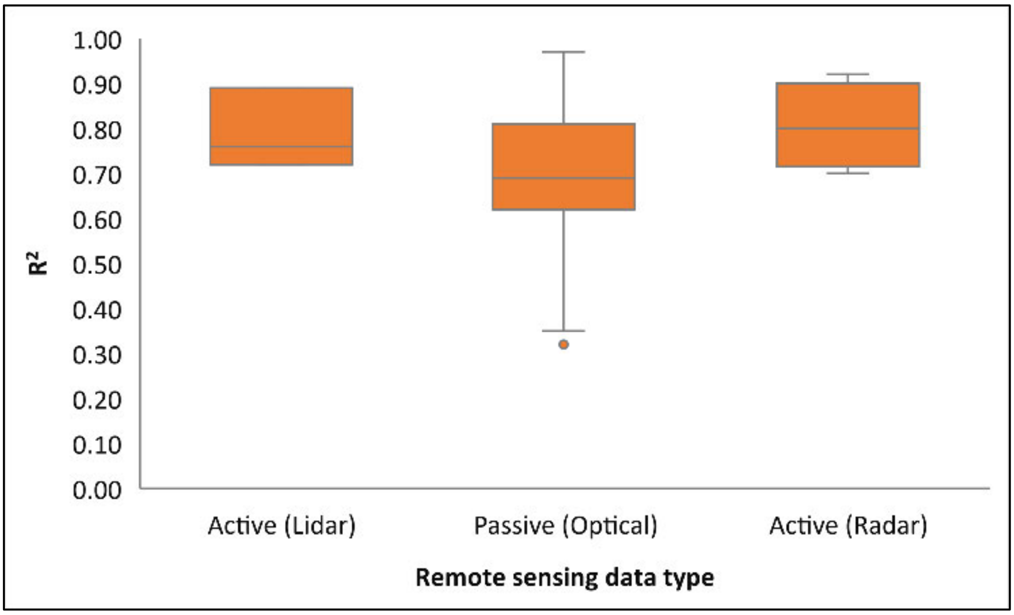

Figure 2. 7 Box plots showing median R^2 by remote sensing data type. The dot shows outlier R^2 values.

From an individual sensor perspective, Table 2.1 shows a variety of sensors that were applied in the AGGB estimation studies. Of the 33 types of sensors applied in the literature under investigation, 72% were optical, 22% were radar and 6% were lidar. The top four sensors that were employed the most were in the optical range and included MODIS, spectrometers, Landsat-8, and Sentinel-2. MODIS and spectroradiometers were the overall most frequently used sensors for AGGB estimation. Within the radar sensor range, ENVISAT ASAR was the most frequently used followed by Sentinel-1 (Table 2.1). Unlike optical sensors, radar sensors such as TerraSAR, CosmoSkyMed, AlosPal-SAR, and Quikscat were utilised in fewer studies. Table 2 shows the studies that applied radar in relation to the grassland type. The savannah grasslands were investigated in three of the studies listed in the table 2.2.

Table 2. 1 Number of publications per individual satellite/instrument.

Sensor	Number of Studies
MODIS	23
Spectroradiometer	23
Landsat 8	18
Sentinel-2	13
AVHRR	4
Landsat-5	4
Landsat-7 ETM$^+$	4
SPOT-5	4
SPOT VGT	4
WorldView-2	4
ENVISAT ASAR	4
UAV	4
WorldView-3	3
PROBA-V	3
Lidar	3
Sentinel-1	3
TerraSAR-X	2
European Remote-Sensing (ERS-1)	2
ALOS PALSAR	2
HJ1B-CCD2	2
RADARSAT	2
COSMO-SkyMed (CSK)	1
Hyperion	1

Indian Remote Sensing	1
MERIS	1
HyMap	1
Apex	1
SSMI	1
Ultrasonic distance sensor	1
QuikScat	1
SPOT-4	1
Digital camera	1
SEA WIND	1

Table 2. 2 An overview of the studies that used radar satellite/instrument to estimate AGGB

Authors	**satellite/instrument.**	**Country**	**Grassland Type**
Sang *et al.* (2014)	ENVISAT-ASAR	China	Steppe
Moreau and Le Toan (2003)	ERS-1	Bolivia	Savannah
Wang *et al.* (2014b)	ENVISAT-ASAR and ALOS POLSAR	Australia	Pampas
Svoray and Shoshany (2002)	ERS-1	Israel	Steppe
Hajj *et al.* (2014)	TERRASAR-X-and COSMO SKYMED	France	Steppe
Schmidt *et al.* (2016)	TERRASAR-X-	Australia	Savannah
(Ali *et al.*, 2017)	TERRASAR-X	Ireland	Steppe
Bao *et al.* (2019)	Sentinel-1	China	Pampas
Naidoo *et al.* (2019)	Sentinel-1	South Africa	Savannah
Braun *et al.* (2018)	ENVISAT-ASAR, ALOS POLSAR and SSMI	Senegal	Savannah
Frolking *et al.* (2005)	SEA WIND	United States	Prairie
Wang *et al.* (2019b)	Sentinel-1	United States	Prairie
(Li and Guo, 2017)	RADARSAT	Canada	Prairie

A further quantitative analysis of the spatial resolution of sensors used in the selected literature shows that the spatial resolution of the sensor utilised in the reviewed literature varied considerably. The majority of the studies employed sensors with spatial resolution >30 metres and 5–30 m and lastly 0–5 m (Figure 2.8). The most widely used sensors were MODIS in the

>30 metres spatial resolution range and Spectroradiometer in the 0–5 m metres spatial resolution range.

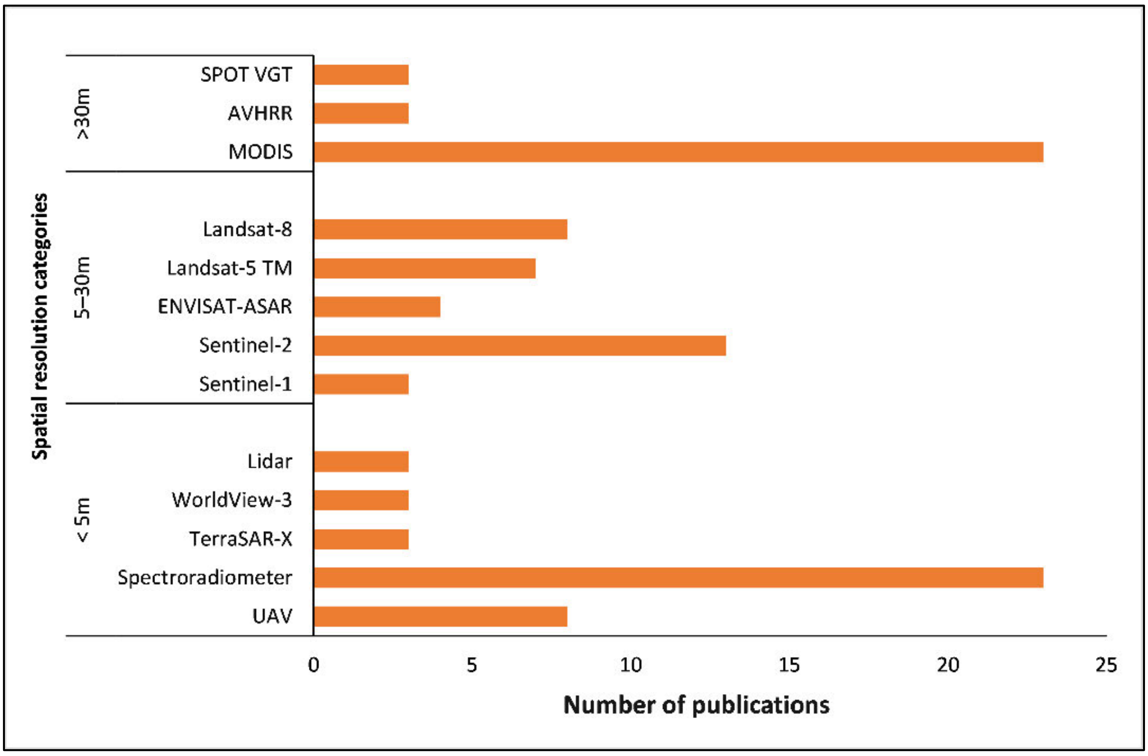

Figure 2. 8 Spatial resolution range associated with the sensors utilised in the reviewed literature. Sensors that were utilised in less than three studies were excluded.

An assessed of whether the spatial resolution influenced the accuracies obtained in these studies was performed. To do this, the sensors were grouped based on their spatial resolution and classified as high (≤5 m), medium (6–30 m), and low (>30 m). As shown in Figure 2.9A, there is a relationship between the spatial resolution and the performance of the model. At a higher spatial resolution, the studies consulted achieved high accuracies, with a mean prediction accuracy of 0.75, while at a lower spatial resolution, low prediction accuracies were reported. In addition, at the medium spatial resolution, the studies obtained highly variable accuracies ranging from 0.28 to 0.75.

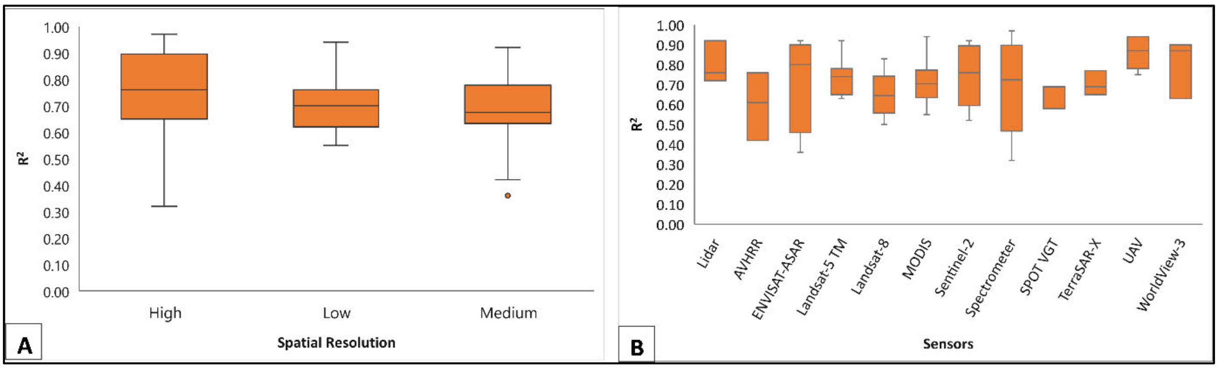

Figure 2. 9 depiction of (A) box plots showing R^2 values for the categories of sensor spatial resolution utilised in the surveyed literature (high ≤ 5 m, medium = 6–30 m and low ≥ 30 m), and (B) box plots showing R^2 values for the sensors utilised in the studies.

Figure 2.9B shows the median R^2 per sensor. Unmanned aerial vehicle (UAV)-based sensors had the highest median coefficient of determination, followed by lidar and Sentinel-3. Generally, sensors with low spatial resolution seem to perform poorly, as also shown in this analysis. Furthermore, the findings from the literature revealed an insignificant variability across the estimates from each sensor class, which had similar minimum RSE values.

2.3.1.5 Predictor Variables Commonly Used in AGGB Studies

Several predictor variables (Table 2.3), such as raw spectral bands (SB), vegetation indices (VI), backscatter coefficients (BC), sward height (SH), fraction of photosynthetically active radiation (FAPAR), agrometeorological variables (rainfall and temperature), topographical variables and a combination of multiple variables (VC) have been widely utilised in remote sensing AGGB. Of the predictor variables (n = 78) (Figure 2.10A), 48% of the studies used vegetation indices as inputs in the developed models for estimating the AGGB, followed by a combination of two or more of the variables mentioned above (e.g., spectral bands + vegetation indices + climate variables). Furthermore, 12% of the studies used spectral bands while 11% of the studies used backscatter coefficients.

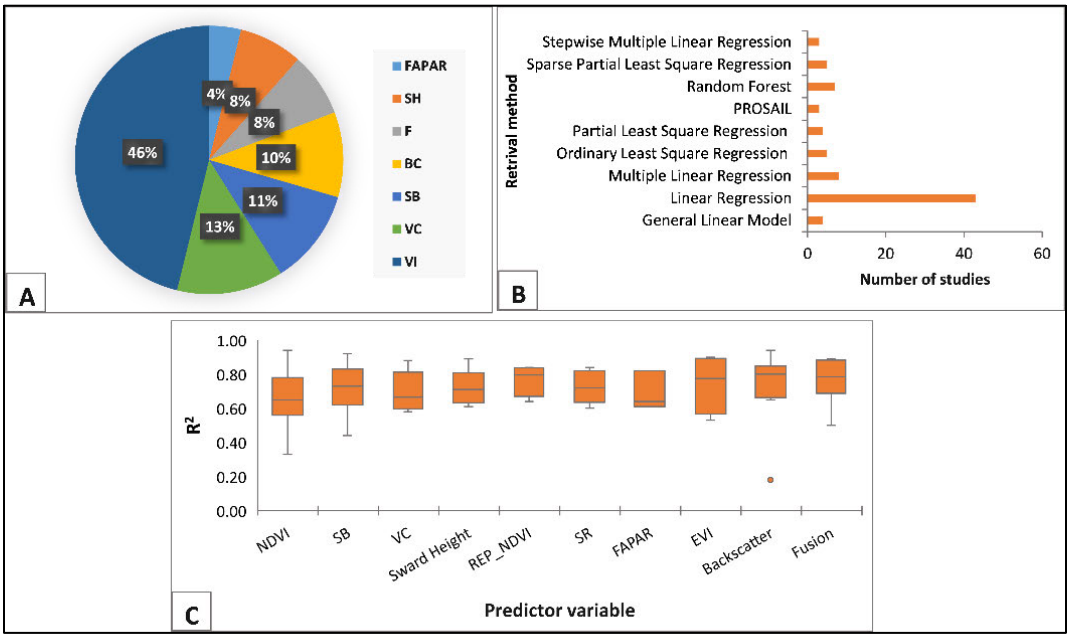

Figure 2. 10 A depiction of (A) predictor variable commonly used for estimating AGGB (SB = spectral bands, VI = vegetation indices, BC = backscatter coefficients, SH = sward height, FAPAR = fraction of photosynthetically active radiation, VC = variable composite, F = Fusion

Table 2. 3 A summary of the predictor variables used in the surveyed literature

Variable Name	Expression/Spectral Bands	Example Study
Backscatter	HH	Ali *et al.* (2017)
Canopy Sward Height	N/A	Wijesingha *et al.* (2019)
Enhanced Vegetation Index (EVI)	$(R_{851} - R_{655})/(R_{851} + 6R_{655} - 7.8R_{482} + 1)$	Meng *et al.* (2020)
Fraction of Photosynthetically Active Radiation (FAPAR)	$FAPAR = 1 - t - r + tr_s$	Schmidt *et al.* (2016)
Modified Soil-Adjusted Vegetation Index (MSAVI)	$[2NIR + 1 - ((2NIR + 1)2 - 8(NIR - R))\ 0.5]/2$	Qi *et al.* (1994)
Normalised Difference Vegetation Index (NDVI)	NDVI = − (infrared band − red band)/(infrared band + red band)	Ikeda *et al.* (1999)
Normalised Band Depth Index (NBDI)	NBDI = BD − Dc/BD + D	Ullah *et al.* (2012)
Ratio Vegetation Index (RVI)	RVI = NIR/Red	Ding *et al.* (2019)
Red Edge-Based NDVI	$(R_{750} - R_{705})/(R_{750} + R_{705})$	Li and Guo (2018)

Red Edge-Based Simple Ratio	$(R_{708} - R_{755})/(R_{708} + R_{755})$	Mutanga and Skidmore (2004)
Simple Ratio (SR)	SR = NIR/Red	Ren and Feng (2015)
Soil-Adjusted Vegetation Index (SAVI)	$1 + L \times (R_{NIR} - R_{RED})/(R_{NIR} + R_{RED}) + L$	Ren and Feng (2015)

R_{NIR} is the reflectance in the Near Infrared band, R_{Red} is the reflectance in the Red band, BD = band depth, Dc = reflectance of the continuum at the same wavelength

The results also revealed nine types of vegetation indices identified in the reviewed literature (Figure 2.10C). Of the vegetation indices, the Normalised Difference Vegetation Index (NDVI) was applied in half of the studies (n = 49), the red edge-based NDVI was the second largest, followed by the Enhanced Vegetation Index (EVI) (8%) and the Modified Soil-Adjusted Vegetation Index (MSAVI) (6%). Although the vegetation indices remain the most widely used input variables in remote sensing-based AGGB retrieval models, the accuracies obtained using these variables in the studies reviewed strongly vary compared to those obtained using backscatter coefficients, sward height and red edge-based NDVI variables (Figure 2.10C).

2.3.1.6 Algorithms Commonly Used in Remote Sensing-Based AGGB Studies

We detected nine regression (9) algorithms in the surveyed literature used to predict AGGB in 82 studies (Figure 10B and 11A). Fifty-two percent (52%) of the studies used linear regression methods. Only a few studies (8%) used machine-learning algorithms such as Random Forest (RF) and 4% used a production efficiency model such as PROSAIL. These models yielded a median R^2 value above 60% showing considerable prediction accuracies in estimating the AGGB (Figure 11B). SPLRS had the highest average prediction accuracy (85%) followed by RF and OLSR.

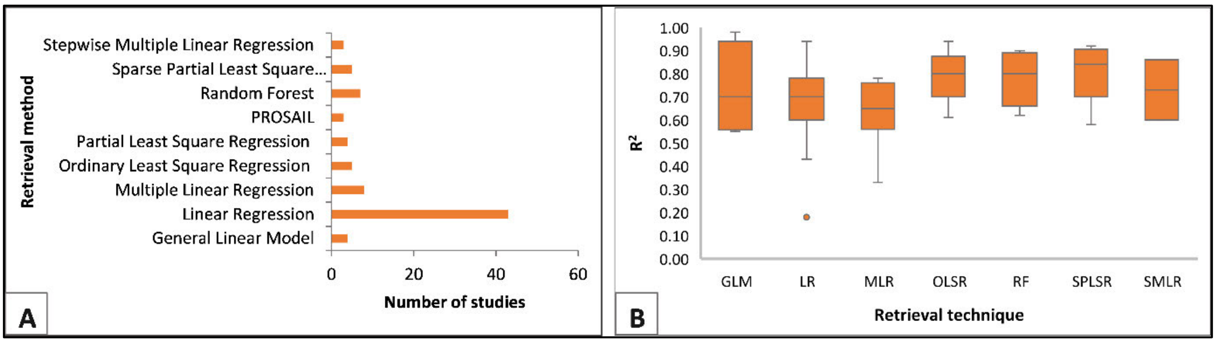

Figure 2. 11 depiction of (A) models commonly used to retrieve AGGB within the reviewed studies, and (B) box plots portraying the R^2 values per model used in the surveyed studies (GLM = General Linear Model, LR = Linear regression, MLR = Multiple linear regression.

2.4 Discussion

2.4.1 Milestones of Remote Sensing-Based AGGB Retrieval

2.4.1.1 Sampling and Analysis Protocol for Training and Validation Purposes

The ability of remote sensing-based algorithms to accurately measure biomass varies with the type of grassland (Su and Bork, 2006). The evidence from the surveyed literature indicates that few studies developed remote sensing-based AGGB models in the savannah grasslands compared to the steppe grasslands. This could be because the steppe is the world's biggest and most extensive grassland biome of the temperate climatic zone (i.e., the Eurasian Steppe) spreading from Hungary to China. However, the savannah ecosystem is characterised by a grass layer that co-exists with scattered trees and shrubs (Mishra and Young, 2020). Surprisingly, savannahs and steppes seem to yield regression qualities that are in close range. The prairie and the Pampas grasslands showed similar trends, implying that the grassland type has little effect on the accuracies of estimates derived from remote sensing data. This is an intriguing observation given that the ecosystems are geographically located in different climate zones (temperate vs. subtropical). On the other hand, accuracies obtained from the Pampas grasslands seem less variable, suggesting some level of consistency across these studies. This could be attributed to the fact that the Pampas grasslands are less heterogeneous and consist of little to no trees (Ricard *et al.*, 2020) compared to savannahs, with a highly variable vegetation structure and species. The co-existence of grasses, trees, and shrubs in savannahs, specifically those on the African continent, present a different set of challenges for remote sensing approaches to retrieve AGGB, hence the inconsistent results detected in the reviewed literature. For example, savannahs in Sudan consist of scattered trees, which form a transition with the Sahara Desert vegetation whereas the savannahs found in Guinea consist of moist semi-deciduous woodlands which form a transition to the evergreen moist forests (Mishra and Young, 2020). Furthermore, fire and herbivory activities pose a further challenge for the remote sensing-based AGGB estimation.

2.4.1.2 Geographical and Temporal Gradients Covered

A country-based analysis showed China as the highest contributor in terms of the output of AGGB estimation research. The dominance of China could be because it occupies a large portion of one of the most extensive Eurasian steppes in the world. Contrarily, developing countries contributed less, specifically those on the African continent. Compared to most developing countries, China strengthened its satellite surveillance capabilities for environmental monitoring with the launch of the HJ 1A/B optical imaging constellations in 2008 (Xue *et al.*, 2008) as well as the HJ 1C radar satellite. Further improvements followed with the launch of the GaoFen-1 and 6 as well as Huanjing-2/A&B (HJ-2/A&B) (Zhong *et al.*, 2021). Some countries in the global south also followed; for example, Egypt launched the Nilesat 101 in 1998 and South Africa launched the SumbadilaSat in 1999 which was only in orbit for 2 years (Oyewole, 2017). Many developing nations find it difficult to maintain their own satellites due to the exorbitant expense of space rocket manufacturing and maintenance (Oyewole, 2017). As a result, there is a lack of sufficient and sustainable access to satellite imagery for this region. Consequently, African countries still outsource remote sensing data from continents in the global north such as North America and Europe. Nevertheless, the dawn of open access policies to data and the Google Earth Engine (GEE) should stimulate the generation of new knowledge in these regions in the near future. A notable limitation of this review is the exclusion of non-English studies, which may have introduced language bias. Consequently, the findings may not fully reflect the global viewpoints, implying that insights from such non-English studies may have been missed. Nevertheless, English remains the predominant language in the domains of remote sensing and ecological monitoring and a considerable proportion of high-impact; peer-reviewed studies in these fields are published in English.

The findings also show a significant uneven trend in the annual production of AGGB retrieval studies across the years during the review period, particularly between the eras 1986–2005 and 2014–2020. The low research yields in the period 1986–2005 could be because fewer sensors (e.g., Landsat and MODIS) were operational. Notably, the reduced publication yield between 2010 and 2013 could be attributed to the limitations in data continuity from satellite missions such as Landsat 7 which experienced a Scan Line Corrector followed by the lack of fully operational successor missions until the launch of Landsat 8 in 2013 and Sentinel-1 in 2014. The steady increase in studies on AGGB after 2014 could be because of the gradual increase in sensor choice. For example, the Sentinel-2 constellation was launched in 2015 and made available at no cost. Nevertheless, an output of fewer than 20 articles per annum is inadequate to conclude on the optimal data, sensors, variables and algorithm for retrieving the AGGB in savannahs. Due to more available sensors (e.g., Landsat-9 and Sentinel-3), the rise is predicted to accelerate exponentially in the next few years. The improved capacity of computer hardware and software such as Google Earth Engine (GEE) that are designed to provide high-performing cloud-based computing, pre-processing algorithms as well as cloud-based storage should further increase the research outputs (Kumar and Mutanga, 2017).

2.4.1.3 Platform and Sensor Configurations

The variability in remote sensing data types and sensors and their associated spatial resolution contributes to some level of uncertainty in the resulting AGGB estimates (Sánchez-Azofeifa *et al.*, 2009). Optical sensors remain the most widely used, particularly from the MODIS and Landsat missions. MODIS and Landsat boast of long-term operations, and the free access policy associated with these constellations (Zhu *et al.*, 2019b) further paved research methods to quantify the AGGB. The pool of studies utilising these sensors in comparison to others provides proof of this. The range of quality checked ready-to-use products (Wang *et al.*, 2005) such as the MOD13Q1 from the MODIS constellation provide further benefits. However, optical sensors such as MODIS have limitations (Zhao *et al.*, 2014; Liang *et al.*, 2016; Zeng *et al.*, 2019), including their failure to collect data on days that are overcast, smoky, or misty (Prudente *et al.*, 2020). The inability of optical sensors to capture data during these adverse weather conditions impedes consistent monitoring, resulting in the omission of useful information for AGGB estimation. The selected studies show that optical sensors have a wide range of accuracies, implying inconsistent predictions.

In contrast, active sensors such as radar and lidar offer alternative datasets for estimating AGGB. These sensors boast about their ability to penetrate clouds and therefore provide data across all weather conditions. Despite this, radar was used in only 13 of the 108 studies reviewed. The low radar usage could be due to the cost tied to the data, therefore, limiting its widespread usage. Furthermore, radar-derived imagery is difficult to interpret for researchers with no prior knowledge and the processing is complex (Tamiminia *et al.*, 2020). Nevertheless, studies reported improved accuracies in AGGB estimates from radar-based models, which were less variable, suggesting some level of consistency and dependable predictions. However, when considering the low-resolution sensors, it is evident that their accuracies might be comparatively lower. However, what is interesting is the remarkably narrower spread of values, excluding a notable outlier. This could indicate a higher level of consistency in accuracy across the low-resolution sensors. Contrarily, the high-resolution sensor shows a diverse range of accuracy values, which is further emphasised by the presence of an outlier. Thus, the concept that a higher resolution guarantees superior results is not a straightforward one. Nevertheless, the higher accuracies obtained for lidar and radar could have been influenced by the smaller number publications that were analysed in this review. The European Space Agency (ESA) currently distributes SAR data (i.e., Sentinel-1) at no cost, to help deal with the access issue. However, few of the reviewed studies tested Sentinel-1, signalling the need for further research. For grass biomass studies, lidar metrics have also been used (Maesano *et al.*, 2020). Unlike the popular optical data which views the two-dimensional (2D) attributes of grasses, lidar can detect three-dimensional (3D) attributes such as canopy height that can be used as proxies when quantifying the AGGB (Sarrazin *et al.*, 2012; Wang *et al.*, 2017). Despite this strength, few of the reviewed studies used lidar. This could be attributed to the fact that they cover small geographical footprints (Xu *et al.*, 2018), which could be expensive to extend in the extensive tropical savannahs.

The review also identified the application of UAVs to a lesser extent in the surveyed studies. For example, only eight studies tested the utility of UAVs for deriving the AGGB. While higher accuracies were observed at the higher spatial resolution, the use of UAVs is limited to field scale (Wang *et al.*, 2022) and their application in larger areas is restricted by their reliance on batteries. Furthermore, UAVs require several continuous flights for complete coverage in larger areas, which results in a large mosaic of tiles that need many hours to process (Calleja *et al.*, 2018). While the use of a single sensor has individual shortfalls, combining sensors provides complimentary strength and fills in missing observations while ensuring data continuity (Maesano *et al.*, 2020). For example, combining Sentinel-1, Landsat-8, and Sentinel-2 showed promising results, with Chen *et al.* (2019a) concluding that the estimation of the AGGB improved by more than 30% relative to the performance of Sentinel-1 alone at low vegetation cover and the optical data of Landsat-8 and Sentinel-2 at high vegetation cover. When combining spectroscopy and small-footprint waveform light detection and ranging (Lidar) data, Sarrazin *et al.* (2012) concluded that fine-scale Lidar may not be able to provide accurate measurement of the herbaceous biomass and highlighted the need for further investigation during the peak growing season. However, in their review, Calleja *et al.* (2018) found that the fusion of optical and UAV data was tested mostly in the ecology and precision agriculture. Few of the reviewed studies pursued fused optical and radar, lidar, UAV or vice versa, and as such, this remains an open field of research.

Few studies used several spatial resolutions associated with sensors to develop AGGB estimation models. Sensors with spatial resolution of 30 m and above were frequently used. This could be due to data accessibility and costs at specific spatial resolution intervals. Sensors with spatial resolution at 15 m and above (AVHRR, SPOT VEGT, MODIS, and Landsat) have been freely accessible for decades compared to their counterparts (Worldview-1/2, TerraSAR, lidar, or UAV-based sensors). Furthermore, earlier missions that piloted the use of remote sensing for monitoring vegetation attributes started with sensors at a spatial resolution of 15 m and above. For example, the Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER) Mission at 15–30 m spatial resolution, the vegetation programme using SPOT VEGETATION at 1100 m, and the MODIS Land Group at 250, 500, and 1000 m spatial resolutions (Sánchez-Azofeifa *et al.*, 2009; Wang *et al.*, 2019a).

Sensors with spatial resolution ranging from 250 to >1000 m are suitable for studies at the national, continental, and global scale. Furthermore, sensors at 250 to >1000 m spatial resolution provide the balance between spatial resolution, surface area coverage and high frequency in data acquisition (Lu, 2006). However, land surfaces are naturally heterogeneous and at this scale, the probability of detecting and capturing minute objects is quite low (Meng *et al.*, 2020). In addition, collecting ground-truthing samples at a 1000 m scale is not time- and cost effective. This presents a large discrepancy between the size of the ground data and the spatial resolution of the image. Spectral mixture techniques were explored as an alternative to overcome the challenge of each pixel being a physical mixture of the multiple components weighted by the dominant area and the mixture spectrum is a linear combination of the end member reflectance spectra (Tompkins *et al.*, 1997). On the bright side, the newly launched Sentinel-3 satellite presents a new dawn of opportunities in estimating and monitoring the AGGB at a larger scale. Compared to its predecessor (MERIS), Sentinel-3, specifically the Operational Land Imager (OLI), consists of six added spectral bands, a large swath width and offers improved signal-to-noise ratio (SNR) for monitoring vegetation condition (Donlon *et al.*, 2012). Furthermore, while other constellations are freely available for academic and research purposes, but limited for commercial users, Sentinel-3 is anticipated to be freely available for both academic research and commercial uses.

Fewer studies explored sensors with spatial resolution ranging between 0 and 5 m and below. Although the spectral differences of grasslands can be captured accurately at a resolution ranging between 0 and 5 m, datasets at this scale are usually not readily accessible. The need for computers with large hard drives for storage further exacerbates the application limitation already associated with such datasets. Furthermore, 0–5 m pixel size is only applicable in localised areas of approximately 100 km^2 and the processing time necessary to cover wider geographical areas is large (Wang *et al.*, 2014a).

2.4.1.4 Predictor Variables

The surveyed literature used several predictor variables as input into the AGGB estimation models. Broadband vegetation indices, particular NDVI, were the most widely used amongst the many variables. The frequent usage of NDVI could be because the index was pioneered specifically to monitor the vegetation conditions, and ultimately the AGGB. Nevertheless, NDVI is sensitive to the soil brightness and cloud shadow (Bannari *et al.*, 1995), rendering it unfit, especially in areas with low grass cover. To reduce the effects of the soil brightness, researchers pioneered the SR, EVI, and SAVI, although only a handful of studies have applied them for estimating the AGGB in the surveyed studies. Furthermore, NDVI is prone to saturation issues (Mutanga and Skidmore, 2004). For example, in areas with high grass cover, NDVI tends to lose sensitivity and therefore produces inaccurate AGGB estimates. Narrowband widths (e.g., red edge position) from hyperspectral data are not prone to saturation issues (Mutanga and Skidmore, 2004). Nevertheless, hyperspectral-derived predictor variables have "The curse of dimensionality" (Timothy *et al.*, 2016), a condition that occurs when analysing high dimensional data that does not occur in low-dimensional spaces.

Other predictor variables including sward height (Wijesingha *et al.*, 2019), FAPAR (Schmidt *et al.*, 2016; Chave *et al.*, 2019), and backscatter coefficients (Butterfield and Malmström, 2009), along with a combination of multiple variables (spectral bands + VI) (Shoko *et al.*, 2018c) were developed. For example, FAPAR has been proven to improve prediction accuracies (Zhao *et al.*, 2014). However, few of the studies reviewed applied these attributes. From a performance perspective, NDVI is prone to obtaining low accuracies. Studies also show that the qualities of NDVI-based regressions as measured by the R^2 vary widely, suggesting inconsistent and unreliable results. These findings are consistent with those of Zhao *et al.* (2014)

and could stem from the fact that most of the studies computed the NDVI from the broadband channels of optical sensors. According to Zhao *et al.* (2007) "Broadband use the mean values of spectral information over broadband widths resulting in the loss of critical information that is available in specific narrow bands". On the other hand, vegetation indices computed from narrow bands such as red edge-based NDVI show better performance. In particular, narrowband vegetation indices can maximise the contrast between grasslands and other land cover categories, therefore providing improved estimations.

The use of radar-derived metrics such as backscatter in savannah grasslands performed better in some studies (Sarrazin *et al.*, 2011; Wang *et al.*, 2014b). For example, (Wang *et al.*, 2014b) demonstrated the capabilities of ENVISAT-ASAR and ALOS POLSAR to measure pasture biomass in Western Australia and obtained good results (i.e., R^2 = 0.92). The low usage of radar-based metrics could be that in addition to the extreme cost, they require some sophisticated software and expert knowledge to interpret both the input and output results. Nevertheless, the launch of Sentinel-1 should extend the predictor variables outside the broadband optical channel to the microwave channels. Furthermore, while the highly priced hyperspectral-based narrow bandwidths provided improved AGGB estimates, Sentinel-2 MSI is also affordable and has robust spectral settings covering the red edge section of the electromagnetic spectrum. For example, Shoko *et al.* (2016) demonstrated the abilities of the red edge-based NDVI derived from Sentinel-2 imagery.

2.4.1.5 Algorithms Developed

An analysis of the studies selected for the review showed that linear regression (LR) techniques were widely applied. This could be because linear equations are inherently simple and so are easy to operate and interpret (Marill, 2004). The computation of linear equations is less time-efficient when compared to other retrieval algorithms. However, LR techniques are parametric and have normality assumptions, a condition that is not always representative of grass biomass on the ground.

Machine learning (ML) algorithms such as Random Forest have been developed and favoured for their non-reliance on assumptions. Furthermore, ML can handle a large number of datasets from heterogeneous landscapes, such as those in savannah ecosystems, which consist of trees, shrubs, and a layer of grasses, making them complex landscapes. However, in this review, fewer studies (8%) tested ML algorithms. Specifically, studies coupling ML algorithms with radar data to estimate the AGGB in savannah landscapes. Compared to conventional algorithms, ML algorithms are non-parametric and therefore independent of assumptions (Barrett *et al.*, 2014). Nevertheless, the low usage of these algorithms could be because they are complex and not easily interpretable (Wadoux *et al.*, 2020). In fact, the majority of the ML algorithms have been categorised as a "black box", meaning that they do not show the user the underlying process behind the prediction or provide features of importance (Wadoux *et al.*, 2020). To a smaller extent, Radiative Transfer models, specifically PROSAIL, have also been applied in the reviewed studies. PROSAIL is a physical-based model that depends on light absorption and scattering to compute vegetation properties. Compared to other statistical-based algorithms and ML models, radiation efficiency models are more robust and transferable as they use physical laws to express the transfer of light within grassland canopies. However, physical-based models lag since different combinations of plant traits can produce the same reflectance spectra. In parallel, physical models tend to perform better in retrieving some vegetation variables (e.g., chlorophyll content and leaf area index), but perform poorly with others (e.g., biomass). Furthermore, if one variable in the model is inaccurate, the whole model is affected (Masenyama *et al.*, 2022).

Today, although in its infancy, deep learning (DL) is emerging as the new generation of algorithms. DL algorithms have successfully delineated agricultural fields (Sharifi *et al.*, 2022), classified satellite imagery (Pritt and Chern, 2017) and mapped soil organic carbon (Odebiri *et al.*, 2021). In parallel, DL algorithms can handle "high dimensional data" and nonlinear relationships compared to their counterparts (Wang *et al.*, 2022). DL algorithms require intensive computation and programming skills to run them, hence their low application. However, there is a need for more experiments and observations, especially for regional and global biomass estimation studies, given the advent of "big data".

From a performance perspective, the ML models yielded median R^2 above 60%, showing considerable prediction accuracies across all the algorithms used in estimating the AGGB. Comparatively, linear regressions were observed to have a wide distribution of accuracies (40% to >80%) in the reviewed studies compared to ML algorithms, which seem to have a narrower distribution of R^2 values. These results support those of Verrelst et al. (Verrelst *et al.*, 2019) who found that the choice of a model for estimating biophysical parameters of grasses has an underlying impact on the accuracy of the results

2.4.1.6 Sampling and Analysis Protocols

In situ data collection remains the most precise method of gathering representative biomass samples (Lu, 2006). Remotely sensed AGGB estimates require a set of ground-truthing data to train and validate the models to be credible and reliable (Eisfelder *et al.*, 2012). Scientists collect in situ grass biomass samples to ensure consistency between what they observe on the satellite image and the ground. Studies frequently harvest (cut and weigh) grass samples for testing and validating remote sensing-based AGGB estimates using $1m^2$ quadrats. They then dry the collected samples mostly for 48 h at 65 degrees Celsius. However, the 1 m2 plot size does not match the spatial resolution of sensors mostly employed to extrapolate these plot-based measurements, producing some uncertainty in the resulting estimates (Eisfelder *et al.*, 2012). Furthermore, (Lu, 2006) pointed that the time-to-time administration and execution of fieldwork activities is costly and tedious. He also observed that fieldwork costs are inversely proportional to the quantity and quality of the validation samples. In order to overcome the mismatch between the pixel size and in situ plot sizes and to compensate for the spatial resolution of the chosen sensors, studies often demarcate sampling plots equal to the spatial resolution of the sensor in less heterogeneous grasslands (Shoko *et al.*, 2018a). Nevertheless, UAVs are emerging as alternatives for validating remote sensing-based AGGB estimates. Recently, portable spectrometers such as Field Spec 3 Max have also served as alternatives for ground-truthing of satellite images.

2.4.2 Research Challenges and Outlook

2.4.2.1 Sampling Research Challenges

The use of remote sensing to estimate the AGGB has shown considerable progress. However, from the reviewed studies, some degree of inconsistency in accuracy exists across the grassland types investigated, the predictor variables used, sensors, algorithms, as well as the sampling protocols adopted. Furthermore, researchers have encountered a myriad of obstacles in the pursuit of these estimates. From a grassland-type perspective, especially those consisting of sparse vegetation, the major concern was the influence of soil background (Ding *et al.*, 2019), shadow from both clouds and topography, species diversity, and vegetation litter, which obscured the spectral response of the grasses (Svoray *et al.*, 2001). In addition, grasslands in semi-arid areas such as the savannah consist of complex phenological stages, canopy structures, and species, compelling the need for sensors with a high temporal resolution. However, most readily available data covering the savannah have very low temporal resolutions. For example, MODIS offers coarse resolution (1000 m) at high temporal resolution (daily) for consistent monitoring purposes and is accessible freely, yet its application to localised studies is limited by its low spatial resolution. In areas with abundant rainfall, vegetation density increases at a faster rate, requiring more frequent observations (Li *et al.*, 2014).

As mentioned above, optical data remains the most widely used data type. However, clouds, topographic shadow, fire smoke, and haze, especially in the tropics and subtropics, with frequent cloud coverage, obscure the spectral response of the targeted vegetation using optical data. This constantly results in observation cut-offs, and consequently the loss of essential information. Radar sensors serve as alternative sources of datasets. Radar data has an important role in AGGB estimation, especially in areas such as the savannah ecosystem with its frequent cloudy conditions. While access to radar data is on the rise, some degree of limitation was reported in the literature, including its sensitivity to vegetation water content, soil moisture, winds, and humidity (Hamdan *et al.*, 2015). Furthermore, radar data analyses involved in pre-processing, removal of noise, and image processing require specialised software, knowledge, and skills to operate the software and ultimately interpret the results. Although radar datasets are available at no cost lately, there is still a relative lack of studies from which to draw concrete conclusions regarding their potential. The same observation can be made about lidar. However, lidar is expensive for applications in the savannah ecosystem.

Another pressing concern in the use of remote sensing to derive AGGB estimates is the issue of resolution (spatial, temporal, or spectral) (Li *et al.*, 2014). Generally, landscapes across the globe are a collage consisting of a diverse range of vegetation classes configured differently (i.e., physiology and morphology), and savannah ecosystems are no exception. Sensors with a high spatial resolution were reported to address the issues around heterogeneity, which usually confuses pixels (Lu, 2006). The major shortcoming associated with higher spatial resolution sensors is their price tag, but observations at lower spatial resolution risk the possibility of having more than one vegetation class in a single pixel (Lu, 2006). The temporal resolution also presents further challenges. For example, monitoring vegetation variables frequently requires sensors with higher temporal resolutions. However, there is a lack of balance between the sensor's spatial and temporal resolution. For example, an increase in the sensor's pixel size implies a decrease in the temporal resolution (Hamdan *et al.*, 2015). In other words, gaining very precise estimates comes at the cost of omitting the time information associated with the observed vegetation variable (Hamdan *et al.*, 2015). Although high-resolution sensors have been proven to provide accurate estimates of the AGGB, they are, however, limited in scope because they do not capture images on a global scale, but rather at small footprints.

Several researchers reported saturation as an issue when retrieving the AGGB using remote sensing data (Mutanga and Skidmore, 2004). For optical data, predictor variables such as the widely used NDVI tend to saturate at dense vegetation. A possible reason for this could be that when the density of vegetation reaches full coverage, the amount of red light that can be absorbed by leaves reaches a peak and, therefore, the sensitivity of reflectance starts diminishing with a rise in biomass (Mutanga and Skidmore, 2004). The red edge position seems capable of overcoming the saturation encountered when using broadband-derived spectra and indices (Mutanga and Skidmore, 2004). For radar data, saturation was obtained above 200 Mg ha^{-1} (Hamdan *et al.*, 2015).

The most notable algorithm used in the consulted studies was statistical regression, specifically the linear function. However, statistical regression algorithms fail to detect the complex relationships that may be associated with the predicted variable without over-fitting the models. Furthermore, these models seem to be limited to the specific sites in which they were developed, hence lacking both temporal and spatial transferability (Eisfelder *et al.*, 2012). Researchers have attempted to transfer the models across sites with no luck in terms of accuracy, and encountering inconsistencies when transferring the models across temporal scales. Other studies highlighted the mismatch between the estimation and in situ measurement scale as a major concern (Pritt and Chern, 2017). For example, 1 m^2 size subplots are normally used to collect samples for tuning the AGGB models and then scaled up to the selected sensor (e.g., 30 m) spatial resolution resulting in uncertain estimates (Bao *et al.*, 2019).

2.4.2.2 Research Outlook

While considerable milestones have been realised in terms of the spatial distribution, sensors, development of predictor variables, and the engineering of extracting models, considerable research opportunities also exist. We provide the following summary:

- Developed countries contributed more research on remote sensing-based AGGB compared to developing countries. As such, more research should be performed in the global south in order to promote an all-inclusive regional reporting.
- Few studies applied remote sensors operating outside the optical channel of the electromagnetic spectrum (i.e., microwave) for retrieving the AGGB. Specifically, radar-derived metrics could add tangible value to the performance of biomass estimation models. The freely available Sentinel-1 offers an opportunity to quantify the AGGB in savannah ecosystems.
- The integration of radar images shows promising results, and a further exploration of the complementary aspect of these sensors should improve the baseline models.
- Although costly, lidar datasets seem promising in terms of accuracy and further studies should explore their full potential in AGGB estimation.
- Despite their limitation, the vegetation indices remain the major predictor variables. Thus, improved accuracies of estimating the AGGB may be realised with the incorporation of supplementary variables such as sward height, FAPAR, agro-meteorological, and topographical variables.
- From a radar perspective, the soil moisture and soil roughness should be taken into consideration during modelling since they contaminate the backscattering processes.
- Many researchers have relied on the less transferable linear regression while machine learning approaches have not been fully explored. However, deep learning algorithms are emerging as the new dawn of algorithms and their utility in AGGB estimation is in its infancy, thereby leaving a gap for further research.
- The mismatch between the estimating and validation scale reduces the accuracy of

estimating the AGGB. The lack of consistency between the *in-situ* measurement and sampling protocol further hinders the comparability across the studies. This signals the need to benchmark the sampling process.

- It is also of interest to explore the use of the newly launched Sentinel-3 and Landsat-9 OLI in quantifying the AGGB in savannah ecosystems.
- Future research on AGGB estimation should focus on the application of multi-source data and multi-temporal data available via cloud-based applications, including GEE, Microsoft Azure and Amazon Web Services (AWS).

2.5 Conclusions

This study synthesised findings of AGGB studies published in the literature to map the current state of the art and provide guidelines for model performance derived from decades of research. Authors have explored remote sensing-based models in steppes, while the African savannah ecosystems have not been fully explored. Savannahs consist of a grass layer that co-exists with scattered trees and shrubs, which may affect the results differently compared to the less heterogeneous steppe landscape. In optical remote sensing, particularly, MODIS and Landsat remain the main source of remotely sensed data for AGGB estimation even though they are affected by soil background and clouds and are prone to saturation issues. Studies using SAR data, especially Sentinel-1, are scarce on a global scale, but the scarcity is more pronounced on the African continent despite being freely available. Vegetation indices, particularly NDVI, are a major AGGB predictor variable. However, the red edge position was pioneered to enhance the estimation of the AGGB and to reduce the saturation effects associated with broad bands. The existence of advanced machine learning algorithms dates back a while, yet researchers still rely on linear regression techniques. The newly developed deep learning algorithms should be explored further. Our results strongly suggest that the remote sensing-based AGGB findings are inconsistent across the consulted studies, making it difficult to draw conclusions on the optimal and universal model that can be used for estimating the AGGB in grasslands. There is, therefore, a need for standardising the remote sensing techniques, including the *in-situ* data collection methods, to ensure transferability of remote sensing-based AGGB models across multiple locations. Overall, this study has comprehensively mapped the potential of remote sensing in retrieving AGGB estimates on a global scale.

2.5 Summary

Due to the shift from cost-tied to open-access data, which led to a surge in research aimed at developing algorithms for estimating aboveground grass biomass (AGGB). This chapter was set to provide a comprehensive synthesis of findings of AGGB studies published in the literature to map the milestones achieved and provide an overview models performance across the grassland types, predictor variables, sensors, algorithms, as well as the sampling protocols adopted. The review uncovered that remote sensing-based models were mostly done in steppes, therefore, suggesting that such studies need to be extended in savannah landscapes with complex vegetation structure, particularly in Africa. These studies based their AGGB models on optical sensors such as MODIS and Landsat. While those that explored SAR, relied on commercial sensors, implying the need to search for alternative and cheap sensors, particularly for the resource-constrained countries with persistent cloud cover. Furthermore, studies still rely on linear regression techniques, indicating the need to explore machine-learning algorithms. Future studies should explore the use of freely available SAR sensors (e.g. Sentinel-1) coupled with scalable machine learning algorithms such as XGBoost in quantifying AGGCS in savannas. Following these findings, the next chapter will explore optimal parameters derived from the freely available Sentinel-1 for estimating AGGCS in savanna ecosystems. This will assist in selecting the most suitable SAR variables to accurately map AGGB.

CHAPTER 3: DETERMINING OPTIMAL SAR PARAMETERS FOR QUANTIFYING ABOVE-GROUND GRASS CARBON STOCK IN SAVANNAH ECOSYSTEMS USING A TREE-BASED ALGORITHM

This chapter is based on:

Remote Sensing in Earth Systems Sciences (2025) 8:251–263
https://doi.org/10.1007/s41976-024-00170-8

RESEARCH

Determining Optimal SAR Parameters for Quantifying Above-Ground Grass Carbon Stock in Savannah Ecosystems Using a Tree-Based Algorithm

Reneilwe Maake[1,2] · Onisimo Mutanga[2] · Johannes George Chirima[1,4] · Mahlatse Kganyago[3]

Received: 5 June 2024 / Revised: 15 November 2024 / Accepted: 28 November 2024 / Published online: 12 December 2024

Abstract

The quantification and monitoring of aboveground grass carbon stock (AGGCS) will inform emission reduction policies and aid in minimising the risks associated with future climate change. This study investigated the sensitivity of Synthetic Aperture Radar (SAR)-derived parameters to predict AGGCS in a savannah ecosystem in Kruger National Park (KNP), South Africa. Particularly, we investigated the capabilities of Sentinel-1 derived parameters, including backscatter coefficients, intensity ratios, normalised radar backscatter, arithmetic computations, and the XGBoost tree-based algorithm, to predict the AGGCS. We further tested if incorporating texture matrices (i.e. Grey Level Co-Occurrence Matrix) can enhance the predictive capability of the models. We found that the linear polarisation (i.e. Vertical Transmit/Vertical Receive (VV) and the intensity ratio (i.e. VH/VV) achieved similar results with Coefficient of Determination ($R^2 = 0.38$), Percent Root Mean Square Error (RMSE% = 31%), Mean Absolute Error (MAE = 6.87) and ($R^2 = 0.37$, RMSE = 31%, MAE = 8.80) respectively. The Radar Vegetation Index (RVI) performed marginally (1%) better ($R^2 = 0.39$, RMSE = 30% and MAE = 6.77) compared to the other variables. Nevertheless, the incorporation texture matrix into the model enhanced prediction capability by approximately 20% (R^2) = 0.60, (RMSE%) = 20%, (MAE) = 3.91). Furthermore, the most influential predictors for AGGCS estimation were Radar Vegetation Index (RVI), Vertical Transmit/Horizontal Receive (VH_{cor}) and Vertical Transmit/Vertical Receive (VV_{cor}) in order of importance. These findings (R^2 values of 0.35–0.39) suggest that SAR data alone does not fully capture the variability in AGGCS, particularly in the complexly configured savannah ecosystems. Nevertheless, the results further suggest that the prediction accuracy of SAR-based AGGC models can be enhanced with the incorporation of texture matrices.

Keywords: *Aboveground grass carbon stock, Savannah ecosystems, Sentinel-1, XGBoost*

3.1 Introduction

The savannah ecosystem is one of the many vital biomes that exist on Earth. It is characterised by a rich biodiversity dominated by grasslands co-existing with patches of woodlands and shrublands (Scholes and Archer, 1997; Higgins *et al.,* 2007). The grass layer serves as a feeding resource for grazing animals, while the woody layer serves as a feeding stock for browsing animals. From a climate perspective, the savannah grass and woody layers account for an estimated 25% of total gross primary production (GPP), rendering this ecosystem a vital sink of carbon (Moore *et al.*, 2016). They help regulate the Earth's climate and maintain a balance of greenhouse gases in the atmosphere. The woody vegetation layer is known for storing large quantities of carbon; however, the sustainability of this capability is threatened by extreme weather conditions such as drought, heat waves, and increased fire activity (Dass *et al.*, 2018). Studies have uncovered a proliferation of mortality in woody plants, which is associated with extreme weather events (Allen *et al.*, 2015), fire (Stevens-Rumann *et al.*, 2018; McDowell *et al.*, 2022), and wildlife interactions (Shannon *et al.*, 2011; Das *et al.*, 2022). In the face of climate change, the grass layer, due to its biological configuration, is emerging as a sustainable terrestrial carbon sink (Dass *et al.*, 2018). The grass layer is resilient to various disturbances, including grazing, fire activity, animal stampedes, and extreme weather conditions (Craine *et al.*, 2013). This implies that the grass layer can sustainably maintain a relatively stable carbon stock over time, an attribute that is vital as we approach the era of above-normal climate change. As such, the quantification and monitoring of aboveground grass carbon stock (AGGCS) will inform emission reduction policies and aid in minimising the risks associated with future climate change. This will align with the urgency emphasised in Sustainable Development Goal 13, which calls for immediate action to address climate change and lessen its repercussions. A further benefit of quantifying AGGCS is the generation of baseline information for scientific research and environmental management. The Paris Agreement is one environmental management initiative mandating countries to disclose their greenhouse gas emissions along with the reduction efforts (Walsh *et al.*, 2017). In addition, statistics on grass carbon stock are essential for meeting reporting demands and ensuring transparency and accountability in climate actions taken by multiple nations.

A myriad of grass carbon stock quantification missions rely on space-borne sensors due to their ability to capture data over large geographical and temporal footprints (Ding *et al.*, 2019; Bindu *et al.*, 2020; Chapungu *et al.*, 2020; Clementini *et al.*, 2020; Guo *et al.*, 2021; Abbass *et al.*, 2022). Primarily, optical sensors have been in use for decades, regardless of their limitations. For example, optical sensors rely on the visible and infrared bands, which are obstructed by clouds and smoke, and their interaction is limited to the canopy structure of the vegetation (Abdi *et al.*, 2022). Consequently, optical sensors are unable to ensure a consistent distribution of data throughout multiple seasons.

While SAR data has been in existence for decades, its application has been hindered by its price tag (Abdi *et al.*, 2022). Sentinel-1 offers new opportunities for quantifying and monitoring above-ground carbon stock beyond the visible range of the electromagnetic spectrum, at no cost, and across smoky, hazy and overcast environmental conditions. Sentinel-1, constructed by the European Space Agency (ESA), is a SAR satellite mission (Potin *et al.*, 2018). The mission comprises a duo of identical sensors, namely Sentinel-1A and Sentinel-1B. These sensors can capture data under several weather conditions, an attribute that ensures a consistent flow of data, allowing for continuous quantification of carbon stock. Furthermore, Sentinel-1 can penetrate beyond the canopy structure of the vegetation, making it particularly valuable for capturing intra-canopy properties of vegetation (Ghasemi *et al.*, 2011). SAR data affords several metrics which can be explored for quantifying above-ground carbon stock, some of which have been documented to be optimal for forests (Ghosh and Behera, 2021; Vatandaşlar and Abdikan, 2022; Zeng *et al.*, 2022) and agricultural fields (Hajj *et al.*, 2014; Wang *et al.*, 2014c; Mandal *et al.*, 2020; Nguyen *et al.*, 2022). Studies showed that incorporating texture matrix with SAR data could improve biomass estimation (Nichol and Sarker, 2010; Naidoo *et al.*, 2016; Liao *et al.*, 2020). However, fewer studies coupled SAR-derived parameters from sentinel-1 with texture matrix for estimating aboveground grass carbon stocks, particularly in savannah ecosystems. The aim of this research is to employ a tree-based learning algorithm to identify the most effective SAR parameters, including texture, for accurately estimating AGGCS in savannah ecosystems.

3.2 Materials and Methods

3.2.1 Test Site Description

An outline and description of the study site is presented in subsection 1.6

3.2.2 Field Above-Ground Grass Carbon Stock Sampling

A grass biomass sampling campaign was conducted from 7–14 February 2020. A thorough outline of how the sites and plots were demarcated as well as how the samples were collected and processed is shown on (Fig. 1.3) as presented in subsection 1.7. A total of 94 plots were sampled to serve as the foundation to answer this question. A demonstration of the approach adopted for field data collection is also shown in Figure 3.1.

Figure 3. 1 A schematic depiction of the approach adopted for field data collection

3.2.3 Data Acquisition

3.3.3.1 Sentinel-1 data acquisition & Pre-Processing

For this objective, we acquired Level 1 Ground Range Detected (GRD) data from Sentinel-1 (S-1) in C-band, which was captured in the Interferometric Wide (IW) Mode on 10th February, 2020. The image was extracted from the Copernicus Open Access Hub, accessible at https://scihub.copernicus. eu/dhus/. An outline of how the satellite image in question was pre-processed is presented in subsection 1.8.

3.2.3.2 Deriving Sentinel-1 predictor variables

To test the sensitivity of SAR in estimating AGGCS, we extracted and calculated several parameters from the Sentinel-1 image, including 2 raw backscatter coefficients, 2 polarisation intensity ratios, 1 radar indices (radar vegetation index), and 2 intensity arithmetic computations, 20 texture matrices as well as 2 geographic coordinates (Latitude and Longitude), as shown in Table 3.1. This amounted to 29 variables.

Table 3. 1 Extracted predictor variables employed in estimating above-ground grass carbon stock

Predictor variable	Expression/Description
Polarisation backscatter coefficients	VH (Vertical Transmit/horizontal receive)
	VV (Vertical Transmit/vertical receive)
Polarisation Intensity ratios	VH/VV
	VV-VH/VV+VH
Radar Indices (RI)	RVI (Radar Vegetation Index)
Arithmetic Intensity Computations	VH-VV
	VV+VH
Texture Matrix	VH/VV_{Con} (Contrast)
	VH/VV_{Dis} (Dissimilarity)
	VH/VV_{Hom} (Homogeneity)
	VH/VV_{ASM} (Agular Second Moment)
	VH/VV_{Ene} (Energy)
	VH/VV_{Max} (Maximum Probability)
	VH/VV_{Ent} (Entropy)
	VH/VV_{Mea} (GLCM Mean)
	VH/VV_{Var} (GLCM Variance)
	VH/VV_{Cor} (GLCM Correlation)

3.3 Statistical Analysis

3.3.1 Regression algorithms

To predict AGGCS using Sentinel-1 derived variables, we employed Extreme Gradient Boosting (XGBoost) to regress field-measured grass carbon stock against Sentinel-1 derived variables. This is an enhanced version of techniques such as Gradient Boosting Machines (GBM) and AdaBoosts, configured with additional features. It is popular for its scalability, efficiency and ability to deliver high performance on a wide range of regression and classification tasks. XGBoost adheres to the structure of gradient boosting where the model is built stage by stage, and each stage focuses on correcting the errors of the previous stage (Chen and Guestrin, 2016). Misclassified points receive increased weight during each iteration, directing subsequent steps to focus on rectifying these misclassifications(Chen and Guestrin, 2016). It typically uses decision trees as base learners. These shallow trees help prevent over fitting and keep the model computationally efficient (Chen and Guestrin, 2016; Beltran *et al.*, 2019). Furthermore, each subsequent tree focuses on rectifying the errors introduced by its predecessors. This is achieved by training each new tree on the residuals, representing the discrepancies between the predicted and actual values of the preceding predictions (Chen and Guestrin, 2016). It uses a weighted sum of the predictions from individual trees, where the weights are determined during training based on the performance of each tree. The optimal parameter identification process is based on a grid search method where the hyper-parameters are selected based on the lowest RMSE from a 10-fold Cross Validation (CV) resampling strategy. More information about how XGBoost works can be found in (Chen and Guestrin, 2016). The decision to choose this algorithm was informed by its performance in previous studies (Kganyago *et al.*, 2022) as well as its ability to handle sparse data which is the case with our samples.

We first regressed field-measured grass carbon stocks against individual Sentinel-1 derived variables to establish their individual sensitivity to predicting AGGCS. We then incorporated texture matrices to assess their combined effect on the prediction accuracy. Incorporating the texture matrices increased the dimension of the feature space, which tends to suffer from multicollinearity. Shrestha (2020) defines multicollinearity as a condition wherein multiple independent variables in a regression analysis are highly correlated with each other. To deal with multicollinearity we performed a pairwise Pearson correlation matrix to assess the relationship between SAR parameters and texture matrices. As in Chen *et al.* (2019a) we selected a threshold of $R \geq 0.8$) as an elimination criterion to select the optimal variables to develop the models. The benefits of removing redundant features from a model is demonstrated in Luo *et al.* (2022). Their results indicated that eliminating redundant parameters improved the accuracy of their AGB prediction models compared to when no feature parameter selection is applied.

3.2.2 Model evaluation

A common method for assessing the performance of a regression or classification model is through an independent test dataset that was not used in the training process (Anguita *et al.*, 2012). However, in situations where the samples are not enough for splitting, the Cross-validation (CV) is an alternative (Isaksson *et al.*, 2008). Thus, we evaluated the Xboost model using a 10-fold cross-validation based on the predictor variables listed in Table 3.1. This technique enables models to be evaluated with the entire training dataset through repeated resampling, which increases the total number of data points available for testing and may help prevent over-fitting (Kohavi, 1995; Anguita *et al.*, 2012). The performance metric is averaged across all k iterations to yield a more reliable evaluation.

The most used the R^2, RMSE% metrics, and the MAE (James, 2013) were selected to evaluate the fitting performance of the model at the five experimental levels. R^2 values fall between 0 and 1. A value of 1 indicates that the model accounts for all variance in the dependent variable, whereas a value of 0 shows no explained variance. The Root Mean Square Error (RMSE), quantifies the absolute magnitude of prediction errors in the same units as the observed carbon stock (g m^{-2}), where lower RMSE values signify predictions that are more closely aligned with the measured data. In this study, RMSE is presented as a percentage (%) relative to the mean observed value, thereby standardising the error metric and enabling scale-independent comparisons across different models. Additionally, the Mean Absolute Error (MAE) represents the average absolute difference between observed and predicted values, serving as an indicator of the typical magnitude of prediction errors. Lower MAE values correspond to greater model accuracy.

These metrics were calculated based on Equations (3.1) – (3.4) as indicated below.

$$R^2 = 1 - \frac{\sum(Y_{measured} - Y_{predicted})^2}{\sum(Y_{measured} - Y_{predicted})^2} \quad (3.1)$$

$$RMSE = \frac{\sqrt{\sum_{i-1}^{N} - (Y_{measured} - Y_{predicted})^2}}{N} \quad (3.2)$$

$$RMSE\% = \frac{\sqrt{\frac{1}{n}\sum_{i-N}^{N} - (Y_{measured} - Y_{\ predicted})^2}}{\bar{Y}} \times 100 \quad (3.3)$$

$$MAE = \sum_{n}^{1} \frac{|Y_{measured} - Y_{predicted}|}{n} \quad (3.4)$$

Where $Y_{measured}$ is the observed AGGCS, $Y_{predicted}$ is the predicted AGGCS, N is the number of observed samples, i is the included predictor variable and $\bar{Y}$ is the mean of the observed AGGCS. The R^2, RMSE and RMSE% equations are adopted from (Dube and Mutanga, 2015), while the MAE is adopted from (Qadeer *et al.*, 2024).

3.4 Results

3.4.1. Exploratory analysis of collected above-ground grass carbon stock samples

Table 3.2 shows the descriptive statistics for the AGGCS in the southern region of Kruger National Park. During the wet season of 2020, the observed AGGCS ranged from 4.5 to 58.2 grams per square meter (g m^{-2}), with a maximum of 58.2 g m^{-2} and an average of 28.7 g m-2. In contrast, the dry season recorded AGGCS values between 11.97 and 99.17 g m^{-2}. The standard deviation was 11.89 g m^{-2} for the wet season and 18.80 g m^{-2} for the dry season.

Table 3. 2 Descriptive statistics of measured aboveground grass carbon stock in Kruger National Park

Variable	No of samples	Min	Max	Mean	StDev[1]
Dry grass carbon stock (g m^{-2})	94	4.5	58.2	28.73	11.89

[1]StDev = Standard Deviation

3.4.2. Prediction Selection for above-ground grass carbon stock prediction

A matrix showing the correlation between Sentinel-1 derived parameters and texture matrices is presented in Figure 3.2 below. As indicated in Figure 3.4 of the 30 variables, 25 were highly correlated (R = ≥ 0.8). The variables that were not highly correlated included RVI, VHcor and VVcor. These variables were used to predict the AGGCS.

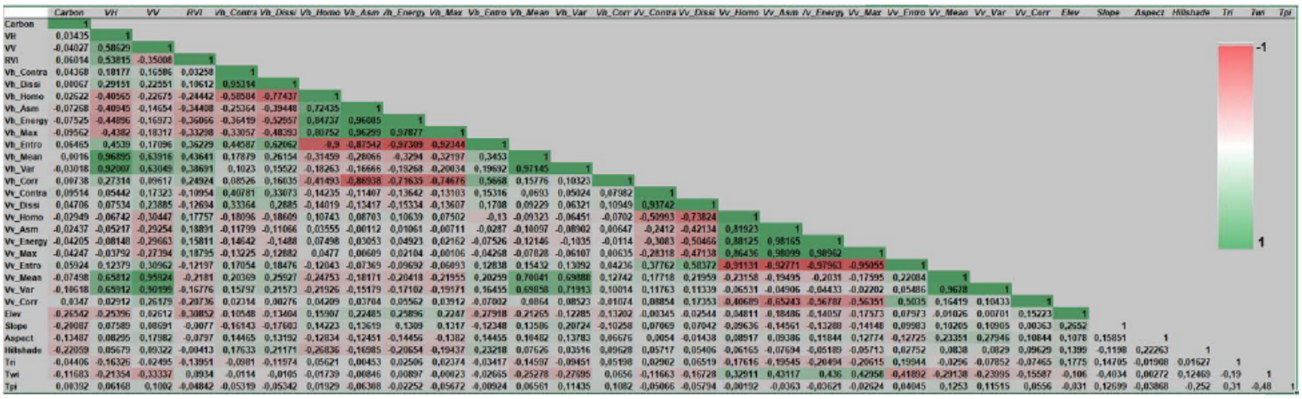

Figure 3. 2 A correlation matrix of Sentinel-1 derived parameters and texture matrices

The predictive capabilities of Sentinel-1 derived parameters coupled with texture matrices and the Extreme Gradient Boosting algorithm from the seven levels of analysis are summarised in Table 3.3 The linear polarisation (i.e. VV) and the intensity ratio (i.e. VH/VV) achieved similar results (R^2 = 0.38, RMSE% = 31, MAE = 6.87) and (R^2 = 0.37, RMSE% = 31, MAE = 6.9) respectively. The Radar Vegetation Index (RVI) performed marginally (≤1%) better (R^2 = 0.39, MAE = 6.77, RMSE% = 30) compared to the other SAR predictors. Arithmetic computations between VH and VV intensities also yielded moderate R^2 values between 0.36-0.37. Meanwhile, combining all SAR variables results in a slightly lower R^2 of 0.36 compared to the individual experiments. Furthermore, an experiment based on texture matrices alone performed within the same range as SAR parameters for both VH and VV (R^2 = 0.39-40, MAE = 5.32-5.21, RMSE% = 30), respectively. Nonetheless, integrating the texture matrix with SAR parameters improved the modelling accuracy (R^2 = 0.60, RMSE% = 20, MAE = 3.97). Figure 3. 3 shows the relationship between observed and predicted above-ground grass carbon stocks based on the XGBoost model integrating texture matrices.

Table 3. 3 Summary of combinations for predicting aboveground grass carbon stock in Kruger National Park

Experiment	Matrices	R^2	RMSE%	MAE
1		**Linear polarisation (LI)**		
	Intensity (VH)	0.38	33.22	6.86
	Intensity (VV)	0.38	31	6.87
2		**Radar Vegetation Index (RVI)**		
	RVI	**0.39**	**30**	**6.77**
3		**Intensity ratios (IR)**		
	VH/VV	0.37	31.03	6.9
	VV-VH/VV+VH	0.37	31.11	6.9
4		**Arithmetic Computations (AC)**		
	VH-VV	0.36	30.84	6.73
	VV+VH	0.37	31.31	6.84
5		**All variables**		
	All optimal S1 variables	0.36	30.84	7.02
6		**Texture Matrices**		
	VH	**0.40**	**30**	**5.32**
	VV	**0.39**	**30**	**5.21**
		Optimal S1 variable + Texture		
7		**0.60**	**20**	**3.97**

Tex = Texture, Topo = Topographic variables, Opt Var = Optimal variables of importance.

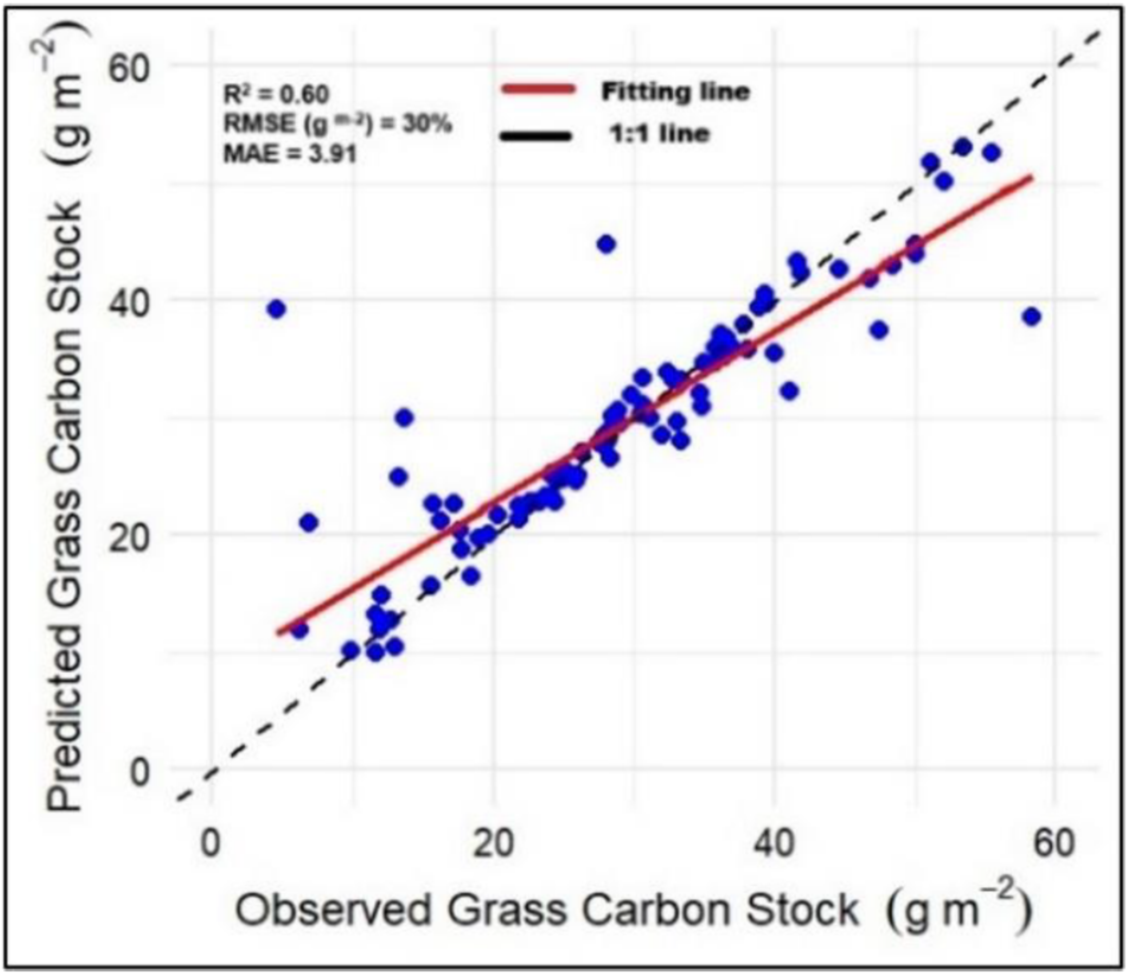

Figure 3. 3 A scatterplot showing the relationship between observed and predicted above-ground grass carbon stocks based on the XGBoost model integrating texture matrices

3.4.4. Important variable selection

Fig. 3.4 illustrates the influential variables selected for predicting AGGCS as measured by their Gain. The Gain metric represents the contribution of each feature to the model's predictive power, with higher values indicating greater importance in making accurate predictions. Results from the correlation matrix, show that of all the 30 variables calculated, only five had a correlation below the set, threshold (i.e. ≥ 0.8). The RVI was the most influential from the SAR parameters whereas the VHcor and VVcor were the least influential from the texture matrices. Overall, the combined input of these parameters optimally predicted (Figure 3.4).

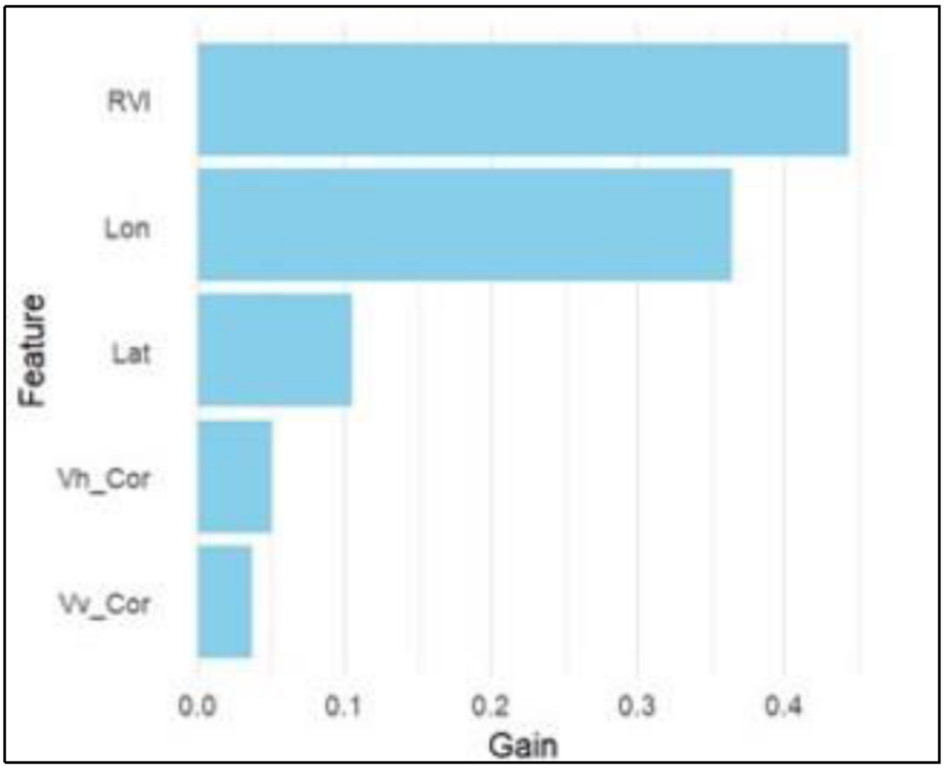

Figure 3. 4 Optimal SAR parameters selected for predicting aboveground grass carbon stock

3.4.5. Comparison of predicted vs observed above-ground grass carbon stock

Figure 3.5 compares the number of sample points (frequency) of the measured and predicted AGGCS. The predicted AGGCS for most classes (i.e. <25 g m^{-2}, 31-35g m-2, 36-40 g m^{-2} and >40 g m^{-2}) are close to the observed values. An exception is the 26-30 g m^{-2} range, which shows a huge underestimate compared to the observed values. This overestimation is more pronounced in the lowest AGGCS class range. The model seems to have obtained a relatively lower error in predicting AGGCS in the 36-40 g m^{-2} class range.

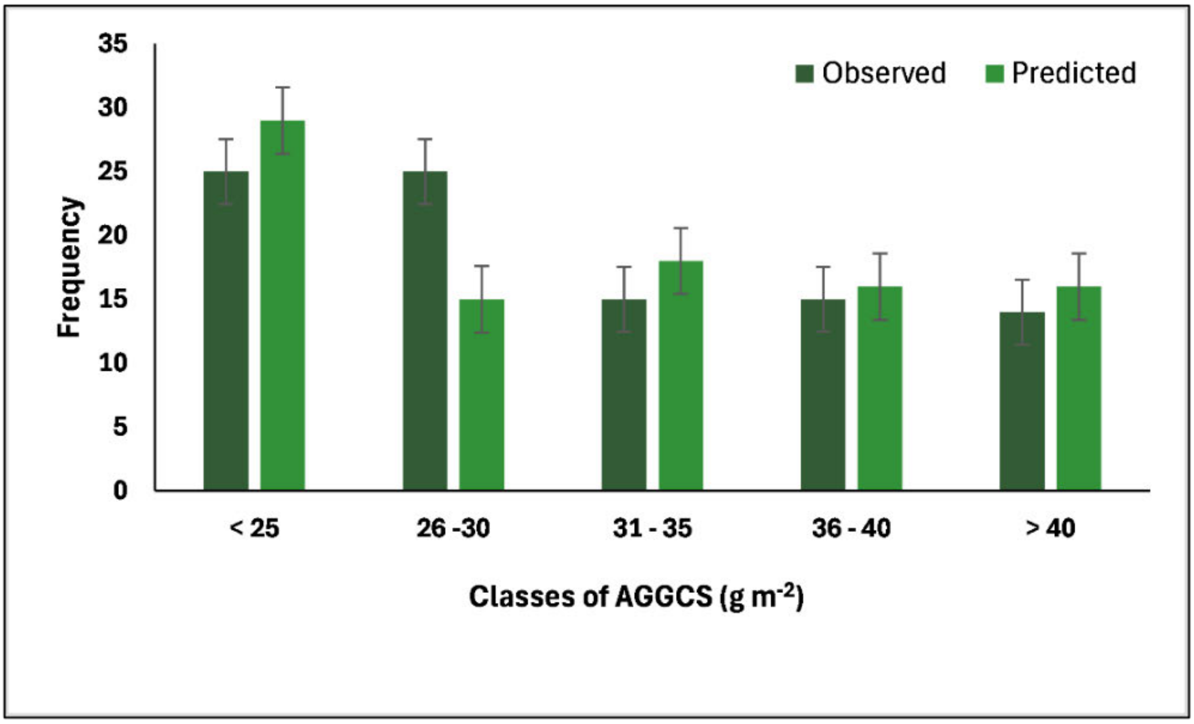

Figure 3. 5 Histogram of measured and predicted aboveground grass carbon stock. Frequency refers to the number of observations per AGGCS range for each category

3.5 Discussion

Precise estimation of climate-regulating ecosystem services in conserved savannah ecosystems such as Kruger National Park is essential for monitoring their contribution to climate regulation and assessing their potential as carbon sinks. This study derived several SAR parameters coupled with texture matrix for quantifying aboveground grass carbon stock in savannah ecosystems using a tree-based algorithm.

3.5.1 Predictive performance of parameters

In this study, a range of SAR-derived parameters including, raw backscatter coefficients (VH, VV), polarisation intensity ratios (VH/VV, VV-VH/VV+VH), radar indices (RVI), and intensity arithmetic computations (VH-VV, VV+VH) were incorporated, with the expectation that they would optimally predict AGGCS. Particularly, the VH/VV Ratio and RVI was anticipated to capture the variability of AGGCS because it separates the cross-polarized (VH) and co-polarized (VV) backscatter signals, which can reflect vegetation density and structure.

High biomass areas often exhibit stronger VH signals due to increased volume scattering (Hamdan *et al.*, 2015) while VV signals can indicate structural orientation in vegetation (Crabbe *et al.*, 2019). The RVI considers various polarization channels (e.g. VH, VV), therefore, combining information from both co-polarized and cross-polarized returns, captures additional insights regarding vegetation structure, density, and volume, as well as water content (Mandal *et al.*, 2020) all of which are closely linked to biomass. However, the predictive power of VH/VV was limited, potentially due to the sparse vegetation in savannah environments, which reduces volume scattering and minimizes the distinction between VH and VV signals. Furthermore, the sparse vegetation is often associated with bare soil background in savannahs (Eisfelder *et al.*, 2012), which may affect backscatter signals where a larger portion of the signal return will emanate from soil than that of grasses. These traits differ from forested ecosystems, where high soil moisture and organic content often amplify the vegetation signal in SAR data (Hamdan *et al.*, 2015).

Similar results were obtained in previous studies, for example (Wang *et al.*, 2013; Voormansik *et al.*, 2020) reported very weak correlation ($R^2 = 0.01$-0.05) when biomass reached a high level. They attributed this weak correlation to rainfall. In this study, AGGCS measurements were done during the wet season. The conditions at the test site at the time of data collection were such that the grasses could have been wet. Generally, wet vegetation tends to exhibit higher backscatter due to increased attenuation and scattering effects. Furthermore, during the rainy season, soil moisture levels typically increase significantly due to precipitation, and this affects the radar signal as it is reported to be sensitive to surface moisture (Sinha *et al.*, 2015a). Water on leaves after rainfall can cause substantial change in SAR backscattering coefficient (Ali *et al.*, 2013). Nasirzadehdizaji *et al.* (2019) also reported weaker correlations ($R^2 = -0.49$) for agricultural fields where the density of the crops was above 70%. Further evidence is provided in Rapiya *et al.* (2023) who observed an unsatisfactory correlation between above-ground biomass (AGB) in natural rangeland using SAR data and optical data. They reported that fluctuation in AGB is most pronounced during periods of peak productivity, with the highest levels observed in late summer, followed by early summer, and gradually decreasing during winter (the cold, dry season). Future studies should factor in these topo-climatic parameters with SAR parameters to better characterise the variability of the AGGCS across different landscapes. However, most available Rainfall data (Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) (Shen *et al.*, 2020). Studies can explore locally available in-situ weather station and soil moisture data for mapping at higher spatial resolutions, 10m in the case of our study.

Beriaux *et al.* (2021) discovered that the correlation between SAR polarimetric parameters and biomass was relatively weak due to the complex interactions between radar signals and forest structure. In this study, AGGCS measurements were done at the peak period where grass growth was at maximum. Compared to dormant grasses, grasses at peak growing stages tend to have taller and denser canopies, which may have an impact on SAR signal penetration depth and scattering processes, and ultimately the accuracy of the AGGCS estimation. Further evidence is provided in Rapiya *et al.* (2023) who observed an unsatisfactory correlation between above-ground biomass (AGB) in natural rangeland using SAR data and optical data. They reported that fluctuation in AGB is most pronounced during periods of peak productivity, with the highest levels observed in late summer, followed by early summer, and gradually decreasing during winter. The transitioning of grasses from one stage to another (i.e. germination, growth, maturity and senescence) may result in variations in vegetation structure. Therefore, acquiring SAR data at different stages of grass growth may capture these temporal variations and provide valuable information for biomass estimation. In addition, savannah ecosystems are complex in nature, consisting of a mixture of grasslands and scattered trees that coexist. The vertical structure and density of grasses and trees may significantly affect SAR signals.

Wang *et al.* (2019b) also reported a moderate correlation between SAR polarimetric parameters and biomass and attributed these findings to factors such as soil moisture. In this study, SAR and field data acquisitions took place in the rainy season. The field conditions at the study site at the time of data collection were such that the soil could have been wet. During the rainy season, soil moisture levels typically increase significantly due to precipitation, and this may have an effect on the radar signals, as it is reported to be sensitive to surface moisture (Ali *et al.*, 2016). We suggest that future studies factor in this soil moisture with SAR parameters to improve the predictive capabilities of the AGGCS models. However, most available Soil moisture products such as the Soil Moisture Active Passive (SMAP) (Chan *et al.*, 2018), Ecosystem Space borne Thermal Radiometer Experiment on Space Station (ECOSTRESS) Cawse-Nicholson and Anderson (2018) and the International Soil Moisture Network (Dorigo *et al.*, 2010) are usually too coarse for localised studies. Studies can explore locally available in-situ soil moisture data for mapping at localised areas.

Generally, if an increase in AGGCS values does not result in a comparable increase in the SAR backscatter values, this implies saturation (Mutanga *et al.*, 2023). In a study focusing on forests, Sinha *et al.* (2015b) observed that the C-band generally saturates at about 70 t ha-1 for tropical forests. In our study, signs of signal saturation were also observed, particularly at 40 g m-2. Nevertheless, actives sensor with high spatial resolution in the L and P spectral bands have been missioned to offer enhanced image contracts due to their multi-looking trait (Mutanga *et al.*, 2023). These sensors can characterize intra-canopy traits and as a result they are able to deal with saturation (Naidoo *et al.*, 2016). However, in many resources constrained savannah regions, such Africa, access to commercial SAR data such as ALOS PALSAR comes at a high cost.

Although the models based solely on SAR parameters were the least effective, when combined with the texture matrices, prediction capability was enhanced. This may be because of the grey-scale properties of images and the spatial position of image pixels, which makes it possible to combine them with backscatter variables to reduce the underestimation or overestimation. Image textures have been reported to be being sensitive to brightness variations arising from canopy structure of vegetation (Haralick *et al.*, 1973). These findings agree with (Naidoo *et al.*, 2016; Cawse-Nicholson and Anderson, 2018; Chen *et al.*, 2019a) who reported improved accuracies when incorporating texture variables. Integrating SAR-derived parameters with texture matrices offers several complementary benefits including advantage on both biophysical and spatial attributes. For example, when SAR parameters explain the structural properties of grasses, texture matrices capture the spatial orientation of these properties. Therefore, providing a composite of variables to account most of the variability in the landscape, which can improve the accuracy of quantifying AGGCS.

Overall, this study underscores the challenges encountered when utilising single source data to estimate AGGCS, especially in complex savannah ecosystems. Nevertheless, the study also highlights potential benefit of integrating multiple parameter from different sources to quantify AGGCS in complex landscapes such as savannahs.

33.7 Conclusion

This study examined the prospect of Sentinel-1-derived matrixes coupled with texture matrices and the XGBoost algorithm for predicting aboveground grass carbon stock in savannah ecosystems. Based on the findings, this study concludes that when utilised separately, SAR parameters and texture matrices tend to underperform (R^2 = 0.36-0.40, RMSE = 30-33%). However, the combined use of SAR parameters and texture matrices results with an R^2 of 0.60 and an RMSE of 20%, suggesting that texture matrices may improve the predictive ability of SAR data for AGGCS estimation. The RMSE values reported are within good predictive accuracy threshold as indicated in Naidoo et al (2019). The RMSE values ranging from 20% to 30% are interpreted as indicative of moderate predictive accuracy. The variables that optimally captured the variability of AGGCS include RVI, VH_{cor} and $VV_{cor\ as}$ well as the latitude and longitude. Overall, the integration of Sentinel-1-derived matrices coupled with texture matrices and the XGBoost algorithm shows potential in predicting aboveground grass carbon stock in savannah ecosystems.

3.6 Summary

A significant constraint on the regular quantification of AGGCS using SAR data in a savannah ecosystem, particularly those in developing nations, is the excessive cost associated with the data and software. The availability of SAR sensors such as Sentinel-1 at no cost, coupled with scalable machine learning algorithms, has been shown to perform well in quantifying AGGCS. Building upon the results outlined in Chapter 2, which highlighted the potential of SAR in estimating AGGCS and noted the limited application of this technology, especially within savannah landscapes, a pertinent inquiry emerged regarding which specific SAR parameter most effectively predicts AGGCS in these ecosystems. This chapter investigated the sensitivity of SAR-derived parameters coupled with texture matrices to predict AGGCS in a savannah ecosystem. RVI performed marginally better compared to the other variables. The incorporation texture matrix into the model enhanced prediction capability, with the most influential predictors being RVI, VH_{cor} and VV_{cor} order of importance. Having uncovered the optimal SAR parameters along with the potential of incorporating texture matrices for AGGCS modelling, subsequent research was targeted towards assessing how these optimal parameters perform within a single season and coupled with topographic variables. The next chapter will test whether incorporating topographic variables, texture matrices coupled with Radar indices, can improve the accuracy of predicting AGGCS in savannah ecosystems.

CHAPTER 4: INTRA-SEASONAL QUANTIFICATION OF ABOVE-GROUND GRASS CARBON STOCK IN SAVANNAH ECOSYSTEMS USING SAR AND TOPOGRAPHIC VARIABLES

This chapter is based on:

Reneilwe Maake*, Onisimo Mutanga, Johannes George Chirima, and Mahlatse Kganyago. "Intra-seasonal quantification of above-ground grass carbon stock in savannah ecosystems using SAR and topographic variables." (2025): Paper orally presented at the Grassland Society of Southern Africa 2025, 21 to 25 July 2025, ANEW Hotel Hilton, KwaZulu-Natal, South Africa. Paper presented by R. Maake.

Abstract

The need for quantifying and monitoring aboveground grass carbon stock (AGGCS), particularly of the grass layer, has gained global traction as many countries face the brunt of climate change. This study tested whether incorporating topographic variables, texture matrices coupled with Radar indices, can improve the accuracy of predicting aboveground grass carbon stocks in savannah ecosystems. With Kruger National Park as the experimental site, fieldwork was conducted in the wet and dry seasons of 2020. We utilised the XGBoost algorithm to predict the aboveground grass carbon stocks and computed the regression coefficient (R^2) and error terms such as Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) expressed as a percentage to assess the model performance. For the wet season, the Radar Vegetation Index (RVI) and texture-based models produced similar results with a ≤ 2% difference. The topographic-based model produced the lowest prediction power (R^2 = 0.25, MAE = 8.9, RMSE% = 40.55) compared to Radar Vegetation Index (RVI) and texture-based models. A similar pattern was observed for the dry season, where the Radar Vegetation Index (RVI) and texture-based models produced similar results with a ≤ 2% margin. Nevertheless, when all the variables were combined, the accuracies improved across both seasons (R^2 = 0.56, MAE = 4.72, RMSE% = 28.05, R^2 = 0.61, MAE = 4.86, RMSE% = 24.66), i.e. by approximately 16% and 21% for the wet and dry seasons, respectively. Nevertheless, the model based on the dry season (R^2 = 0.61, MAE = 4.86, RMSE% = 24.66) explained 61 % of the variation compared to the wet season model. In conclusion, integrating RVI, texture, and topographic variables shows potential in predicting AGGCS in the complex Savannah ecosystems.

Keywords: *Aboveground Grass Carbon Socks, Savannah Ecosystems, Sentinel-1, Texture, Topographic, season*

4.1 Introduction

The aboveground carbon stock is unevenly distributed in both space (Shiyomi *et al.*, 1998; Klaus *et al.*, 2016; Liang *et al.*, 2016) and temporal dimensions (Jin *et al.*, 2014). The need for quantifying and monitoring aboveground grass carbon stock (AGGCS) has gained global traction as many countries face the brunt of climate change. Understanding the distribution of AGGB across different spatial and temporal dimensions is crucial, especially for facilitating the comprehensive implementation of international greenhouse gas inventory programs as well as Goal 13 of the Sustainable Development Goals (SDGS). For example, Goal 13 of the Sustainable Development Goals (SDGS) calls for immediate action to combat climate change and mitigate its impacts. In response to these calls, researchers and policymakers are searching for ways to limit the accumulation of carbon dioxide (CO_2) in the atmosphere.

As a terrestrial ecosystem, the savannah consists of the grass and woody layers that serve as a sink for atmospheric CO_2 (Kuplich *et al.*, 2005). The grass and woody layers eliminate the atmospheric CO_2 through photosynthesis and accumulate it in their biomass (both below and aboveground). They account for an estimated 25% of total gross primary production (GPP) (Moore *et al.*, 2016). Nonetheless, natural and anthropogenic disturbances such as fires and grass harvesting release large amounts of stored carbon into the atmosphere and consequently, transform these pools into sources of greenhouse gases. This alarms the need to devise effective approaches to quantify and monitor AGGCS. Such information is essential for myriad activities ranging from greenhouse gas inventories, terrestrial carbon accounting, as well as climate change modelling (Nichol and Sarker, 2010). In addition, statistics on AGGCS are essential for ensuring an up-to-date flow of information for reporting purposes as required by international agreements such as the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol, as well as the Paris Agreement (Nichol and Sarker, 2010; Abbass *et al.*, 2022).

The ability of the grass and woody layers to capture or release carbon changes with seasons. For example, carbon sequestration reaches highest levels during the peak growth season and declines as senescence begins. Intra-seasonal estimation of carbon stocks involves assessing the quantity of carbon stored in grasses at different temporal windows within a single season. From a climate change perspective, this approach offers insights into carbon changes over time, reflecting the impacts of short-term environmental conditions and management strategies on

carbon stocks. From a grazing perspective, Kearney *et al.* (2022) and Jansen *et al.* (2021) emphasise that within-season biomass estimation is valuable for guiding grazing management practices in the subsequent years. Furthermore, quantification and monitoring carbon stocks well within the season can serve as an early warning to indicate carbon stock stress factors such as drought, excessive grazing, or pest infestations. While many studies use single-date data, within-season estimation has received less attention, although it can significantly contribute to a mechanistic understanding of ecosystem dynamics (Maestre and Cortina, 2002)
Coupled with machine learning, several studies have demonstrated great potential for using Remote sensing as a proxy for quantifying AGGCS (Ding *et al.*, 2019; Chapungu *et al.*, 2020; Gao *et al.*, 2020; Meng *et al.*, 2020; Alvarez-Mendoza *et al.*, 2022; Chang *et al.*, 2024; Li *et al.*, 2024; Netsianda and Mhangara, 2025). Nevertheless, many carbon stock quantification studies were based on spectral reflectance and vegetation indices extracted from optical sensors (Chapungu *et al.*, 2020) and hyperspectral (Liu *et al.*, 2024). The observational drawback of optical data is limited to cloud-free conditions, shallow penetration depth of the vegetation canopy, and less sensitivity to senescent vegetation. The Active sensors, such as the Synthetic Aperture Radar (SAR), offer alternative data with benefits ranging from all weather conditions, day and night observation (Singh *et al.*, 2021), to high penetration depth of the vegetation structure (El Hajj et al., 2018). Studies have demonstrated the usefulness of the backscattering coefficients and polarisation metrics extracted from SAR (Naidoo *et al.*, 2015; Berninger *et al.*, 2018). Nevertheless, the common observational drawback of both Optical and SAR sensors is that they saturate at certain maximum biomass levels (i.e. of 350–400 ton/ha for the C-band (Kasischke *et al.*, 2002) and at ± 0.3 g.cm^{-2} from multispectral bands (Mutanga *et al.*, 2012). Therefore, incorporating ancillary variables in SAR-based models could provide unprecedented opportunities for improving AGGCS estimates.

Ancillary data, such as topographic (i.e. altitude and slope) (Dieleman *et al.*, 2013) and texture matrices, affect the distribution of carbon stock. For example, elevation is observed to influence the zonal climate conditions along the vertical gradient (He *et al.*, 2012), whereas the position of the slope and its orientation are expected to affect both hydrological processes and the intensity of solar radiation (Zhu *et al.*, 2019a). This, in turn, affects the distribution of carbon stocks. Considering this, several studies have also demonstrated that incorporating such topographic variables into remote sensing models can effectively improve the prediction accuracies (Xie *et al.*, 2009; Jia *et al.*, 2024; Yao and Ren, 2024). While some of these studies were tested in forested ecosystems (Yao and Ren, 2024), those that tested this concept on grassland were conducted in the homogeneous temperate ecosystems (Xie *et al.*, 2009; Jia *et al.*, 2024). Therefore, it is worth testing how incorporating terrain variables in estimating AGGCS will perform in the savannah ecosystem.

Nevertheless, studies have attempted to quantify seasonal variation of carbon stocks (Folega *et al.*, 2015; Nyamugama and Kakembo, 2015; Thapa *et al.*, 2015; Musthafa and Singh, 2022). However, most studies focus on annual snapshots carbon stocks. In savannah ecosystems, however, grass growth and carbon dynamics change rapidly within a season due to variable rainfall and grazing pressures. There is a need for higher temporal resolution in estimating aboveground carbon stocks to better capture these intra-seasonal dynamics and support timely ecological or land management decisions. Studies have incorporated texture matrix with SAR-based derivatives (Kuplich *et al.*, 2005; Cutler *et al.*, 2012; Sarker *et al.*, 2012; Naidoo *et al.*, 2016; Liao *et al.*, 2020) and topographic variables (Bispo *et al.*, 2014) separately. However, the Synthetic Aperture Radar (SAR) and topographic data are often insufficiently integrated, despite their potential to improve predictions by capturing structure, moisture, and terrain effects. There is a need to evaluate and integrate SAR data (which penetrates cloud cover) and topographic variables to improve the accuracy and robustness of grass carbon stock models. This study seeks to quantify the intra-seasonal variability of AGGCS within savannah ecosystems by utilizing SAR data, texture matrices, and topographic variables. Furthermore, it aims to assess the extent to which the integration of these datasets enhances the predictive power of carbon stock modelling.

4.2 Materials and methods

4.2.1 Experimental site description

An outline and description of the study site is presented in subsection 1.6

*4.2.*2 Field above-ground grass carbon stock sampling

Fieldwork was conducted in February for the wet season and July for the dry season of 2020. Specifically, grass samples were measured and collected in February (wet season) and July (dry season). In total, 94 plots were sampled to serve as the foundation to answer this question. An outline of how the sites and plots were demarcated as well as how the samples are processed is presented in subsection 1.7. Figure 4.1 shows the condition of the study site during the wet and dry seasons of 2020 a) the wet season b) the dry season.

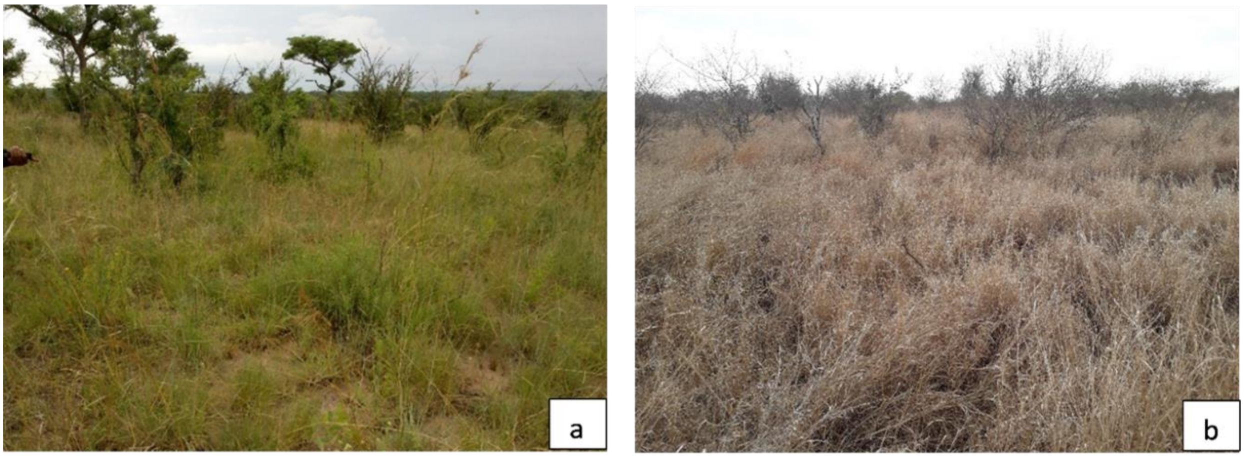

Figure 4. 1 Selected photos showing the condition of the study site during the wet and dry seasons a) the wet season of 2020 b) the dry season of 2020.

4.2.3 Data acquisition & pre-processing

4.2.3.1 Sentinel-1 data acquisition

Sentinel-1 images were collected for the wet and dry seasons of 2020. Specifically, two images were downloaded from the Copernicus Open Access Hub, accessible at https://scihub.copernicus.eu/dhus/. This means one in February (wet season) and another in July (dry season) of 2020 respectively. An outline of how the satellite image in question was pre-processed is presented in subsection 1.8.

4.2.3.2 Digital Elevation Model (DEM) Acquisition

To account for the influence of the terrain on the accuracy of the model, we downloaded the Shuttle Radar Topography Mission's (SRTM) Digital Elevation Model (DEM) from the United States Geological Survey (USGS). The SRTM mission offers high-resolution (i.e. 30 m) DEM on a near-global scale (Yang *et al.*, 2011). Figure 4.2 shows a graphic depiction of the Sentinel-1 pre-processing workflow that was adopted in the study. As shown in Figure 4.2, the image was first resampled to a 10 m spatial resolution to align it with the spatial resolution of the Sentinel-1 image. We then extracted topographic variables (Table 4.1) using the ArcGIS 10.8 software.

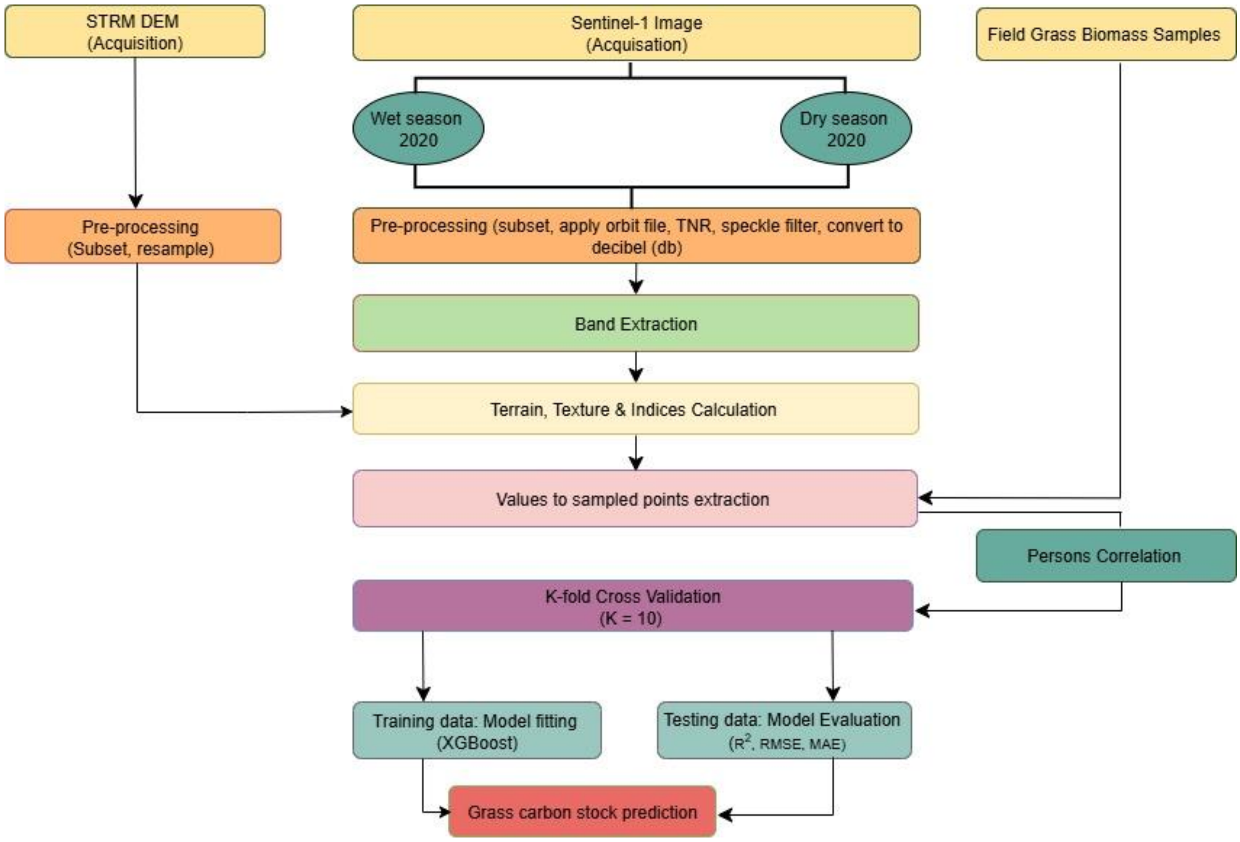

Figure 4. 2 A depiction of the Sentinel-1 pre-processing workflow used in the study

4.2.4 Deriving predictor variables

Selecting appropriate input variables is essential for optimal model performance. This selection process is important to enable the model to detect the potential relationships between the dependent and independent variables in the training process (Osman *et al.*, 2021). For this study, we utilised the VH and VV channels from Sentinel-1 to compute the radar vegetation index (i.e. RVI) and the texture matrices based on the Grey-level co-occurrence matrix (GLCM) method (Table 4.1). Following the methodology of Haralick *et al.* (1973), the GLCM parameters were determined using a 3×3 processing window.

For the topographic variables, we relied on the DEM to extract the elevation parameter and subsequently computed six additional topographic variables, including the slope, aspect, hill shade, Topographic Roughness Index, Topographic Wetness Index as well as the Topographic

Position Index, which are detailed in Table 4.1. The choice of these variables was informed by their effective use in the literature (Salinas-Melgoza *et al.*, 2018).

Table 4. 1 Selected predictor variables utilised in this study.

Predictor variable	Expression	Description
Radar Indices (RI)	RVI	Radar Vegetation Index
Texture Matrix	VH/VV_{Con}	Contrast
	VH/VV_{Dis}	Dissimilarity
	VH/VV_{Hom}	Homogeneity
	VH/VV_{ASM}	Angular Second Moment
	VH/VV_{Ene}	Energy
	VH/VV_{Max}	Maximum Probability
	VH/VV_{Ent}	Entropy
	VH/VV_{Mea}	GLCM Mean
	VH/VV_{Var}	GLCM Variance
	VH/VV_{Cor}	GLCM Correlation
Topographic Matrices	H	Elevation
	Slope	Slope
	Aspect	Aspect
	Hillshade	Hillshade
	TRI	Topographic Roughness Index
	TWI	Topographic Wetness Index
	TPI	Topographic Position Index

4.3 Statistical analysis

4.3.1 Regression algorithms

The prediction of AGGCS based on the above-tabulated variables relied on the Extreme Gradient Boosting (XGBoost). The XGBoost is a highly effective and efficient machine-learning algorithm that operates on the gradient boosting framework. It is commonly employed for supervised learning tasks such as regression, classification, and ranking, given its scalability and capacity to manage intricate datasets with minor adjustments (Chen *et al.*, 2019a). The core concept of a boosting algorithm is integrating the outputs of the weak learners consecutively for achieving better performance (Kavzoglu and Teke, 2022) . A more inclusive elaboration on how the XGBoost functions can be found in Chen and Guestrin (2016). The XGBoost algorithm has been successfully employed in predicting groundwater Osman *et al.* (2021) estimating grass carbon stocks (Maake *et al.*, 2025), crop biochemical parameters (Kganyago *et al.*, 2022) and mapping forest above-ground biomass (Tamiminia *et al.*, 2022).

Coalescing several parameters (i.e. raw bands, texture matrices and topographic variables) during modelling tends to introduce multicollinearity. Multicollinearity appears when two or more independent variables in a multiple regression model are highly correlated (Daoud, 2017; Shrestha, 2020). A pairwise Pearson correlation matrix was ran to assess the magnitude of collinearity. A threshold of $R \geq 0.8$ was selected to eliminate variables unsuitable for subsequent model configurations (Luo *et al.*, 2022). Eliminating variables that contribute less to the accuracy of a prediction model has proven effective in Luo *et al.* (2022). In this study, a subset of variables showing the least correlation with the dependent variable from a pool of 33 variables was selected for further analysis.

4.3.2 Model Evaluation

A common method for assessing the performance of a regression or classification model is through an independent test dataset that was not used in the training process (Anguita *et al.*, 2012). However, in situations where the samples are not enough for splitting, the Cross-validation (CV) is an alternative (Collart and Guisan, 2023). Thus, we evaluated the Xboost model using a 10-fold cross-validation based on the predictor variables listed in Table 4.1. This technique enables models to be evaluated with the entire training dataset through repeated resampling, which increases the total number of data points available for testing and may help prevent over-fitting (Kohavi, 1995; Anguita *et al.,* 2012). The performance metric is averaged across all k iterations to yield a more reliable evaluation.

The most used coefficient of determination (R^2), the root mean squared error (RMSE) metrics, and the Mean Absolute Error (MAE) (James, 2013) were selected to evaluate the fitting performance of the model at the five experimental levels. R^2 values fall between 0 and 1. A value of 1 indicates that the model accounts for all variance in the dependent variable, whereas a value of 0 shows no explained variance. The RMSE on the other hand, calculates the absolute magnitude of prediction errors in the same units (i.e. as the observed carbon stock (g m^{-2}), with lower values indicating predictions closer to measured values. The RMSE in this study is expressed as a percentage (%) to standardise the relative to the mean observed value, providing a scale-independent measure of error that facilitates comparisons models In addition, the MAE provides the average of the absolute difference between the observed value and the predicted values. It quantifies the typical magnitude of prediction errors, with lower MAE values indicating higher model accuracy. These matrices were calculated based on Equations (4.5) – (4.8) as indicated below.

$$R^2 = 1 - \frac{\sum(Y_{measured} - Y_{predicted})^2}{\sum(Y_{measured} - Y_{predicted})^2} \tag{4.5}$$

$$RMSE = \frac{\sqrt{\sum_{i-1}^{N} - (Y_{measured} - Y_{predicted})^2}}{n} \tag{4.6}$$

$$RMSE\% = \frac{\sqrt{\frac{1}{n}\sum_{i-N}^{N} - (Y_{measured} - Y_{\ predicted})^2}}{\bar{Y}} \times 100 \tag{4.7}$$

$$MAE = \sum_{n}^{1} \frac{|Y_{measured} - Y_{predicted}|}{n} \tag{4.8}$$

Where $Y_{measured}$ is the observed AGGCS, $Y_{predicted}$ is the predicted AGGCS, n is the number of observed samples, i is the included predictor variable and $\bar{Y}$ is the mean of the observed AGGCS. The R^2, RMSE and RMSE% equations are adopted from (Dube and Mutanga, 2015), while the MAE is adopted from (Qadeer *et al.*, 2024).

4.4 Results

4.4.1 Exploratory analysis of the collected above-ground grass carbon stock samples

Table 4.2 shows the descriptive statistics for the AGGCS in the southern region of Kruger National Park. During the wet season of 2020, the observed AGGCS ranged from 4.5 to 58.2 grams per square meter (g m^{-2}), with a maximum of 58.2 g m^{-2} and an average of 28.7 g m^{-2}. In contrast, the dry season recorded AGGCS values between 11.97 and 99.17 g m^{-2}. The results show that the AGGCS in the wet season recorded a moderate standard deviation (11.89 g m^{-2}) ranging from 4.5 to 58.2 g m^{-2}, suggesting slight right-skewness and moderate variability compared to the dry season that exhibits higher variability (StDev = 18.80 g m^{-2}, range 11.97–99.17 g m^{-2}).

Table 4. 2 Descriptive statistics of AGGCS for the wet and dry seasons of 2020

Period	Variable	No of samples	Min	Max	StDev[1]
Wet season (2020)	Dry grass carbon stock (g m^{-2})	94	4.5	58.2	11.89
Dry season (2020)	Dry grass carbon stock (g m^{-2})	84	11.97	99.17	18.80

[1]StDev = Standard Deviation

4.4.2 Correlation analysis for AGGCS Prediction

The linear correlation between field AGGCS and the 33 pool of extracted variables is presented in Figure 4.2. Of the 33 variables, 24 were highly correlated (R = ≥ 0.8). The variables that were not highly correlated included RVI, VH_{cor} and VV_{cor}, Slope, Aspect, Hill shade, TRI, TWI, TPI which were incorporated into the model to predict the AGGCS.

	Carbon	VH	VV	RVI	Vh_Contra	Vh_Dissi	Vh_Homo	Vh_Asm	Vh_Energy	Vh_Max	Vh_Entro	Vh_Mean	Vh_Var	Vh_Corr
Carbon	1													
VH	0,03435	1												
VV	-0,04027	0,58629	1											
RVI	0,06014	0,53815	-0,35008	1										
Vh_Contra	0,04368	0,18177	0,16586	0,03258	1									
Vh_Dissi	0,00067	0,29151	0,22551	0,10612	0,95314	1								
Vh_Homo	0,02622	-0,40565	-0,22675	-0,24442	-0,58584	-0,77437	1							
Vh_Asm	-0,07268	-0,40945	-0,14654	-0,34408	-0,25364	-0,39448	0,72435	1						
Vh_Energy	-0,07525	-0,44896	-0,16973	-0,36066	-0,36419	-0,52957	0,84737	0,96085	1					
Vh_Max	-0,09562	-0,4382	-0,18317	-0,33298	-0,33057	-0,48393	0,80752	0,96299	0,97877	1				
Vh_Entro	0,06465	0,4539	0,17096	0,36229	0,44587	0,62062	-0,9	-0,87542	-0,97309	-0,92344	1			
Vh_Mean	0,0016	0,96895	0,63916	0,43641	0,17879	0,26154	-0,31459	-0,28066	-0,3294	-0,32197	0,3453	1		
Vh_Var	-0,03018	0,92007	0,63049	0,38691	0,1023	0,15522	-0,18263	-0,16666	-0,19268	-0,20034	0,19692	0,97145	1	
Vh_Corr	0,00738	0,27314	0,09617	0,24924	0,08526	0,16035	-0,41493	-0,86938	-0,71635	-0,74676	0,5668	0,15776	0,10323	1
Vv_Contra	0,09514	0,05442	0,17323	-0,10954	0,40781	0,33073	-0,14235	-0,11407	-0,13642	-0,13103	0,15316	0,0693	0,05024	0,07982
Vv_Dissi	0,04706	0,07534	0,23885	-0,12694	0,33364	0,2885	-0,14019	-0,13417	-0,15334	-0,13607	0,1708	0,09229	0,06321	0,10949
Vv_Homo	-0,02949	-0,06742	-0,30447	0,17757	-0,18096	-0,18609	0,10743	0,08703	0,10639	0,07502	-0,13	-0,09323	-0,06451	-0,0702
Vv_Asm	-0,02437	-0,05217	-0,29254	0,18891	-0,11799	-0,11066	0,03555	-0,00112	0,01061	-0,00711	-0,0287	-0,10097	-0,08902	0,00647
Vv_Energy	-0,04205	-0,08148	-0,29663	0,15811	-0,14642	-0,1488	0,07498	0,03053	0,04923	0,02162	-0,07526	-0,12146	-0,1035	-0,0114
Vv_Max	-0,04247	-0,03792	-0,27394	0,18795	-0,13225	-0,12882	0,0477	0,00609	0,02104	-0,00106	-0,04268	-0,07828	-0,06107	0,00635
Vv_Entro	0,05924	0,12379	0,30962	-0,12197	0,17054	0,18476	-0,12043	-0,07369	-0,09692	-0,06093	0,12838	0,15432	0,13092	0,04236
Vv_Mean	-0,07498	0,65812	0,95924	-0,2181	0,20369	0,25927	-0,24753	-0,18171	-0,20418	-0,21955	0,20259	0,70041	0,69888	0,12742
Vv_Var	-0,10618	0,65912	0,90199	-0,16776	0,15797	0,21573	-0,21926	-0,15179	-0,17102	-0,19171	0,16455	0,69858	0,71913	0,10014
Vv_Corr	0,0347	0,02912	0,26179	-0,20736	0,02314	0,00276	0,04209	0,03704	0,05562	0,03912	-0,07002	0,0864	0,08523	-0,01074
Elev	-0,26542	-0,25396	0,02612	-0,30852	-0,10548	-0,13404	0,15907	0,22485	0,25896	0,2247	-0,27918	-0,21265	-0,12285	-0,13202
Slope	-0,20087	0,07589	0,08691	-0,0077	-0,16143	-0,17603	0,14223	0,13619	0,1309	0,1317	-0,12348	0,13586	0,20724	-0,10258
Aspect	-0,13487	0,08295	0,17982	-0,0797	0,14465	0,13192	-0,12834	-0,12451	-0,14456	-0,1382	0,14455	0,10482	0,13783	0,06676
Hillshade	-0,22059	0,05679	0,09322	-0,00413	0,17633	0,21171	-0,26836	-0,16985	-0,20654	-0,19437	0,23218	0,07626	0,03516	0,09628
Tri	-0,04406	-0,16326	-0,02495	-0,13951	-0,0981	-0,11974	0,05621	0,00453	0,02506	0,02374	-0,03417	-0,14597	-0,09451	0,05198
Twi	-0,11683	-0,21354	-0,33337	0,0934	-0,0114	-0,0105	-0,01739	-0,00846	0,00897	-0,00023	-0,02665	-0,25278	-0,27695	0,0656
Tpi	0,00392	0,06168	0,1002	-0,04842	-0,05319	-0,05342	0,01929	-0,06308	-0,02252	-0,05672	-0,00924	0,06561	0,11435	0,1082

	Vv_Contra	Vv_Dissi	Vv_Homo	Vv_Asm	Vv_Energy	Vv_Max	Vv_Entro	Vv_Mean	Vv_Var	Vv_Corr	Elev	Slope	Aspect	Hillshade	Tri	Twi	Tpi
Vv_Contra	1																
Vv_Dissi	0,93742	1															
Vv_Homo	-0,50993	-0,73824	1														
Vv_Asm	-0,2412	-0,42134	0,81923	1													
Vv_Energy	-0,3083	-0,50466	0,88125	0,98165	1												
Vv_Max	-0,28318	-0,47138	0,86436	0,98099	0,98962	1											
Vv_Entro	0,37762	0,58372	-0,91131	-0,92771	-0,97963	-0,95055	1										
Vv_Mean	0,17718	0,21959	-0,23158	-0,19495	-0,2031	-0,17595	0,22084	1									
Vv_Var	0,11763	0,11339	-0,06531	-0,04906	-0,04433	-0,02202	0,05486	0,9678	1								
Vv_Corr	0,08854	0,17353	-0,40689	-0,65243	-0,56787	-0,56351	0,5035	0,16419	0,10433	1							
Elev	-0,00345	-0,02544	-0,04811	-0,18486	-0,14057	-0,17573	0,07973	-0,01026	0,00701	0,15223	1						
Slope	0,07069	0,07042	-0,09636	-0,14561	-0,13288	-0,14148	0,09983	0,10205	0,10905	0,00363	0,2652	1					
Aspect	0,0054	-0,01438	0,08917	0,09386	0,11844	0,12774	-0,12725	0,23351	0,27946	0,10844	0,1078	0,15851	1				
Hillshade	0,05717	0,05406	-0,06165	-0,07694	-0,05189	-0,05713	0,02752	0,0838	0,0829	0,09629	0,1399	-0,1198	0,22263	1			
Tri	0,02902	0,06519	-0,17616	-0,19545	-0,20494	-0,20615	0,19944	-0,0296	-0,07852	-0,07465	0,1775	0,14705	-0,01908	0,01627	1		
Twi	-0,11663	-0,16728	0,32911	0,43117	0,436	0,42958	-0,41892	-0,29138	-0,23995	-0,15587	-0,106	-0,4034	0,00272	0,12469	-0,19	1	
Tpi	-0,05066	-0,05794	-0,00192	-0,0363	-0,03621	-0,02624	0,04045	0,1253	0,11515	0,0556	-0,031	0,12699	-0,03868	-0,252	0,31	-0,48	1

Figure 4. 3 A correlation matrix of Sentinel-1 derived parameters and texture matrices. (2025)

4.4.3 Predictive performance of parameters

Table 4.3 summarises the predictive capabilities associated with parameters derived from the Sentinel-1 image and the DEM, in conjunction with the Extreme Gradient Boosting algorithm across the five levels of analysis for both the wet and the dry seasons. For the wet season, the

Radar Vegetation Index (RVI) and texture-based models produced similar results with a ≤ 2% margin. The topography-based model demonstrated the lowest predictive performance (R^2 = 0,39, MAE = 8.9, RMSE = 17.20) compared to the Radar Vegetation Index (RVI)-based (R^2 = 0.39, MAE = 4.02, RMSE = 11.38) compared and $VH_{texture}$-based models (R^2 = 0.40, MAE = 5.31, RMSE = 23.43) (for the wet season. A similar pattern was observed for the dry season, where the Radar Vegetation Index (RVI) and texture-based models produced similar results with a ≤ 2% margin. Nevertheless, when all the variables were combined, the accuracies improved across both seasons (R^2 = 0.56, MAE = 4.72, RMSE = 0.03, R^2 = 0.61, MAE = 4.86, RMSE = 0.03), i.e. by approximately 16% and 21% for the wet and dry seasons, respectively. Nevertheless, the model based on the dry season (R^2 = 0.61, MAE = 4.86, RMSE = 0.03) explained 61 % of the variation compared to the wet season model.

Table 4. 3 Summary of combinations for predicting aboveground grass carbon stock in Kruger National Park

Temporal Window	Predictor Variable	R^2	RMSE (g. m^2)	RMSE%	MAE (g. m^{-2})
Wet season (2020)					
	RVI	0.39	11.38	30	7.02
	$VH_{Texture}$	0.40	23.43	30.84	5.32
	$VV_{Texture}$	0.39	24.22	30.88	5.21
	Topo	0.25	**17.20**	40.55	8.9
	RVI + Tex + Topo	0.56	12.46	28.05	4.72
	Opt Var	**0.65**	**7.41**	**25.49**	**3.76**
Dry season (2020)					
	RVI	0.40	7.20	30	5.32
	$VH_{Texture}$	0.43	17.35	30.33	5.32
	$VV_{Texture}$	0.42	16.60	30.22	5.32
	Topo	0.23	11.89	40.22	0.14
	RVI + Tex + Topo	0.61	11.20	24.66	4.86
	Opt Var	**0.78**	**4.58**	**15.95**	**3.23**

[1]RVI = Radar Vegetation Index; $VH_{tetxure}$ = Vertical Transmit and Horizontal Receive Texture, $VV_{tetxure}$ = Vertical Transmit and Vertical Receive Texture; Topo = Topographic variables, Opt Var = Optimised Variables

4.4.4 Important variable selection

The variables that had the greatest impact on predicting AGGCS, evaluated based on their Gain are shown in Figure 4.4. The Gain metric quantifies the contribution of each feature to the model's predictive capability, with high values signifying a greater significance in facilitating accurate predictions. For the wet season, the RVI, TRI, TPI and slope were the top four most influential, respectively, while RVI, TRI, Slope and Aspect contributed the most to the predictive power for the dry season. The variables that contributed the least to the accuracy included hill shade, VH_{cor} and VV_{cor} for both the wet and dry seasons. Radar vegetation indexes and terrain factors were included in most of the AGB models for across the two seasons, although the order of selected variables differed. This result indicated that multiple features, rather than a single feature, influenced AGB estimation.

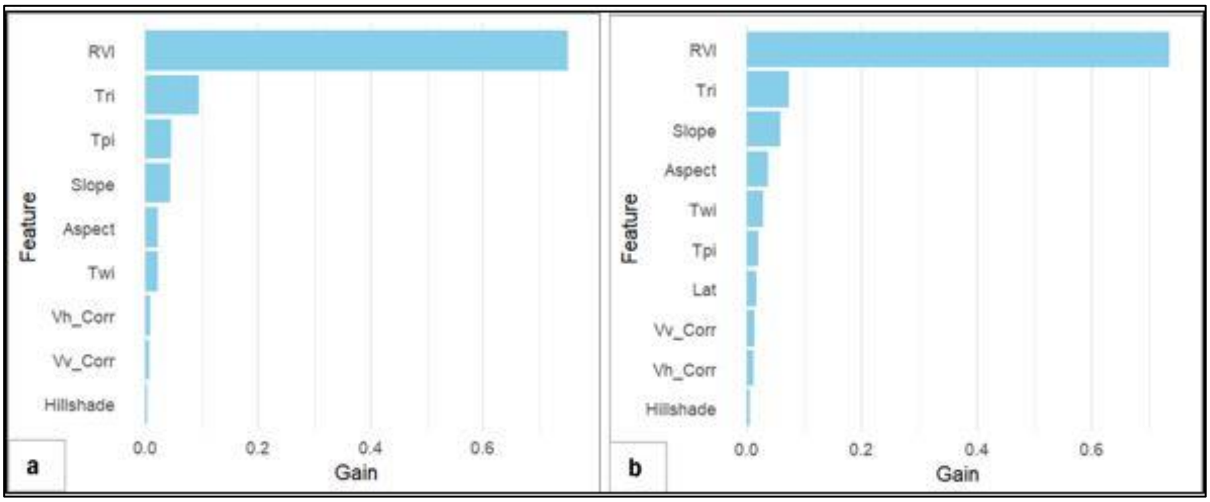

Figure 4. 4 Feature of importance for predicting aboveground grass carbon stock, variables: a) wet season, b) dry season

4.4.5 Performance of Optimised Models by Feature Relevance

Variable importance aims to evaluate how much certain features enhance the overall predictive power of the model (Williamson *et al.*, 2021). As such, the model was re-run using the variable that had the greatest impact on predicting AGGCS. To do this, the XGBoost algorithm was used to fit the dataset based on the selected optimal parameters (i.e. RVI, TRI, TPI and Slope) for the wet season and (i.e. RVI, TRI, Slope and Aspect) for the dry season. As shown on table 4.3, the dry season-based model demonstrated the highest predictive power (R^2 = 0.78, RMSE = 4.5, RMSE% = 15.95 and MAE = 17.12) compared to the wet season-based model (R^2 = 0.65, RMSE = 7.41, RMSE% = 25.49 and MAE = 3.76). Overall, improvement in the predictive power of the models based on the selected optimal parameters across the wet and dry seasons was achieved. The scatter plot depicted in Figure 4.5 illustrates the wet and dry season predictive models based on the selected optimal parameters.

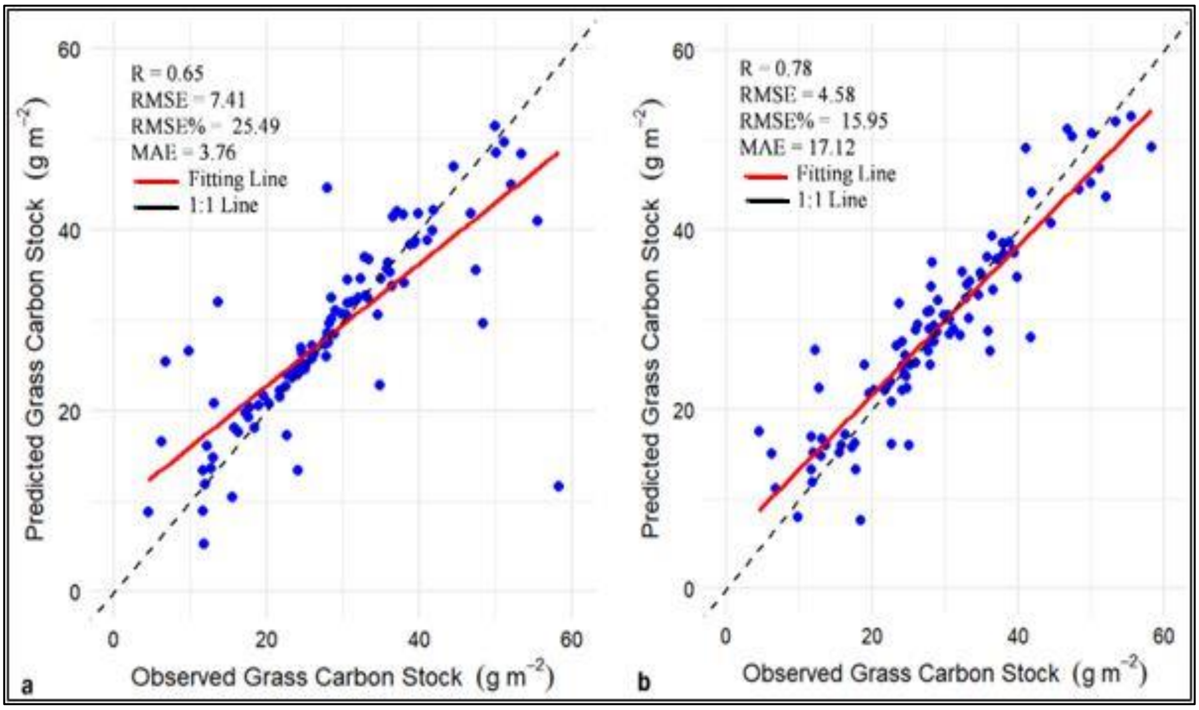

Figure 4. 5 Scatterplots showing the relationship between observed and predicted AGGCS based on the XGBoost model integrating texture matrices and topographic variables (a) wet season (b) dry season

4.4.6 Comparison of predicted vs observed AGGCS

The spatial distribution analysis of AGGCS across the wet and the dry season was conducted using these models based on the selected optimal parameters. Fig. 4.6 illustrates the comparison between the frequency of sample points for the measured and predicted AGGCS across the seasons. For the wet season, most categories (specifically <25 g m^{-2}, 31-35 g m^{-2}, 36-40 g m^{-2}, and >40 g m^{-2}), the predicted AGGCS values are in close range to the observed ones. However, the 26-30 g m^{-2} category shows a significant underestimation when compared to the observed data. The model appears to have achieved a relatively lower error rate in predicting AGGCS for the 36-40 g m^{-2} class range. Meanwhile, the dry season model generally overestimates AGGCS in most categories (specifically <20 g m^{-2}, 30-40 g m^{-2}, and 51-60 g m^{-2}), and underestimates in the 20-30 g m^{-2}, 41-50 g m^{-2} and ≥70 g m^{-2}. Overall, the wet season shows predictions that are more consistent across lower AGGCS categories while the dry season shows more variability in predictions, especially in both the lower and higher-class ranges.

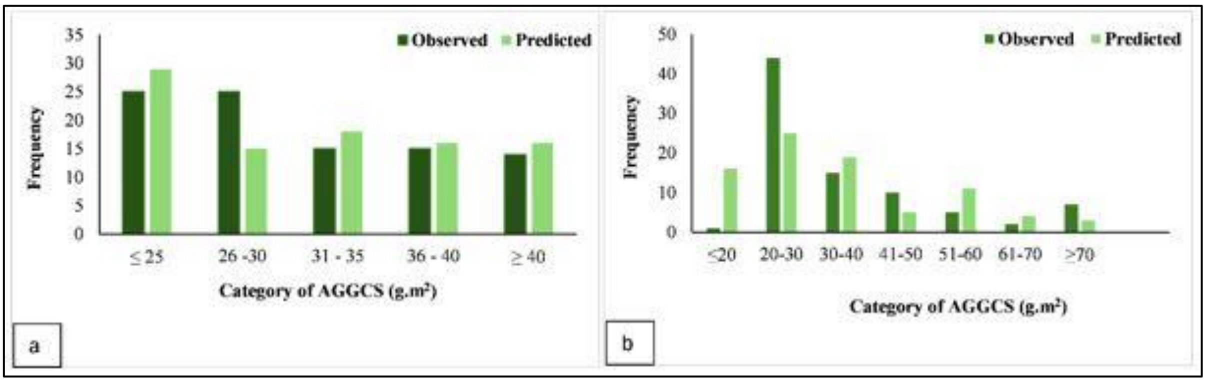

Figure 4. 6 Histograms of observed vs predicted AGGCS, (a) wet season, (b) dry season. Frequency refers to the number of observations per AGGCS range for each category

4.5 Discussion

Carbon stock changes with the change in seasons. As such, strengthening our ability to monitor and quantify these changes well within the season provides a snapshot of short-term related changes, offering near real-time information on variations in AGGCS. Remote sensing techniques offer near-real-time monitoring and quantification. In this study, a combination of radar indices (RVI), texture matrices, and topographic variables were tested to predict AGGCS in savannah ecosystems within a single season cycle, particularly in the wet and dry seasons.

4.5.1 Predictive performance of parameters

The findings suggest that integrating RVI, texture, and topographic variables improves the predictive capabilities of the ExBoost model across all seasons compared to the performance of RVI, texture metrics topographic variables as stand-alone. Savanna ecosystems are heterogeneous, consisting of a transition from high-density grass areas to low-density grass areas (Higgins *et al.*, 2007) with a soil background. Thus, it is expected that incorporating these datasets hold great potential in enhancing the predictive power of the models as they bring complementary sensitivities to account for the heterogeneity associated with the savannah ecosystem that each variable may struggle to detect as a stand-alone predictor. For example, SAR parameters capture the structural properties of grasses, texture matrices capture the spatial geometry of these grasses and have been reported to be sensitive to brightness variations arising from the canopy structure of vegetation (Haralick *et al.*, 1973) while topographic variables capture physical properties.

This result corroborates with those of Zhou *et al.* (2013) who mapped aboveground forest biomass using Quickbird imagery and found that the combination of vegetation indices, texture, and topography yielded an even better R^2 value of 0.74 with the lowest RMSE (17.21 t/ha) and bias (−1.85 t/ha). In another study mapping grassland in in China, Yao and Ren (2024) found that a random forest (RF) model, which included data on topography, climate, and remote sensing, demonstrated strong simulation accuracy, with a training dataset showing R^2 = 0.73, RMSE = 44.85 g.m^2, and MAE = 30.92 g m^2. Using Landsat-8, a study by Kelsey and Neff (2014) also found that models that incorporate image texture variables show a stronger correlation with biomass compared to those that rely solely on physical and spectral variables. Our results are consistent with those of Sarker *et al.* (2012), who reported that a goodness of fit improved to 0.91 (adjusted R^2) and an RMSE of 26.95 t/ha for biomass levels up to 532 t/ha using the ratio of texture parameters of C-HV/C-HH.

Overall, the results suggest the dry season-based model could only explain 61% of the variation, which is an acceptable accuracy but not a substantial improvement. This unsubstantial improvement could probably be due to the complex configuration of the savannah ecosystem. The results also suggest that SAR data might have a limited role in estimating biophysical traits of grasses in heterogeneous landscapes such as the savannah. In uniform areas such as agricultural fields and forests, where backscattering mechanisms tend to be alike, SAR data tends to provide more valuable insights into vegetation properties.

4.5.2 The influence of selected features on AGGCS models

The XGBboost algorithm makes provisions for effective metrics of relative variable importance. The results from the feature importance analysis suggest that the top four influential features included RVI, TPI and TRI and RVI, TPI and Aspect for the wet and dry season, respectively. The RVI and TPI’s contributions to the predictions were consistent across the two seasons. RVI is computed from the VV and VH polarisations of the SAR backscatter and is useful for assessing vegetation structure. For the wet season, grasses grow quickly, resulting in substantial biomass increases. As such, RVI, as computed from the VV and VH polarisations, effectively captures the physiological and structural characteristics of the grasses. For the dry season, even though grasses senescence, RVI remains sensitive to residual biomass and structural characteristics of the vegetation, though the correlation is not as strong as it is during the wet season. These results corroborate the results of Chang *et al.* (2024), who integrated RVI with optical index to estimate biomass in grazing lands across the contiguous United States. They found that RVI exhibited the highest correlation with AGB, and the nonlinear model based on RVI achieved an R^2 of 0.66 with measured dry aboveground biomass. The TRI contributed more among the topographic variables. This could be because the TRI identifies landscape features that influence grass distribution and microclimates Avohou and Sinsin (2009). For example, areas with greater ruggedness might exhibit varying soil depths and water retention abilities, resulting in localised variations in carbon stock.

Meanwhile, TPI affects accessibility, erosion patterns, and habitat complexity. For example, valleys associated with low TPI values usually accumulate water and nutrients, stimulating higher grass biomass and carbon stocks. For the dry season, TPI was less influential since the availability of water decreases, and the influence of topographical factors on biomass growth becomes less significant. Nevertheless, slope was more influential in the dry season. Slope explains the physical processes, such as runoff and erosion that influence the distribution of carbon stocks. Studies by found the opposite. Among the topographic features, including slope, aspect and elevation, aspect has the lowest correlation with the biomass of a forest in this study. The above results highlight that the importance of input variables for AGGCS prediction depended not only on the texture, topographic and vegetation indices or a combination of such variables, but also on seasons.

4.6 Conclusions

In this study, the incorporation of radar indices, topographic variables, and texture matrices coupled with the XGBoost algorithm were tested on whether they can enhance the quantification of AGGCS in the complex savannah ecosystems. Based on the findings, this study concludes that:

- The performance of RVI, texture metrics topographic variables as stand-alone tends to produce weak correlations.
- Integrating RVI, texture, and topographic variables improves the predictive capabilities of the XGBoost model across all seasons
- The variables that optimally captured the variability of AGGCS include RVI (spectral) and TPI and TRI features (Terrain).
- Overall, the integration of Sentinel-1-derived matrices coupled with texture matrices and the XGBoost algorithm shows potential in predicting aboveground grass carbon stock in savannah ecosystems.

4.7 Summary

The Algorithms that are developed and validated within a single year tend to detect anomalies specific to that year. This can result in over-fitting, which subsequently diminishes their performance when applied to data from different years, resulting in over-fitting and underperformance in subsequent years. The last objective of this thesis was to model the effect of seasons on the predictive capability of the XGBoost algorithm. This was done to test the temporal stability and transferability of the models and parameters within the season and across years. The optimised integrated SAR, texture and topographic features were fitted into the XGBoost algorithm for the wet and dry conditions of the years 2020 and 2021. We found that the contribution of SAR, texture and topographic parameters varied between the wet and dry seasons in both years. Notably, the predictive capabilities of the models also differed within the seasons and across the years. Overall, the XGBoost model that integrated all the variables (topographic, texture and SAR indices) showed the highest predictive power in both the wet and dry seasons across both years. Notably, the model developed for the dry season exhibited superior predictive power compared to that of the wet season. The findings underscore the impact of environmental and phenological factors on the predictive power of remote sensing-based models for carbon stock estimations. The last chapter concludes the thesis, situates the research findings within a broader context, and presents recommendations for future research.

CHAPTER 5: INTER-SEASONAL QUANTIFICATION OF ABOVE-GROUND GRASS CARBON STOCK IN SAVANNAH ECOSYSTEMS USING SAR DATA

Abstract

Global warming is one of the most significant global challenges of modern societies, with its impacts spreading across ecological, social, and economic spheres. Co-dominating the savannah landscape, the grass layer can either serve as a net carbon source or sink. As a carbon source, the grass layer generally respires approximately 50% of the Gross Primary Productivity (GPP). However, this rate is elevated to above normal by anthropogenic activities. Thus, quantifying and monitoring terrestrial carbon stocks is essential. This study assessed the impact of seasons on the predictive capabilities of coalesced Synthetic Aperture Radar (SAR) indices, texture matrices and topographic variables in quantifying aboveground grass carbon stocks in savannah ecosystems. Grass biomass samples were collected in the dry and wet seasons of 2020 and 2021 in Kruger National Park (KNP). The study applied on XGBoost machine leaning to investigate relationships between grass samples and predictor variables. The model performance was evaluated using coefficient (R^2) and error measures such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) expressed as a percentage to assess the model performance. The findings indicate that the contribution of SAR-based indices, texture and topographic parameters to the predictive power of the XGBoost model varied between the wet and dry seasons in both years. Notably, the predictive capabilities of the models also differed within the seasons and across the years. Notably, when tested individually, the RVI, $VH_{texture}$ and $VV_{texture}$ moderately (R^2 ranging 35-41, RMSE 23-30%, MAE of 5-18), explained variability in AGGCS for the wet seasons of 2020 and 2021. When testing the combination of all the predictor variables, the predictive capability of the XGBoost improved by over 16% and 20% for the wet seasons of 2020 and 2021, respectively. Overall, the XGBoost model that integrated all the variables (topographic, texture and SAR indices) showed the highest predictive power in the wet and dry seasons across both years. Nevertheless, the model calibrated with the dry season data (R^2 = 0.86, MAE = 2.8, RMSE = 13.30%) explained 61 % of the variation compared to the wet season model. The findings underscore the impact of environmental and phenological factors on the predictive power of remote sensing-based models for carbon stock estimations.

Keywords: Aboveground Grass Carbon Stocks, Savannah Ecosystems, Sentinel-1, Texture, Topographic variables, Machine learning, Inter-annual

5.1 Introduction

Global warming is one of the most significant global challenges of modern societies, with its impacts spreading across ecological, social, and economic spheres. Its effects are global, yet certain countries and ecosystems experience more of the negative impacts associated with this phenomenon. Worldwide, terrestrial ecosystems function as a net sink for atmospheric carbon dioxide (CO_2); however, the extent of this sink is still uncertain (Grace *et al.*, 2006). Savannahs are one of the biomes for the sequestration of atmospheric carbon through gross primary productivity (GPP), contributing over 20% of the global GPP annually (Grace *et al.*, 2006; Beer *et al.*, 2010). This productivity is a product of the interplay between tree layer and grass layer that co-dominate the savannah landscapes (Moore *et al.*, 2016).

Co-dominating the savannah landscape, the grass layer can either serve as a net carbon source or sink. As a carbon source, the grass layer generally respires approximately 50% of the GPP (Ma *et al.*, 2007; Chapin III *et al.*, 2011). In savannah landscapes, the respiration process is shaped by natural processes such as grazing by herbivores (Bristow *et al.*, 2016; Moore *et al.*, 2016) and wildfires (Beringer *et al.*, 2015). The rate at which the grass layer respires carbon back into the atmosphere is elevated to above-normal by anthropogenic activities such as land use change (Moore *et al.*, 2016). For example, records show that CO_2 concentrations have increased from 278 parts per million (ppm) approximately sixty years ago to 420.54 ppm as of April 2023 (Lan, 2023). As such, efforts are needed to normalise the above-normal respiration rates to maintain a balance between grasses as a carbon source or sink in savannah ecosystems.

The status of the grass layer as a carbon source or sink is further shaped by seasons (Xu and Baldocchi, 2004). This is influenced by fluctuations in temperature, precipitation, and environmental factors that shape grass growth, respiration, and the decomposition processes (Xu and Baldocchi, 2004; Ma *et al.*, 2007; Scott *et al.*, 2023; Yang *et al.*, 2025). As such, the inter-annual quantification of aboveground grass carbon stocks (AGGCS) is the first step in shaping our knowledge of the grass layer as a carbon source or sink. This is essential for predicting the ecosystem's response to environmental pressures. The above requires a comprehensive approach that integrates natural processes, sustainable land management, and climate change mitigation strategies.

Nevertheless, few studies track inter-annual variability in AGGCS, which fluctuates by ±15% due to climate shifts (Wang *et al.*, 2024). The escalating pace of climate change underscores the urgency of accurately quantifying carbon stocks in savannah ecosystems. Furthermore, the grass layer plays a crucial role in regulating hydrological processes by influencing water infiltration, runoff, and groundwater recharge. Thus, the health and integrity of the grass layer are pivotal for the overall functioning and resilience of savannah ecosystems, making its conservation and management imperative for biodiversity conservation, sustainable land use, CO_2 mitigation and ecosystem stability. Such precise information is indispensable for various purposes, including efficient carbon accounting and management strategies, conservation efforts aimed at preserving biodiversity, and the promotion of sustainable land use practices.

Conventional remote sensing approaches for estimating AGGCS predominantly depend on optical-based indices such as Normalised Difference Vegetation Index (NDVI) (Berger *et al.*, 2020; Dube *et al.*, 2021; Zumo *et al.*, 2022; Peng *et al.*, 2023; Wu *et al.*, 2023; Zhao *et al.*, 2023; Netsianda and Mhangara, 2025). However, optical-based indices are limited for seasonal monitoring activities. For example, in the wet season, optical sensors are obscured by clouds resulting in data gaps. In addition, most grasses undergo senescence in the dry season which optical-based indices struggle to detect. Furthermore, heterogeneously configured savannahs limits the effective application of optical-based-indices (Abdi *et al.*, 2022). On the contrary, Radar sensors are insensitive to cloud (Ghasemi *et al.*, 2011), a strength which is valuable for seasonal landscape monitoring in tropical savannahs. Despite their utility, both Optical and Synthetic Aperture Radar (SAR) sensors exhibit a shared limitation in observational capacity, as they tend to saturate at specific maximum biomass thresholds, which are approximately 350–400 tons per hectare for C-band sensors (Kasischke *et al.*, 2002) and around ± 0.3 g/cm² for multispectral bands (Mutanga *et al.*, 2012).

To deal with these limitations, several studies have employed a multi-parameter coalescing approach (Kuplich *et al.*, 2005; Cutler *et al.*, 2012; Liu *et al.*, 2019; Yao and Ren, 2024; Maake *et al.*, 2025). While the majority of these investigations were carried out in forest environments (Holmström and Fransson, 2003; Goh *et al.*, 2013; Dusseux *et al.*, 2014). There is a scarcity of research focussing on grasses in savannah ecosystems (Hong and Wdowinski, 2014). Therefore, it is equally important to explore the potential of Synthetic Aperture Radar

(SAR) data with additional parameters, such as texture and topography, for the estimation of carbon stocks in savannah ecosystems.

Attempts have been made to characterise the seasonal variations of AGGCS with a single SAR pass in the peak growing season with poor results (Kuplich *et al.*, 2005; Shang *et al.*, 2009). However, most of these studies where targeting forests (Santos *et al.*, 2003; Mitchard *et al.*, 2009; Kumar *et al.*, 2015) (Santos *et al.*, 2003; Mitchard *et al.*, 2009; Kumar *et al.*, 2015; Omar *et al.*, 2017). Meanwhile, studies that attempted to characterize temporal dynamics of AGGCS using SAR data were conducted outside the sub-Saharan Africa with different environmental settings (Schmidt *et al.*, 2016; Ali *et al.*, 2017; Rapiya *et al.*, 2023). With the exception of Rapiya *et al.* (2023), these studies have not attempted to quantify temporal dynamics of grass stocks across inter-annual wet and dry seasons. Furthermore, even Rapiya *et al.* (2023)'s study was conducted along a climate, soil and topographic gradient that significantly differs from the conditions of the savannah examined in this study. As such, the capability of Sentinel-1 for quantifying the seasonal variability of carbon stocks across the wet and dry season in savannah ecosystems is yet to be fully explored. The primary objective of this study is to assess the impacts of seasons on the predictive performance of coalesced SAR indices, texture and topographic variables in quantifying AGGCS in Savannah ecosystems.

5.2 Materials and methods

5.2.1 Experimental site description

An outline and description of the study site is presented in subsection 1.6

5.2.2 Field above-ground grass carbon stock sampling

A grass biomass sampling campaign was conducted from in the wet and dry seasons of 2020 and 2021 respectively. A thorough outline of how the sites and plots were demarcated, as well as how the samples were collected and processed, is presented in subsection 1.7. The conditions of the study site during the period investigated are presented in Figure 5.1

Figure 5. 1 Selected photos showing the condition of the study site during the wet and dry seasons (a) the wet season of 2020 (b) the dry season of 2020 (c) the wet season of 2021 (d) the dry season of 2021

5.3 Image acquisition & Pre-processing

5.3.1 Sentinel-1 data acquisition

Sentinel-1 images were collected for the wet and dry seasons of 2020 and 2021 respectively. Specifically, two images were downloaded from the Copernicus Open Access Hub, accessible at https://scihub.copernicus.eu/dhus/. This means two in February (wet season) 2020 and 2021, and two in July (dry season) 2020 and 2021 respectively. A thorough outline of how the images were collected are pre-processed is presented in subsection 1.8. Figure 5.2 presents the overall workflow followed to retrieve AGGCS using Sentinel-1 SAR data, topographic variables, and field biomass samples. As shown in Figure 5.2, the process began with the acquisition and pre-processing of SRTM DEM and Sentinel-1 images for both wet and dry seasons of the years 2020-2021, including subsetting, orbit correction, speckle filtering, and conversion to decibels. Topographic variables and SAR-derived features (bands, textures, and indices) are then computed and spatially extracted to match the locations of the field biomass samples. Pearson correlation is used to identify relevant predictor variables. The extracted datasets were then subjected to a 10-fold cross-validation method, where part of the data is used for model training (XGBoost) and the remainder for model testing. The Model performance was evaluated using R^2, RMSE (expressed as percentage), and MAE. The process ended with the spatial prediction of AGGCS.

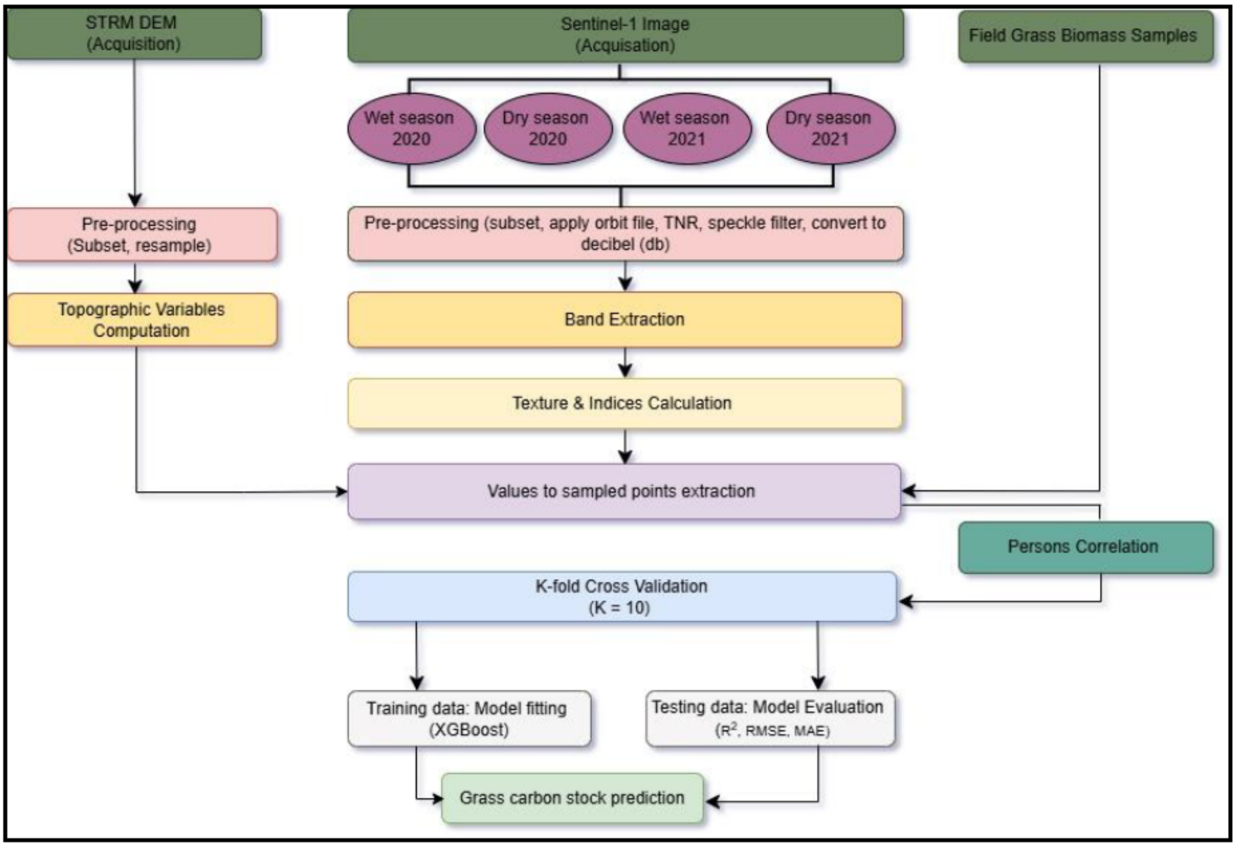

Figure 5. 2 A diagrammatic representation of the Sentinel-1 pre-processing workflow adopted in this study

5.3.2 Digital Elevation Model (DEM) Acquisition

The Shuttle Radar Topography Mission's (SRTM) Digital Elevation Model (DEM) with a spatial resolution of 30 m was downloaded from the USGS. The SRTM mission offers high-resolution DEM on a near-global scale (Yang *et al.*, 2011). As shown in Figure 4, the image was first resampled to a 10 m spatial resolution to align it with the spatial resolution of Sentinel-1. We then extracted topographic variables (Table 5.1) using the ArcGIS 10.8 software. The choice of incorporating DEM was informed by previous studies that demonstrate acceptable accuracies when incorporated with SAR data (Shao and Zhang, 2016; Musthafa and Singh, 2022)

5.3.3 Deriving predictor variables

Based on the Sentinel-1 VH and VV channels, we calculated the Radar Vegetation Index (RVI) and 20 texture matrices following the Gray-level co-occurrence matrix (GLCM) method (Table 5.1). As in Haralick *et al.* (2007), the GLCM parameters were computed in a 3 × 3 processing window. Based on the DEM, we extracted and calculated 7 topographic variables.

Table 5. 1 A selection of variables used in estimating above-ground grass carbon stock

Predictor variable	Expression	Description
Radar Indices (RI)	RVI	Radar Vegetation Index
Texture Matrix	VH/VV_{Con}	Contrast
	VH/VV_{Dis}	Dissimilarity
	VH/VV_{Hom}	Homogeneity
	VH/VV_{ASM}	Angular Second Moment
	VH/VV_{Ene}	Energy
	VH/VV_{Max}	Maximum Probability
	VH/VV_{Ent}	Entropy
	VH/VV_{Mea}	GLCM Mean
	VH/VV_{Var}	GLCM Variance
	VH/VV_{Cor}	GLCM Correlation
Topographic Matrices	H	Elevation
	Slope	Slope
	Aspect	Aspect
	Hillshade	Hillshade
	TRI	Topographic Roughness Index
	TWI	Topographic Wetness Index
	TPI	Topographic Position Index

5.3.1 Regression algorithms

The reliability of regression tasks is often based on several assumptions, including linearity (Daoud, 2017). Predicting AGGCS is often associated with nonlinear relationships among predictors, which require algorithms with capabilities to detect such relationships. This study utilised Extreme Gradient Boosting (XGBoost) to explore the relationship, if any, between field-measured grass carbon stock and the selected predictor variables. Compared to other gradient boosting frameworks such as AdaBoosts, XGBoost offers extra functionalities, making it more appealing for regression and classification tasks. Some of these additional functions include regularization for better generalization, feature importance analysis, sparse data handling, diverse tasks handling, Central processing Unit (CPU) Acceleration and large dataset handling. Furthermore, XGBoost is configured to improve the effectiveness and efficiency of predictions by merging several weak learners (decision trees) into a robust learner (Chen and Guestrin, 2016). A grid search method is employed to select the ideal parameter, where the hyperparameters are selected based on the lowest Root Mean Squared Error (RMSE) from a k-fold cross-validation (CV) resampling strategy. Chen and Guestrin (2016) Chen and Guestrin (2016) provide a comprehensive description of how XGBoost works. Several studies (Kganyago *et al.*, 2022) obtained promising results when utilising this model.

Fitting a model with multiple predictors, particularly radar indices, texture and terrain variables, is often associated with multicollinearity. To deal with this, we subject the predictor variables to a pairwise Pearson correlation. We then set a threshold of $R \geq 0.8$ as a condition to filter variables to fit the XGBoost model. The advantages of eliminating redundant features from a model are illustrated in the work of Luo *et al.* (2022). The findings of their study demonstrated that the removal of redundant parameters enhanced the accuracy of their AGB prediction models in comparison to scenarios where feature parameter selection was not implemented.

5.3.2 Optimisation of Predictor Variables

To improve the precision and computational efficiency of the model, a method by Yao and Ren (2024) was adopted and furthermore, 7 variables were examined by constructing various combinations of these variables, to assess the predictive power and ultimately select the optimal subset that demonstrated the highest predictive power for developing the distribution maps and scatterplots.

5.3.3. Model evaluation

To evaluate how the XGBoost algorithm performs, the 10-fold CV method with 10 iterative validation steps was utilised. In this workflow, the parameter was set to 10 iterations (i.e., validation steps). Each iteration trains the model on 90% of the data and tests it on the remaining 10%, rotating systematically through all 10-folds. The choice of 10 iterations is based on good performance obtained from previous studies (Kganyago *et al.* 2022). This validation method enhances the reliability of assessing a model's performance by fitting the model several times using different training and testing sets each time (Zhang and Liu, 2023), therefore helping to reduce the impact of variability in the data (Wong, 2015). Furthermore, this method is valuable when working with limited samples. The predictor variables listed in Table 5.1 were utilised in the regression analysis. The analysis was performed at multiple experimental levels (see Table 5.3), based on different configurations of the predictor variables to find the optimal Sentinel-1-derived variable suitable for estimating above-ground grass biomass in savannah ecosystems.

To ensure the accuracy and reliability of the model, we computed the coefficient of determination (R^2), the root mean squared error (RMSE) expressed as a percentage, and the Mean Absolute Error (MAE) (James, 2013) to evaluate the performance of the XGBoost. R^2 ranges between 0 and 1. An R^2 value of 1 indicates that the model perfectly explains the variance in the dependent variable, while a value of 0 indicates that the model does not explain any of the variance. These metrics were calculated based on the Equations (5.9) – (5.12) as indicated below.

$$R^2 = 1 - \frac{\sum(Y_{measured} - Y_{predicted})^2}{\sum(Y_{measured} - Y_{predicted})^2} \quad (5.9)$$

$$RMSE = \frac{\sqrt{\sum_{i-1}^{N} - (Y_{measured} - Y_{predicted})^2}}{n} \quad (5.10)$$

$$RMSE\% = \frac{\sqrt{\frac{1}{n}\sum_{i-N}^{N} - (Y_{measured} - Y_{\ predicted})^2}}{\bar{Y}} \times 100 \quad (5.11)$$

$$MAE = \sum_{n}^{1} \frac{|Y_{measured} - Y_{predicted}|}{n} \quad (5.12)$$

Where $Y_{measured}$ is the observed AGGCS, $Y_{predicted}$ is the predicted AGGCS, *n* is number of observed samples, i is the included predictor variable and $\bar{Y}$ is the mean of the observed

AGGCS. The R^2, RMSE and RMSE% equations are adopted from (Dube and Mutanga, 2015), while the MAE is adopted from (Qadeer *et al.*, 2024).

5. 4 Results and discussion

5.4.1 Descriptive statistics of observed AGGCS

Table 5.2 contains the descriptive statistics of the observed AGGCS across the wet and dry seasons of 2020 and 2021, respectively. For the wet season of 2020 and 2021, the observed AGGCS ranged between a minimum of 4.5 and 8.98 grams per square metre (g m^{-2}), with the maximum ranging between 58.2 g m^{-2} and 69.98. The average was 28.73 and 26.89 for the wet seasons of 2020 and 2021, respectively. The standard deviation for the wet season was 11.98 and 11.26 g $m^{-2,}$ across the years 2020 and 2021, respectively. In contrast, the dry season recorded minimums of 11.97 and 13.73 and maximums of 99.17 and 90.11 for the years 2020and 2021, respectively. The standard deviation for the dry season ranged between 18.08 and 16.66 across 2020 and 2021, respectively

Table 5. 2 Descriptive statistics of measured aboveground grass carbon stock in Kruger National Park

Season	Variable	Samples	Min	Max	Mean	StDev[1]
Wet (2020)	Dry AGGCS (g m^{-2})	94	4.5	58.2	28.73	11.98
Dry (2020)	Dry AGGCS (g m^{-2})	84	11.97	99.17	34.32	18.08
Wet (2021)	Dry AGGCS (g m^{-2})	90	8.98	69.98	26.89	11.26
Dry (2021)	Dry AGGCS (g m^{-2})	91	13.73	90.11	36.51	16.66

[1]StDev = Standard Deviation

5.4.2 Predictor Selection and Optimisation for AGGCS Prediction

Figure 5.3 shows a matrix of the correlation between RVI, texture matrices and topographic variables. As indicated in Figure 5, 24 of the 33 variables were highly correlated (R = ≥ 0.8). Specifically, the RVI exhibited low correlations. Among the 20 texture variables, VH_{Cor}, VV_{Cor} demonstrated low correlations whereas Slope, Hill shade, TRI, TPI and TWI demonstrated low correlations among the 7 topographic variables. Overall, a total of 7 variables, including RVI, VH_{Cor},VV_{Cor}, Slope, Hill shade, TRI, TPI and TWI were used in subsequent AGGCS predictions.

	Carbon	VH	VV	RVI	Vh_Contra	Vh_Dissi	Vh_Homo	Vh_Asm	Vh_Energy	Vh_Max	Vh_Entro	Vh_Mean	Vh_Var	Vh_Corr	Vv_Contra	Vv_Dissi	Vv_Homo	Vv_Asm	Vv_Energy	Vv_Max	Vv_Entro	Vv_Mean	Vv_Var	Vv_Corr	Elev	Slope	Aspect	Hillshade	Tri	Twi	Tpi
Carbon	1																														
VH	0,03435	1																													
VV	-0,04027	0,58629	1																												
RVI	0,06014	0,53815	-0,35008	1																											
Vh_Contra	0,04368	0,18177	0,16586	0,03258	1																										
Vh_Dissi	0,00067	0,29151	0,22551	0,10612	0,95314	1																									
Vh_Homo	0,02622	-0,40565	-0,22675	-0,24442	-0,58584	-0,77437	1																								
Vh_Asm	-0,07268	-0,40945	-0,14654	-0,34408	-0,25364	-0,39448	0,72435	1																							
Vh_Energy	-0,07525	-0,44896	-0,16973	-0,36066	-0,36419	-0,52957	0,84737	0,96085	1																						
Vh_Max	-0,09562	-0,4382	-0,18317	-0,33298	-0,33057	-0,48393	0,80752	0,96299	0,97877	1																					
Vh_Entro	0,06465	0,4539	0,17096	0,36229	0,44587	0,62062	-0,9	-0,87542	-0,97309	-0,92344	1																				
Vh_Mean	0,0016	0,96895	0,63916	0,43641	0,17879	0,26154	-0,31459	-0,28066	-0,3294	-0,32197	0,3453	1																			
Vh_Var	-0,03018	0,92007	0,63049	0,38691	0,1023	0,15522	-0,18263	-0,16666	-0,19268	-0,20034	0,19692	0,97145	1																		
Vh_Corr	0,00738	0,27314	0,09617	0,24924	0,08526	0,16035	-0,41493	-0,86938	-0,71635	-0,74676	0,5668	0,15776	0,10323	1																	
Vv_Contra	0,09514	0,05442	0,17323	-0,10954	0,40781	0,33073	-0,14235	-0,11407	-0,13642	-0,13103	0,15316	0,0693	0,05024	0,07982	1																
Vv_Dissi	0,04706	0,07534	0,23885	-0,12694	0,33364	0,2885	-0,14019	-0,13417	-0,15334	-0,13607	0,1708	0,09229	0,06321	0,10949	0,93742	1															
Vv_Homo	-0,02949	-0,06742	-0,30447	0,17757	-0,18096	-0,18609	0,10743	0,08703	0,10639	0,07502	-0,13	-0,09323	-0,06451	-0,0702	-0,50993	-0,73824	1														
Vv_Asm	-0,02437	-0,05217	-0,29254	0,18891	-0,11799	-0,11066	0,03555	-0,00112	0,01061	-0,00711	-0,0287	-0,10097	-0,08902	0,00647	-0,2412	-0,42134	0,01923	1													
Vv_Energy	-0,04205	-0,08148	-0,29663	0,15811	-0,14642	-0,1488	0,07498	0,03053	0,04923	0,02162	-0,07526	-0,12146	-0,1035	-0,0114	-0,3083	-0,50466	0,88125	0,98165	1												
Vv_Max	-0,04247	-0,03792	-0,27394	0,18795	-0,13225	-0,12882	0,0477	0,00609	0,02104	-0,00106	-0,04268	-0,07828	-0,06107	0,00635	-0,28318	-0,47138	0,86436	0,98099	0,98962	1											
Vv_Entro	0,05924	0,12379	0,30962	-0,12197	0,17054	0,18476	-0,12043	-0,07369	-0,09692	-0,06093	0,12838	0,15432	0,13092	0,04236	0,37762	0,58372	-0,91131	-0,92771	-0,97963	-0,95055	1										
Vv_Mean	-0,07498	0,65812	0,95924	-0,2181	0,20369	0,25927	-0,24753	-0,18171	-0,20418	-0,21955	0,20259	0,70041	0,69888	0,12742	0,17718	0,21959	-0,23158	-0,19495	-0,2031	-0,17595	0,22084	1									
Vv_Var	-0,10618	0,65012	0,90199	-0,16776	0,15797	0,21573	-0,21926	-0,15179	-0,17102	-0,19171	0,16455	0,69858	0,71913	0,10014	0,11763	0,11339	-0,06531	-0,04906	-0,04433	-0,02202	0,05486	0,9678	1								
Vv_Corr	0,0347	0,02912	0,26179	-0,20736	0,02314	0,00276	0,04209	0,03704	0,05562	0,03912	-0,07002	0,0864	0,08523	-0,01074	0,08854	0,17353	-0,40689	-0,65243	-0,56787	-0,56351	0,5035	0,16419	0,10433	1							
Elev	-0,26542	-0,25396	0,02612	-0,30852	-0,10548	-0,13404	0,15907	0,22485	0,25896	0,2247	-0,27918	-0,21265	-0,12285	-0,13202	-0,00345	-0,02544	-0,04811	-0,18486	-0,14057	-0,17573	0,07973	-0,01026	0,00701	0,15223	1						
Slope	-0,20087	0,07589	0,08691	-0,0077	-0,16143	-0,17603	0,14223	0,13619	0,1309	0,1317	-0,12348	0,13586	0,20724	-0,10258	0,07069	0,07042	-0,09636	-0,14561	-0,13288	-0,14148	0,09983	0,10205	0,10905	0,00363	0,2652	1					
Aspect	-0,13487	0,08295	0,17982	-0,0797	0,14465	0,13192	-0,12834	-0,12451	-0,14456	-0,1382	0,14455	0,10482	0,13783	0,06676	0,0054	-0,01438	0,08917	0,09386	0,11844	0,12774	-0,12725	0,23351	0,27946	0,10844	0,1078	0,15851	1				
Hillshade	-0,22059	0,05679	0,09322	-0,00413	0,17633	0,21171	-0,26836	-0,16985	-0,20654	-0,19437	0,23218	0,07626	0,03516	0,09628	0,05717	0,05406	-0,06165	-0,07694	-0,05189	-0,05713	0,02752	0,0838	0,0829	0,09629	0,1399	-0,1198	0,22263	1			
Tri	-0,04406	-0,16326	-0,02495	-0,13951	-0,0981	-0,11974	0,05621	0,00453	0,02506	0,02374	-0,03417	-0,14597	-0,09451	0,05198	0,02902	0,06519	-0,17616	-0,19545	-0,20494	-0,20615	0,19944	-0,0296	-0,07852	-0,07465	0,1775	0,14705	-0,01908	0,01627	1		
Twi	-0,11683	-0,21354	-0,33337	0,0934	-0,0114	-0,0105	-0,01739	-0,00846	0,00897	-0,00023	-0,02665	-0,25278	-0,27695	0,0656	-0,11663	-0,16728	0,32911	0,43117	0,436	0,42958	-0,41892	-0,29138	-0,23995	-0,15587	-0,106	-0,4034	0,00272	0,12469	-0,19	1	
Tpi	0,00392	0,06168	0,1002	-0,04842	-0,05319	-0,05342	0,01929	-0,06308	-0,02252	-0,05672	-0,00924	0,06561	0,11435	0,1082	-0,05066	-0,05794	-0,00192	-0,0363	-0,03621	-0,02624	0,04045	0,1253	0,11515	0,0556	-0,031	0,12899	-0,03868	-0,252	0,31	-0,48	1

-1

1

Figure 5. 3 A correlation coefficient matrix between the predictor variables (i.e. RVI, texture matrices and topographic variables)

5.4.3 Predictive performance of parameters

Table 5.3 provides a summary of the predictive capabilities of the radar indices, texture and topographic variables tested in this study to predict AGGCS. When tested individually, the RVI, $VH_{texture}$ and $VV_{texture}$ moderately (R^2 ranging 35-41, RMSE% 23-30, MAE of 5-18), explained variability in AGGCS for the wet seasons of 2020 and 2021, therefore contributing almost similarly to the model. In contrast, the topographic variables yielded lower (R^2 values between 0.25-0.28, RMSE% 30 and MAE) predictive power. Meanwhile, the XGBoost results obtained for the dry seasons show similar trends with the RVI, $VH_{texture}$ and $VV_{texture}$ moderately (R^2 ranging 40-45, RMSE% 7.6 -30, MAE 4.35-12.87) explaining the variability in AGGCS. When testing the combination of all the predictor variables, the predictive capability of the XGBoost improved by over 16% and 20% for the wet seasons of 2020 and 2021, respectively. In contrast, the XGBoost results obtained for the dry seasons improved by approximately 25% and 35% for the dry seasons of 2020 and 2021, respectively. Overall, the XGBoost model that integrated all the variables (topographic, texture and SAR indices) showed the highest predictive power in the wet and dry seasons across both years.

Table 5. 3 Summary of combinations for predicting aboveground grass carbon stock in Kruger National Park

Temporal Window	Predictor Variable	R^2	RMSE (g m^{-2})	RMSE%	MAE
Wet season (2020)					
	RVI	0.39	11.38	30	7.02
	$VH_{Texture}$	0.40	23.43	30.84	5.32
	$VV_{Texture}$	0.39	24.22	30.88	5.21
	Topo	0.37	17.20	40.55	9.71
	RVI + Tex + Topo	0.56	12.46	28.05	4.72
	Opt Var	**0.65**	**7.41**	**25.49**	**3.76**
Dry season (2020)					
	RVI	0.40	7.20	30	5.32
	$VH_{Texture}$	0.43	17.35	30	5.32
	$VV_{Texture}$	0.42	16.60	30	5.32
	Topo	0.51	11.89	49.24	0.14
	RVI + Tex + Topo	0.61	11.20	24.66	4.86
	Opt Var	**0.78**	**4.58**	**15.95**	**3.23**
Wet season (2021)					
	RVI	0.41	7.11	30.84	8.95
	$VH_{Texture}$	0.39	23.40	30.96	18.35
	$VV_{Texture}$	0.37	24.4	40.25	18.12
	Topo	0.28	23.40	43.49	18.13
	RVI + Tex + Topo	0.60	0.02	32.7	3.38
	Opt Var	**0.67**	**5.19**	**19.29**	**2.92**
Dry season (2021)					
	RVI	0.45	7.16	30.70	4.35
	$VH_{Texture}$	0.44	17.24	30.40	12.87
	$VV_{Texture}$	0.43	16.55	40.25	12.58
	Topo	0.28	0.04	40.95	6.31
	RVI + Tex + Topo	0.75	0.03	18.76	4.82
	Opt Var	**0.86**	**4.79**	**13.30**	**2.81**

[1]RVI = Radar Vegetation Index; $VH_{tetxure}$ = Vertical Transmit and Horizontal Receive Texture, $VV_{tetxure}$ = Vertical Transmit and Vertical Receive Texture; Topo = Topographic variables, Opt Var = Optimised Variables

5.4.4 Important variable

The XGBoost algorithm has an add-in function to assess the variable importance. This function assists in quantifying the contribution that each predictor variable adds towards the model's performance using the "gain", "frequency", and "cover" as indicators (Hastie *et al.*, 2009). Figure. 5.4 illustrate the selected variables that contributed the most in predicting AGGCS as measured by their Gain. For the wet season of 2020, the AGGCS was optimally predicted with RVI, TRI, TPI, Slope, Aspect, TWI, VH_{cor}, VV_{cor} and Hill hade in the order of importance. Meanwhile, RVI, TRI, Slope, Aspect, TWI, TPI, Lat, VV_{cor}, VH_{cor}, and Hill shade optimally predicted AGGCS in the dry season of 2020 in the order of importance. For the wet season of 2021, the most influential variables included VH_{co}r, VV_{co}r, Aspect, Lat, Slope, RVI, TWI, TPI, TRI and Hill shade in the order of importance. Contrarily, RVI, Slope, TRI, Aspect, VH_{cor}, TPI, TWI, Lat, VV_{cor} and Hill shade optimally predicted AGGCS in the dry season of 2021, respectively. Overall, AGGCS was optimally predicted based on RVI, Slope, TRI, Aspect, VH_{cor}, TPI, TWI, Lat, VV_{cor} and Hill shade.

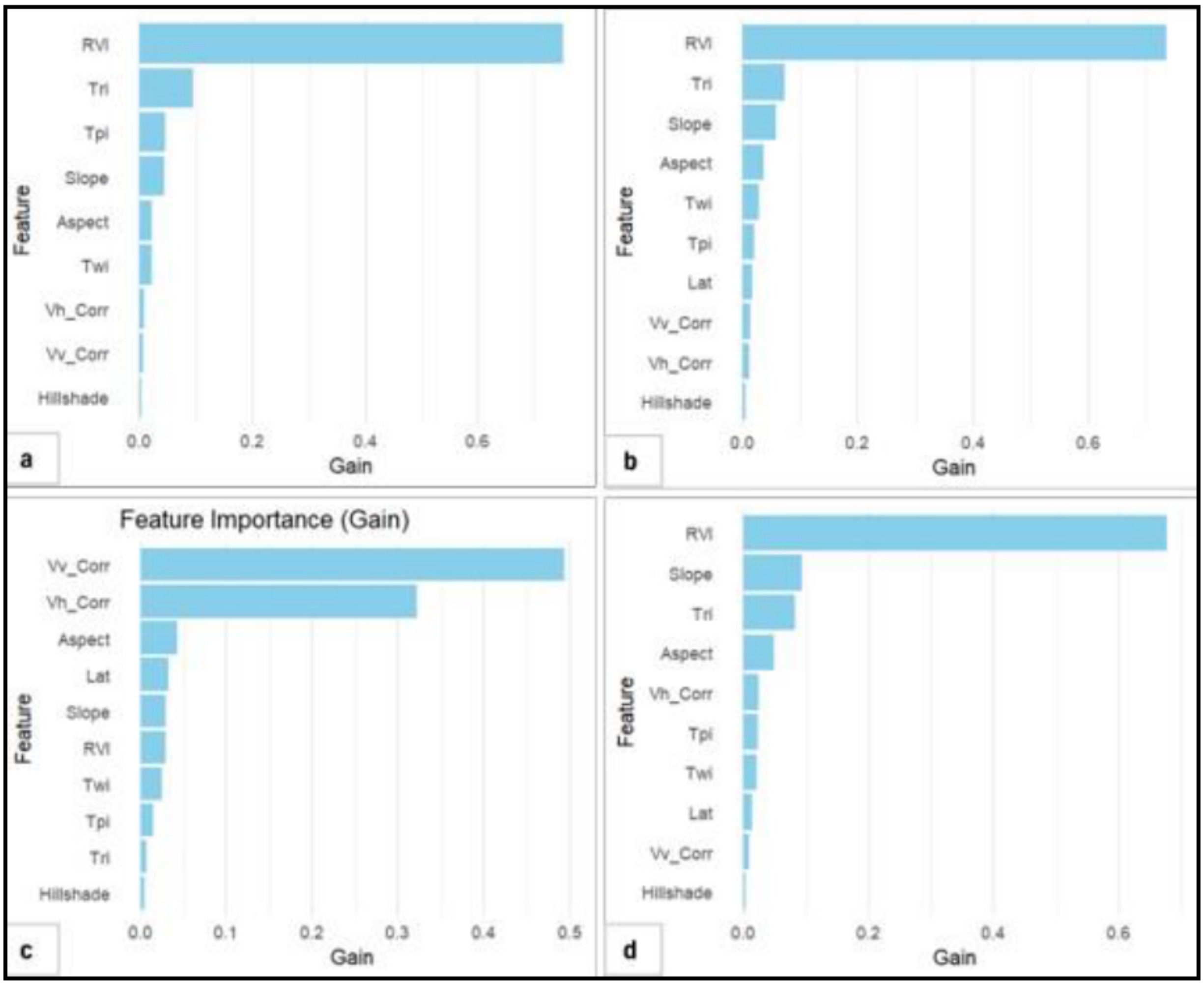

Figure 5. 4 Variable importance scores of selected predictors utilised in predicting AGGCS: a) 2020 wet season, b) 2020 dry season, c) 2021 wet season, and d) 2021 dry season.

5.4.5 Performance of optimised models by feature relevance

With "Gain" as an indicator of the predictor variable contribution to the predictive power of the XGBoost algorithm, the results from the variable importance revealed 4 variables that contributed more to the predictive power of the XGBoost algorithm across the wet and dry seasons of 2020 and 2021 consecutively. These 4 variables were then used to develop models for the 4 seasons. Figure 5.5 shows the scatter plots of models fitted with the top 4 variables ranked in order of importance. Specifically, the RVI, TRI, TPI, and Slope and VV_{cor}, VH_{corr}, Aspect and Lat for the wet season of 2020 and 2021, respectively. Meanwhile, the RVI, TRI, Slope, and Aspect and for the dry season. As presented in Table 5.3, the model developed for the dry season of 2020 and 2021 exhibited superior predictive capability, with a coefficient of determination (R^2) of 0.78 and 0.86, a root mean square error (RMSE) of 4.5 and 4.9, RMSE% of 15.95 and 13.30, and a mean absolute error (MAE) of 17.12 and 2.81 respectively, in contrast to the wet season model, which recorded an R^2 of 0.65 and 0.67, an RMSE of 7.41 and 5.19, a percentage RMSE of 25.49 and 25.49, and an MAE of 3.76 and 3.23. Overall, there was a notable enhancement in the predictive performance of the models based on the selected optimal parameters across both the wet and dry seasons of 2020 and 2021, respectively.

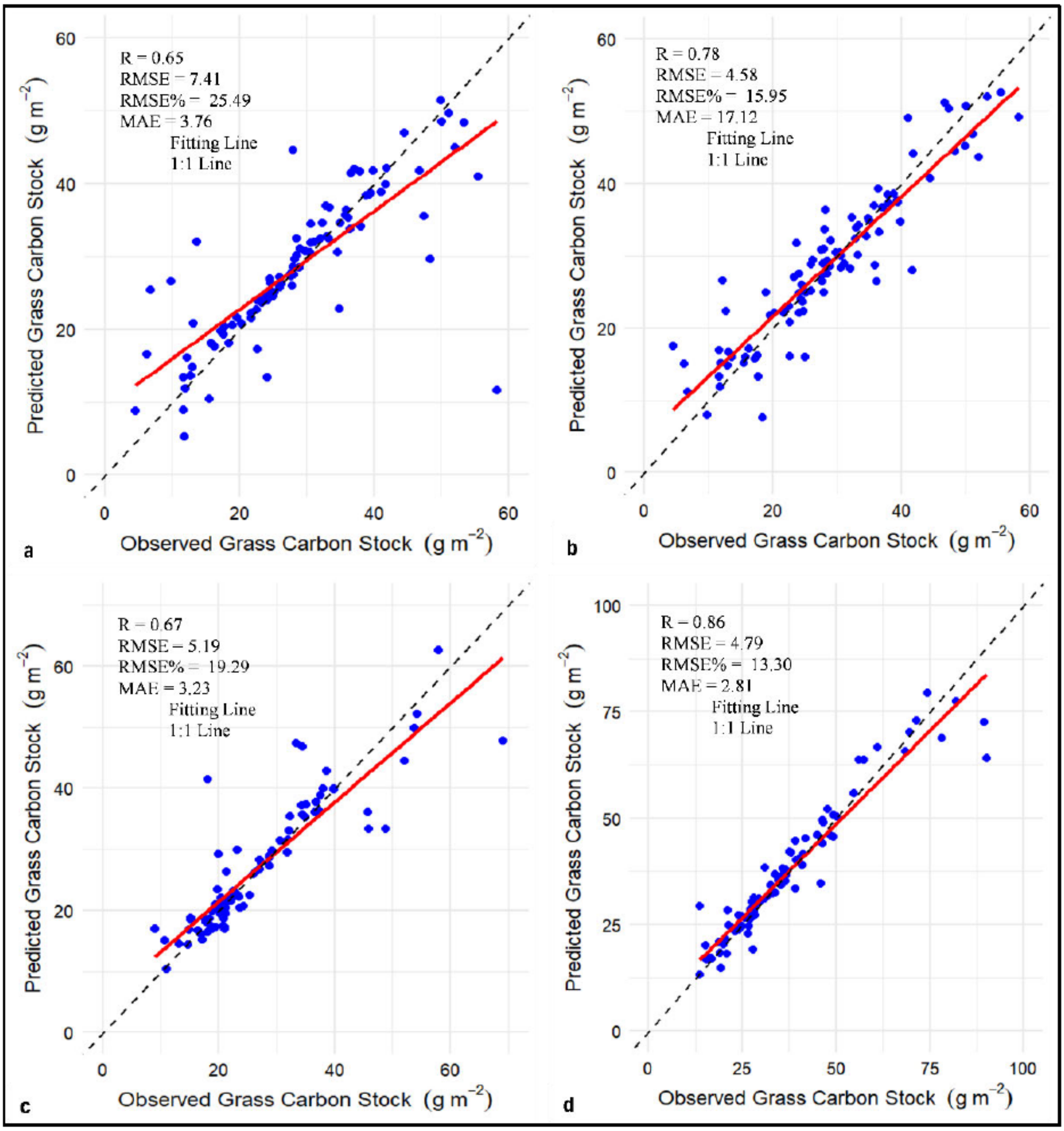

Figure 5. 5 Scatterplots showing the relationship between observed and predicted AGGCS based on the XGBoost model integrating optimal variables: a) 2020 wet season, b) 2020 dry season, c) 2021 wet season, and d) 2021 dry season.

5.4.6 Spatial distribution maps for optimal predictor combinations and XGBoost model

The spatial distribution of AGGCS was mapped utilising the optimal models across the two wet and dry seasons of 2020 and 2021, respectively. Figure 5.6 illustrates the spatial distribution of the estimated AGGCS during these seasons across the study site. The spatial distribution of AGGCS across the four models demonstrated a high degree of variability. All models indicated highest AGGCS values in the dry seasons of 2020 and 2021 western part of the study site, with AGB values ranging from 186 to 357 g.m². Conversely, the lowest AGB values were recorded in wet seasons, where values ranged from 17 to 110 g.m². The results obtained from the XGBoost analysis indicated a consistently lower value across the seasons. In contrast, the underestimations observed in the models were predominantly evident in the high AGGCS regions of the south of Kruger National Park. This assessment is further supported by the ensemble standard deviation of estimates derived from the six machine learning models employed in the study. In the wet season, AGGCS covers large parts of the study site, especially on plains and open grasslands, compared to the dry season, where patches of high AGGCS are restricted to isolated areas.

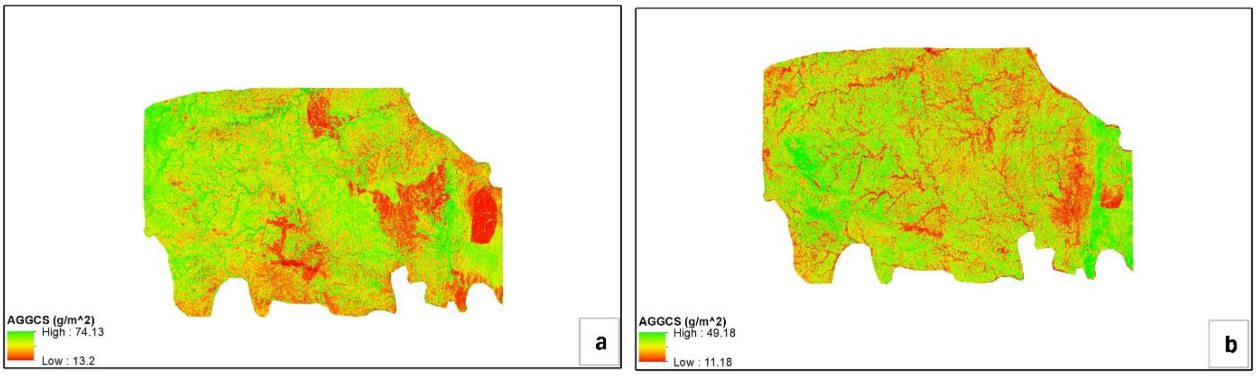

Figure 5. 6 Spatial distribution of predicted AGGCS based on optimal models: a) 2020 dry season, b) 2021 dry season.

5.5 Discussion

In this study, a combination of selected parameters, including Radar Indices, texture matrix and terrain variables, where regressed with observed field AGGCS samples using the XGBoost ensemble algorithm. Particularly, we aimed to assess the impact of seasons on the sensitivity of SAR indices coupled with a texture matrix and topographic variables for quantifying AGGCS in savannah ecosystems. The findings show that among the variables utilised in the study, RVI and TRI consistently contributed the most in modelling AGGCS across the wet and dry seasons of 2020 and the dry season of 2021, respectively. All models indicated highest AGGCS values in the dry seasons of in western part of the study site, with AGB values ranging from 186 to 357 $g.m^2$ and the lowest AGB values in wet seasons, where values ranged from 17 to 110 $g.m^2$. The findings of the XGBoost models based on the combination of all the variables outperformed those that employed a single-variable model. Overall, the models based on the dry season data consistently outperformed the models fitted with the wet-season data within and across the years under investigation.

5.5.1 Comparing predictor variables in quantifying AGGCS

The findings of this study show that different combinations of predictors (i.e. radar indices, texture, and topography) contributed differently across the wet and dry seasons of 2020 and 2021, respectively. This observation was expected because the wet and dry seasons consist of distinct environmental conditions, and this variability tends to be more or less in certain years than others. For example, when quantifying AGGCS in the wet season, the top four predictor variables that contributed the most to the predictive power of the XGBoost algorithm included RVI, TRI, TPI and Slope for 2020, whereas $VH_{cor,}$ VV_{cor}, Aspect and latitude were the most influential for 2021. The same was observed for the dry season of both years, where RVI, TRI, Slope and Aspect were the most influential predictor variables for 2020, and RVI, TRI, Aspect and Slope contributed the most for 2021. This could be primarily due to the unique contextual factors and processes occurring in those years, which shape the relevance and impact of those predictors. Another possible explanation could be soil moisture retention in specific geologies or soil. The findings of this study corroborate with those of Kanja *et al.* (2025) , who evaluated multi-seasonal SAR and optical imagery for above-ground biomass estimation using the national forest inventory of Zambia reported various contributing factors amongst polarisation channels, texture and spectral variables across the rainy, hot season and annual. In another study utilising seasonal Landsat 8 images for forest aboveground biomass estimation, Li *et al.* (2019), reported that the selected spectral variables used for AGB models of different months and different forest types varied across species.

Nevertheless, among the variables utilised in the study, RVI and TRI consistently contributed the most in modelling AGGCS across the wet and dry seasons of 2020 and the dry season of 2021, respectively. For RVI, this could be attributed to a correlation with various aspects of grass structure, such as leaf area index and plant height, as well as with biochemical properties, including lignin and cellulose (Wu *et al.*, 2024). For example, in the wet season, RVI avoids saturation issues common in dense vegetation by using a ratio of co-polarised (VV) to cross-polarised (VH), which remains responsive to high biomass levels (Mandal *et al.*, 2020). In contrast, in the dry season, RVI leverages its sensitivity to vegetation structures rather than greenness, enabling biomass estimation even when grass is senescent or dormant (Li and Guo, 2017). Studies in semi-arid grasslands show RVI accurately estimates dry biomass by mitigating soil background interference (Wu *et al.*, 2024).

Nevertheless, RVI moderately explained variability in AGGCS across the seasons. This could be because the C-band has a shorter wavelength, which tends to be limited when mapping a wider range of biomass values (Chang *et al.*, 2024). Nonetheless, active sensors capturing images within the L and P spectrum have been commissioned to provide an improved image bandwidth that can capture a wider range of AGGCS values (Santi *et al.*, 2017). These sensors are capable of characterising intra-canopy features, thereby enabling them to effectively address issues related to saturation (Chang *et al.*, 2022; Mutanga *et al.*, 2023). The consistent contribution of the TPI to the predictive power of the XGBoost could be because it plays a crucial role in elucidating the factors that govern the spatial distribution of carbon stocks (Römkens *et al.*, 2002). TRI reflects physical landscape complexity; it correlates with vegetation structural attributes, such as canopy height and density, which exhibit minimal variation between wet and dry seasons. This structural aspect of biomass remains relatively stable and thus TRI serves as a reliable proxy for biomass estimation year-round (Anees *et al.*, 2024).

5.5.2 The Predictive Capabilities of the Optimised AGGCS Models

This study evaluated the capability of three groups of predictors (i.e., SAR indices, Texture matrices and topographic variables) when used alone or in combination for AGGCS quantification. The results revealed low performance when the XGBoost algorithm was fitted with these variables independently, with the XGBoost model based on the topographic variables performing the lowest. This could be explained by their intrinsic weaknesses as standalone predictors. For example, RVI is influenced by soil and vegetation moisture, which is prone to noise in wet seasons. Texture, on the other hand, tends to depend on scale; the 3 x 3 window sizes utilised in this study could have been coarse to enable consistent AGGCS across the models. The predictive power of the topographic-based XGBoost was the lowest of all the models, and this could be because the topographic indicators are derived from the STRM DEM with a spatial resolution of 30 m, which is coarser than the spatial resolution of the Radar indices and texture matrices. However, the XGBoost models based on the combination of all the variables outperformed those that employed a single-variable model. The results revealed a further improvement when the XGBoost model was fitted with the optimised combination of parameters. This is supported by higher R^2 values and lower RMSE values, which indicate that they fit the data better and explain more variation in AGGCS than less optimised variables. Their integration in XGBoost leverages complementary strengths (i.e. SAR's structural sensitivity, texture's spatial context, and topography's environmental stability) to achieve robust AGGCS quantification across seasons.

The findings of this study corroborate with those of Li *et al.* (2020), who achieved improved results when combining optimal features derived from Gao Fen 7 (GF-7) stereo imagery, Sentinel-1, Sentinel-2 and the Advanced Land Observing Satellite digital elevation model (ALOS DEM) in estimating forest biomass. In the temperate forests of the Himalayan region, Anees *et al.* (2024) examined the relationship between predictors from Landsat-9 and several topographical features also reported that the integration of vegetation-based indices, including TNDVI, NDVI, RVI, and GNDVI, further refines AGB mapping precision. A study by Liang *et al.* (2022) evaluated the performance of spectral information and texture features derived from UAV-based high-resolution RGB imagery with different textural parameter settings for estimating AGB of rubber plantation. They found that spectral information coupled with texture features derived from suitable GLCM parameter selection has great potential in improving the quantification of AGB in rubber plantations. This finding agrees with (Nichol and Sarker, 2010; Naidoo *et al.*, 2016; Chen *et al.*, 2019b) who achieved enhanced predictive powers when incorporating texture variables

5.5.3 Impact of seasons on SAR-based AGGCS quantification

The findings of this study show that the predictive power of the XGBoost algorithm as well as that the contribution of SAR, texture and topographic parameters varied across the wet and the dry seasons of 2020 and 2021, respectively. This attributed to the changes in environmental conditions including differences in precipitation intensity, soil moisture and phonological stages of the grasses (vegetative vs senescence). This, in turn, could affect the Synthetic Aperture Radar (SAR) signal and the contributions of various parameters. Our findings also show that the models developed for the dry season exhibited superior yet varying predictive power compared to that of the wet season. This could be because diminished dense vegetation as well as the soil moisture, which often obscure the sensor signals during the wet season. The result of these study agree with of (Kanja *et al.*, 2025) who evaluated the ability to map the AGB of the Miombo woodlands in Zambia across four agro-ecological zones using both multi-seasonal SAR (Sentinel-1A) and optical (Landsat-8 OLI) imagery and reported that rainy season images were the least accurate in predicting forest biomass. Others also show that models trained and validated using data from the same year may inadvertently learn patterns that reflect year-specific anomalies, resulting in overfitting and reduced performance when applied to data from other years (Valbuena *et al.*, 2017; Filippelli *et al.*, 2024).

5.6 Conclusions

Utilising in situ AGB samples collected across the wet and dry seasons of 2020 and 2021, we extracted 33 predictor variables derived from various sources, including Sentinel-1, Texture, STRM DEM to assess the impact of seasons on the combined predictive capability coalesced SAR indices, texture and topographic variables in estimating AGGCS at a 10 m spatial resolution. Based on the findings, it is conclude that:

- Season play a significant role in shaping the predictive capabilities as well as the contribution of predictor variables
- Standalone SAR indices, texture and topographic variables prduce weak correlations
- The multi-parameter approach is optimal for estimating AGGCS savanna ecosystems
- The inclusion of the 33 variables does not significantly improve the performance of the XGBoost algorithms in modelling AGGCS in the complex savanna ecosystems, but optimisation of variables does,

- The variables that optimally captured the variability of AGGCS includes Overall, RVI, Slope, TRI, Aspect, VHcor, TPI, TWI, Lat, VVcor and Hillshade
- The predictions based on the dry season model explains much of the variability in AGGCS compared to the wet season model

Overall, the integration of Sentinel-1-derived matrices, texture matrices and topographic variables coupled with XGBoost algorithm shows potential in predicting above-ground grass carbon stock in savannah ecosystems.

5.7 Summary

Algorithms that are developed and validated within a single year tend to detect anomalies specific to that year. This can result in over-fitting, which subsequently diminishes their performance when applied to data from different years, resulting in over-fitting and underperformance in subsequent years. The last objective of this thesis was to model the effect of seasons on the predictive capability of the XGBoost algorithm. This was done to test the temporal stability and transferability of the models and parameters within the season and across years. The optimised integrated SAR, texture and topographic features were fitted into the XGBoost algorithm for the wet and dry conditions of the years 2020 and 2021. The findings indicate that the contribution of SAR, texture and topographic parameters varied between the wet and dry seasons in both years. Notably, the predictive capabilities of the models also differed within the seasons and across the years. Overall, the XGBoost model that integrated all the variables (topographic, texture and SAR indices) showed the highest predictive power in both the wet and dry seasons across both years. Notably, the model developed for the dry season exhibited superior predictive power compared to that of the wet season. The findings underscore the impact of environmental and phenological factors on the predictive power of remote sensing-based models for carbon stock estimations. The last chapter concludes the thesis and situates the research findings within a broader context and presents recommendations for future research.

CHAPTER 6: THE APPLICATION OF SAR DATA FOR QUANTIFYING AGGC IN SAVANNA ECOSYSTEMS: A SYNTHESIS

6.1 Introduction

The world is currently facing an unprecedented challenge due to global warming, which has manifested more severely than previously expected. These sudden and alarming global changes underscore the importance of quantifying and monitoring above-ground grass carbon stock (AGGCS), especially in savanna ecosystems, which cover a quarter of the earth's surface (Scurlock and Hall, 1998). This approach is critical for climate change initiatives intended to enhance both mitigation and adaptation strategies. Understanding the distribution of AGGB across different spatial scales and seasons is crucial, especially for facilitating a comprehensive implementation of international greenhouse gas inventory programs as well as Goal 13 of the Sustainable Development Goals (SDGS), which calls for societies to take urgent action to combat climate change and its impacts. However, a wall-to-wall quantification and monitoring of AGGB necessitates the implementation of advanced technologies and methodologies that are specifically designed to adapt to the rapid dynamics of global changes. In this regard, Earth observation sensors, with improved spatial and temporal configurations, provide valuable imagery that can facilitate the achievement of this objective.

However, the application of earth observation sensors in seasonal quantification and monitoring of AGGCS, particularly in the tropical savanna, is hindered by persistent cloud cover; meanwhile, the wide adoption of SAR data is hindered by the cost associated with the acquisition of the images. Nevertheless, over 70% of the world's savannas are concentrated in the global south, particularly within the tropics (Mishra and Young, 2020), where resource constraints and frequent cloud cover impede the ability to conduct long-term monitoring of AGGCS. Nevertheless, the Sentinel-1 mission is a game changer for monitoring and quantifying AGGCS, particularly in the African savanna ecosystem, as it offers unprecedented data that cuts across all weather conditions, day and night observation, as well as cloud and within-canopy vegetation penetration. However, in landscapes with intricate topography, such as mountainous areas, radar signals are influenced by variations in the terrain, resulting in alterations in the intensity of radar echoes (Van Zyl, 1993). These terrain-induced effects can impede the precise assessment of carbon stocks, particularly in steeply sloped environments (Xu *et al.*, 2025). Another challenge with SAR-based models for quantifying and monitoring

AGGCS is the issue of signal saturations (Kumar *et al.*, 2015; Sinha *et al.*, 2015b). These challenges underscore the necessity for complementary data and improved processing methodologies to enhance the precision and dependability of AGGCS estimations derived from Synthetic Aperture Radar (SAR), particularly within sparse, dense and topographically intricate savannah ecosystems. This thesis, therefore, leveraged the freely available Sentinel-1 SAR data coupled with advanced machine learning algorithms to model AGGCS in the complex savannah ecosystems. The quantification of AGGCS in these ecosystems, and particularly on the grass layer, is essential to foster an all-inclusive and regionally representative reporting to the global carbon budget. This chapter provides an overview of the research results, links between chapters and conclusions based on the objectives outlined in subsection 1.3.

6.1.1 Reviewing the progress and challenges in the retrieval of AGGCS in Savannah ecosystems using remote sensing

The adoption of open-access policies for some remote sensing data has sparked an increase in research focused on developing algorithms for estimating aboveground grass biomass (AGGB). This surge in studies underscores the need for a comprehensive synthesis of published findings to evaluate methodological advancements, identify limitations in remote sensing-based AGGB estimations, and highlight future research needs, especially for the complex savannah ecosystems. The importance of synthesising findings from multiple independent studies is emphasised in Guzzo *et al.* (1987). This section, therefore, provides an overview of the methodological progress and limitations derived from the published body of literature on the application of remote in AGGCS estimations and further suggests new directions. To achieve this, this study utilised electronic databases such as the Web of Science, IEEE Explorer, Scopus, and Google Scholar, employing both qualitative and quantitative techniques. The findings indicate that of the publications reviewed ($n = 108$), only 19% of the studies developed remote sensing-based models for estimating AGGCS in savanna ecosystems, indicating a notable underrepresentation of this ecological system. The underrepresentation becomes more pronounced when focusing on the specific sensors employed, where a limited number of studies developed SAR-based models for estimating AGGCS in savanna ecosystems. Furthermore, of the few studies that developed SAR-based models for quantifying AGGCS in savanna ecosystems, the majority relied on linear models that failed to capture the complexity of vegetation within these ecosystems. Overall, the findings from this chapter highlight the progress made in the application of remote sensing for estimating AGGCS while underscoring the persistent challenges and opportunities for future research, particularly using SAR data products. Driven by the availability of resources, addressing these gaps was the forefront of this thesis to contribute to the AGGCS algorithm's advancement while ensuring a balanced representation in AGGB estimations across ecosystems, vegetation types and regions of the world. As such, these findings served as the framework for defining the scope and guiding the methodological choice for the entire thesis. The subsequent chapters were expanded based on these insights. Specifically, the observed underrepresentation of SAR-based models tailored towards quantifying AGGCS in savannah ecosystems in the global south directed this study towards addressing this ecological and geographical imbalance. As such, subsequent chapters were diverted towards developing algorithms for quantifying AGGCS in the savannah ecosystem. Furthermore, the freely available Sentinel-1 was underutilised, prompting this study to broaden its scope to explore alternative predictors beyond the traditional optical-based vegetation indices. In addition, it was observed that integrating multiple variables and sensors yields improved predictive power. As such, subsequent chapters were tailored towards leveraging the complementary strengths of SAR by incorporating texture and topographic variables with SAR data to improve the predictive power. Lastly, the savannah ecosystem's vegetation composition (i.e. dominant grass layer coexisting with the woody layer), introduces complex relationships that are not sufficiently captured by linear models. This complexity necessitates the need to test advanced machine learning algorithms in the chapters to better account for these complex relationships.

6.1.2 Determine the optimal SAR parameters suitable for quantifying AGGCS in Savannah ecosystems

Synthetic Aperture Radar (SAR) offers various metrics that can be utilised for quantifying AGGCS, some of which have been reported as optimal for forest (Ghosh and Behera, 2021; Vatandaşlar and Abdikan, 2022; Zeng *et al.*, 2022). Additionally, studies have demonstrated that integrating texture variables with SAR parameters can enhance the model's ability to predict carbon stocks (Nichol and Sarker, 2010; Naidoo *et al.*, 2016; Liao *et al.*, 2020). The approach can surpass the challenges encountered when predicting carbon stocks in savanna ecosystems, which consist of a complex vegetation structure that traditional optical sensors struggle to penetrate. Nevertheless, fewer studies have tested the combined use of parameters derived from Sentinel-1 and texture matrices for predicting AGGCS in complexly configured savanna ecosystems. This chapter of the thesis investigated the sensitivity of SAR-derived parameters coupled with texture matrices to predict AGGCS in a savannah ecosystem. To do this, various parameters were derived from Sentinel-1, including backscatter coefficients, intensity ratios, normalised radar backscatter, and arithmetic computations. The XGBoost tree-based algorithm was employed to predict the AGGCS. The findings indicate that parameters derived from a single-date Sentinel-1 image alone were inadequate for quantifying AGGCS in a savannah ecosystem. However, when incorporating texture matrices with the SAR parameters, improved predictive power was observed. Overall, of all the groups of SAR and texture-based variables, RVI, VHcor, and VVcor as well as the latitude and longitude in order of importance, optimally captured the variability of AGGCS in savannah ecosystems. These findings agree with the hypothesis set for this chapter that the combination of specific SAR-derived and texture parameters improves the accuracy of quantifying AGGCS in Savannah ecosystems than others. The findings of this study underscore the integration of Sentinel-1-derived matrices coupled with texture matrices and the XGBoost algorithm as the optimal approach for predicting AGGCS in the complexly configured savannah ecosystems. These findings serve as a baseline against which future estimates in similar or different ecological settings can be compared.

6.1.3 Evaluating the sensitivity of SAR-derived indices coupled with topo-edaphic and texture parameters in quantifying within-season variation of above-ground grass carbon stocks in savannah ecosystems

The capacity of the grass layer to sequester or emit carbon fluctuates with seasons. Studies show that carbon sequestration reaches its highest levels during the peak growth season and declines as senescence begins (Maschler *et al.*, 2022; Zhu *et al.*, 2024). This trend is driven by physiological shifts in vegetation, where peak photosynthetic activity maximises carbon uptake, followed by reduced assimilation and increased carbon release during senescence (Liu *et al.*, 2025). As such, assessing the quantity of carbon stored in grasses at different temporal windows within a single season is vital. This can serve as an early warning to indicate carbon stock stress from factors such as drought, excessive grazing, or pest infestations. However, a significant constraint on the regular quantification of AGGCS within the African savannas, including

South Africa, is the prohibitive cost of data. Furthermore, the heterogeneous nature of prompt sensors with sub-meter resolutions to capture these dynamics (Abdi *et al.*, 2022). , which is almost impossible with the traditional and freely available optical sensors. Nevertheless, courtesy of the open access policy associated with the Sentinel-1 mission, we can map the distribution of AGCS at 10m spatial resolution at no cost. This chapter expands on the scope of the preceding chapter by incorporating topographic variables with the optimal SAR and texture parameters investigated above to assess how they would perform across the wet and dry seasons of a single year.

The findings of this chapter indicate that, as stand-alone predictors, SAR indices, texture and topographic variables produce weak correlations, suggesting that this approach struggles to effectively explain the variability in AGGCS in the complex savannas. Improved predictive powers of the XGBoost model were observed with the coalescing SAR, texture, and topographic across the wet and the dry of one season cycle. The variables that optimally captured the variability of AGGCS include RVI (spectral) and TPI and TRI features (Terrain). Overall, the integration of Sentinel-1-derived matrices coupled with texture metrics topographic variables and the XGBoost algorithm shows potential in predicting AGGCS in savannah ecosystems. These findings concur with the hypothesis that "within-seasonal variations in above-ground biomass in savanna ecosystems can be accurately estimated with the combined use of SAR-derived indices, texture and topographic variables". Although conducted in different setups, the potential of multi-parameter coalescing has achieved improved results (Sarker *et al.*, 2012; Zhou *et al.*, 2013; Kelsey and Neff, 2014; Yao and Ren, 2024). This approach provides insights that can inform grazing and fire management regimes in the subsequent years

6.1.4 Evaluating the utility of SAR data in quantifying the inter-annual variation of AGGCS in Savanna ecosystems

From an optical sensing perspective, mapping grass carbon stocks across years using optical sensors faces significant challenges, particularly within the savanna ecosystem in the tropics, which experiences frequent cloud cover. Specifically, in the rainy season, clouds tend to obscure optical sensors in capturing critical information, therefore, creating data gaps during optimal sequestration periods. While microwave sensors offer an alternative to fill the data gaps, they also suffer from saturation issues (Kasischke *et al.*, 2002). Studies have suggested employing a multi-parameter approach to overcome these challenges. Prompted by the findings from the previous chapter, this section of the thesis broadened its scope to explore the performance of the optimal coalesced SAR indices, texture and topographic variables across the wet and dry seasons of multiple years in predicting AGGCS. The findings of this study show that different combinations of predictors (i.e radar indices, texture, and topography) contributed differently across the wet and dry seasons of 2020 and 2021, respectively. Nevertheless, among the variables utilised in the study, RVI and TRI consistently contributed the most in modelling AGGCS across the wet and dry seasons of 2020 and the dry season of 2021, respectively. However, the XGBoost models based on the combination of all the variables outperformed those that employed a single-variable model. Overall, the dry season-based AGGCS models consistently outperformed the wet-season-based models within and across the years under investigation. Providing a baseline for long-term monitoring and policy interventions. The study reveals significant seasonal variations in the estimates of AGGCS across and within and across the years. These variations are closely related to changes in environmental conditions including differences in precipitation intensity, soil moisture and phenological stages of the grasses (vegetative vs senescence). These findings confirm with the hypothesis outlined in section 1.4 for this chapter, that the predictive capability of SAR-derived indices coupled with topography and texture measures in quantifying AGGCS within Savanna ecosystems is affected by seasonal variation This, in turn, could affect the Synthetic Aperture Radar (SAR) signal and the contributions of various parameters. The result of this study agrees with those of Kanja *et al.* (2025); Valbuena *et al.* (2017) and Filippelli *et al.* (2024).

6.2 Challenges and future research

Despite the demonstrated efficacy of SAR-based models in various remote sensing research areas, including AGGCS mapping, these models encounter several challenges that constrain their overall utility, and these are outlined below:

- Variables such as topographic, rainfall, soil moisture and temperature influences the distribution and abundance of AGGCS across various landscapes. In this research, efforts were undertaken to integrate soil moisture and precipitation data from established sources, specifically the Soil Moisture Active Passive (SMAP) and the Climate Hazards Group Infrared Precipitation with Station data (CHIRPS).

Nonetheless, the spatial resolution of these datasets, which is 9 km for SMAP and 1 km for CHIRPS, proved insufficient for the scale of our mapping, which is set at 10 m. Future studies, should explore locally available weather and soil moisture parameters from ground-based weather stations.

- The extrapolation of predictions derived from remote sensing-based AGGCS models is a prevalent and complex endeavour within the field of remote sensing. Nevertheless, it is crucial for facilitating mapping efforts at broader scales, including regional, landscape, and national levels. The other source of error relates to the size of a measurement plot used in this study, which only represents a small fraction of the area covered by a sentinel-1 and the DEM pixel, and the value obtained from the plot may not truly reflect the total biomass value within that pixel of the Sentinel 1 image. Future studies should enlarge the number of field samples of the same grasses to capture enough variability and ensure representativeness. Furthermore, the scale mismatch between DEM, which is captured at a spatial resolution of 30m, compared to the Sentinel-1 image, which is captured at 10m spatial resolution, was also a limiting factor. Although the study resampled the DEM to a 10 m spatial resolution, this process may have introduced some level of error. When feasible, it is advisable to utilise DEM products that possess a higher spatial resolution, such as 10 meters or finer, to align with the resolution of Sentinel-1 data. This approach minimises the necessity for resampling. Future studies can improve the results by using LIDAR, which offers terrain attributes at higher spatial resolution compared to SRTM Digital Elevation Models (DEMs), which may be advantageous for SAR-based AGGCS models. The only setback of LIDAR is the price tag as well as computational demand for large-scale mapping. Similarly, studies can also opt for Unmanned Aerial systems, which is relatively cheaper but also limited in terms of geographical footprint.
- This study was limited to the use of SAR data, as no cloudless optical images were available for the wet season but also demonstrated potential in coalescing multiple parameters; future studies should explore the combined use of optical and SAR data in such an ecosystem. Combining C-band SAR with optical products (e.g., Sentinel-2 or Landsat surface reflectance, vegetation indices), L-band SAR (e.g., ALOS-2 PALSAR-2, SAOCOM), and space borne LIDAR (e.g. GEDI footprint height metrics or ICESat-2 ATL08 canopy parameters) can improve structural characterization, reduce saturation,

and enhance sensitivity to biomass gradients.

- This study relied on C-band Sentinel-1 image, which has shown unsatisfactory results as a stand-alone predictor; future studies should explore other sensors within the P and the L bands of the active electromagnetic spectrum. These longer wavelengths coupled with these sensors provide improved canopy penetration and stronger sensitivity to structural components relevant to AGGCS estimation in the heterogeneous savannah ecosystems. However, in many resource-constrained savannah regions, such as Africa, access to commercial SAR data boarding the P and L- band sensors, such as ALOS PALSAR is hindered by the high cost. Nevertheless, future SAR missions such as BIOMASS and emerging drone-based SAR technologies such as SAOCOM, NISAR, ALOS-4, and Tandem-L missions bring prospects to improve and strengthen seasonal and inter-annual AGGCS mapping savannah ecosystems.

6.3 Recommendations

- It is essential to develop standardised protocols for model evaluations of matrix data to improve the comparability of findings across different studies.
- The validation of remote sensing-based algorithms is critical remote sensing community. Currently, data collection protocols are subjective; it is recommended that comprehensive protocols for the collection of ground-truth data be established to guarantee both consistency and reliability in the results obtained.
- Encourage the establishment of collaborative partnerships among researchers, academic institutions, and policymakers to facilitate the exchange of knowledge, tools, and resources.
- Enhance the utilisation of publicly accessible SAR datasets, including Sentinel-1, for the estimation of AGGCS at an operational level.
- Enhance policies that guarantee the ongoing availability and accessibility of both remote sensing datasets and ground observation data for validating remote sensing models, to support research and inform decision-making processes in developing regions

- The existing frameworks for reporting carbon emissions are tailored for forests, and overlooking contributions from grasses. It is recommended that an all-inclusive global framework be developed for reporting of AGGCS statistics to ensure accuracy for carbon accounting and climate reporting systems.
- Involve local communities in carbon monitoring initiatives to promote sustainable practices and improve ground-truth data collection.
- Establish financial mechanisms, such as carbon credits or payments for ecosystem services, to incentivise conservation and restoration efforts that enhance grass carbon stocks in savannahs.

6.4 Conclusions

The process of mapping AGGCS is undergoing significant transformation, fuelled by advancements in technology, the emergence of novel data sources and algorithms, as well as the increasing demand for information that is both precise and relevant in shaping decision-making. The core of this study was to explore the potential of the freely available SAR (i.e. Sentinel-1) data coupled with a scalable XGBoost algorithm in AGGCS within and across seasons in the complex savannah ecosystems. The other objective assessed whether incorporating topo-edaphic variables and texture parameters can enhance the quantification of AGGCS in these complex savannah ecosystems. The findings in this study highlight the role of the grass layer in the global carbon cycle and the conditions that compel the system to serve as a carbon source or a sink. Moreover, the findings of this study strongly position remote sensing coupled with scalable machine learning algorithms as techniques to reliably and effectively quantify and monitor AGGC over short and long-term scales, particularly the incorporation of multiple parameters to deal with the complex configuration of vegetation in savannah ecosystems. Overall, this study contributes to our understanding of remote sensing-based methodologies for quantifying AGGCS, particularly the complexly configured savannahs. It demonstrates the potential of SAR coalesced with texture and topographic variables in improving the predictive capabilities of AGGCS models. By analysing the patterns of AGGCS across the wet and dry seasons within and across the two years, the study has shed insights into the temporal variability of carbon stocks in savanna ecosystems. The results provide baseline statistics against which future studies can compare and for long-term monitoring. From the policy lens, these estimates emanating from this study can be used for national and regional carbon budgeting and reporting. Land-use managers can benefit from the patterns in the distribution of AGGCS when designing land-use regimes, including fire control and ecosystem service restorations.

6.5 References

't Mannetje, L. (2000). Measuring biomass of grassland vegetation. Field and laboratory methods for grassland and animal production research, CABI Publishing Wallingford UK: 151-177.

Abbass, K., Qasim, M. Z., Song, H., Murshed, M., Mahmood, H. and Younis, I. (2022). "A review of the global climate change impacts, adaptation, and sustainable mitigation measures." Environmental Science and Pollution Research **29**(28): 42539-42559.

Abdi, A. M., Brandt, M., Abel, C. and Fensholt, R. (2022). "Satellite remote sensing of savannas: Current status and emerging opportunities." Journal of Remote Sensing.

Adjorlolo, C. and Mutanga, O. (2013). "Integrating remote sensing and geostatistics to estimate woody vegetation in an african savanna." Journal of Spatial Science **58**(2): 305-322.

Adjorlolo, C., Mutanga, O., Cho, M. and Ismail, R. (2012). "Challenges and opportunities in the use of remote sensing for c3 and c4 grass species discrimination and mapping." African Journal of Range & Forage Science **29**(2): 47-61.

Ali, C. F., Dwyer, E., Green, S., Ali, I., Cawkwell, F. and Dwyer, E. (2017). "Modeling managed grassland biomass estimation by using multitemporal remote sensing data-a machine learning approach " IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **10**(7): 3254-3264.

Ali, I., Cawkwell, F., Dwyer, E. and Green, S. (2016). "Modeling managed grassland biomass estimation by using multitemporal remote sensing data—a machine learning approach." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **10**(7): 3254-3264.

Ali, I., Schuster, C., Zebisch, M., Förster, M., Kleinschmit, B. and Notarnicola, C. (2013). "First results of monitoring nature conservation sites in alpine region by using very high resolution (vhr) x-band sar data." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **6**(5): 2265-2274.

Allen, C. D., Breshears, D. D. and McDowell, N. G. (2015). "On underestimation of global vulnerability to tree mortality and forest die-off from hotter drought in the anthropocene." Ecosphere **6**(8): 1-55.

Alvarez-Mendoza, C. I., Guzman, D., Casas, J., Bastidas, M., Polanco, J., Valencia-Ortiz, M., Montenegro, F., Arango, J., Ishitani, M. and Selvaraj, M. G. (2022). "Predictive modeling of above-ground biomass in brachiaria pastures from satellite and uav imagery using machine learning approaches." Remote Sensing **14**(22): 5870.

Anees, S. A., Mehmood, K., Khan, W. R., Sajjad, M., Alahmadi, T. A., Alharbi, S. A. and Luo, M. (2024). "Integration of machine learning and remote sensing for above ground biomass estimation through landsat-9 and field data in temperate forests of the himalayan region." Ecological Informatics **82**: 102732.

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L. and Ridella, S. (2012). The'k'in k-fold cross validation. ESANN.

Avohou, H. T. and Sinsin, B. (2009). "The effects of topographic factors on aboveground biomass production of grasslands in the atacora mountains in northwestern benin." Mountain Research and Development **29**(3): 250-254.

Bai, Y. and Cotrufo, M. F. (2022). "Grassland soil carbon sequestration: Current understanding, challenges, and solutions." Science **377**(6606): 603-608.

Bannari, A., Morin, D., Bonn, F. and Huete, A. R. (1995). "A review of vegetation indices." Remote sensing reviews **13**: 95-120.

Bao, N., Li, W., Gu, X. and Liu, Y. (2019). "Biomass estimation for semiarid vegetation and mine rehabilitation using worldview-3 and sentinel-1 sar imagery." Remote Sensing **11**(23): 2855.
Bargiel, D. (2017). "A new method for crop classification combining time series of radar images and crop phenology information." Remote sensing of environment **198**: 369-383.
Barrachina, M., Cristóbal, J. and Tulla, A. F. (2015). "Estimating above-ground biomass on mountain meadows and pastures through remote sensing." International Journal of Applied Earth Observation and Geoinformation **38**: 184-192.
Barrett, B., Nitze, I., Green, S. and Cawkwell, F. (2014). "Assessment of multi-temporal, multi-sensor radar and ancillary spatial data for grasslands monitoring in ireland using machine learning approaches." Remote sensing of environment **152**: 109-124.
Basuki, T. M., Skidmore, A. K., van Laake, P. E., van Duren, I. and Hussin, Y. A. (2012). "The potential of spectral mixture analysis to improve the estimation accuracy of tropical forest biomass." Geocarto International **27**(4): 329-345.
Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D. and Bonan, G. B. (2010). "Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate." Science **329**(5993): 834-838.
Beltran, J. C., Valdez, P. and Naval, P. (2019). Predicting protein-protein interactions based on biological information using extreme gradient boosting. 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), IEEE.
Berger, C., Lux, H., Urban, M., Schmullius, C., Baade, J., Thiel, C., Wigley-Coetsee, C. and Smit, I. (2020). Annual grass biomass mapping with landsat-8 and sentinel-2 data over kruger national park, south africa. IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE.
Beriaux, E., Jago, A., Lucau-Danila, C., Planchon, V. and Defourny, P. (2021). "Sentinel-1 time series for crop identification in the framework of the future cap monitoring." Remote Sensing **13**(14): 2785.
Beringer, J., Hutley, L. B., Abramson, D., Arndt, S. K., Briggs, P., Bristow, M., Canadell, J. G., Cernusak, L. A., Eamus, D. and Edwards, A. C. (2015). "Fire in australian savannas: From leaf to landscape." Global Change Biology **21**(1): 62-81.
Berninger, A., Lohberger, S., Stängel, M. and Siegert, F. (2018). "Sar-based estimation of above-ground biomass and its changes in tropical forests of kalimantan using l-and c-band." Remote Sensing **10**(6): 831.
Bindu, G., Rajan, P., Jishnu, E. and Joseph, K. A. (2020). "Carbon stock assessment of mangroves using remote sensing and geographic information system." The Egyptian Journal of Remote Sensing and Space Science **23**(1): 1-9.
Bispo, P., Santos, J., Valeriano, M., Touzi, R. and Seifert, F. (2014). "Integration of polarimetric palsar attributes and local geomorphometric variables derived from srtm for forest biomass modeling in central amazonia." Canadian Journal of Remote Sensing **40**(1): 26-42.
Bouvet, A., Mermoz, S., Le Toan, T., Villard, L., Mathieu, R., Naidoo, L. and Asner, G. P. (2018). "An above-ground biomass map of african savannahs and woodlands at 25 m resolution derived from alos palsar." Remote sensing of environment **206**: 156-173.
Braun, A., Wagner, J. and Hochschild, V. (2018). "Above-ground biomass estimates based on active and passive microwave sensor imagery in low-biomass savanna ecosystems." Journal of Applied Remote Sensing **12**(4): 046027.

Bristow, M., Hutley, L. B., Beringer, J., Livesley, S. J., Edwards, A. C. and Arndt, S. K. (2016). "Quantifying the relative importance of greenhouse gas emissions from current and future savanna land use change across northern australia." Biogeosciences **13**(22): 6285-6303.
Butterfield, H. and Malmström, C. (2009). "The effects of phenology on indirect measures of aboveground biomass in annual grasses." International Journal of Remote Sensing **30**(12): 3133-3146.
Calleja, J. F., Pagés, O. R., Díaz-Álvarez, N., Peón, J., Gutiérrez, N., Martín-Hernández, E., Relea, A. C., Melendi, D. R. and Álvarez, P. F. (2018). "Detection of buried archaeological remains with the combined use of satellite multispectral data and uav data." International Journal of Applied Earth Observation and Geoinformation **73**: 555-573.
Cawse-Nicholson, K. and Anderson, M. C. (2018). "Ecosystem spaceborne thermal radiometer experiment on space station (ecostress) mission." California Institute of Technology.
Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S. and Entekhabi, D. (2018). "Development and assessment of the smap enhanced passive soil moisture product." Remote sensing of environment **204**: 931-941.
Chang, G. J., Oh, Y. and Shoshany, M. (2022). "Biomass estimation along a climatic gradient using multi-frequency polarimetric radar vegetation index." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **3**: 369-374.
Chang, J. G., Kraatz, S., Anderson, M. and Gao, F. (2024). "Enhanced polarimetric radar vegetation index and integration with optical index for biomass estimation in grazing lands across the contiguous united states." Remote Sensing **16**(23): 4476.
Chapin III, F. S., Matson, P. A. and Vitousek, P. (2011). Principles of terrestrial ecosystem ecology, Springer Science & Business Media.
Chapungu, L., Nhamo, L. and Gatti, R. C. (2020). "Estimating biomass of savanna grasslands as a proxy of carbon stock using multispectral remote sensing." Remote Sensing Applications: Society and Environment **17**: 100275.
Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., Armston, J., Baker, T. R., Coomes, D. and Disney, M. (2019). "Ground data are essential for biomass remote sensing missions." Surveys in Geophysics **40**: 863-880.
Chen, J., Gu, S., Shen, M., Tang, Y. and 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**(24): 6497-6517.
Chen, L., Wang, Y., Ren, C., Zhang, B. and Wang, Z. (2019a). "Optimal combination of predictors and algorithms for forest above-ground biomass mapping from sentinel and srtm data." Remote Sensing **11**(4): 414.
Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.
Chen, X., Hutley, L. B. and Eamus, D. (2003). "Carbon balance of a tropical savanna of northern australia." Oecologia **137**: 405-416.
Chen, Y., Guerschman, J. P., Cheng, Z. and Guo, L. (2019b). "Remote sensing for vegetation monitoring in carbon capture storage regions: A review." Applied energy **240**: 312-326.
Clementini, C., Pomente, A., Latini, D., Kanamaru, H., Vuolo, M. R., Heureux, A., Fujisawa, M., Schiavon, G. and Del Frate, F. (2020). "Long-term grass biomass estimation of pastures from satellite data." Remote Sensing **12**(13): 2160.
Collart, F. and Guisan, A. (2023). "Small to train, small to test: Dealing with low sample size in model evaluation." Ecological Informatics **75**: 102106.

Crabbe, R. A., Lamb, D. W., Edwards, C., Andersson, K. and Schneider, D. (2019). "A preliminary investigation of the potential of sentinel-1 radar to estimate pasture biomass in a grazed pasture landscape." Remote Sensing **11**(7): 872.
Craine, J. M., Ocheltree, T. W., Nippert, J. B., Towne, E. G., Skibbe, A. M., Kembel, S. W. and Fargione, J. E. (2013). "Global diversity of drought tolerance and grassland climate-change resilience." Nature Climate Change **3**(1): 63-67.
Cutler, M., Boyd, D. S., Foody, G. M. and 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.
D'Onofrio, D., Baudena, M., Lasslop, G., Nieradzik, L. P., Wårlind, D. and von Hardenberg, J. (2020). "Linking vegetation-climate-fire relationships in sub-saharan africa to key ecological processes in two dynamic global vegetation models." Frontiers in Environmental Science **8**: 136.
Daoud, J. I. (2017). Multicollinearity and regression analysis. journal of physics: Conference series, IOP Publishing.
Das, A. A., Thaker, M., Coetsee, C., Slotow, R. and Vanak, A. T. (2022). "The importance of history in understanding large tree mortality in african savannas." Ecography **2022**(1).
Dass, P., Houlton, B. Z., Wang, Y. and Warlind, D. (2018). "Grasslands may be more reliable carbon sinks than forests in california." Environmental Research Letters **13**(7): 074027.
de Araujo, L. S., dos Santos, J. R., da Costa Freitas, C. and Xaud, H. A. M. (1999). The use of microwave and optical data for estimating aerial biomass of the savanna and forest formations at roraima state, brazil. IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293), IEEE.
de Lima, D. O., Lorini, M. L. and Vieira, M. V. (2018). "Conservation of grasslands and savannas: A meta-analysis on mammalian responses to anthropogenic disturbance." Journal for Nature Conservation **45**: 72-78.
Debastiani, A. B., Sanquetta, C. R., Dalla Corte, A. P., Pinto, N. S. and Rex, F. E. (2019). "Evaluating sar-optical sensor fusion for aboveground biomass estimation in a brazilian tropical forest." Annals of Forest Research **62**(2): 109-122.
Dieleman, W. I., Venter, M., Ramachandra, A., Krockenberger, A. K. and Bird, M. I. (2013). "Soil carbon stocks vary predictably with altitude in tropical forests: Implications for soil carbon storage." Geoderma **204**: 59-67.
Ding, L., Li, Z., Wang, X., Yan, R., Shen, B., Chen, B. and Xin, X. (2019). "Estimating grassland carbon stocks in hulunber china, using landsat8 oli imagery and regression kriging." Sensors **19**(24): 5374.
Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H. and Mavrocordatos, C. (2012). "The global monitoring for environment and security (gmes) sentinel-3 mission." Remote sensing of environment **120**: 37-57.
Dorigo, W., Hahn, S., Hohensinn, R., Paulik, C., Wagner, W., Drusch, M. and van Oevelen, P. (2010). The international soil moisture network-a data hosting facility for in situ soil moisture measurements in support of smos cal/val. EGU General Assembly Conference Abstracts.
Dube, T. and Mutanga, O. (2015). "Evaluating the utility of the medium-spatial resolution landsat 8 multispectral sensor in quantifying aboveground biomass in umgeni catchment, south africa." ISPRS Journal of Photogrammetry and Remote Sensing **101**: 36-46.

Dube, T., Shoko, C. and Gara, T. W. (2021). "Remote sensing of aboveground grass biomass between protected and non-protected areas in savannah rangelands." African Journal of Ecology **59**(3): 687-695.
Dusseux, P., Corpetti, T., Hubert-Moy, L. and Corgne, S. (2014). "Combined use of multi-temporal optical and radar satellite images for grassland monitoring." Remote Sens **6**: 6163-6182.
Eisfelder, C., Kuenzer, C. and Dech, S. (2012). "Derivation of biomass information for semi-arid areas using remote-sensing data." International Journal of Remote Sensing **33**(9): 2937-2984.
FAO (2010). Challenges and opportunities for carbon sequestration in grassland systems. A technical report on grassland management and climate change mitigation. Integrated Crop Management, Food and Agriculture Organization of the United Nations. **9:** 1-57.
Fiala, K. (2010). "Belowground plant biomass of grassland ecosystems and its variation according to ecological factors." Ekológia (Bratislava) **29**(2): 182-206.
Filippelli, S. K., Schleeweis, K., Nelson, M. D., Fekety, P. A. and Vogeler, J. C. (2024). "Testing temporal transferability of remote sensing models for large area monitoring." Science of Remote Sensing **9**: 100119.
Flombaum, P. and Sala, O. (2007). "A non-destructive and rapid method to estimate biomass and aboveground net primary production in arid environments." Journal of Arid Environments **69**(2): 352-358.
Folega, F., Woegan, Y. A., Marra, D., Wala, K., Batawila, K., Seburanga, J. L., Zhang, C.-y., Peng, D.-l., Zhao, X.-h. and Akpagana, K. (2015). "Long term evaluation of green vegetation cover dynamic in the atacora mountain chain (togo) and its relation to carbon sequestration in west africa." Journal of Mountain Science **12**: 921-934.
Frolking, S., Fahnestock, M., Milliman, T., McDonald, K. and Kimball, J. (2005). "Interannual variability in north american grassland biomass/productivity detected by seawinds scatterometer backscatter." Geophysical research letters **32**(21).
Gao, J., Liang, T., Liu, J., Yin, J., Ge, J., Hou, M., Feng, Q., Wu, C. and Xie, H. (2020). "Potential of hyperspectral data and machine learning algorithms to estimate the forage carbon-nitrogen ratio in an alpine grassland ecosystem of the tibetan plateau." ISPRS Journal of Photogrammetry and Remote Sensing **163**: 362-374.
Gao, T., Xu, B., Yang, X., Jin, Y., Ma, H., Li, J. and Yu, H. (2013). "Using modis time series data to estimate aboveground biomass and its spatio-temporal variation in inner mongolia's grassland between 2001 and 2011." International Journal of Remote Sensing **34**(21): 7796-7810.
Gertenbach, W. D. (1983). "Landscapes of the kruger national park." Koedoe **26**(1): 9-121.
Ghasemi, N., Sahebi, M. R. and Mohammadzadeh, A. (2011). "A review on biomass estimation methods using synthetic aperture radar data." International Journal of Geomatics and Geosciences **1**(4): 776-788.
Ghosh, P. and Mahanta, S. (2014). "Carbon sequestration in grassland systems." Range Management and Agroforestry **35**(2): 173-181.
Ghosh, S. M. and Behera, M. D. (2021). "Aboveground biomass estimates of tropical mangrove forest using sentinel-1 sar coherence data-the superiority of deep learning over a semi-empirical model." Computers & Geosciences **150**: 104737.
Goh, J. Y., Miettinen, J., Chia, A. S., Chew, P. T. and Liew, S. C. (2013). "Biomass estimation in humid tropical forest using a combination of alos palsar and spot 5 satellite imagery." Asian Journal of Geoinformatics **13**(4): 1-10.

Grace, J., José, J. S., Meir, P., Miranda, H. S. and Montes, R. A. (2006). "Productivity and carbon fluxes of tropical savannas." Journal of Biogeography **33**(3): 387-400.
Guo, L., Sun, X., Fu, P., Shi, T., Dang, L., Chen, Y., Linderman, M., Zhang, G., Zhang, Y. and Jiang, Q. (2021). "Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas." Geoderma **398**: 115118.
Guzzo, R. A., Jackson, S. E. and Katzell, R. A. (1987). "Meta-analysis analysis." Research in organizational behavior **9**(1): 407-442.
Hajj, M. E., Baghdadi, N., Belaud, G., Zribi, M., Cheviron, B., Courault, D., Hagolle, O. and Charron, F. (2014). "Irrigated grassland monitoring using a time series of terrasar-x and cosmo-skymed x-band sar data." Remote Sensing **6**(10): 10002-10032.
Hamdan, O., Hasmadi, I. M., Aziz, H. K., Norizah, K. and Zulhaidi, M. H. (2015). "L-band saturation level for aboveground biomass of dipterocarp forests in peninsular malaysia." Journal of Tropical Forest Science: 388-399.
Haralick, R. M., Shanmugam, K. and Dinstein, I. H. (1973). "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics(6): 610-621.
Haralick, R. M., Shanmugam, K. and Dinstein, I. H. (2007). "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics(6): 610-621.
Hastie, T., Tibshirani, R. and Friedman, J. (2009). The elements of statistical learning, Citeseer.
He, Z., Zhao, W., Liu, H. and Tang, Z. (2012). "Effect of forest on annual water yield in the mountains of an arid inland river basin: A case study in the pailugou catchment on northwestern china's qilian mountains." Hydrological Processes **26**(4): 613-621.
Higgins, S. I., Bond, W. J., February, E. C., Bronn, A., Euston-Brown, D. I., Enslin, B., Govender, N., Rademan, L., O'Regan, S. and Potgieter, A. L. (2007). "Effects of four decades of fire manipulation on woody vegetation structure in savanna." Ecology **88**(5): 1119-1125.
Holmström, H. and Fransson, J. E. (2003). "Combining remotely sensed optical and radar data in k nn-estimation of forest variables." Forest Science **49**(3): 409-418.
Hong, S. H. and Wdowinski, S. (2014). "Multitemporal multitrack monitoring of wetland water levels in the florida everglades using alos palsar data with interferometric processing." IEEE Geoscience and Remote Sensing Letters **11**(8): 1355-1359.
Houghton, R., Hall, F. and Goetz, S. J. (2009). "Importance of biomass in the global carbon cycle." Journal of Geophysical Research: Biogeosciences **114**(G2).
Huang, C. y. and Geiger, E. L. (2008). "Climate anomalies provide opportunities for large-scale mapping of non-native plant abundance in desert grasslands." Diversity and Distributions **14**(5): 875-884.
Ikeda, H., Okamoto, K. and Fukuhara, M. (1999). "Estimation of aboveground grassland phytomass with a growth model using landsat tm and climate data." International Journal of Remote Sensing **20**(11): 2283-2294.
IPCC (2006). Guidelines for national greenhouse gas inventories: Agriculture, forestry and other land use. Japan, Institute for Global Environmental Strategies **4**.
Isaksson, A., Wallman, M., Göransson, H. and Gustafsson, M. G. (2008). "Cross-validation and bootstrapping are unreliable in small sample classification." Pattern recognition letters **29**(14): 1960-1965.
James, G. (2013). An introduction to statistical learning, springer.
Janecke, B. B. (2020). "Vegetation structure and spatial heterogeneity in the granite supersite, kruger national park." Koedoe **62**(2): 1-12.

Jansen, V., Kolden, C., Schmalz, H., Karl, J. and Taylor, R. (2021). "Using satellite-based vegetation data for short-term grazing monitoring to inform adaptive management." RANGELAND ECOLOGY & MANAGEMENT **76**: 30-42.
Jia, Z., Zhang, Z., Cheng, Y., Borjigin, S. and Quan, Z. (2024). "Grassland biomass spatiotemporal patterns and response to climate change in eastern inner mongolia based on xgboost model estimates." Ecological Indicators **158**: 111554.
Jiao, C., Yu, G., He, N., Ma, A., Ge, J. and Hu, Z. (2017). "Spatial pattern of grassland aboveground biomass and its environmental controls in the eurasian steppe." Journal of Geographical Sciences **27**: 3-22.
Jin, Y., Xu, B., Yang, X., Li, J., Wang, D. and Ma, H. (2011). "Remote sensing dynamic estimation of grass production in xilinguole, inner mongolia." Sci. Sin. Vitae **41**(12): 1185.1195.
Jin, Y., Yang, X., Qiu, J., Li, J., Gao, T., Wu, Q., Zhao, F., Ma, H., Yu, H. and Xu, B. (2014). "Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, northern china." Remote Sensing **6**(2): 1496-1513.
Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M. R., Kuemmerle, T., Meyfroidt, P. and Mitchard, E. T. (2016). "A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring." Remote Sensing **8**(1): 70.
Kanja, K., Zhang, C. and Atkinson, P. M. (2025). "Evaluating multi-seasonal sar and optical imagery for above-ground biomass estimation using the national forest inventory of zambia." International Journal of Applied Earth Observation and Geoinformation **139**: 104494.
Kasischke, E. S., Christensen, N. L. and Bourgeau-Chavez, L. L. (2002). "Correlating radar backscatter with components of biomass in loblolly pine forests." IEEE Transactions on Geoscience and Remote Sensing **33**(3): 643-659.
Kavzoglu, T. and Teke, A. (2022). "Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (xgboost) and natural gradient boosting (ngboost)." Arabian Journal for Science and Engineering **47**(6): 7367-7385.
Kawamura, K., Akiyama, T., Yokota, H., Tsutsumi, M., Yasuda, T., Watanabe, O., Wang, G. and Wang, S. (2005). "Monitoring of forage conditions with modis imagery in the xilingol steppe, inner mongolia." International Journal of Remote Sensing **26**(7): 1423-1436.
Kearney, S. P., Porensky, L. M., Augustine, D. J., Gaffney, R. and Derner, J. D. (2022). "Monitoring standing herbaceous biomass and thresholds in semiarid rangelands from harmonized landsat 8 and sentinel-2 imagery to support within-season adaptive management." Remote sensing of environment **271**: 112907.
Kelsey, K. C. and Neff, J. C. (2014). "Estimates of aboveground biomass from texture analysis of landsat imagery." Remote Sensing **6**(7): 6407-6422.
Kemp, D. and Michalk, D. (2007). "Towards sustainable grassland and livestock management." The Journal of Agricultural Science **145**(6): 543-564.
Kganyago, M., Adjorlolo, C., Sibanda, M., Mhangara, P., Laneve, G. and Alexandridis, T. (2022). "Testing sentinel-2 spectral configurations for estimating relevant crop biophysical and biochemical parameters for precision agriculture using tree-based and kernel-based algorithms." Geocarto International: 1-25.
Khan, K., Khan, S. N., Ali, A., Khokhar, M. F. and Khan, J. A. (2025). "Estimating aboveground biomass and carbon sequestration in afforestation areas using optical/sar data fusion and machine learning." Remote Sensing **17**(5): 934.

Klaus, V. H., Boch, S., Boeddinghaus, R. S., Hölzel, N., Kandeler, E., Marhan, S., Oelmann, Y., Prati, D., Regan, K. M. and Schmitt, B. (2016). "Temporal and small-scale spatial variation in grassland productivity, biomass quality, and nutrient limitation." Plant Ecology **217**: 843-856.
Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai, Montreal, Canada.
Kruger, A., Makamo, L. and Shongwe, S. (2002). "An analysis of skukuza climate data." Koedoe **45**(1): 87-92.
Kumar, L. and Mutanga, O. (2017). Remote sensing of above-ground biomass, MDPI. **9:** 935.
Kumar, L., Sinha, P., Taylor, S. and Alqurashi, A. F. (2015). "Review of the use of remote sensing for biomass estimation to support renewable energy generation." Journal of Applied Remote Sensing **9**(1): 097696-097696.
Kuplich, T. M., Curran, P. J. and Atkinson, P. M. (2005). "Relating sar image texture to the biomass of regenerating tropical forests." International Journal of Remote Sensing **26**(21): 4829-4854.
Lan, X., Tans, P. and Thoning, K. (2023). "Trends in globally-averaged co2 determined from noaa global monitoring laboratory measurements." (No Title).
Lan, X., Tans, P. and K.W. Thoning (2023). "Trends in globally-averaged co2 determined from noaa global monitoring laboratory measurements." Version 2023-07. Retrieved 4 August 2023, from https://gml.noaa.gov/ccgg/trends/gl_data.html.
Laurin, G. V., Chen, Q., Lindsell, J. A., Coomes, D. A., Del Frate, F., Guerriero, L., Pirotti, F. and Valentini, R. (2014). "Above ground biomass estimation in an african tropical forest with lidar and hyperspectral data." ISPRS Journal of Photogrammetry and Remote Sensing **89**: 49-58.
Li, G., Wang, J., Wang, Y., Wei, H., Ochir, A., Davaasuren, D., Chonokhuu, S. and Nasanbat, E. (2019). "Spatial and temporal variations in grassland production from 2006 to 2015 in mongolia along the china-mongolia railway." Sustainability (Switzerland) **11**(7).
Li, H., Li, F., Xiao, J., Chen, J., Lin, K., Bao, G., Liu, A. and Wei, G. (2024). "A machine learning scheme for estimating fine-resolution grassland aboveground biomass over china with sentinel-1/2 satellite images." Remote sensing of environment **311**: 114317.
Li, M., Zang, S., Zhang, B., Li, S. and Wu, C. (2014). "A review of remote sensing image classification techniques: The role of spatio-contextual information." European Journal of Remote Sensing **47**(1): 389-411.
Li, Y., Li, M., Li, C. and Liu, Z. (2020). Forest aboveground biomass estimation using landsat 8 and sentinel-1a data with machine learning algorithms. Sci rep 10: 9952.
Li, Z. and Guo, X. (2017). "Can polarimetric radarsat-2 images provide a solution to quantify non-photosynthetic vegetation biomass in semiarid mixed grassland?" Canadian Journal of Remote Sensing **43**(6): 593-607.
Li, Z. and Guo, X. (2018). "Non-photosynthetic vegetation biomass estimation in semiarid canadian mixed grasslands using ground hyperspectral data, landsat 8 oli, and sentinel-2 images." International Journal of Remote Sensing **39**(20): 6893-6913.
Liang, T., Yang, S., Feng, Q., Liu, B., Zhang, R., Huang, X. and Xie, H. (2016). "Multi-factor modeling of above-ground biomass in alpine grassland: A case study in the three-river headwaters region, china." Remote sensing of environment **186**: 164-172.
Liang, Y., Kou, W., Lai, H., Wang, J., Wang, Q., Xu, W., Wang, H. and Lu, N. (2022). "Improved estimation of aboveground biomass in rubber plantations by fusing spectral and textural information from uav-based rgb imagery." Ecological Indicators **142**: 109286.

Liao, Z., He, B. and Quan, X. (2020). "Potential of texture from sar tomographic images for forest aboveground biomass estimation." International Journal of Applied Earth Observation and Geoinformation **88**: 102049.
Liu, Y., Fan, Y., Feng, H., Chen, R., Bian, M., Ma, Y., Yue, J. and Yang, G. (2024). "Estimating potato above-ground biomass based on vegetation indices and texture features constructed from sensitive bands of uav hyperspectral imagery." Computers and Electronics in Agriculture **220**: 108918.
Liu, Y., Gong, W., Xing, Y., Hu, X. and Gong, J. (2019). "Estimation of the forest stand mean height and aboveground biomass in northeast china using sar sentinel-1b, multispectral sentinel-2a, and dem imagery." ISPRS Journal of Photogrammetry and Remote Sensing **151**: 277-289.
Liu, Z., Fu, Y. H., Chen, D., Crowther, T. W., Ding, R., Hao, F., Vitasse, Y., Zhang, X., Zhou, S. and Zohner, C. M. (2025). "Increased early-season productivity drives earlier peak of vegetation photosynthesis across the northern hemisphere." Communications Earth & Environment **6**(1): 157.
Lorber, A. and Kowalski, B. R. (1988). "The effect of interferences and calibration design on accuracy: Implications for sensor and sample selection." Journal of chemometrics **2**(1): 67-79.
Low, A. B., Rebelo, A. G. (1998). Vegetation of south africa, lesotho and swaziland: A companion to the vegetation map of south africa, lesotho and swaziland. Pretoria, South Africa., Department of Environmental Affairs and Tourism.
Lu, D. (2006). "The potential and challenge of remote sensing-based biomass estimation." International Journal of Remote Sensing **27**(7): 1297-1328.
Luo, S., Jiang, X., He, Y., Li, J., Jiao, W., Zhang, S., Xu, F., Han, Z., Sun, J. and Yang, J. (2022). "Multi-dimensional variables and feature parameter selection for aboveground biomass estimation of potato based on uav multispectral imagery." Frontiers in Plant Science **13**: 948249.
Ma, S., Baldocchi, D. D., Xu, L. and Hehn, T. (2007). "Inter-annual variability in carbon dioxide exchange of an oak/grass savanna and open grassland in california." Agricultural and Forest Meteorology **147**(3-4): 157-171.
Ma, W., Fang, J., Yang, Y. and Mohammat, A. (2010). "Biomass carbon stocks and their changes in northern china's grasslands during 1982–2006." Science China Life Sciences **53**: 841-850.
Maake, R., Mutanga, O., Chirima, G. and Sibanda, M. (2023). "Quantifying aboveground grass biomass using space-borne sensors: A meta-analysis and systematic review." Geomatics **3**(4): 478-500.
Maake, R., Mutanga, O., Chirima, J. G. and Kganyago, M. (2025). "Determining optimal sar parameters for quantifying above-ground grass carbon stock in savannah ecosystems using a tree-based algorithm." Remote Sensing in Earth Systems Sciences **8**(1): 251-263.
MacFadyen, S., Hui, C., Verburg, P. H. and Van Teeffelen, A. J. (2019). "Spatiotemporal distribution dynamics of elephants in response to density, rainfall, rivers and fire in kruger national park, south africa." Diversity and Distributions **25**(6): 880-894.
Madundo, S. D., Mauya, E. W. and Kilawe, C. J. (2023). "Comparison of multi-source remote sensing data for estimating and mapping above-ground biomass in the west usambara tropical montane forests." Scientific African **21**: e01763.
Maesano, M., Khoury, S., Nakhle, F., Firrincieli, A., Gay, A., Tauro, F. and Harfouche, A. (2020). "Uav-based lidar for high-throughput determination of plant height and above-ground biomass of the bioenergy grass arundo donax." Remote Sensing **12**(20): 3464.

Maestre, F. T. and Cortina, J. (2002). "Spatial patterns of surface soil properties and vegetation in a mediterranean semi-arid steppe." Plant and soil **241**: 279-291.
Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M. E., García-Santos, G., Fernandes, R. and Berger, M. (2012). "Sentinels for science: Potential of sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land." Remote sensing of environment **120**: 91-101.
Mandal, D., Kumar, V., Ratha, D., Dey, S., Bhattacharya, A., Lopez-Sanchez, J. M., McNairn, H. and Rao, Y. S. (2020). "Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 sar data." Remote sensing of environment **247**: 111954.
Marill, K. A. (2004). "Advanced statistics: Linear regression, part i: Simple linear regression." Academic emergency medicine **11**(1): 87-93.
Maschler, J., Keller, J., Bialic-Murphy, L., Zohner, C. M. and Crowther, T. W. (2022). "Carbon source reduction postpones autumn leaf senescence in a widespread deciduous tree." Frontiers in Plant Science **13**: 868860.
Masenyama, A., Mutanga, O., Dube, T., Bangira, T., Sibanda, M. and Mabhaudhi, T. (2022). "A systematic review on the use of remote sensing technologies in quantifying grasslands ecosystem services." GIScience & Remote Sensing **59**(1): 1000-1025.
Mavridou, A., Pappa, O., Papatzitze, O., Dioli, C., Kefala, A. M., Drossos, P. and Beloukas, A. (2018). "Exotic tourist destinations and transmission of infections by swimming pools and hot springs—a literature review." mdpi.com.
McDowell, N. G., Sapes, G., Pivovaroff, A., Adams, H. D., Allen, C. D., Anderegg, W. R., Arend, M., Breshears, D. D., Brodribb, T. and Choat, B. (2022). "Mechanisms of woody-plant mortality under rising drought, co2 and vapour pressure deficit." Nature Reviews Earth & Environment **3**(5): 294-308.
McKinley, D. C., Ryan, M. G., Birdsey, R. A., Giardina, C. P., Harmon, M. E., Heath, L. S., Houghton, R. A., Jackson, R. B., Morrison, J. F. and Murray, B. C. (2011). "A synthesis of current knowledge on forests and carbon storage in the united states." Ecological Applications **21**(6): 1902-1924.
Meng, B., Liang, T., Yi, S., Yin, J., Cui, X., Ge, J., Hou, M., Lv, Y. and 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing **13**: 2986-2995.
Mishra, N. B. and Young, K. R. (2020). Savannas and grasslands. Terrestrial ecosystems and biodiversity, CRC Press**:** 235-247.
Mistry, J. (2000). "Savannas." Progress in Physical Geography **24**(4): 601-608.
Mitchard, E. T., Saatchi, S. S., Woodhouse, I. H., Nangendo, G., Ribeiro, N., Williams, M., Ryan, C. M., Lewis, S. L., Feldpausch, T. and Meir, P. (2009). "Using satellite radar backscatter to predict above-ground woody biomass: A consistent relationship across four different african landscapes." Geophysical research letters **36**(23).
Moore, C. E., Beringer, J., Evans, B., Hutley, L. B., McHugh, I. and Tapper, N. J. (2016). "The contribution of trees and grasses to productivity of an australian tropical savanna." Biogeosciences **13**(8): 2387-2403.
Mordelet, P. and Le Roux, X. (2006). Tree/grass interactions. Lamto, Springer**:** 139-161.
Moreau, S. and Le Toan, T. (2003). "Biomass quantification of andean wetland forages using ers satellite sar data for optimizing livestock management." Remote sensing of environment **84**(4): 477-492.

Musthafa, M. and Singh, G. (2022). "Improving forest above-ground biomass retrieval using multi-sensor l-and c-band sar data and multi-temporal spaceborne lidar data." Frontiers in Forests and Global Change **5**: 822704.
Mutanga, O., Adam, E. and Cho, M. A. (2012). "High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm." International Journal of Applied Earth Observation and Geoinformation **18**: 399-406.
Mutanga, O., Masenyama, A. and Sibanda, M. (2023). "Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects." ISPRS Journal of Photogrammetry and Remote Sensing **198**: 297-309.
Mutanga, O. and Skidmore, A. K. (2004). "Narrow band vegetation indices overcome the saturation problem in biomass estimation." International Journal of Remote Sensing **25**(19): 3999-4014.
Myeong, S., Nowak, D. J. and Duggin, M. J. (2006). "A temporal analysis of urban forest carbon storage using remote sensing." Remote sensing of environment **101**(2): 277-282.
Naidoo, L., Mathieu, R., Main, R., Kleynhans, W., Wessels, K., Asner, G. and Leblon, B. (2015). "Savannah woody structure modelling and mapping using multi-frequency (x-, c-and l-band) synthetic aperture radar data." ISPRS Journal of Photogrammetry and Remote Sensing **105**: 234-250.
Naidoo, L., Mathieu, R., Main, R., Wessels, K. and Asner, G. P. (2016). "L-band synthetic aperture radar imagery performs better than optical datasets at retrieving woody fractional cover in deciduous, dry savannahs." International Journal of Applied Earth Observation and Geoinformation **52**: 54-64.
Naidoo, L., Van Deventer, H., Ramoelo, A., Mathieu, R., Nondlazi, B. and 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.
Nasirzadehdizaji, R., Balik Sanli, F., Abdikan, S., Cakir, Z., Sekertekin, A. and Ustuner, M. (2019). "Sensitivity analysis of multi-temporal sentinel-1 sar parameters to crop height and canopy coverage." Applied Sciences **9**(4): 655.
Netsianda, A. and Mhangara, P. (2025). "Aboveground biomass estimation in a grassland ecosystem using sentinel-2 satellite imagery and machine learning algorithms." Environmental Monitoring and Assessment **197**(2): 138.
Nguyen, T. T., Pham, T. D., Nguyen, C. T., Delfos, J., Archibald, R., Dang, K. B., Hoang, N. B., Guo, W. and Ngo, H. H. (2022). "A novel intelligence approach based active and ensemble learning for agricultural soil organic carbon prediction using multispectral and sar data fusion." Science of The Total Environment **804**: 150187.
Nichol, J. E. and Sarker, M. L. R. (2010). "Improved biomass estimation using the texture parameters of two high-resolution optical sensors." IEEE Transactions on Geoscience and Remote Sensing **49**(3): 930-948.
Nickmilder, C., Tedde, A., Dufrasne, I., Lessire, F., Tychon, B., Curnel, Y., Bindelle, J. and Soyeurt, H. (2021). "Development of machine learning models to predict compressed sward height in walloon pastures based on sentinel-1, sentinel-2 and meteorological data using multiple data transformations." Remote Sensing **13**(3): 408.
Numata, I., Roberts, D. A., Chadwick, O. A., Schimel, J., Sampaio, F. R., Leonidas, F. C. and Soares, J. V. (2007). "Characterization of pasture biophysical properties and the impact of grazing intensity using remotely sensed data." Remote sensing of environment **109**(3): 314-327.

Nyamugama, A. and Kakembo, V. (2015). "Estimation and monitoring of aboveground carbon stocks using spatial technology." South African Journal of Science **111**(9-10): 01-07.
O'Connor, B., Dwyer, E., Cawkwell, F. and Eklundh, L. (2012). "Spatio-temporal patterns in vegetation start of season across the island of ireland using the meris global vegetation index." ISPRS Journal of Photogrammetry and Remote Sensing **68**: 79-94.
Odebiri, O., Odindi, J. and Mutanga, O. (2021). "Basic and deep learning models in remote sensing of soil organic carbon estimation: A brief review." International Journal of Applied Earth Observation and Geoinformation **102**: 102389.
Olsson, E. G. A. and Ouattara, S. (2013). "Opportunities and challenges to capturing the multiple potential benefits of redd+ in a traditional transnational savanna-woodland region in west africa." Ambio **42**: 309-319.
Omar, H., Misman, M. A. and Kassim, A. R. (2017). "Synergetic of palsar-2 and sentinel-1a sar polarimetry for retrieving aboveground biomass in dipterocarp forest of malaysia." Applied Sciences **7**(7): 675.
Osborne, C. P., Charles-Dominique, T., Stevens, N., Bond, W. J., Midgley, G. and Lehmann, C. E. (2018). "Human impacts in african savannas are mediated by plant functional traits." New Phytologist **220**(1): 10-24.
Osman, A. I. A., Ahmed, A. N., Chow, M. F., Huang, Y. F. and El-Shafie, A. (2021). "Extreme gradient boosting (xgboost) model to predict the groundwater levels in selangor malaysia." Ain Shams Engineering Journal **12**(2): 1545-1556.
Ouchi, K. (2013). "Recent trend and advance of synthetic aperture radar with selected topics." Remote Sensing **5**(2): 716-807.
Oyewole, S. (2017). "Space research and development in africa." Astropolitics **15**(2): 185-208.
Parr, C. L., Lehmann, C. E., Bond, W. J., Hoffmann, W. A. and Andersen, A. N. (2014). "Tropical grassy biomes: Misunderstood, neglected, and under threat." Trends in ecology & evolution **29**(4): 205-213.
Parton, W., Scurlock, J., Ojima, D., Gilmanov, T., Scholes, R. J., Schimel, D. S., Kirchner, T., Menaut, J. C., Seastedt, T. and García Moya, E. (1993). "Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide." Global Biogeochemical Cycles **7**(4): 785-809.
Paruelo, J. M., Pineiro, G., Baldi, G., Baeza, S., Lezama, F., Altesor, A. and Oesterheld, M. (2010). "Carbon stocks and fluxes in rangelands of the rio de la plata basin." RANGELAND ECOLOGY & MANAGEMENT **63**(1): 94-108.
Peng, J., Zeiner, N., Parsons, D., Féret, J.-B., Söderström, M. and Morel, J. (2023). "Forage biomass estimation using sentinel-2 imagery at high latitudes." Remote Sensing **15**(9): 2350.
Piao, S., Fang, J., Zhou, L., Tan, K. and Tao, S. (2007). "Changes in biomass carbon stocks in china's grasslands between 1982 and 1999." Global Biogeochemical Cycles **21**(2).
Pillar, V. D. and Overbeck, G. E. (2025). "Nature conservation policies are biased toward forests and neglect grassy ecosystems worldwide." Science **388**(6747): eadx7441.
Pohl, C. and Van Genderen, J. L. (1998). "Review article multisensor image fusion in remote sensing: Concepts, methods and applications." International Journal of Remote Sensing **19**(5): 823-854.
Potin, P., Rosich, B., Miranda, N., Grimont, P., Shurmer, I., O'Connell, A., Krassenburg, M. and Gratadour, J.-B. (2018). Sentinel-1 constellation mission operations status. IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE.
Pritt, M. and Chern, G. (2017). Satellite image classification with deep learning. 2017 IEEE applied imagery pattern recognition workshop (AIPR), IEEE.

Prudente, V. H. R., Martins, V. S., Vieira, D. C., e Silva, N. R. d. F., Adami, M. and Sanches, I. D. A. (2020). "Limitations of cloud cover for optical remote sensing of agricultural areas across south america." Remote Sensing Applications: Society and Environment **20**: 100414.
Qadeer, A., Shakir, M., Wang, L. and Talha, S. M. (2024). "Evaluating machine learning approaches for aboveground biomass prediction in fragmented high-elevated forests using multi-sensor satellite data." Remote Sensing Applications: Society and Environment **36**: 101291.
Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H. and Sorooshian, S. (1994). "A modified soil adjusted vegetation index." Remote sensing of environment **48**(2): 119-126.
Rapiya, M., Ramoelo, A. and Truter, W. (2023). "Seasonal evaluation and mapping of aboveground biomass in natural rangelands using sentinel-1 and sentinel-2 data." Environmental Monitoring and Assessment **195**(12): 1544.
Redjadj, C., Duparc, A., Lavorel, S., Grigulis, K., Bonenfant, C., Maillard, D., Saïd, S. and Loison, A. (2012). "Estimating herbaceous plant biomass in mountain grasslands: A comparative study using three different methods." Alpine Botany **122**: 57-63.
Ren, H. and Feng, G. (2015). "Are soil-adjusted vegetation indices better than soil-unadjusted vegetation indices for above-ground green biomass estimation in arid and semi-arid grasslands?" Grass and Forage Science **70**(4): 611-619.
Ren, H. and Zhou, G. (2012). "Estimating senesced biomass of desert steppe in inner mongolia using field spectrometric data." Agricultural and Forest Meteorology **161**: 66-71.
Ricard, M. F., Berhongaray, G. and Viglizzo, E. F. (2020). The argentine pampas: A novel ecosystem at the crossroad. Earth systems and environmental sciences, Elsevier**:** 117–127.
Römkens, M. J., Helming, K. and Prasad, S. (2002). "Soil erosion under different rainfall intensities, surface roughness, and soil water regimes." Catena **46**(2-3): 103-123.
Salinas-Melgoza, M. A., Skutsch, M. and Lovett, J. C. (2018). "Predicting aboveground forest biomass with topographic variables in human-impacted tropical dry forest landscapes." Ecosphere **9**(1): e02063.
Sánchez-Azofeifa, G. A., Castro-Esau, K. L., Kurz, W. A. and Joyce, A. (2009). "Monitoring carbon stocks in the tropics and the remote sensing operational limitations: From local to regional projects." Ecological Applications **19**(2): 480-494.
Sang, H., Zhang, J., Lin, H. and Zhai, L. (2014). "Multi-polarization asar backscattering from herbaceous wetlands in poyang lake region, china." Remote Sensing **6**(5): 4621-4646.
Santi, E., Paloscia, S., Pettinato, S., Fontanelli, G., Mura, M., Zolli, C., Maselli, F., Chiesi, M., Bottai, L. and Chirici, G. (2017). "The potential of multifrequency sar images for estimating forest biomass in mediterranean areas." Remote sensing of environment **200**: 63-73.
Santos, J. R., Freitas, C. C., Araujo, L. S., Dutra, L. V., Mura, J. C., Gama, F. F., Soler, L. S. and Sant'Anna, S. J. (2003). "Airborne p-band sar applied to the aboveground biomass studies in the brazilian tropical rainforest." Remote sensing of environment **87**(4): 482-493.
Sarker, M. L. R., Nichol, J., Iz, H. B., Ahmad, B. B. and Rahman, A. A. (2012). "Forest biomass estimation using texture measurements of high-resolution dual-polarization c-band sar data." IEEE Transactions on Geoscience and Remote Sensing **51**(6): 3371-3384.
Sarrazin, M., Van Aardt, J., Asner, G., McGlinchy, J., Messinger, D. and 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**(6): 653-665.
Sarrazin, M. J. D., Aardt, J. A. N. V., Asner, G. P., Mc Glinchy, J., Messinger, D. W. and Wu, J. (2012). "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**(6): 653-665.
Schmidt, K. S. and Skidmore, A. K. (2001). "Exploring spectral discrimination of grass species in african rangelands." International Journal of Remote Sensing **22**(17): 3421-3434.
Schmidt, M., Pringle, M., Devadas, R., Denham, R. and Tindall, D. (2016). "A framework for large-area mapping of past and present cropping activity using seasonal landsat images and time series metrics." Remote Sensing **8**(4): 312.
Scholes, R. J. and Archer, S. R. (1997). "Tree-grass interactions in savannas." Annual review of Ecology and Systematics **28**(1): 517-544.
Scott, R. L., Johnston, M. R., Knowles, J. F., MacBean, N., Mahmud, K., Roby, M. C. and Dannenberg, M. P. (2023). "Interannual variability of spring and summer monsoon growing season carbon exchange at a semiarid savanna over nearly two decades." Agricultural and Forest Meteorology **339**: 109584.
Scurlock, J. and Hall, D. (1998). "The global carbon sink: A grassland perspective." Global Change Biology **4**(2): 229-233.
Shang, J., McNairn, H., Champagne, C. and Jiao, X. (2009). Application of multi-frequency synthetic aperture radar (sar) in crop classification. Advances in geoscience and remote sensing, IntechOpen.
Shannon, G., Thaker, M., Vanak, A. T., Page, B. R., Grant, R. and Slotow, R. (2011). "Relative impacts of elephant and fire on large trees in a savanna ecosystem." Ecosystems **14**: 1372-1381.
Shao, Z. and Zhang, L. (2016). "Estimating forest aboveground biomass by combining optical and sar data: A case study in genhe, inner mongolia, china." Sensors **16**(6): 834.
Sharifi, A., Mahdipour, H., Moradi, E. and Tariq, A. (2022). "Agricultural field extraction with deep learning algorithm and satellite imagery." Journal of the Indian Society of Remote Sensing: 1-7.
Shen, Z., Yong, B., Gourley, J. J., Qi, W., Lu, D., Liu, J., Ren, L., Hong, Y. and Zhang, J. (2020). "Recent global performance of the climate hazards group infrared precipitation (chirp) with stations (chirps)." Journal of Hydrology **591**: 125284.
Shiyomi, M., Okada, M., Takahashi, S. and Tang, Y. (1998). "Spatial pattern changes in aboveground plant biomass in a grazing pasture." Ecological research **13**: 313-322.
Shoko, C., Mutanga, O. and Dube, T. (2016). "Progress in the remote sensing of c3 and c4 grass species aboveground biomass over time and space." ISPRS Journal of Photogrammetry and Remote Sensing **120**: 13-24.
Shoko, C., Mutanga, O. and Dube, T. (2018a). "Determining optimal new generation satellite derived metrics for accurate c3 and c4 grass species aboveground biomass estimation in south africa." Remote Sensing **10**(4): 564.
Shoko, C., Mutanga, O., Dube, T. and Slotow, R. (2018b). "Characterizing the spatio-temporal variations of c3 and c4 dominated grasslands aboveground biomass in the drakensberg, south africa." International Journal of Applied Earth Observation and Geoinformation **68**(September 2017): 51-60.
Shoko, C., Mutanga, O., Dube, T. and Slotow, R. (2018c). "Characterizing the spatio-temporal variations of c3 and c4 dominated grasslands aboveground biomass in the drakensberg, south africa." International Journal of Applied Earth Observation and Geoinformation **68**: 51-60.
Shrestha, N. (2020). "Detecting multicollinearity in regression analysis." American Journal of Applied Mathematics and Statistics **8**(2): 39-42.

Sibanda, M., Mutanga, O. and Rouget, M. (2015). "Examining the potential of sentinel-2 msi spectral resolution in quantifying above ground biomass across different fertilizer treatments." ISPRS Journal of Photogrammetry and Remote Sensing **110**: 55-65.
Sibanda, M., Mutanga, O. and Rouget, M. (2016). "Comparing the spectral settings of the new generation broad and narrow band sensors in estimating biomass of native grasses grown under different management practices." GIScience & Remote Sensing **53**(5): 614-633.
Singh, P., Diwakar, M., Shankar, A., Shree, R. and Kumar, M. (2021). "A review on sar image and its despeckling." Archives of Computational Methods in Engineering **28**: 4633-4653.
Sinha, S., Jeganathan, C., Sharma, L. K. and Nathawat, M. S. (2015a). "A review of radar remote sensing for biomass estimation." International Journal of Environmental Science and Technology **12**: 1779-1792.
Sinha, S., Santra, A., Das, A., Sharma, L., Mohan, S., Nathawat, M., Mitra, S. S. and Jeganathan, C. (2019). "Regression-based integrated bi-sensor sar data model to estimate forest carbon stock." Journal of the Indian Society of Remote Sensing **47**(9): 1599-1608.
Sinha, S., Sharma, L., Jeganathan, C., Nathawat, M., Das, A. and Mohan, S. (2015b). "Efficacy of insar coherence in tropical forest remote sensing in context of redd." International Journal of Advancement in Remote Sensing, GIS and Geography **3**(1a): 38-46.
Sola, I., García-Martín, A., Sandonís-Pozo, L., Álvarez-Mozos, J., Pérez-Cabello, F., González-Audícana, M. and Llovería, R. M. (2018). "Assessment of atmospheric correction methods for sentinel-2 images in mediterranean landscapes." International Journal of Applied Earth Observation and Geoinformation **73**: 63-76.
Stelmaszczuk-Górska, M. A., Urbazaev, M., Schmullius, C. and Thiel, C. (2018). "Estimation of above-ground biomass over boreal forests in siberia using updated in situ, alos-2 palsar-2, and radarsat-2 data." Remote Sensing **10**(10): 1550.
Stevens-Rumann, C. S., Kemp, K. B., Higuera, P. E., Harvey, B. J., Rother, M. T., Donato, D. C., Morgan, P. and Veblen, T. T. (2018). "Evidence for declining forest resilience to wildfires under climate change." Ecology letters **21**(2): 243-252.
Su, J. and Bork, E. (2006). "Influence of vegetation, slope, and lidar sampling angle on dem accuracy." Photogrammetric Engineering and Remote Sensing **72**(11): 1265-1274.
Svoray, T. and Shoshany, M. (2002). "Sar-based estimation of areal aboveground biomass (aab) of herbaceous vegetation in the semi-arid zone: A modification of the water-cloud model." International Journal of Remote Sensing **23**(19): 4089-4100.
Svoray, T., Shoshany, M., Curran, P. J., Foody, G. M. and Perevolotsky, A. (2001). "Relationship between green leaf biomass volumetric density and ers-2 sar backscatter of four vegetation formations in the semi-arid zone of israel." International Journal of Remote Sensing **22**(8): 1601-1607.
Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C. M. and Johnson, L. (2022). "Evaluating pixel-based and object-based approaches for forest above-ground biomass estimation using a combination of optical, sar, and an extreme gradient boosting model." ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **3**: 485-492.
Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S. and Brisco, B. (2020). "Google earth engine for geo-big data applications: A meta-analysis and systematic review." ISPRS Journal of Photogrammetry and Remote Sensing **164**: 152-170.
Tanase, M. A., Panciera, R., Lowell, K., Hacker, J. and Walker, J. P. (2013). Estimation of forest biomass from l-band polarimetric decomposition components. 2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS, IEEE.

Thapa, R. B., Motohka, T., Watanabe, M. and Shimada, M. (2015). "Time-series maps of aboveground carbon stocks in the forests of central sumatra." Carbon Balance and Management **10**: 1-13.
Theron, E. J., van Aardt, A. C. and Du Preez, P. J. (2020). "Vegetation distribution along a granite catena, southern kruger national park, south africa." Koedoe: African Protected Area Conservation and Science **62**(2): 1-11.
Timothy, D., Onisimo, M. and Riyad, I. (2016). "Quantifying aboveground biomass in african environments: A review of the trade-offs between sensor estimation accuracy and costs." Tropical Ecology **57**(3): 393-405.
Tompkins, S., Mustard, J. F., Pieters, C. M. and Forsyth, D. W. (1997). "Optimization of endmembers for spectral mixture analysis." Remote sensing of environment **59**(3): 472-489.
Torres, R., Snoeij, P., Davidson, M., Bibby, D. and Lokas, S. (2012). The sentinel-1 mission and its application capabilities. 2012 IEEE International Geoscience and Remote Sensing Symposium, IEEE.
Trollope, W., Potgieter, A. and Zambatis, N. (1989). "Assessing veld condition in the kruger national park using key grass species." Koedoe **32**(1): 67-93.
Ullah, S., Si, Y., Schlerf, M., Skidmore, A. K., Shafique, M. and Iqbal, I. A. (2012). "Estimation of grassland biomass and nitrogen using meris data." International Journal of Applied Earth Observation and Geoinformation **19**: 196-204.
UNFCCC (2007). Report of the conference of the parties on its thirteenth session. Bali**:** 1-59.
Valbuena, R., Hernando, A., Manzanera, J., Görgens, E., Almeida, D., Mauro, F., García-Abril, A. and Coomes, D. (2017). "Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation." Ecological Modelling **366**: 15-26.
Van Zyl, J. J. (1993). "The effect of topography on radar scattering from vegetated areas." IEEE Transactions on Geoscience and Remote Sensing **31**(1): 153-160.
Vatandaşlar, C. and Abdikan, S. (2022). "Carbon stock estimation by dual-polarized synthetic aperture radar (sar) and forest inventory data in a mediterranean forest landscape." Journal of Forestry Research **33**(3): 827-838.
Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J. P., Lewis, P., North, P. and Moreno, J. (2019). "Quantifying vegetation biophysical variables from imaging spectroscopy data: A review on retrieval methods." Surveys in Geophysics **40**(3): 589-629.
Voormansik, K., Zalite, K., Sünter, I., Tamm, T., Koppel, K., Verro, T., Brauns, A., Jakovels, D. and Praks, J. (2020). "Separability of mowing and ploughing events on short temporal baseline sentinel-1 coherence time series." Remote Sensing **12**(22): 3784.
Wadoux, A. M.-C., Minasny, B. and McBratney, A. B. (2020). "Machine learning for digital soil mapping: Applications, challenges and suggested solutions." Earth-Science Reviews **210**: 103359.
Walsh, B., Ciais, P., Janssens, I. A., Penuelas, J., Riahi, K., Rydzak, F., Van Vuuren, D. P. and Obersteiner, M. (2017). "Pathways for balancing co2 emissions and sinks." Nature communications **8**(1): 14856.
Wang, C., Price, K. P., Merwe, D., An, N. and Wang, H. (2014a). "Modeling above-ground biomass in tallgrass prairie using ultra-high spatial resolution suas imagery." PE&RS, Photogrammetric Engineering & Remote Sensing **80**(12): 1151-1159.
Wang, C., Price, K. P., Van Der Merwe, D., An, N. and Wang, H. (2014b). "Modeling above-ground biomass in tallgrass prairie using ultra-high spatial resolution suas imagery." Photogrammetric Engineering and Remote Sensing **80**(12): 1151-1159.

Wang, D., Xin, X., Shao, Q., Brolly, M., Zhu, Z. and Chen, J. (2017). "Modeling aboveground biomass in hulunber grassland ecosystem by using unmanned aerial vehicle discrete lidar." Sensors **17**(1): 180.
Wang, G., Liu, S., Liu, T., Fu, Z., Yu, J. and Xue, B. (2019a). "Modelling above-ground biomass based on vegetation indexes: A modified approach for biomass estimation in semi-arid grasslands." International Journal of Remote Sensing **40**(10): 3835-3854.
Wang, J., Rich, P. M., Price, K. P. and Dean Kettle, W. (2005). "Relations between ndvi, grassland production, and crop yield in the central great plains." Geocarto International **20**(3): 5-11.
Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R. B. and Chang, Q. (2019b). "Estimating leaf area index and aboveground biomass of grazing pastures using sentinel-1, sentinel-2 and landsat images." ISPRS Journal of Photogrammetry and Remote Sensing **154**: 189-201.
Wang, X., Ge, L. and Li, X. (2013). "Pasture monitoring using sar with cosmo-skymed, envisat asar, and alos palsar in otway, australia." Remote Sensing **5**(7): 3611-3636.
Wang, X., Ge, L., Li, X. and Gherardi, S. (2014c). "The feasibility of using envisat asar and alos palsar to monitor pastures in western australia." Photogrammetric Engineering & Remote Sensing **80**(1): 43-57.
Wang, Y., Wang, M., Zhang, J., Wu, Y. and Zhou, Y. (2024). "Assessment of carbon stocks and influencing factors in terrestrial ecosystems based on surface area." iScience **27**(12).
Wang, Z., Ma, Y., Zhang, Y. and Shang, J. (2022). "Review of remote sensing applications in grassland monitoring." Remote Sensing **14**(12): 2903.
White, R. P., Murray, S., Rohweder, M., Prince, S. and Thompson, K. (2000). Grassland ecosystems, World Resources Institute Washington, DC, USA.
Wijesingha, J., Moeckel, T., Hensgen, F. and Wachendorf, M. (2019). "Evaluation of 3d point cloud-based models for the prediction of grassland biomass." International Journal of Applied Earth Observation and Geoinformation **78**: 352-359.
Williamson, B. D., Gilbert, P. B., Carone, M. and Simon, N. (2021). "Nonparametric variable importance assessment using machine learning techniques." Biometrics **77**(1): 9-22.
Wong, T.-T. (2015). "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation." Pattern recognition **48**(9): 2839-2846.
Wu, N., Crusiol, L. G. T., Liu, G., Wuyun, D. and Han, G. (2023). "Comparing the performance of machine learning algorithms for estimating aboveground biomass in typical steppe of northern china using sentinel imageries." Ecological Indicators **154**: 110723.
Wu, N., Zhang, C., Wei, Z., Shi, R., Zhu, F. and Liu, S. (2024). "Assessment of the impact of coastal wetland saltmarsh vegetation types on aboveground biomass inversion." Remote Sensing **16**(24): 4762.
Xie, Y., Sha, Z., Yu, M., Bai, Y. and Zhang, L. (2009). "A comparison of two models with landsat data for estimating above ground grassland biomass in inner mongolia, china." Ecological Modelling **220**(15): 1810-1818.
Xu, K., Su, Y., Liu, J., Hu, T., Jin, S., Ma, Q., Zhai, Q., Wang, R., Zhang, J. and Li, Y. (2020). "Estimation of degraded grassland aboveground biomass using machine learning methods from terrestrial laser scanning data." Ecological Indicators **108**: 105747.
Xu, L. and Baldocchi, D. D. (2004). "Seasonal variation in carbon dioxide exchange over a mediterranean annual grassland in california." Agricultural and Forest Meteorology **123**(1-2): 79-96.

Xu, Q., Man, A., Fredrickson, M., Hou, Z., Pitkänen, J., Wing, B., Ramirez, C., Li, B. and Greenberg, J. A. (2018). "Quantification of uncertainty in aboveground biomass estimates derived from small-footprint airborne lidar." Remote sensing of environment **216**: 514-528.
Xu, W., Cheng, Y., Luo, M., Mai, X., Wang, W., Zhang, W. and Wang, Y. (2025). "Progress and limitations in forest carbon stock estimation using remote sensing technologies: A comprehensive review." Forests **16**(3): 449.
Xue, Y., Li, Y., Guang, J., Zhang, X. and Guo, J. (2008). "Small satellite remote sensing and applications–history, current and future." International Journal of Remote Sensing **29**(15): 4339-4372.
Yang, C., Tian, Y., Cui, J., He, G., Li, J., Li, C., Duan, H., Wei, Z., Yan, L. and Xia, X. (2025). "Co 2 flux characteristics of the open savanna and its response to environmental factors in the dry–hot valley of jinsha river, china." Biogeosciences **22**(8): 2097-2114.
Yang, L., Meng, X. and Zhang, X. (2011). "Srtm dem and its application advances." International Journal of Remote Sensing **32**(14): 3875-3896.
Yang, Y., Fang, J., Ma, W., Guo, D. and Mohammat, A. (2010). "Large-scale pattern of biomass partitioning across china's grasslands." Global Ecology and Biogeography **19**(2): 268-277.
Yao, Y. and Ren, H. (2024). "Estimation of grassland aboveground biomass in northern china based on topography-climate-remote sensing data." Ecological Indicators **165**: 112230.
Zambatis, N., Zacharias, P., Morris, C. and Derry, J. (2006). "Re-evaluation of the disc pasture meter calibration for the kruger national park, south africa." African Journal of Range and Forage Science **23**(2): 85-97.
Zeng, N., Ren, X., He, H., Zhang, L., Zhao, D., Ge, R., Li, P. and Niu, Z. (2019). "Estimating grassland aboveground biomass on the tibetan plateau using a random forest algorithm." Ecological Indicators **102**: 479-487.
Zeng, P., Zhang, W., Li, Y., Shi, J. and Wang, Z. (2022). "Forest total and component above-ground biomass (agb) estimation through c-and l-band polarimetric sar data." Forests **13**(3): 442.
Zhang, X. and Liu, C.-A. (2023). "Model averaging prediction by k-fold cross-validation." Journal of Econometrics **235**(1): 280-301.
Zhao, D., Huang, L., Li, J. and Qi, J. (2007). "A comparative analysis of broadband and narrowband derived vegetation indices in predicting lai and ccd of a cotton canopy." ISPRS Journal of Photogrammetry and Remote Sensing **62**(1): 25-33.
Zhao, F., Xu, B., Yang, X., Jin, Y., Li, J., Xia, L., Chen, S. and Ma, H. (2014). "Remote sensing estimates of grassland aboveground biomass based on modis net primary productivity (npp): A case study in the xilingol grassland of northern china." Remote Sensing **6**(6): 5368-5386.
Zhao, X., Wu, B., Xue, J., Shi, Y., Zhao, M., Geng, X., Yan, Z., Shen, H. and Fang, J. (2023). "Mapping forage biomass and quality of the inner mongolia grasslands by combining field measurements and sentinel-2 observations." Remote Sensing **15**(8): 1973.
Zhong, B., Yang, A., Liu, Q., Wu, S., Shan, X., Mu, X., Hu, L. and Wu, J. (2021). "Analysis ready data of the chinese gaofen satellite data." Remote Sensing **13**(9): 1709.
Zhou, J.-J., Zhao, Z., Zhao, Q., Zhao, J. and Wang, H. (2013). "Quantification of aboveground forest biomass using quickbird imagery, topographic variables, and field data." Journal of Applied Remote Sensing **7**(1): 073484-073484.
Zhu, M., Feng, Q., Zhang, M., Liu, W., Qin, Y., Deo, R. C. and Zhang, C. (2019a). "Effects of topography on soil organic carbon stocks in grasslands of a semiarid alpine region, northwestern china." Journal of Soils and Sediments **19**: 1640-1650.

Zhu, N., Wang, J., Luo, D., Wang, X., Shen, C. and Wu, N. (2024). "Using eddy covariance observations to determine the carbon sequestration characteristics of subalpine forests in the qinghai–tibet plateau." Biogeosciences **21**(15): 3509-3522.
Zhu, Z., Woodcock, C. E., Rogan, J. and Kellndorfer, J. (2012). "Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using landsat and sar data." Remote sensing of environment **117**: 72-82.
Zhu, Z., Wulder, M. A., Roy, D. P., Woodcock, C. E., Hansen, M. C., Radeloff, V. C., Healey, S. P., Schaaf, C., Hostert, P., Strobl, P., Pekel, J. F., Lymburner, L., Pahlevan, N. and Scambos, T. A. (2019b). "Benefits of the free and open landsat data policy." Remote sensing of environment **224**: 382-385.
Zolkos, S. G., Goetz, S. J. and Dubayah, R. (2013). "A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing." Remote sensing of environment **128**: 289-298.
Zumo, I. M., Hashim, M. and Hassan, N. (2022). "Mapping grass above-ground biomass of grazing-lands using satellite remote sensing." Geocarto International **37**(16): 4843-4856.