Leveraging Remotely Sensed Data and Machine Learning for Quantifying Soil Organic Carbon Stocks in Woody-Encroached Areas of Bisley Nature Reserve

By

Sfundo Mthiyane

218030355

Submitted in the fulfilment of the academic requirements for the degree of

Master of Science in Environmental Sciences, in the College of Agriculture,

Engineering, and Science, University of KwaZulu-Natal

Pietermaritzburg, South Africa

November 2024

Abstract

Woody encroachment has emerged as a significant driver of land cover change in grasslands, with profound effects on Soil Organic Carbon (SOC). SOC is an important indicator of soil fertility, and thus crucial for grassland productivity. Previously, woody encroachment has been reported as a primary source of SOC alteration in grasslands. However, there are still debates and uncertainties on whether this phenomenon amplifies or reduces SOC sequestration. Therefore, it is necessary to further evaluate SOC accumulation in grasslands affected by proliferation of woody plants. Remote sensing offers freely available and cost-effective data with improved spatial and spectral resolution to quantify SOC. In this regard, the current study aimed to evaluate the role of remote sensing in quantifying the spatial variability of SOC across a woody encroached Bisley Nature Reserve. The first objective focused on quantifying the spatial variability of SOC stocks in both pristine and woody-encroached grasslands using PlanetScope spectral bands and vegetation indices. At a depth of 0-30 cm, the study found that landscapes dominated by woody encroachment exhibited higher SOC values compared to pristine grasslands. Using Deep Neural Networks, a combination of PlanetScope spectral bands and vegetation indices model achieved acceptable accuracy ($R^2 = 0.64$) for quantifying SOC stocks at this depth. Interestingly, NDVI was the most important variable for estimating SOC within woody encroached grassland. However, to fully understand the dynamics of SOC accumulation and its vertical distribution across different soil depths, it was necessary to expand the analysis. Hence the second objective extended the investigation by utilizing a Random Forest algorithm and integrating additional remotely sensed data to model SOC stocks at multiple soil depths (0-30 cm, 30-60 cm, and 60-100 cm). This approach provided a more comprehensive view of SOC variability, revealing a higher concentration of SOC in the top 30 cm compared to deeper layers. By incorporating topographic variables, Synthetic Aperture Radar (SAR), Sentinel-2, and PlanetScope data, the model produced higher accuracy for deeper soil layers, with R^2 values of 0.76 at 60 cm and 0.79 at 100 cm. SAR data enhanced the model by offering insights into subsurface conditions. These findings underscore the necessity of investigating SOC at different depths to fully capture its spatial distribution and highlight the potential of remote sensing and machine learning to improve SOC mapping accuracy across woody-encroached grasslands.

Keywords: Soil Organic Carbon (SOC); PlanetScope; Woody Encroachment; Grassland; Machine Learning

Preface

This research was conducted at the School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa, from February 2023 to October 2024, under the supervision of Prof O. Mutanga, Prof J. Odindi and Dr T.N Matongera.

I declare that this document presents the author's original work which has never been submitted to any institution as any other form qualification fulfilment. The document presents my original work, and whenever contributions of others are involved, acknowledgements are made.

Sfundo Mthiyane Signed [redacted] …… Date…19/11/2024…….

As the candidate's supervisor, I certify the mentioned statement and have approved this thesis for submission.

Supervisor(s) Signature: Prof O. Mutanga…… [redacted] .. Date:…18/11/2024.

Prof J. Odindi…… [redacted] …Date:……....18/11/2024.

Dr T.N Matongera … [redacted] ….. Date: ……19/11/2024.

Plagiarism Declaration

I Sfundo Mthiyane, declare that:

1. The research reported in this thesis, except where otherwise indicated is my original research.

2. This thesis has not been submitted for any degree or examination at any other institution.

3. This thesis does not contain other person's data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted:

   a. Their words have been re-written, and the general information attributed to them has been referenced.
   b. Where their exact words have been used, their writing has been placed in italics inside quotation marks and referenced

5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the references section.

Signed:…  …… Date…18/11/2024……..

Dedication

This dissertation is dedicated to my beloved Mother, Nomusa Mathomo Mnguni.

Acknowledgements

Firstly, I would like to thank God almighty for bringing me this far, indeed it is all because of his love and grace. I would like to appreciate Cave of Adullum Apostolic Center, my church, for the prayers and support. To my family, you have always believed in me, thank you for your support and prayers.

I would like to thank the School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, for providing a platform for me to pursue my studies.

I would like to send my heart-felt appreciation to Dr Trylee Matongera for believing and trusting in me. Thank you for your guidance, I am grateful for the opportunity and know that under your supervision, I will continue to grow and excel in my career.

To my supervisors, Prof O Mutanga, Prof J Odindi, and Dr T Matongera, I would like to send my gratitude and appreciation for your guidance, assistance and supervision throughout this dissertation. Thank you for trusting and believing in me, your knowledge and expertise have contributed greatly to the completion of this research. Thank you, Prof Mutanga, for all your financial support, this has played a huge role in the completion of this research in record time.

I would like to return my appreciation to my friends and colleagues, Celuxolo Dlamini, Ntuthuko Mncwabe, Bheka Mlambo, Sinethemba Ndlovu, Siphamandla Maseko, Kuhlekonke Mathenjwa, Wandile Khumalo, and Mbali Manyaka, for their assistance with field work and data collection. I would like to thank you also for your encouragement, emotional and financial support. I would like to thank Dr Collins Matiza for assistance with python coding.

I would like to extend my appreciation to the department senior students and staff, namely, Anita Masenyama, Dr Mthembeni Mngadi, Mr. Brice Gijbertsen, Ms. Andile Mshengu, Adeola Arougundade, Ms. Shanitha Ramroop, Dr Nel, Prof Hill, Prof Jemma, and Mr. Donavin Devos, for their valuable administrative, logistics, and technical assistance.

Table of Contents

List of Tables

List of Figures

Abstract

Woody encroachment has emerged as a significant driver of land cover change in grasslands, with profound effects on Soil Organic Carbon (SOC). SOC is an important indicator of soil fertility, and thus crucial for grassland productivity. Previously, woody encroachment has been reported as a primary source of SOC alteration in grasslands. However, there are still debates and uncertainties on whether this phenomenon amplifies or reduces SOC sequestration. Therefore, it is necessary to further evaluate SOC accumulation in grasslands affected by proliferation of woody plants. Remote sensing offers freely available and cost-effective data with improved spatial and spectral resolution to quantify SOC. In this regard, the current study sought to evaluate the role of remote sensing in quantifying the spatial variability of SOC across a woody encroached Bisley Nature Reserve. The first objective focused on quantifying the spatial variability of SOC stocks in both pristine and woody-encroached grasslands using PlanetScope spectral bands and vegetation indices. At a depth of 0-30 cm, the study found that landscapes dominated by woody encroachment exhibited higher SOC values compared to pristine grasslands. Using Deep Neural Networks, a combination of PlanetScope spectral bands and vegetation indices model achieved acceptable accuracy ($R^2 = 0.64$) for quantifying SOC stocks at this depth. Interestingly, NDVI was the most important variable for estimating SOC within woody encroached grassland. However, to fully understand the dynamics of SOC accumulation and its vertical distribution across different soil depths, it was necessary to expand the analysis. Hence the second objective extended the investigation by utilizing a Random Forest algorithm and integrating additional remotely sensed data to model SOC stocks at multiple soil depths (0-30 cm, 30-60 cm, and 60-100 cm). This approach provided a more comprehensive view of SOC variability, revealing a higher concentration of SOC in the top 30 cm compared to deeper layers. By incorporating topographic variables, Synthetic Aperture Radar (SAR), Sentinel-2, and PlanetScope data, the model produced higher accuracy for deeper soil layers, with R^2 values of 0.76 at 60 cm and 0.79 at 100 cm. SAR data enhanced the model by offering insights into subsurface conditions. These findings underscore the necessity of investigating SOC at different depths to fully capture its spatial distribution and highlight the potential of remotely sensed data and machine learning to improve SOC mapping accuracy across woody-encroached grasslands.

Keywords: Soil Organic Carbon (SOC); PlanetScope; Woody Encroachment; Grassland; Machine Learning

Chapter One: General Introduction

1.1 Introduction

Soils are major reservoirs of terrestrial organic carbon, sequestrating approximately 1500 Pg of global Soil Organic Carbon (SOC), equivalent to the sum total of the biotic pool of 600 Pg and atmospheric pool of 750 Pg (McGrath & Zhang, 2003). Hence, a small alteration in SOC pool can potentially influence the atmospheric carbon pool and thus contribute to climate change and global warming (Lamichhane et al., 2019; Odebiri, Mutanga, & Odindi, 2022). SOC stocks are largely stored in numerous terrestrial ecosystems around the globe, including grasslands. Grassland soils are linked with high SOC sequestration triggered by high above and below ground litter rate and decomposition (Jackson et al., 2007). SOC is a key indicator of grassland fertility and quality, mitigating greenhouse gas emissions and promoting plant productivity (Gerke, 2022a). Relatively, alteration of SOC is classified as soil degradation and can have severe implications to grassland productivity. In grasslands, woody encroachment is a documented major land-cover change that affect SOC distribution, particularly in tropic and subtropical regions (Liu et al., 2011; O. W. Van Auken, 2009). In South Africa, woody encroachment is prominent and affects approximately 10-20 million hectares of the terrestrial ecosystems, about 70% being grasslands (Venter et al., 2021). Given the extent of woody encroachment into South African grasslands, particularly within protected areas, there is growing interest in understanding how this phenomenon can influence climate systems and regional biogeochemical cycles through its impact on SOC stocks.

The dominance of woody encroachment has significantly contributed to the degradation of grasslands, thus altering SOC distribution. Previous studies have discovered that woody proliferation in grasslands often alters SOC stocks, however, there is no consensus regarding whether woody encroachment amplify or reduce SOC concentration (Jackson et al., 2007; Jackson et al., 2002; Li et al., 2016). Interestingly, according to Li et al. (2016),landscapes dominated by woody encroachment tend to exhibit high SOC values compared to those dominated by pristine grasses. In contrast, Qiu et al. (2012) and Coetsee et al. (2013) established that areas exposed to woody encroachment are subjected to low SOC distribution compared to undisturbed grasses. These inconsistencies clearly indicate a significant regional variation in the influence of woody proliferation on SOC stocks (Li et al., 2016). Hence, the extent at which woody encroachment affect SOC largely remains unclear. Although it is widely believed that SOC is more concentrated at a depth of 30 cm, Lorenz and Lal (2005) notes that

there is a chance for improved sequestration of SOC in deeper soil depths within woody encroached landscapes. Information on the dynamics of SOC accumulation in deeper soil horizons is therefore necessary for better contextualization of the influence of woody proliferation to grassland SOC.

As aforementioned, grasslands are one of the largest sequesters of SOC. Also, SOC is crucial for grassland productivity and climate change mitigation (Odebiri, Mutanga, & Odindi, 2022). Changes to SOC accumulation in grasslands can potentially affect carbon storage and productivity. Due to the inconstancies in past literature regarding the extent at which woody encroachment influence SOC, there is a need for comprehensive evaluation of SOC after proliferation of woody plants. This information is crucial for accurate monitoring of SOC in grasslands affected by woody encroachment. SOC records in grasslands are vital for tracking vegetation productivity at an ecosystem level. Previously, studies evaluating SOC distribution in deeper soils of woody encroached grasslands have largely utilized traditional methods that include direct field measurements and wet soil analysis (Angelopoulou et al., 2019; Li et al., 2016; Mureva et al., 2018). Unfortunately, traditional methods are costly, labour-intensive, time-consuming and difficult to execute over large heterogeneous spatial extents. Therefore, cost-efficient and reliable methods are needed for improved SOC monitoring. Remote sensing is well established as a reproducible, rapid, efficient, and environmentally friendly means of quantifying SOC within grasslands affected by woody encroachment (Odebiri et al., 2020; Zhou et al., 2020). Due to cutting-edge spectral information associated with remotely sensed data, there is potential for an improved quantification of SOC in grasslands affected by woody proliferation (Bhunia et al., 2019; Odebiri, Mutanga, & Odindi, 2022). The improved spatial resolution of these satellites minimizes challenges that are linked with mixed pixels, soil-background interference, and shadow casting, all of which can potentially affect the accuracy of SOC mapping in complex heterogeneous grasslands. Additionally, widely accessible medium-resolution sensors, such as PlanetScope, with their finer spectral and spatial resolutions, have a potential to estimate SOC in grasslands (Koparan, 2019).

PlanetScope is a high spatial resolution multispectral sensor with advanced sensing capabilities that have not been fully explored for assessing SOC stocks accumulation within woody encroached grasslands (Andreatta et al., 2022). The sensor is characterized by eight spectral bands, including red-edge region of the electromagnetic spectrum. The region is highly sensitive to vegetation biomass which in turn inform the distribution of SOC (Li et al., 2020; Matiza et al., 2024). Additionally, PlanetScope acquires spectral data at a high (3m) spatial resolution that is critical for quantifying SOC stocks at an ecosystem level. Spectral vegetation indices derived from PlanetScope data promise to improve SOC modelling (Koparan, 2019; Tan et al., 2023). However, the rich information of the sensor has not been fully explored when estimating SOC accumulation across woody encroached grasslands. Furthermore, the evaluation of SOC accumulation at deeper soil depths is necessary for improved understanding of the implication of woody encroachment in grasslands. Consequently, the aforementioned mentioned characteristics of PlanetScope imagery present an opportunity for further evaluation of SOC assimilation at deeper soil horizons of a woody encroached grasslands. Therefore, the current study aims at evaluating remote sensing applications for quantifying the spatial variability of SOC across a woody encroached Bisley Nature Reserve.

1.2 Aims and Objectives

The overall aim of the research was to apply remote sensing techniques to quantify the spatial distribution of SOC stocks across a woody-encroached grassland.

The following specific objectives were set:

- To quantify the spatial variability of SOC across a pristine and woody encroached grassland using PlanetScope image data.
- To estimate soil organic carbon (SOC) at multiple depths (30 cm, 60 cm, and 100 cm) in woody-encroached grasslands by integrating Sentinel-1 (S1), Sentinel-2 (S2), PlanetScope (PS) satellite imagery, and topographic variable.

1.3 Significance of the Study

In grasslands, SOC is essential for stabilizing the structure of the soil, improving water holding capacity, and a paramount indicator of soil fertility. Interestingly, soils with elevated SOC amounts are associated with high nutrient enrichment and improved vegetation productivity. Therefore, any alternation of SOC accumulation in grasslands could potentially affect the soil

nutrient enrichment, thus affecting productivity. Woody encroachment is a well-documented phenomenon affecting grassland vegetation. Notably, grasslands are degrading due to proliferation of woody plants. Due to the debates that exist in the literature about the implications of woody encroachment on SOC, there is a necessity for further assessment of SOC after woody invasion. Also, employing remote sensing and machine learning techniques can immensely improve our knowledge and understanding of SOC distribution in woody encroached grasslands. Additionally, improved capabilities of remote sensing technology promise to yield accurate and trustworthy results explaining SOC stocks distribution in grasslands affected by woody invasion. This study will enhance our understanding on SOC accumulation at deeper soil depths of a woody encroached grassland. It will provide critical information about SOC distribution after woody encroachment, that can be utilized by environmental managers to make informed decisions on the conservation of Bisley Nature Reserve.

1.4 Description of Study Site

The study was conducted at Bisley Valley Nature Reserve (29° 39' 53" S; 30° 23' 32" E), approximately 7.8 km Southwest of Pietermaritzburg Central Business District (Figure 1.1). The reserve is approximately 3.5 km^2 and is covered by the Ngongoni Veld and Hinterland Thornveld of the sub-escarpment savanna bioregion in KwaZulu-Natal (Kraai et al., 2023). The nature reserve experiences an average annual precipitation of 694 mm, and hot summers and mild winters with an average temperature of 26 °C and17 °C, respectively. The flora in the reserve includes *Vachellia nilotica* and *Vachellia sieberiana* trees, along with common grasses such as *Eragrostus curvula* and *Panicum maximum* (Ward et al., 2017). The nature reserve has wildlife that includes zebra, giraffe, impala, and various bird species. The area's terrain is characterized by uneven topography with altitude ascending from 700 m to 830 m above sea level and predominated by the dolerite rock type with some shale outcrops, especially in the low-lying areas. The study area is characterized by a variety of soil types, namely Mispah, Dundee, Glenrosa, Clovely, and Avalon (Jaganyi, 1998). Among these, Mispah soils predominate, covering most of the landscape. Mispah soils are typically shallow and rocky, often forming over hard rock or partially weathered material, which limits their effective depth (Otto, 2019). This shallow depth contributes to low organic matter content, as the restricted rooting zone constrains the accumulation and retention of organic residues (Rossouw, 2016; Van Huyssteen et al., 2021). Furthermore, these soils exhibit moderate vulnerability to erosion, particularly under improper land management or when protective vegetation cover is reduced

(Bühmann et al., 2006). In contrast, the Dundee, Glenrosa, Clovely, and Avalon soils, although present, are less extensive. They vary in their mineral composition, texture, and organic matter content, potentially creating small-scale heterogeneity across the landscape (Gaston et al., 1996; Moberly & Meyer, 1984). For instance, Glenrosa and Clovely soils often exhibit slightly higher organic content and better drainage characteristics than Mispah soils, while Avalon soils tend to be deeper and can support more robust vegetation under favourable conditions (Zablotowicz et al., 2000). However, due to the prevalence of Mispah soils, the overall capacity of the area to store and maintain soil organic carbon is limited, making soil organic carbon an important indicator of land degradation or improvement in local land management. In the past 10 years, no formal management practices have been applied in the nature reserve to maintain the grass dominance. The negligence of the reserve has favoured the dominance of woody vegetation which has significantly altered the soil's physical and chemical characteristics, including SOC stock distribution. However, in the recent past, several control practices have been adopted to minimize the dominance of woody vegetation in the reserve. Hence, whereas most landscapes are dominated by woody vegetation, there are several pristine grassland patches across the reserve.

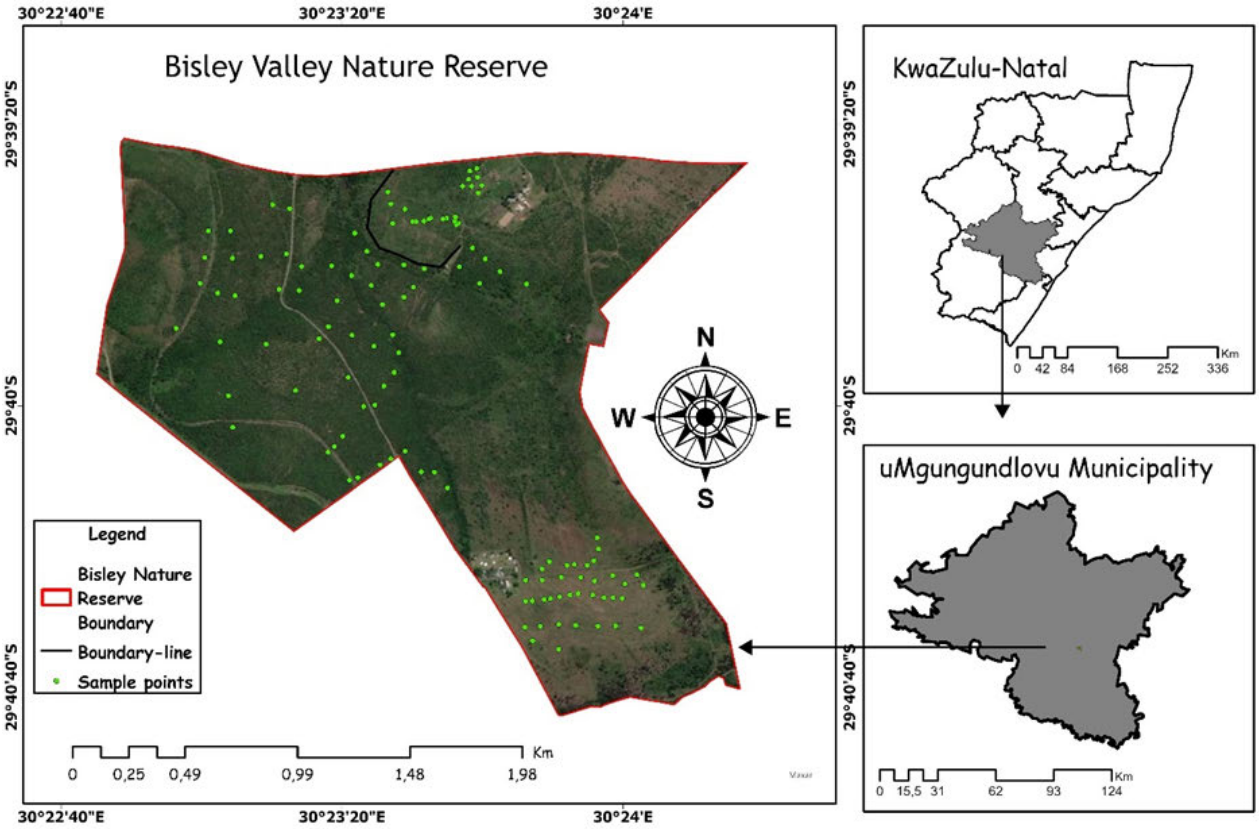

Figure 1.1: Location of Bisley Nature Reserve in KwaZulu Natal Province of South Africa.

1.4 Structure of the Thesis

The dissertation includes four chapters. In chapter One a background, that outlines the aim, main objectives, research questions, hypothesis and the significance of this research, is outlined. Chapter Two is based on quantifying the spatial variability of Soil Organic Carbon (SOC) across a woody encroached Bisley Nature Reserve using PlanetScope data. The study highlights the robustness of remote sensing techniques in quantifying SOC variability within woody encroached grasslands. The chapter compares SOC stocks variability between woody encroached and pristine grassland patches within Bisley Nature Reserve. Chapter Three focuses on evaluating the spatial distribution of SOC at different soil depths of woody encroached grassland by integrating SAR, Sentinel 2 and PlanetScope data. Chapter Four is a synthesis of the whole thesis that highlights important issues addressed by study, literature gaps, findings of the research, closing remarks, and recommendations for future research.

Chapter Two: Quantifying the influence of woody encroachment on soil organic carbon across Bisley Nature Reserve using PlanetScope data

This chapter is based on:

Mthiyane, S., Mutanga, O., & Matongera, T. N., Odindi, J. (Submitted for a book chapter). Quantifying the influence of woody encroachment on soil organic carbon across Bisley Nature Reserve using PlanetScope data.

Abstract

Globally, grasslands constitute one of the largest land ecosystems that sequestrate substantial amounts of Soil Organic Carbon (SOC) and provide various ecological services such as regulating the climate, improving water quality, and supporting biodiversity. Woody encroachment has pronounced impacts on grassland ecosystems, exerting substantial influence on their structure, ecological dynamics, and SOC distribution. To the best of our knowledge, concerns and discrepancies about the influence of woody proliferation on SOC within protected grasslands remain largely underexplored. Remote sensing offers cost-effective, time-efficient, and environmentally friendly means to quantify SOC distribution and variability in grasslands. Consequently, this study sought to quantify the spatial variability of SOC across a pristine and woody encroached grassland using PlanetScope image data. Employing a stratified sampling technique, 254 samples were collected, and the Loss-on-Ignition procedure adopted to determine SOC. The PlanetScope spectral bands and derived vegetation indices were utilized to quantify SOC variability between woody encroached and pristine grassland. The Deep Neural Network (DNN) model was trained based on 10-fold cross-validation with SOC values and 20 predictors. To improve the prediction accuracy of the model, hyper-parameter tuning was employed. The findings indicate a higher concentration of SOC in woody encroached vegetated areas compared to pristine grasslands. The DNN model produced acceptable accuracies with Root Mean Square Error (RMSE) of 1.90 t/ha and R^2 value of 0.64. The study provides a framework for continuous monitoring of SOC stocks in protected grasslands.

Keywords: Grassland, Woody Encroachment, Soil Organic Carbon, PlanetScope, Deep Neural Network

2.1 Introduction

Grasslands comprise over 40% of the Earth's terrestrial ecosystems worldwide (Wang & Fang, 2009). As the predominant ecosystem in temperate regions (Loveland & Webb, 2003; Merino et al., 2004), grasslands are essential for providing unique and significant ecological services, such as reducing soil erosion, preserving biodiversity, and enhancing food production (Bugalho & Abreu, 2008; Carlier et al., 2009). Furthermore, grasslands constitute one of the most significant repositories of Soil Organic Carbon (SOC), crucial for providing various ecological services such as regulating the climate, improving water quality, and supporting biodiversity (Sommer & Bossio, 2014; Vågen & Winowiecki, 2013). In grasslands, SOC helps to improve soil structure, retain water, and provide plant nutrients (Conant & Paustian, 2002; Gerke, 2022b); characteristics that are critical for increasing the resilience of soils towards droughts and floods. In this regard, a decline in SOC can trigger a loss of soil quality and productivity. The accumulation and loss of SOC in grasslands are primarily influenced by land cover characteristics and change (Ratajczak et al., 2012; Wieczorkowski & Lehmann, 2022). Woody proliferation is a natural phenomenon that has been dominant in several grassland ecosystems across the globe for the past 200 years (Browning et al., 2008; Saintilan & Rogers, 2015). According to Ratajczak et al. (2012) and Ge and Zou (2013), the proliferation of woody plants is a key phenomenon causing landscape degradation and significantly affect the ecology, land surface, and soil biophysical characteristics of grasslands.

Understanding the trade-offs and adverse influence of woody encroachment is important because various natural grasslands across South Africa are degrading due to woody proliferation (Ratajczak et al., 2012; O. W. Van Auken, 2009). Woody proliferation affects about 15 million hectares of grasslands in South Africa (Mureva et al., 2018). The study by Gartzia et al. (2014) notes that invasion of woody plants amplifies water and nutrient balance changes at various spatial and temporal extents, which are primary resources for numerous animal species thriving in grasslands. The study by Ratajczak et al. (2012) showed that the proliferation of woody plants has a negative influence on species (grazers) abundance in grassland ecosystems, particularly in areas receiving high precipitation. Their study observed a decline in species richness and abundance in grasslands affected by woody proliferation. According to Mochi et al. (2022), woody encroachment minimizes food source availability for numerous grazers within the grassland. Although the productivity of herbivore biomass can be maintained in areas influenced by woody proliferation, the herbivore assemblage re-structuring causes intense changes in their average body size (Smit & Prins, 2015). Moreover, woody

proliferation reduces the amount of water available in the soil and promotes surface runoff, which may potentially result in seasonal drought, consequently influencing grassland productivity (Schreiner-McGraw et al., 2020).

Studies reveal that a transition from a grass-dominated to a woody-dominated ecosystem may potentially impact SOC sequestration (Li et al., 2016; Mureva et al., 2018). Jackson et al. (2002) observed that the increase of woody plants significantly enhances the levels of soil organic carbon (SOC) stocks in grassland areas, both regionally and globally. However, Li et al. (2016) note that there are still uncertainties on the potential influence of woody proliferation on SOC at local scales. Whereas Alberti et al. (2011) and Jackson et al. (2002) suggest that woody encroachment increases the concentration of SOC stocks at an ecosystem level, others (Mureva et al., 2018; Zhou et al., 2017) indicate that encroachment of woody plants reduces SOC stocks in grasslands at a local scale. These inconsistencies can be linked to various environmental factors that include historical land-use patterns and soil biophysical attributes (Bahri et al., 2022). Furthermore, the proliferation of woody vegetation could amplify the spatial heterogeneity of soil characteristics (Liu et al., 2011), complicating SOC monitoring in grasslands. Consequently, there is a necessity for further evaluation of the relationship between SOC stocks and woody encroachment at an ecosystem level. Additionally, it is essential to create a reliable approach for measuring SOC stocks in localized grasslands. This is useful to environmental managers in formulating effective environmental management strategies and policies.

Historically, SOC monitoring relied primarily on conventional methods that includes wet soil analysis, laboratory procedures, and field measurements (Angelopoulou et al., 2019; Mzinyane et al., 2015; Odebiri et al., 2020). Although these methods offer good accuracy, drawbacks that include time-consuming processes, intensive labour, and high costs make their adoption challenging at large spatial extents (Bartholomeus et al., 2011; Hernandez-Clemente et al., 2023; Wang et al., 2018; Xu & Zhai, 2023). Recognizing these shortcomings, there is need for rapid, reliable and cost-effective strategies for quantifying and monitoring SOC in grasslands (Eggleston et al., 2006; Odebiri et al., 2020). Hence, remote sensing technology has become a reliable tool offering cost-effective, time-efficient, and environmentally friendly approaches to assess SOC variability at various spatial extents (Guo et al., 2021; Odebiri, Mutanga, Odindi, et al., 2022). Notably, new generation commercial sensors including WorldView series have been extensively utilized for SOC monitoring in different ecological systems (Broderick et al., 2015; Thaler et al., 2019). These sensors consist of unique band settings that are strategically

positioned within the red-edge region crucial for improved vegetation spectral sensitivity (Rapinel et al., 2014). Unfortunately, despite their robustness in quantifying SOC, these sensors may not be appropriate for monitoring SOC stocks in developing regions such as Southern Africa (Mngadi et al., 2021). This is because these sensors are associated with small swath width, high acquisition cost and are not readily available (Matongera et al., 2017). Consequently, SOC monitoring is still not clearly understood in many heterogeneous landscapes, particularly in Sub-Saharan grasslands. These limitations have necessitated the adoption of freely accessible multispectral sensors that provide improved spatial, spectral, and temporal resolution characteristics, which are critical for accurately quantifying SOC (Bhunia et al., 2019; Mutanga et al., 2016; Odebiri et al., 2020). These advancements promise to revolutionize SOC monitoring, offering vital insights for sustainable land management practices.

Global interest in remote sensing methods for SOC research has been sparked by new innovations in reliable, open-source sensor technology, like the PlanetScope Multispectral sensor (Tan et al., 2023; K. Wang et al., 2021). The sensor has gained popularity for mapping and assessing SOC stocks due to its cutting-edge sensing properties (Andreatta et al., 2022; M. H. Koparan et al., 2022; Marta, 2018). It consists eight spectral bands of the electromagnetic spectrum including the red-edge regions that are particularly responsive to biophysical characteristics of green vegetation (Frazier & Hemingway, 2021; Roy et al., 2021). According to Odebiri et al. (2020), green biomass is a key indicator of SOC stocks as its accumulation is highly affected by its litter and decomposition rate. Studies have utilized red-edge wavebands to assess and predict SOC stocks at various landscapes (Castaldi et al., 2019; M. H. Koparan et al., 2022; Xie et al., 2022). Furthermore, the PlanetScope satellite is characterized by an exceptionally fine spatial resolution of 3 meters (Francini et al., 2020). The sensor's finer pixel size is valuable for capturing finer details, which is crucial for mapping the variability of SOC, particularly at small spatial extents. While the characteristics of PlanetScope data to SOC mapping are crucial, its heterogeneity and complexity call for a fast and reliable analytical technique that can successfully quantify SOC stocks to various extents are adopted. To date, Deep Learning (DL) techniques have gained popularity in numerous remote sensing studies attributed to their ability to handle big and complex datasets (Abdar et al., 2021; Zhang et al., 2019; Zhou et al., 2017). The DL models can automatically extract abstract and invariant components from remotely sensed data, hence have been widely adopted to map complex non-linear data, including SOC stocks (Zhang et al., 2019).

Although numerous studies evaluating the influence of woody proliferation on SOC at a national and regional scale have employed diverse remote sensing techniques, information about the influence of this phenomenon on SOC at an ecosystem level is still lacking, particularly in protected areas (Jackson et al., 2002; Odebiri et al., 2021). Moreover, literature has revealed the success of spectral vegetation indices derived from the red-edge band region of PlanetScope sensor in quantifying SOC stocks in complex landscapes. Previous studies suggest that sensors with explicit spatial and spectral resolution may offer enhanced information, essential for regular monitoring of SOC in grasslands affected by woody proliferation (Tan et al., 2023). Therefore, the finer spatial and spectral characteristics, including red-edge region of freely available PlanetScope are critical for SOC quantification within woody encroached landscapes with limited spatial extent. Furthermore, current uncertainties that exist in the literature about the influence of woody proliferation on SOC pool advocates for further assessment of the relationship that exist between SOC and woody encroachment. This can be achieved by assessing the variability of SOC stocks between pure grasses and encroached grasslands. Therefore, this study sought to quantify the spatial variability of SOC stocks between a pristine and woody encroached grassland using PlanetScope spectral bands and vegetation indices.

2.2 Materials and Methods

2.2.1 Field data collection

Soil samples were collected across the study area using a stratified random sampling approach from September 21st and 23rd, 2023 (during the spring season). Prior knowledge of the site facilitated the delineation of two homogeneous strata, namely pristine grassland and woody-encroached grassland (Figure 2.1). Within the woody-encroached stratum, 15 transects were established, each measuring 110 m in length. Along each transect, eight sampling plots (5 m² each) were marked at 10 m intervals, resulting in a total of 120 sampling plots in that stratum. At each plot, soil samples were collected at a depth of 30 cm using a soil auger, following the procedure recommended by Vågen and Winowiecki (2013). The same sampling protocol was applied in the pristine grassland stratum, ensuring consistency in data collection. A Trimble Handheld GPS with sub-meter accuracy was used to record the coordinates of each sampling point for subsequent spatial analysis. In total, 240 soil samples were collected and transported to the laboratory for further analysis of soil organic carbon and other relevant parameters.

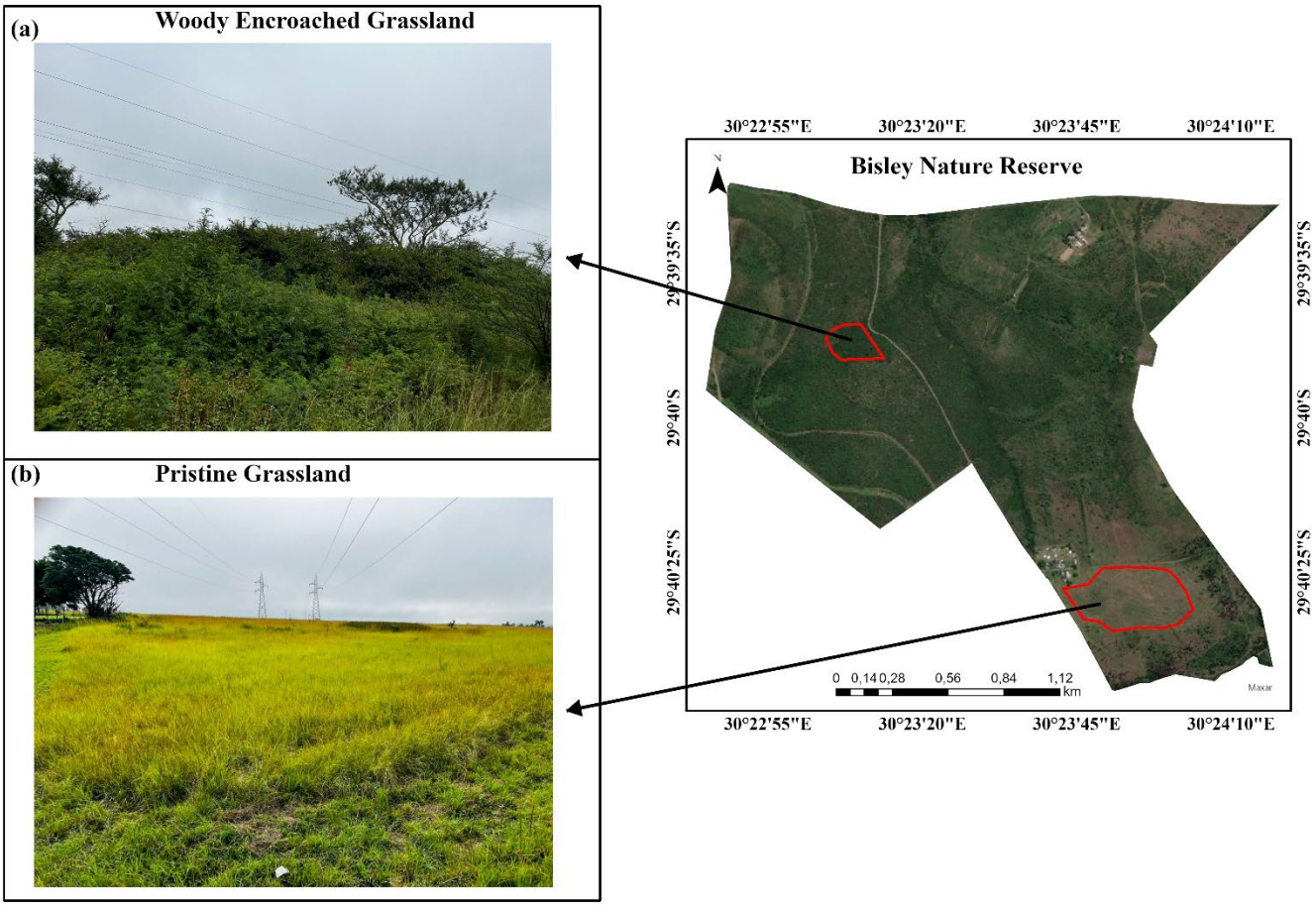

Figure 2.1: Imagery of a woody encroached (a) and pristine (b) grasslands in Bisley Nature Reserve.

Furthermore, to clearly visualize encroached and pristine landscapes within Bisley Nature Reserve, a supervised landcover classification was executed through Google Earth Engine (GEE) platform. The study site was separated into three landcover classes namely, woody encroached grassland, pristine grassland and other landcover classes that include buildings, roads, water bodies and bare soils. A total of 65 presence locations were collected for each landcover class using the Trimble Handheld GPS with sub-meter accuracy, 70% of the collected points were used for training the model and 30% for testing.

2.2.2 Laboratory analysis

In the laboratory, the soil was air-dried for three days to eliminate excess water and sieved through a 2-mm sieve. Subsequently, the samples were oven-dried at 105 °C for 24 hours (Konen et al., 2002). Then, the soil samples were analysed to determine the SOC concentration. The Loss-On-Ignition technique was adopted to determine SOC as described by Schulte and Hopkins (1996). A muffle furnace (Model Exation 1200-30 L) was used to combust soil samples at the temperature of 360 ◦C for two hours. An estimation of the SOC percentage from the Losson-On-Ignition method (SOCLOI) was calculated using equation 1 Schulte and Hopkins (1996).

$$SOCLOI = \left[\frac{soil\ weight\ before\ ignition - soil\ weigh\ after\ ignition}{soil\ weight\ before\ ignition}\right] * 100 \quad \text{..... Equation (1)}$$

2.2.3 Image data acquisition and pre-processing

A PlanetScope Multispectral image of the area with less than 10% cloud cover was acquired on the 1st of September 2023 from the Planet Labs Explorer (https://www.planet.com). The study utilized one PlanetScope Ortho Scene product at Level 3B. Product Level 3B data are geometrically corrected by sensor telemetry and scaled, orthorectified, projected, and modelled to surface reflectance by the vender (Žížala et al., 2019). Moreover, Planet offers two primary radiometry options for PlanetScope scenes, i.e., visual, and analytic products. Analytic products are available with the option for delivery as scaled top of atmosphere radiance or surface reflectance and are highly recommended for modelling applications (Roy et al., 2021). They consist of eight spectral bands namely, Coastal Blue (431-452 nm), Blue (465-515 nm), Green 1 (513-549 nm), Green 2 (547-583 nm), Yellow (600-620 nm), Red (650-680 nm), Red-

Edge (697-713 nm), and Near infrared (845-885 nm), with a spatial resolution of 3 m (Zhao & Liu, 2022). The sensor covers the Red-Edge region of the electromagnetic spectrum, which is critical for SOC stock monitoring (Andreatta et al., 2022). Moreover, PlanetScope is characterized by a daily temporal resolution at nadir, which is essential for frequent monitoring of the SOC, especially for localized landscapes (Shimizu et al., 2020). Subsequently, the PlanetScope spectral bands, integrated with vegetation indices, were utilized as predictor variables.

1.2.4 Image classification

Furthermore, PlanetScope image was utilized for conducting landcover classification. The satellite images were classified using Random Forest classier. The RF is an ensemble learning technique that quantifies a response variable through combining multiple decision trees. According to Odebiri et al. (2020), the RF classifier is time efficient and requires less manual intervention when compared to other classifiers such as Support Vector Machine.

2.2.5 Spectral vegetation indices

The study employed various spectral band combinations derived from PlanetScope imagery to generate widely used spectral vegetation indices for the quantification of SOC, as outlined in Table 2.1. Evidence from prior literature supports the utility of vegetation indices for enhancing model performance in quantifying the accumulation of SOC across diverse landscapes (Odebiri, Mutanga, & Odindi, 2022; Odebiri et al., 2020; K. Wang et al., 2021). Therefore, the derived vegetation indices were integrated with corresponding spectral bands to develop an optimized model for the accurate quantification of SOC.

Table 2.1: Spectral indices derived from PlanetScope and their formulae.

Indices	Formulae	References
Normalized Difference Vegetation Index (NDVI)	(NIR-RED)/(NIR+RED)	(Rouse et al., 1974)
Enhance Vegetation Index (EVI)	2.5*((NIR-RED)/(NIR+6*RED-7.5BLUE+1))	(Huete et al., 1999)
Modified Soil Adjusted Vegetation Index (MSAVI)	2*((NIR+1-$\sqrt{(2 * NIR +)^2 - 8(NIR - RED)}$)/2	(Qi et al., 1994)
Green Normalized Difference Vegetation Index (GNDVI)	(NIR-GREEN)/(NIR+GREEN)	(Ahamed et al., 2011)
Coloration Index (CI)	(RED - BLUE)/RED	(Escadafal et al., 1994)

Photosynthetic Vigour Ratio (PVR)	GREEN − RED/GREEN + RED	(Metternicht, 2003)
Greenness Index (GI)	GREEN/RED	(Main et al., 2011)
Anthocyanin Reflectance Index (ARI)	(1/GREEN) − (1/RED_EDGE)	(Gitelson et al., 2001)
Transformed Vegetation Index (TVI)	$\sqrt{(NDVI) + 0.5}$	(Hunt Jr et al., 2011)
Renormalized Difference Vegetation Index (RDVI)	(NIR-RED)/(NIR+RED)$^{0.5}$	(Ehammer et al., 2010)
Enhanced Vegetation Index 2 (EVI_2)	2.4*(NIR – RED)/(NIR+RED+1)	(Miura et al., 2008)
Chlorophyll Index Green (CIG)	(NIR/GREEN) −1	(Ahamed et al., 2011)
Simple Ratio Red/NIR Ratio Vegetation Index (SRRE)	RED/NIR	(Bannari et al., 1995)
Blue Normalized Difference Vegetation Index (BNDVI)	(NIR-BLUE)/(NIR+BLUE)	(Hancock & Dougherty, 2007)
Plant Pigment Ratio (PPR)	(GREEN – BLUE)/ (GREEN + BLUE)	(Metternicht, 2003)
Infrared Percentage Vegetation Index (IPVI)	[((NIR/NIR + RED)/2) *(NDVI + 1)]	(Kooistra et al., 2003)
Log Ratio (LogR)	Log (NIR/RED)	(Henrich et al., 2011)
Modified Normalized Difference Vegetation Index (MNDVI)	(NIR−Red)/(NIR+Red−2*Blue)	(Huete et al., 1997)
Normalized Difference Red-Edge (NDRE)	(NIR – REDEDGE)/ (NIR + REDEDGE)	(Barnes et al., 2000)
Adjusted Vegetation Index (SAVI)	(NIR − Red)/ (NIR+ Red + L) *(1+L)	(Ahamed et al., 2011)

2.2.6 Deep Neural Networks (DNN)

Recently, DNN has gained popularity in modelling and predicting SOC distribution in various ecosystems ascribed to the model's ability to explain complex relationships that exist between explanatory and target variables (García-Martínez et al., 2020; Odebiri, Mutanga, & Odindi, 2022). DNN model is a collection of multiple neurons placed in a specific sequence of numerous layers (Montavon et al., 2018; Odebiri, Mutanga, & Odindi, 2022). Neurons from one layer are linked to the next layer up to the final estimated neuron. In essence, neurons receive activations from the previous layer (Larochelle et al., 2009). The layers that contain a network of neurons produce a complex nonlinear mapping from the input to the output using backpropagation (Montavon et al., 2018). Additionally, to avoid model overfitting, the dropout regularization technique can be applied to hidden layers. An example of the DNN architecture is shown in Figure 2.2.

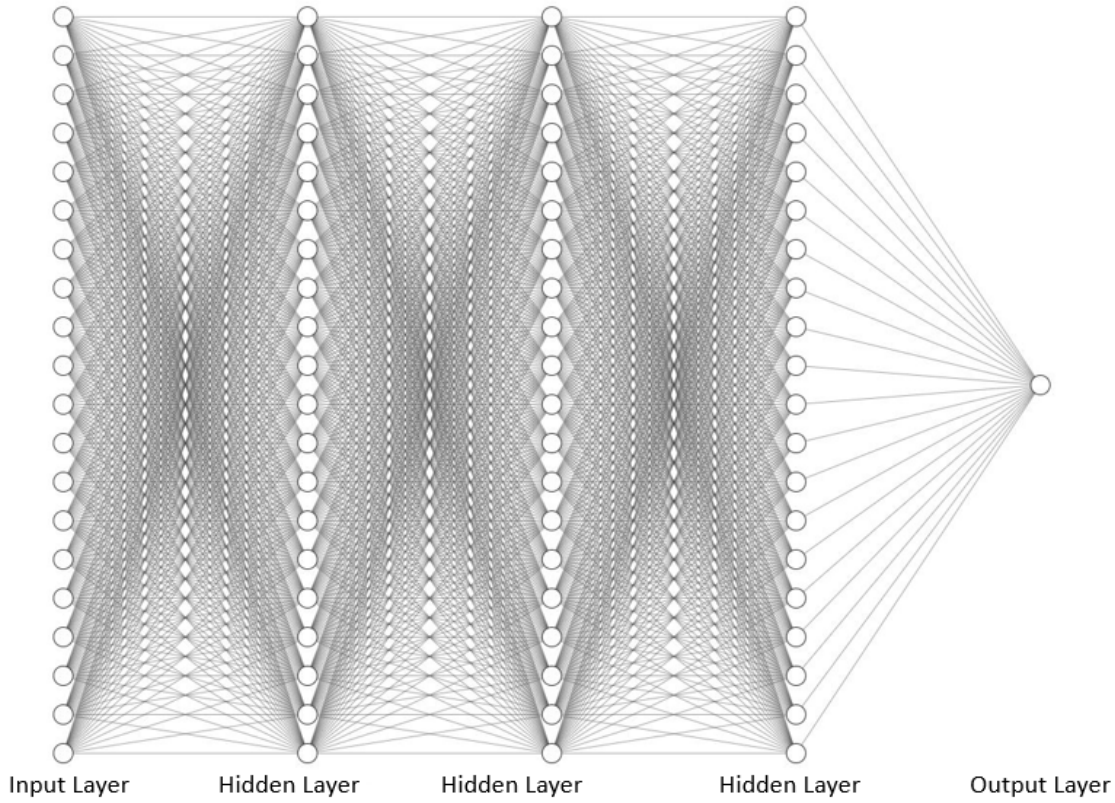

Figure 2.2: A graphical example of Deep Neural Network (DNN) architecture.

The python programming language was used to develop the DNN model. A combination of 14 vegetation indices and eight PlanetScope spectral bands were used as input variables to build the model. The input variables were separated into ten even parts and used for training and testing the model. To yield the best output for the model, the hyperparameter tuning technique, where hyperparameters were randomly optimized repeatedly until the best results were achieved was adopted. A summary of hyper-parameters used to develop the model is presented in Table 2.2.

Table 2.2: Deep Neural Networks (DNN) defined hyper-parameters.

Algorithm	Hyper-parameters	Parameter as used	Parameter Description
DNN	Hidden Layer	2	Number of hidden layers
	Node Size	100	Number of nodes (neurons) in the hidden layer
	Network weight initialization	normal	Initialized weights of the layers from input to output
	Learning rate	0.001	Adjust the weights of the network
	Dropout regularization	0.5	Neurons randomly dropped during training to reduce overfitting
	Epochs	80	Number of training iteration

2.2.7 Feature selection

Generally, regression analysis is challenged by multicollinearity, which is a statistical phenomenon in which two or more predictors are moderately or highly correlated. Hence, it is advisable that the model is developed using the least correlated predictors (Ismail & Mutanga, 2010; Odebiri et al., 2020). To counteract the challenge of multicollinearity, a Pearson correlation test was executed. Therefore, one predictor between two highly correlated predictor was removed from the input dataset using the method suggested by Behnamian et al. (2017). The predictor variables that have proven in the past research to be the best predictors for SOC stocks, particularly in grasslands, were selected for predicting SOC distribution.

2.2.8 Model evaluation metrics

To evaluate the DNN model, three commonly used metrics namely, coefficient of determination (R^2), Root Mean Squared Error (RMSE), and Lin's Concordance Correlation Coefficient (LCCC) were reported. RMSE demonstrates the difference between the observed and estimated SOC values. The coefficient of determination measures the performance of the model in estimating the SOC and ranges between 0 and 100%, while LCCC reports on how well bivariate pairs of the observed SOC conform to predicted SOC values (King & Chinchilli, 2001; Odebiri, Mutanga, & Odindi, 2022). These metrics can be expressed as;

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{O,i} - X_{p,i})^2}{n}} \quad \text{.....Equation(2)}$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(X_0 - P)^2}{\sum_{i=1}^{n}(X_o - O')^2}\right] \quad \text{.....Equation(3)}$$

$$LCCC = \frac{2r\sigma_0\sigma_p}{\sigma_p^2 + \sigma_p^2 + [O' - p']^2} \quad \text{..... Equation(4)}$$

Where X_o and X_p represent the measured and predicted SOC values, while n signify the total number of observations. Furthermore, σ_0 and σ_p are the respective variance of the observed and predicted value, while o' signifies the averages of the measured and predicted SOC.

Additionally, 10-fold cross-validation was used to ensure that each dataset was used at least once to avoid bias for DNN algorithms. The best-fitted model normally produces a higher R^2 and LCCC, and a lower RMSE. Hence, the performance of the model was identified using the three-accuracy metrics. Furthermore, variable importance ranking was executed using the relative variable importance and SHapely Additive exPlanations (SHAP), respectively, to identify the most significant predictors of SOC. A digital SOC map of the study area was developed using the best predictor variables derived from the DNN model.

Furthermore, the accuracy assessment for landcover classification was conducted within the GEE and the confusion matrix method adopted. The model employed the overall accuracy, producer's accuracy, and user's accuracy to evaluate RF classifier. Accuracy assessment was executed by comparing training dataset with testing dataset as suggested by Nasiri et al. (2022) and Mahdianpari et al. (2020).

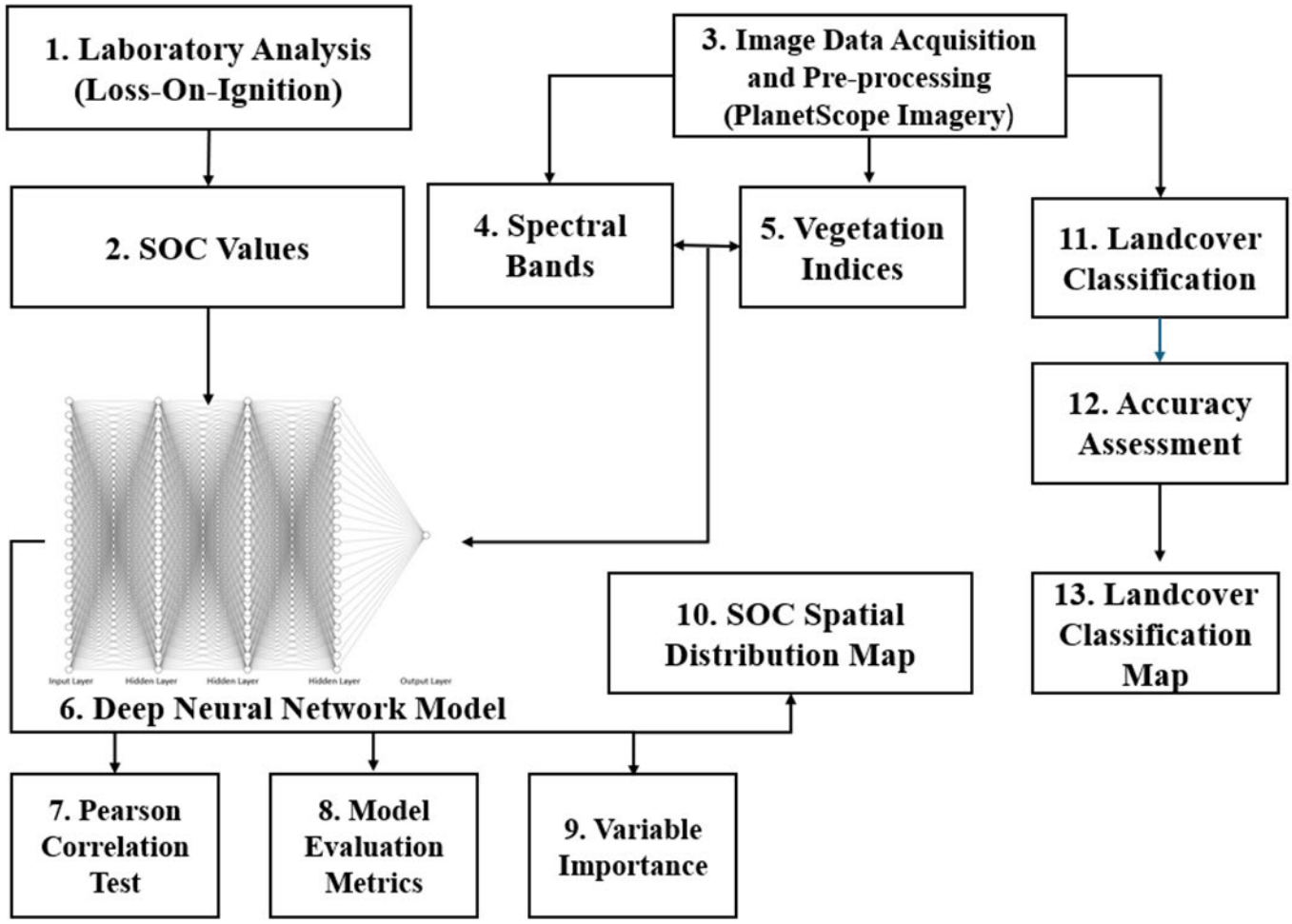

Figure 2.3: Summary of procedures followed in quantifying SOC.

2.2.9 Statistical analysis

The Statistical Package for the Social Science (SPSS version 28.0) software was utilized to perform a paired t-test to investigate the level of significance of SOC distribution between pristine and woody encroached grassland. The measures of central tendency, particularly mode, mean, maximum, minimum, standard deviation and standard error, were executed using SPSS. Microsoft excel was utilized to investigate the Coefficient of Variation (CV), skewness, and kurtosis for woody and pristine grassland. Also, a box and whisker plot were executed on Microsoft excel.

2.3 Results

2.3.1 Descriptive statistics

Table 2.3 shows the descriptive statistics for both pristine and woody encroached grasslands. The paired t -test between woody and pristine grassland ($p = 0.05$) indicate that there is a significant difference in SOC distribution between the two sites. In addition, for pristine grassland, data showed strong negative skewness (-0.75) and kurtosis (1.34), satisfying the normality rule (Hair Jr et al., 2021), while woody encroached grassland SOC data indicated a strong positive skewness (0.75) and kurtosis (1.33). Moreover, the CV for pristine and woody encroached grassland was 11.89%. The CV value suggests that there was a variation between pristine and encroached grasslands (Figure 2.4).

Table 2.3: Descriptive statistics of the dataset.

Landcover	Minimum Value (%)	Maximum Value (%)	Mean Value (%)	Standard deviation (%)	Standard error (%)
Pristine Grassland	3.6	9.4	6.6	1.46	0.19
Woody Encroached Grassland	2.4	13.1	7.56	2.76	0.31

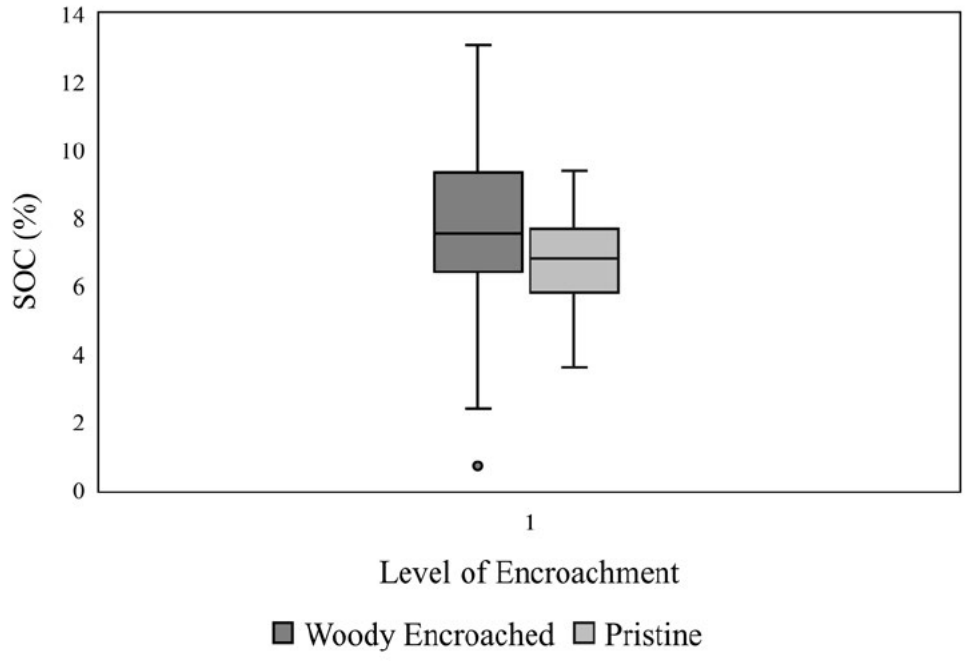

Figure 2.4: The distribution of SOC between a pristine and woody encroached grassland.

According to the Pearson correlation test, all spectral bands exhibited high correlation values, between 0.9 and 1 (Figure 2.5). The Red-edge band was utilized for building the DNN model. Moreover, all the vegetation indices that exhibited high correlation values were removed. Therefore, only twenty vegetation indices were used to build the best model as detailed in Table 2.1.

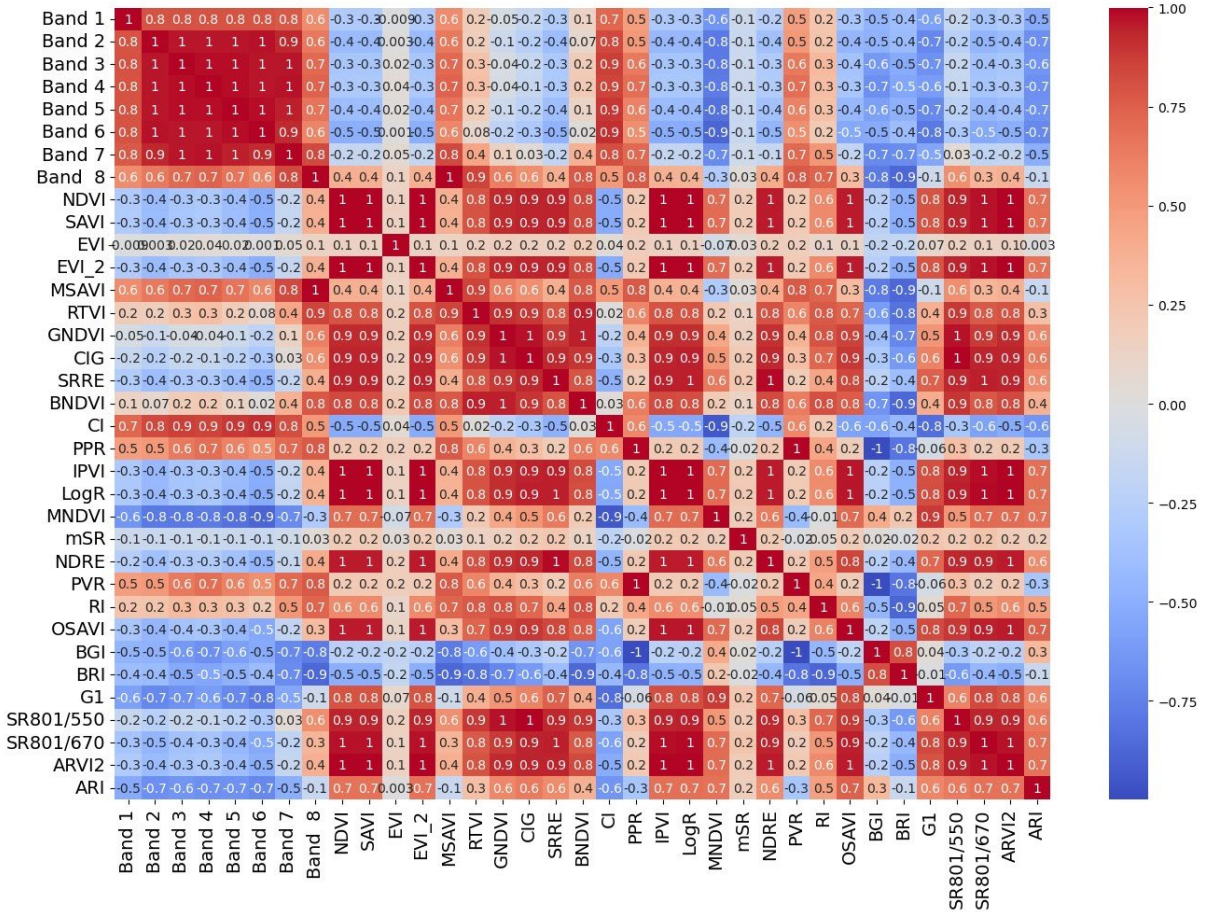

Figure 2.5: Pearson correlation test of predictor variables.

2.3.2 Evaluation and performance of model

The DNN model performed well in quantifying SOC across a woody encroached grassland, with an R^2 value of 63.78, and an RMSE of 1.901 t/ha. Figure 2.6 illustrates the correlation between the observed and predicted SOC derived from DNN model. Furthermore, the train and test dataset showed a good model fit and generation with training data exhibiting an R^2 value of 67,02, RMSE of 1.36 t/ha, and LCCC value of 90.67, while testing data exhibited an R^2

value of 63.78, RMSE of 1.91 t/ha, and LCC of 80.60. Moreover, Table 2.4 explains the performance of RF classifier.

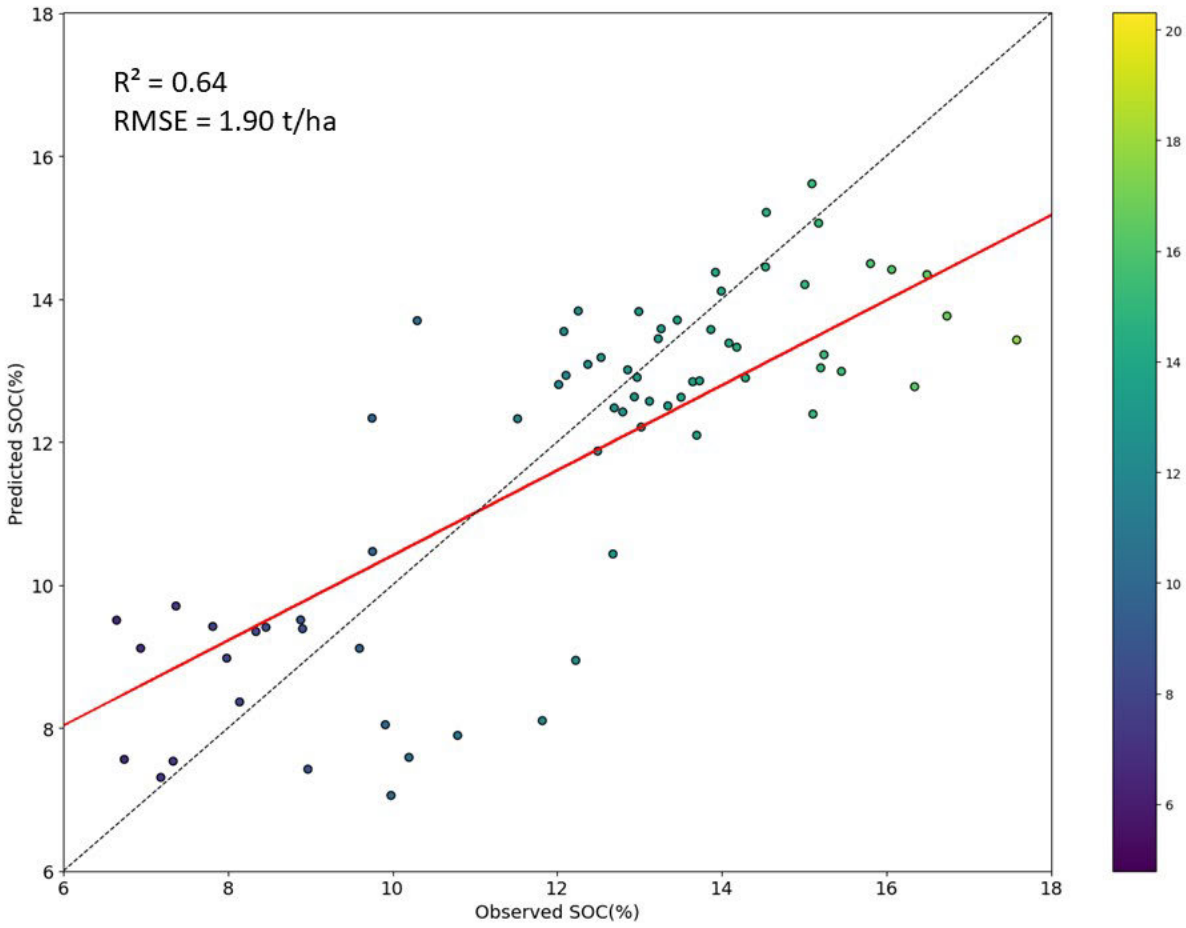

Figure 2.6: Relationship between predicted and measured SOC using Deep Neural Network (DNN).

Table 2.4: The accuracy assessment of RF classifier.

Classifier	Overall accuracy (%)	Producer accuracy (%)	Kappa accuracy (%)	Consumer accuracy (%)
RF	0.95	0.88	0.92	0.96

2.3.3 Variable importance assessment of model

The results in Figure 2.7 illustrate the importance and contribution of each predictor variables using SHapely Additive exPlanations (SHAP) for DNN. The best three explanatory variables

for the DNN model were NDVI, MSAVI and TVI while LogR, EVI and PPR were the least important variables. The Figure further indicate that NDRE and MNDVI exhibited similar contributions towards quantifying SOC.

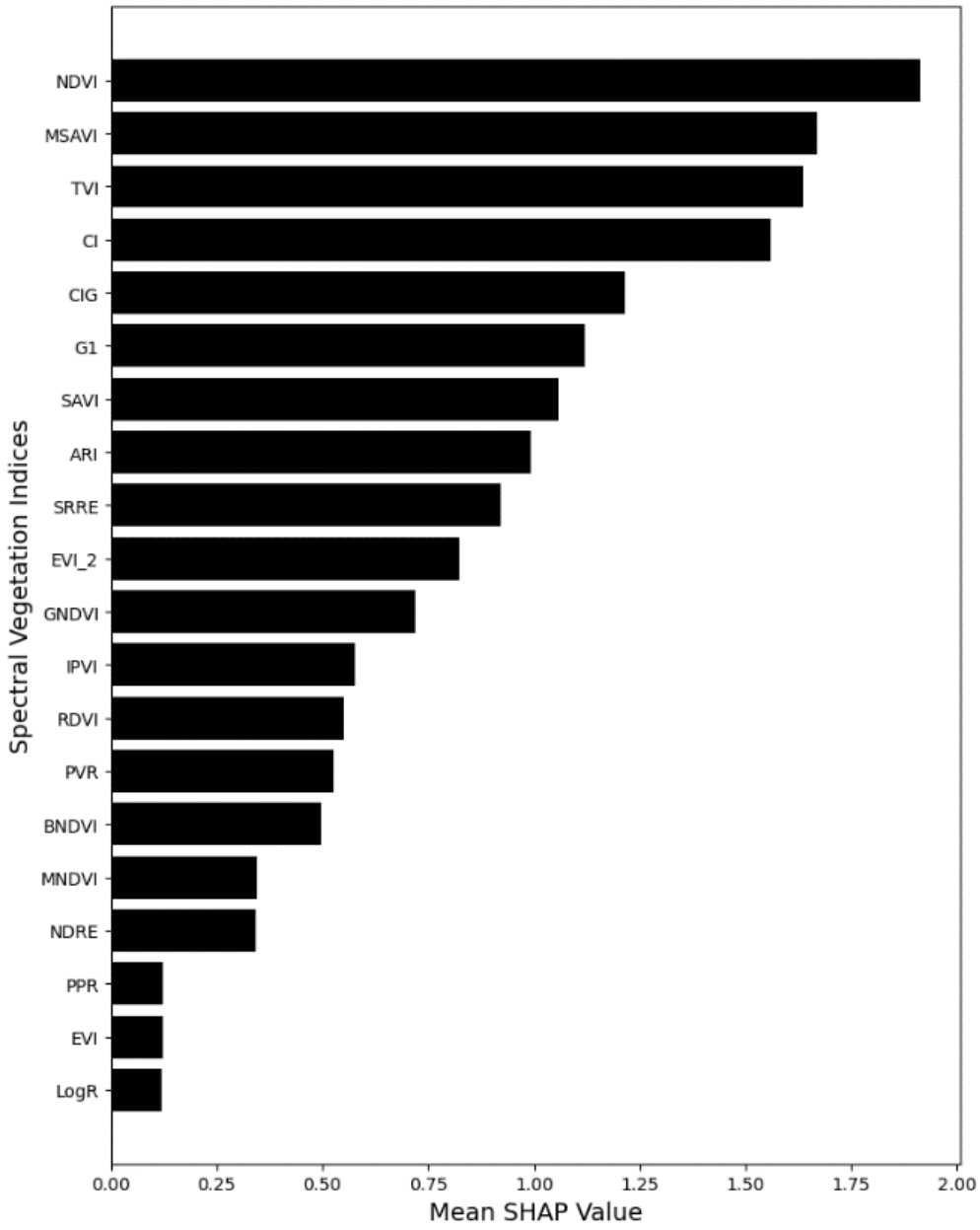

Figure 2.7: Importance ranking of variables used for quantifying SOC within Bisley Nature Reserve.

2.3.4 Spatial distribution of SOC

Figure 2.8 (a) depicts the distribution of SOC stocks in both woody encroached and pristine landscapes, respectively, within the Bisley Nature Reserve, while Figure 2.8 (b) explains different land cover classes within the study area. According to Figure 2.8 (a), woody

encroached landscapes are dominated by medium and high concentrations of SOC, while low concentrations of SOC stocks were observed to be dominant in pristine grasslands. Furthermore, SOC concentration is observed to be higher along the river channels which are located across the central parts of the reserve. The map further reveals that woody encroached landscapes located central, West and South of the reserve are dominated by high concentration of SOC. The SOC spatial distribution maps clearly reveal that there is a noticeable variability of SOC distribution between pristine and woody encroached landscapes. For instance, the distribution of SOC is generally higher across woody encroached grasslands in comparison to pristine grasslands which exhibited low SOC concentrations.

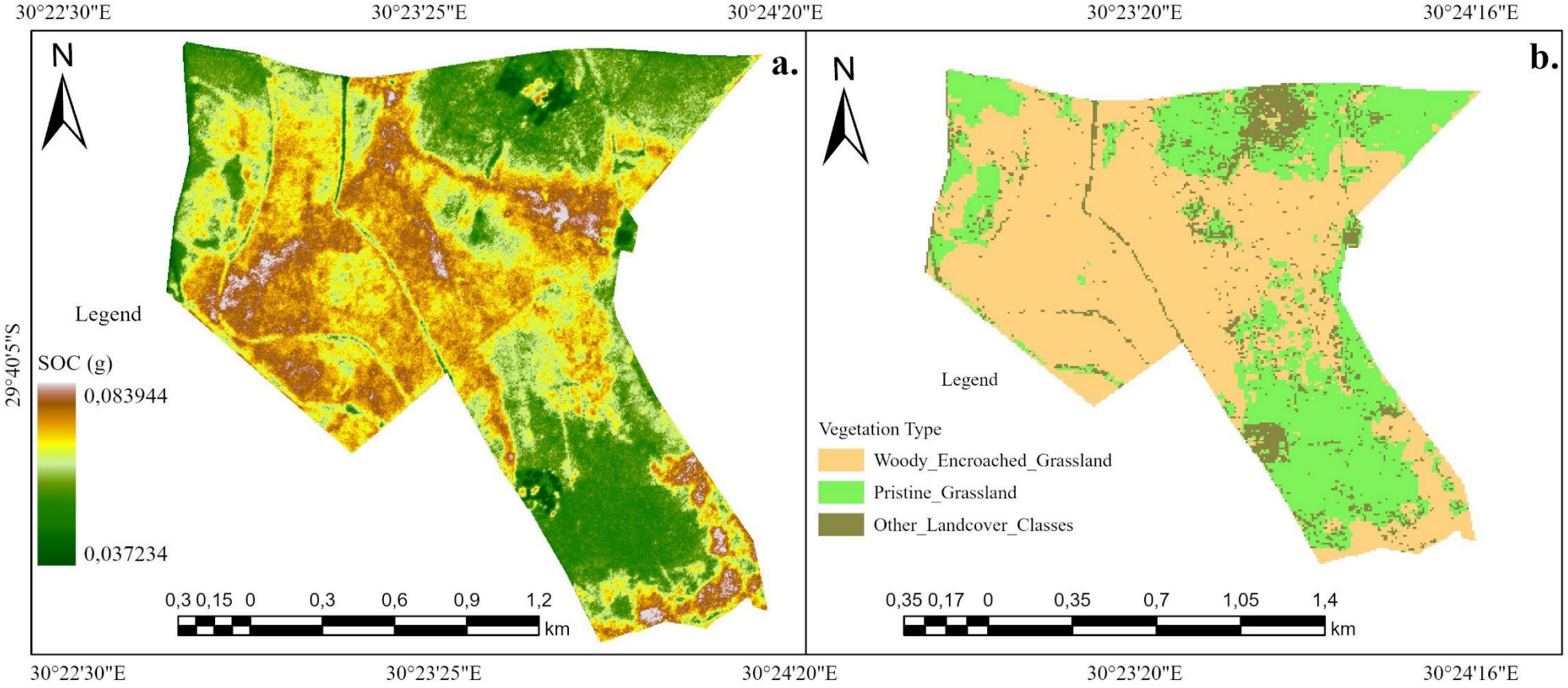

Figure 2.8: The spatial distribution of SOC for Deep Neural Network (DNN)(a.) and landcover classification map (b.) using PlanetScope MSI data.

2.4 Discussion

This chapter quantified the spatial variability of SOC stocks between a pristine and woody encroached grassland using PlanetScope's spectral bands and vegetation indices. The chapter examined the impact of woody vegetation proliferation on SOC stocks within grasslands using the DNN model. The findings show that woody encroached areas have a high concentration of SOC stock compared to pristine grasslands (Figure 2.8 a). A two-sample t-test confirmed this difference ($p < 0.05$), indicating a statistically significant increase in SOC stock under woody encroachment. According to Fantappiè et al. (2010) and Rasse et al. (2005), SOC accumulation is highly influenced by the surface litter and root biomass. Therefore, areas with high litter biomass are expected to exhibit high SOC concentrations. The study by Yang et al. (2021) established SOC to be concentrated in areas with high shrub encroachment compared to pure grasslands, and suggested that woody encroached landscapes are associated with high mean litter and root biomass. As a result, SOC distribution is expected to be more in grasslands with a proliferation of woody vegetation compared to pure grasslands (Figure 2.3). Similarly, Odebiri et al. (2020) established that SOC stock is commonly concentrated in areas with higher vegetation biomass compared to areas with less vegetation cover.

Our findings further shows that woody encroached grassland has the highest maximum percentage of SOC (13.1%) when compared to pristine grassland (9.4%) (Figure 2.3). Several regional studies corroborate our observations that woody encroachment often increases SOC. For instance, in semi-arid savannas of Africa, Sankaran et al. (2004) and Barger et al. (2011) reported that shrub invasion can elevate carbon sequestration by increasing both above- and belowground biomass, particularly under intermediate rainfall regimes where woody species thrive without outcompeting herbaceous vegetation entirely. This is consistent with the notion that the additional litter and root inputs from woody plants augment organic matter accumulation, enhancing soil carbon pools. In the southwestern United States, mesquite (Prosopis spp.) and other shrubs have encroached into grasslands over the past century, leading to a restructuring of ecosystems and shifts in SOC dynamics (Barger et al., 2011; Jackson et al., 2002). Similar to our findings, these studies demonstrated that SOC content tends to increase under shrub canopies compared to adjacent grass-dominated areas, primarily due to greater litterfall and more extensive root systems. However, it is also noted that the net effect of woody encroachment on ecosystem SOC can vary with precipitation gradient, soil type, and management history (Scott et al., 2006).

Interestingly, in some arid systems where woody species have replaced perennial grasses entirely, reduced herbaceous cover can lead to soil erosion, partially offsetting gains in SOC (Eldridge & Soliveres, 2015). Likewise, in Australian rangelands, Eldridge et al. (2011) reported that shrub encroachment can increase SOC in the topsoil layer, although the magnitude of this effect depends on grazing intensity, fire regimes, and other land-use pressures. Where woody encroachment is managed or occurs alongside moderate grazing, the combination of woody litter input and reduced soil disturbance often promotes organic matter accumulation. Furthermore, Li et al. (2016) reported that in semi-arid and humid regions, leguminous shrubs encroachment in grasslands have greatly amplified SOC pool, while non-legumes have shown to have less impact on SOC stocks, particularly in topsoil. The overall consensus across various regional studies, including those from African savannas, the southwestern United States, and Australian rangelands, is that woody encroachment can enhance SOC stocks, at least in the upper soil layers, by adding substantial organic inputs from litterfall and belowground biomass.

Woody encroached areas are also associated with high soil metabolism due to high litter fall and dead matter accumulation, which increases the concentration of SOC (Mayer et al., 2020; Odebiri et al., 2020). Consequently, forested regions possess elevated levels of SOC in comparison to pristine grasslands. Furthermore, areas with high soil moisture exhibit high vegetation biomass (Lozano-Parra et al., 2018; Ruiz-Sinoga et al., 2011). Hence, areas with high soil moisture are associated with high SOC compared to dry areas. Figure 2.8 reveals that areas along the river channels have high concentrations of SOC. This aligns with the findings of Evans et al. (2011), which indicate a robust positive correlation between SOC and soil moisture. However, the study by Kerr and Ochsner (2020) noted that the correlation between SOC and soil moisture can vary depending on other soil characteristics including soil temperature, inter-site vegetation variability, and soil physical attributes. According to Chen et al. (2016), soil moisture promotes vegetation growth, productivity, and decomposition, which enhance SOC sequestration. The dominant soil types in this study area are Mispah, Dundee, Glenrosa, Clovely, and Avalon soils. These soils were dry and shallow, thus explaining low SOC distribution across the grassland. However, the results presented by the current study may not be conclusive because all the soil samples were obtained at a depth of 30 cm. According to Zhou et al. (2017), there is a significant different between SOC located in topsoil and SOC in deeper layers of the soil profile. Therefore, it is recommended that further investigations about the spatial variability of SOC in woody encroached grasslands be done, considering the depth.

The DNN model produced acceptable accuracy (R^2 = 0.64) for quantifying SOC stocks. This is consistent with Odebiri, Mutanga and Odindi (2022) who sought to compare the performance of the DL approach with commonly used classical models in SOC stock mapping. Their study established that the DNN model outcompete classical models when estimating SOC stocks. According to K. Wang et al. (2021), DL approaches commonly outperform classical models when dealing with large and complicated datasets. Unlike other classical models, the DL models can read and understand intricate relationships between many ecological parameters and utilize feature representations that are solely acquired from the data (Odebiri, Mutanga, & Odindi, 2022). Additionally, the DNN model is characterized by many interconnected hidden layers and neurons that enable the model to extract numerous representative characteristics from the input dataset. The results of the current study were influenced by the attributes of the nature reserve and non-linearity of the dataset. The field data was collected from a localized encroached grassland and the dataset used to build the DNN model was complex. As a result, the DNN was the best model for quantifying SOC stocks across a woody encroached grassland.

Among the PlanetScope-derived vegetation indices, the most important index was NDVI. According to Odebiri et al. (2020), there is a high correlation between SOC and NDVI, hence the index is the most important variable when estimating SOC, particularly within vegetated landscape. The NDVI can better minimize the influence of topography and possess a good classification capability when dealing with heterogenous landscapes (Barati et al., 2011; Zhang et al., 2019). The study by Yang et al. (2020) notes that NDVI can also be utilized as a primary determinant of vegetation health, which will automatically inform the spatial distribution of SOC in grassland ecosystems. Spectral indices including MSAVI, TVI and CI were also ranked high by the model when quantifying SOC. Yet, PPR, EVI and LogR had less significant contribution to the DNN model. The spectral bands of PlanetScope were combined with derived spectral vegetation indices to build the best model for quantifying SOC stocks. Previous literature suggests that red-edge spectral bands are the most valuable for quantifying SOC stocks (Andreatta et al., 2022; Mngadi et al., 2021; Odebiri et al., 2020). The red-edge wavelength regions of PlanetScope imagery provide important reflectance information, essential for accurate SOC mapping. These regions, being highly sensitive to vegetation characteristics, offer reliable means to assess SOC distribution. The sensor's remarkable spatial resolution of three meters contributes significantly to the wealth of information available for SOC quantification at an ecosystem level.

While few studies have endeavoured to quantify SOC using the spectral information of PlanetScope, especially in the Sub-Saharan African, noteworthy contributions have been made. For instance M. H. Koparan et al. (2022) explored the capability of PlanetScope imagery in estimating and mapping SOC stocks. Their findings concluded that models built from PlanetScope-derived data exhibited excellent accuracies compared to models from other multispectral sensors. Additionally, it is advisable to conduct further investigations into the utilization of PlanetScope imagery bands rather than red-edge regions, particularly in the context of SOC mapping studies. The outcomes of the current study hold practical significance for the uMsunduzi Municipality on insights into SOC distribution within Bisley Nature Reserve. Additionally, nature reserve managers and indeed similar invaded landscapes can benefit from this study's insights for continuous monitoring of SOC stock.

1.5 Conclusion

The proliferation of woody plants in grassland ecosystems alters the SOC stock cycle. This study employed a remote sensing algorithm (DNN) and PlanetScope MSI data to quantify the spatial variability of SOC within a woody encroached and pristine grassland. The chapter established that areas exposed to woody encroachment exhibit high SOC concentration. The DNN model produced acceptable accuracy when quantifying the spatial variability of SOC within woody encroached grasslands at a localized scale. Utilizing the SHAP technique, the study identified the most crucial predictor variables for SOC estimation. Notably, NDVI, MSAVI, TVI, and CI emerged as the most significant variables. The developed model is a cost-effective and time-efficient tool for monitoring SOC in both pristine and encroached grasslands. It is recommended that future studies explore the distribution of SOC in areas experiencing woody proliferation. Furthermore, the impact of bioclimatic, soil depth, and topographic residuals on SOC should be investigated using remotely derived information, particularly within woody encroached grasslands.

Chapter Three: Modelling Soil Organic Carbon at Multiple Depths in Woody Encroached Grasslands Using Integrated Remotely Sensed and Physical Variables Data

This chapter is based on:

Mthiyane, S., Mutanga, O., & Matongera, T.N., Odindi, J. (Published). Modelling Soil Organic Carbon at Multiple Depths in Woody Encroached Grasslands Using Integrated Remotely Sensed Data. *Environmental Monitoring and Assessment,* Manuscript ID: 2dc1b69d-576e-437c-a3f0-9ed2cdad840b v1.0.

Abstract

Woody plants encroachment into grasslands has considerable hydrological and biogeochemical consequences to grassland soils that include altering the Soil Organic Carbon (SOC) pool. In this regard, continuous SOC stock assessment and evaluation at deeper soil depths of woody encroached grasslands is essential for informed management and monitoring of the phenomenon. Due to high litter biomass and deep root structures, woody encroached landscapes have been suggested to alter the accumulation of SOC at deeper soil layers, however, the level at which woody encroachment sequester SOC within localized protected grasslands is still poorly understood. Recently, driven by improved spatial and spectral data characteristics and availability of cost-effective and environmentally friendly data, remote sensing techniques and approaches have gained traction in SOC analysis. Hence, the objective of this study was to estimate SOC at various depths (30 cm, 60 cm, and 100 cm) in a woody-encroached grassland by integrating Sentinel-1 (S1), Sentinel-2 (S2), PlanetScope (PS) satellite imagery, and topographic variables. SOC was quantified from 360 field-collected soil samples using the loss-On-Ignition (LOI) method and spatial distribution of SOC across the Bisley Nature Reserve modelled using the Random Forest (RF) algorithm. Results show that the fusion of topographic variables, Synthetic Aperture Radar (SAR), and PlanetScope data effectively modelled SOC stocks at all investigated soil depths, with high R^2 values of 0.79 and RMSE of 0.254 t/ha. The distribution of SOC was established to be high at 30 cm compared to 60 cm and 100 cm depths. Notably, the horizontal reception (VH), Slope, Topographic Weightiness Index (TWI), Band 11 and vertical reception (VV) were optimal predictors of SOC in woody encroached landscapes. These findings underscore the value of fused spectral data with topographic variables and the RF model for precise SOC modelling in woody encroached ecosystems. This study's insights are pivotal for designing cost and labour effective assessment and monitoring methodology for effective management and monitoring of SOC in woody encroached environments.

Keywords: Woody Encroachment, Soil Organic Carbon, Random Forest, Remote Sensing.

3.1 Introduction

Globally, the conversion of mesic grasslands into woody dominated ecosystems has increased substantially in the past century (Ratajczak et al., 2012; O. W. Van Auken, 2009). The commonly documented drivers of woody proliferation are fire suppression, overgrazing, nutrients availability, climate change and global carbon dioxide enrichment (Kgosikoma & Mogotsi, 2013; Ratajczak et al., 2012). Encroachment of woody plant into grasslands has considerable hydrological and biogeochemical consequences to grassland soils (Honda & Durigan, 2016; Stevens et al., 2017). It reduces the quantity of productive grazing landscapes for wild herbivores and livestock (Aweto, 2024; Ding & Eldridge, 2024; Pinheiro et al., 2022) and reduces the frequency and intensity of grazing and fire, which are key factors maintaining grassland diversity (Ratajczak et al., 2012; Smit & Prins, 2015) A study by McKinley et al. (2008) notes that woody proliferation have pronounced impact on below ground nitrogen and carbon pools in grasslands, while (Belay & Kebede, 2010; Liu et al., 2011; Mureva et al., 2018) note that woody encroachment impact Soil Organic Carbon (SOC), particularly at an ecosystem level. Alberti et al. (2011), (Chiti et al., 2017) and Zhou et al. (2017) note that woody encroachment substantially alters the spatial pattern and variability of surficial SOC stocks. However, to date, there is a dearth in literature that has adopted remotely sensed datasets and approaches to establish the distribution of SOC at deeper soil depths, particularly within woody encroached grasslands.

A study by Zhou et al. (2017) quantified SOC at 120 cm depth and established a high concentration in woody invaded areas compared to pristine grasses. Their study also observed an increase in SOC at deeper soil layers, particularly in woody dominated landscapes. According to Rumpel and Kögel-Knabner (2011), SOC concentration is expected to be higher in deeper soils of woody vegetated landscapes due to higher root litter and decomposition. However, most studies that have investigated SOC in deeper soils of woody encroached grasslands have typically utilized traditional methods that rely on field-based observation that are expensive, labour intensive and time consuming (Blaser et al., 2014; Mureva et al., 2018). Advancements in satellite technology and data storage have tremendously revolutionized SOC stocks modelling over the last few decades. Remotely sensed data are pivotal tool for SOC modelling SOC as they offer cost effective, labour-efficient and time saving data (Odebiri et al., 2021). Hence, the availability of remotely sensed data promise to revolutionize SOC quantification in deeper soil depths of woody invaded grasslands (Liu et al., 2011; Odebiri et al., 2024; Zhou et al., 2017).

Previous studies investigating the influence of woody invasion on SOC distribution have primarily used multispectral optical sensors including Landsat Mission and MODIS (Venter et al., 2021). However, these sensors are limited by their vulnerability to atmospheric interference leading to decreased accuracy when quantifying SOC in deeper soil layers. Consequently, Synthetic Aperture Radar (SAR) technology has gained traction for monitoring SOC stocks due to its ability to provide weather independent and vegetation sensitive images that are crucial for quantifying SOC in woody invaded landscapes. The European Space Agency (ESA) provides many freely available remotely sensed data with improved spatial resolution. The noticeable advantages of SAR data include the improved spatial and spectral resolutions that can best capture soil-vegetation relationship, making them a feasible and attractive option for quantifying SOC stocks (Zhou et al., 2020). These unique features are important for robust quantification of SOC stocks at deeper soil depths. Additionally, SAR images are not restricted to the visible and infrared portions of the electromagnetic spectrum but possess radar sensors that can be utilized to detect SOC distribution within vegetated landscapes. Also, S2 provides images with thirteen bands that cover the visible near infra-red to short wave infra-red spectral range that is paramount for SOC estimation. Studies by Yang and Guo (2019) and Morais et al. (2023) have reported on the capability of SAR data in modelling SOC stocks in grasslands at different soil layers.

Regardless of these advantages, the capability of SAR data for modelling SOC within woody encroached grasslands has not been fully explored. This is because areas with small geographical extents require images with exceptionally high spatial resolutions, such as PlanetScope (PS) multispectral sensor characterized by 3 m spatial resolution (M. Koparan et al., 2022; Koparan, 2019). Previous literature shows that a combination of SAR, S2 and PS data is useful in improving the detection of above ground biomass, which eventually informs SOC distribution in deeper soils. According to Koparan (2019), local mapping of SOC stocks requires images which finer spatial resolution, such as PS. However, the sensor is only limited to eight bands, and commonly affected by atmospheric interference, hence image scenes might not always be available. Therefore, combining SAR and S2 spectral information with PS data can potentially produce better results when estimating SOC within localized landscapes affected by proliferation of woody plants.

Recently, Machine Learning (ML) algorithms have become popular for quantifying relationships between SOC and remotely sensed data (Odebiri et al., 2021; Yang et al., 2021; Zhou et al., 2020). Previous literature has proven the value of Random Forest (RF) regression

model to quantify the distribution of SOC stocks within localized heterogeneous ecosystems (Grimm et al., 2008; Pouladi et al., 2019), including woody encroached landscapes (Venter et al., 2021). RF estimates a response variable based on a series of explanatory variable through building a set of regression trees and averaging the outcome. The outcome of all individual trees is averaged to make a single prediction. RF is easy to implement, and can handle a significant amount of training dataset (Odebiri et al., 2020). Furthermore, RF has an ability to model non-linear relationships by using both continuous and categorical predictor variables, which is fundamental for precise quantification of SOC in woody encroached landscapes. To further improve the accuracy of the RF model, it is fundamental to include topographic factors as one of the predictors of SOC. Environmental factors, such as slope, elevation and topographic wetness index, can be combined with remotely sensed data to best model the accumulation of SOC stocks in sub-surface soils (Zhou et al., 2020).

Most studies investigating the implication of woody invasion to SOC accumulation have been conducted on surface soils at a depth of 0-30 cm, particularly in grasslands (Liu et al., 2011; Throop & Archer, 2008). Fewer studies have sought to evaluate the accumulation of SOC at deeper soil depth of woody encroached grasslands using remotely sensed data. According to Odebiri et al. (2024), deeper soils (>30 cm) sequestrate more than half of the total SOC pool. Hence, it is imperative that the extent at which woody encroachment affect SOC is assessed. The lack of comprehensive and conclusive information on the influence of woody proliferation on SOC stocks present an opportunity for further assessment of the phenomenon, particularly in localized grasslands. Consequently, the objective of this study was to estimate soil organic carbon (SOC) at various depths (30 cm, 60 cm, and 100 cm) in a woody-encroached grassland by integrating Sentinel-1 (S1), Sentinel-2 (S2), PlanetScope (PS) satellite imagery, and topographic variables. The study coupled RF, remotely sensed data and topographic variables to model SOC distribution at different soil depths of Bisley Nature Reserve affected by a proliferation of woody vegetation on a grassland. The study also evaluated the spatial distribution of SOC from a wood encroached to a pristine grassland.

3.2 Materials and Methods

3.2.1 Field Data Collection

Field sampling was conducted at Bisley Nature Reserve between the 18th and 20th of March 2024. The study employed a purposive sampling strategy to establish three transects that were 622 m long measured from a pristine grassland to a woody encroached grassland. This was done to clearly visualize the transition of vegetation biomass from pristine grassland to woody encroached grassland. At each transect, a total of 40 plots (15m^2 each) were established. Within each plot, 1x1 metre square quadrant was placed every 15 m, and three soil samples collected per quadrant at different soil depths. As a result, a total of 40 soil samples were collected for each depth (30 cm, 60 cm and 100 cm) using a handheld MAC AFRIC 68 CC EARTH Auger Drill. Additionally, utilizing the Trimble Handheld Global Positioning System with a sub meter accuracy, Global Positioning System (GPS) locations for each plot were recorded. A total of 360 soil samples were collected and subsequently dispatched to the laboratory for analysis.

3.2.2 Laboratory Analysis

The soil samples were air dried for three days on trays to remove excess moisture. Samples were then sieved through a 2mm sieve and oven dried for 24 hours at 105 °C before analysis. For SOC determination, 1 g of soil were added to glass beakers, placed in a muffle furnace at 360 °C for 2 hours, cooled in desiccators at room temperature, and weighted. SOC was measured as the difference between soil mass before ignition and soil mass after ignition, divided by the soil mass before ignition (Schulte & Hopkins, 1996). All calculated values were tabled in Excel for statistical analysis.

3.2.3 Image Acquisition and Preprocessing

A combination of radar (SAR), S2 and PS satellite images were used in this study. The spatial and spectral information of each sensor is presented in Table 3.1. PS MSI level-3B imagery used in this research was captured on the 8th of March 2024 and acquired from Planet website (https://www.planet.com/markets/education-and-research/) on the 4th of March 2024 with 5% cloud cover. No image preprocessing was employed for PS imagery as it is supplied atmospherically and radiometrically corrected. PS is characterized by high spatial resolution (3m) and red-edge band that is critical for SOC monitoring, especially in deeper layers of the soil profile (Matiza et al., 2024).

Table 3.1: Spectral and spatial information of PS, S2 and SAR data.

	PlanetScope	**Sentinel 1**	**Sentinel 2**
Resolution	3 m	10 m	10 m, 30 m & 60 m
Band type	Coastal Blue (431-452 nm) Blue (465-515 nm) Green I (513-549 nm) Green II (547-583 nm) Yellow (600-620 nm) Red (650-680 nm) Red edge (697-713 nm) Near-Infrared (845-885 nm)	C-band (VH)	B1 (443 nm), B2 (490 nm), B3 (560 nm), B4 (665 nm), B5 (705 nm), B6 (740 nm), B7 (783 nm), B8 (842 nm), B8a (865 nm), B9 (940 nm), B10 (1375 nm), B11 (1610 nm), B12 (2190 nm)

SAR data used for quantifying SOC captured on the 05th of March 2024 was downloaded from ESA (https://dataspace.copernicus.eu/) on the 15th of April 2024. SAR products include both Sentinel-1A and Sentinel-1B. According to Geudtner et al. (2014), Sentinel-1A operate in four unique imaging modes which provide C-band with different resolutions. SAR technology is a type of radar that is utilized to build two- or three-dimensional reconstructions of features, including landscapes. This is an active sensor that sends pulses of energy and reads information from the energy reflected by the target object. The sensor provides continuous images that are not affected by clouds, smoke, and haze. SAR offers dual polarisation capability with vertical transmission and either horizontal reception (VH) or vertical reception (VV), and rapid product delivery (Torres et al., 2017). These attributes are critical for continuous monitoring of earth phenomena, including SOC stocks. The Ground Range Detected (GRD) format S1 image of interferometric wide swath mode was pre-processed with ESA's Sentinel-1 Toolbox in the software SNAP (version 6.4.5). To focus on the study location's pristine and woody encroached grasses, the data was split into sub-swaths. The image size was also reduced to better processing efficiency. S1 processing routine also incorporated thermal noise removal, Range-Doppler Terrain Correction and geocoding.

S2 image (Level 2B) captured on 15th of April 2024 was downloaded from the Copernicus open access hub (https://dataspace.copernicus.eu/) on the 18th of April 2024. The sensor's image is characterized by 13 spectral bands with unique spatial resolutions (Table 3.1). S2 bands range

from visible light and near infrared to short-wave infrared. Due to these spectral characteristics, S2 data has an ability to effectively quantify SOC within heterogenous landscapes (Castaldi et al., 2019). S2 image was atmospherically corrected using Sen2Cor processor (V.2.5.5) plugin in SNAP (version 6.4.5) software. The image was resampled to 10m to maintain the key attributes of S2 data to maximum extent.

3.2.4 Remotely Sensed Spectral Vegetation Indices

Through different band combinations from S1, S2, and PS images, relevant spectral vegetation indices were generated. Previous studies suggest that spectral vegetation indices are paramount predictors of SOC distribution because SOC cannot be directly learned through modelling techniques (Odebiri, Mutanga, & Odindi, 2022; Shafizadeh-Moghadam et al., 2022; Zhou et al., 2020). X. Wang et al. (2021) notes that there is a pronounced relationship between vegetation indices and SOC stocks. As a result, five popularly utilized spectral vegetation indices that have been proven to best predict SOC distribution, were generated from PS, S1 and S2, together with spectral bands and used to predict SOC distribution (Table 3.2).

Table 3.2: S1 & S2 derived spectral vegetation indices.

Satellite	Index	Formula	Reference
S2	Normalized Difference Vegetation Index (NDVI)	(B8-B4)/(B8+B4)	(Rouse et al., 1974)
	Green Normalized Difference Vegetation Index (GNDVI)	(B9-B3)/(B9+B3)	(Ahamed et al., 2011)
	Red-Edge 1	(Red-Edge/Red)	(Cloutis et al., 1996)
	Modified Soil Adjusted Vegetation Index (MSAVI)	$29+1-1*(\sqrt{(29*B1)^2-8(B9-B5)})/2$	(Wu et al., 2007)
S1	Radar Vegetation Index (RVI)	(8*VH)/ (HH + VV + 2*VH)	(Kim et al., 2011)

3.2.5 Topographic Variables

Literature shows that there is an association between topographic factors and SOC distribution (Tu et al., 2018; Yu et al., 2020; Zhou et al., 2020). These include Slope, Aspect, Elevation and Topographic Wetness Index (TWI). Topographic variables were extracted from a SRTM Digital

Elevation Model (DEM) image with 30 m spatial resolution acquired from Earth Explore online platform. Using ArcGIS pro (version 3.0), topographic covariates residuals were extracted and later combined with spectral bands and vegetation indices to quantity SOC distribution in different soil layers.

3.2.6 Modelling Approach

Each observation in our statistical analysis included spectral band values, spectral vegetation indices, and SOC measurements at the plot level that were taken from the relevant plot locations in PS, SAR and S2 images, and topographic variables. The study employed regression analysis to determine relationship between predictor variables and the field measured SOC. Additionally, using Microsoft Excel and IBM SPPS (version 29), a *t*-test was executed to determine the level of significance between SOC located at the depth of 0-30 cm, 30-60 cm, and 60-100 cm.

3.2.7 Random Forest Regression Model

The RF algorithm, a non-parametric tree-based machine learning approach, has gained popularity in monitoring and predicting various environmental parameters (Mngadi et al., 2021; Odebiri et al., 2020). As highlighted by Nabiollahi et al. (2019) and Sreenivas et al. (2014), a distinctive feature of RF is its ability to handle both small and large datasets, making it easy to comprehend and implement. The algorithm initiates by generating a range of bootstrap samples from the original dataset and then randomly samples different predictors during the training phase. Each bootstrap creates a regression tree, and the algorithm selects the best split among numerous variables (Singh et al., 2017). Notably, it constructs multiple uncorrelated trees for training, utilizing a random subset comprising two-thirds of all samples, while leaving one-third for validation. Crucially, the optimization of node-size, *mtry*, and *ntree* is essential for achieving optimal predictive performance. The Root Mean Square Error (RMSE) of the training datasets was used to optimize the *mtry* and *ntree* in order to determine the optimal RF model for quantifying the distribution of SOC (Table 3.3).

Table 3.3: Random Forest (RF) defined hyper-parameters.

Algorithm	Hyper-parameters	Parameter as used	Parameter Description
RF	*Mtry*	25	Number of input variables
	Ntree	41	Number of trees

	Node size	1	Number of observations
	Random State	42	Possible combinations of the dataset

3.2.8 Predictor Variable Selection

Literature suggests that as the number of input variables increase, RF becomes more complex (Odebiri et al., 2020). For RF model, it is advisable to use a smaller number of predictors to minimize complexity and chances of multicollinearity. Therefore, to prevent multicollinearity, a Pearson correlation test was executed. As a results, one predictor between two highly correlated predictor was eliminated from the input datasets using Out-Of-Bag (OOB) method suggested by Behnamian et al. (2017) . The OOB error rate method is commonly used for eliminating all the input variables with less or no contribution towards the accuracy of the model. The backward elimination method was adopted until the best number of predictors was obtained. Consequently, a total of twenty optimal predictor variables were utilized to generate the final digital SOC distribution map at all selected depths.

3.2.9 Accuracy Assessment

Using a 70–30 holdout validation technique, the RF model's efficacy in measuring SOC in a woody encroached landscape was evaluated. The total input dataset (N = 120) was separated into 70% (n = 84) for training and 30% (n = 36) for testing the datasets. To reduce the probability of sampling bias, a 10-fold cross validation was adopted. The Root Mean Square Error (RMSE) and R^2 were presented in order to evaluate the model's overall performance. RMSE report the difference that exist between the observed and estimated SOC values. The R^2, always between 0 to 100 %, indicate the percentage of the response variable variation that is described by the model. Additionally, the significance of each predictor variable to the overall performance of the RF model was assessed using Shapely Additive Explanations (SHAP)

3.3 Results

3.3.1 Summary Statistics

The summary statistics describing the predicted SOC data at different soil depths are reported in Table 4. SOC data for the three soil depths indicated a normal distribution. The SOC stock varied between 0.037 and 0.095 g in the three soil layers. The Coefficient of Variation (CV) was noted to be low for all the soils depths (< 17 %). Results show that the distribution of sample values were negatively skewed for both 30 cm and 100 cm and positively skewed for 60 cm (Table 3.4). The results further show that SOC distribution decreased with increasing soil depths and is mostly concentrated at the 30 cm depth (Figure 3.1).

Table 3.4: Basic statistics of SOC at different soil depths.

Depth (cm)	Minimum (g)	Mean (g)	Maximum (g)	Standard deviation (g)	Standard Error	Skewness	CV (%)	Kurtosis
0-30	0.057	0.077	0.095	0.009	0.002	0.309	11,69	-0.158
30-60	0.046	0.061	0.069	0.006	0.001	-1.231	9,84	1.269
60-100	0.037	0.049	0.061	0.008	0.002	0.302	16,33	-1.330

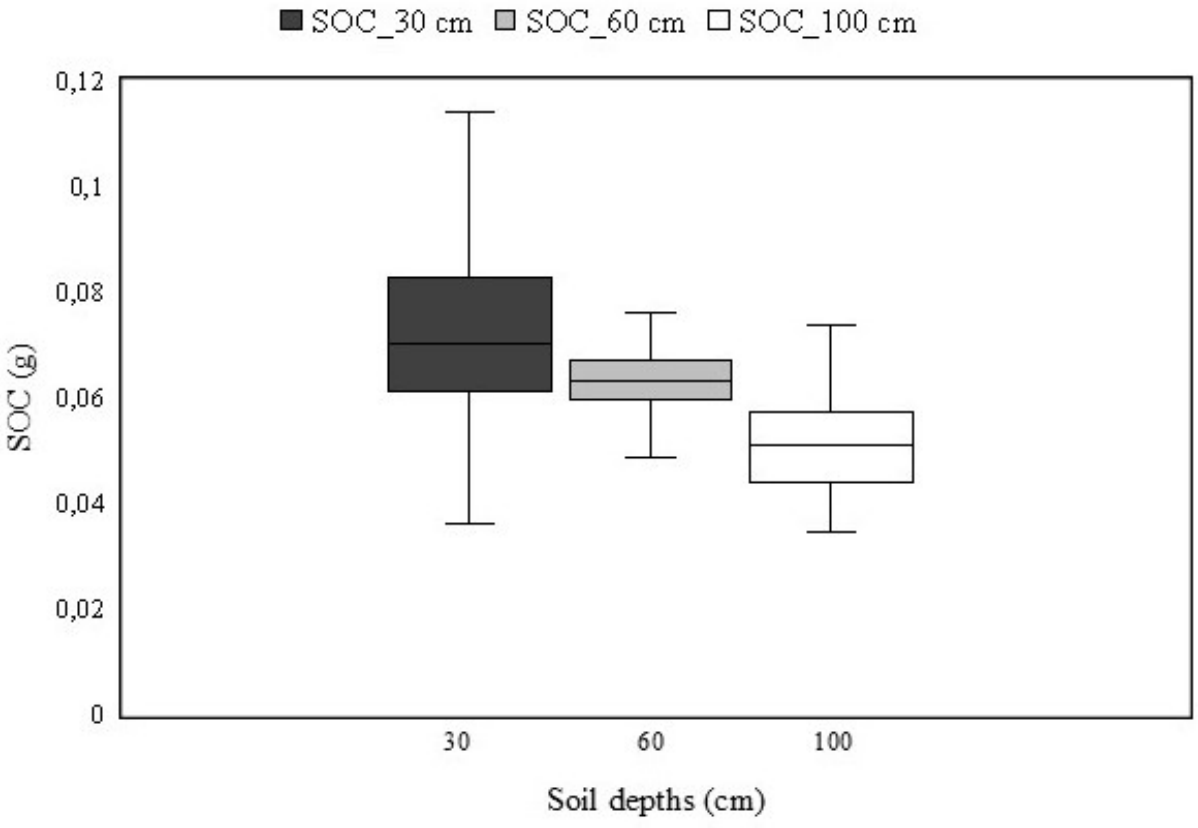

Figure 3.1: Soil Organic Carbon distribution at three different soil depths (30 cm, 60 cm & 100 cm).

3.3.2 Horizontal distribution of SOC

The Figure 3.2 shows the distribution of SOC stock at horizontal distance from the boundary, a centre between pristine and woody encroached landscape, towards either woody encroached or pristine grassland for all the investigated soil depths. Figure 3.2 (a), suggest that the amount of SOC increases from the boundary towards the centre of a woody encroached landscape and decreases from the centre toward the edges of the encroached landscape. A similar trend is observed when moving from the boundary towards the centre of a pristine landscape, however, SOC continue to increase towards the edges (Figure 3.2.a). For Figure 3.2 (b), SOC distribution decreases from the boundary to the centre of the pristine grassland and further decrease toward the edges. These results demonstrate the spatial variability of SOC stocks between a woody encroached and pristine grassland.

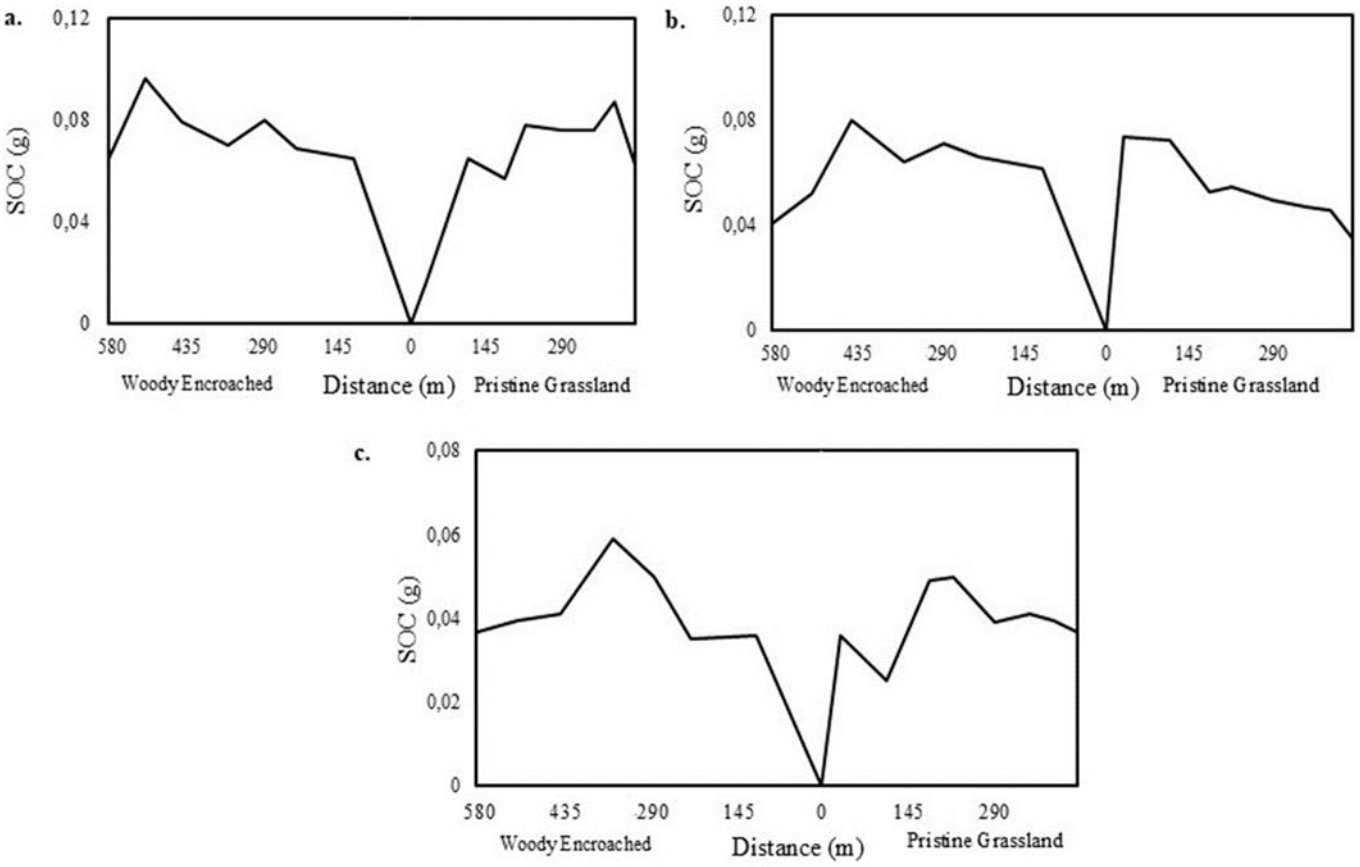

Figure 3.2: The horizontal distribution of SOC from the boundary towards centre of Woody encroached or pristine landscape. (a) SOC distribution at the depth of 30 cm, (b) 60 cm, and (c) 100 cm.

3.3.3 Evaluation and Performance of Model

The Random Forest (RF) model, utilizing spectral vegetation indices and satellite bands, demonstrated strong performance in modelling the spatial distribution of soil organic carbon (SOC) in deeper soil depths of woody-encroached Bisley Nature Reserve exhibiting R^2 of 0.79 and RMSE of 0.254 t/ha. Figure 3.3 presents the correlation between the predicted and observed SOC values as modelled by the RF algorithm.

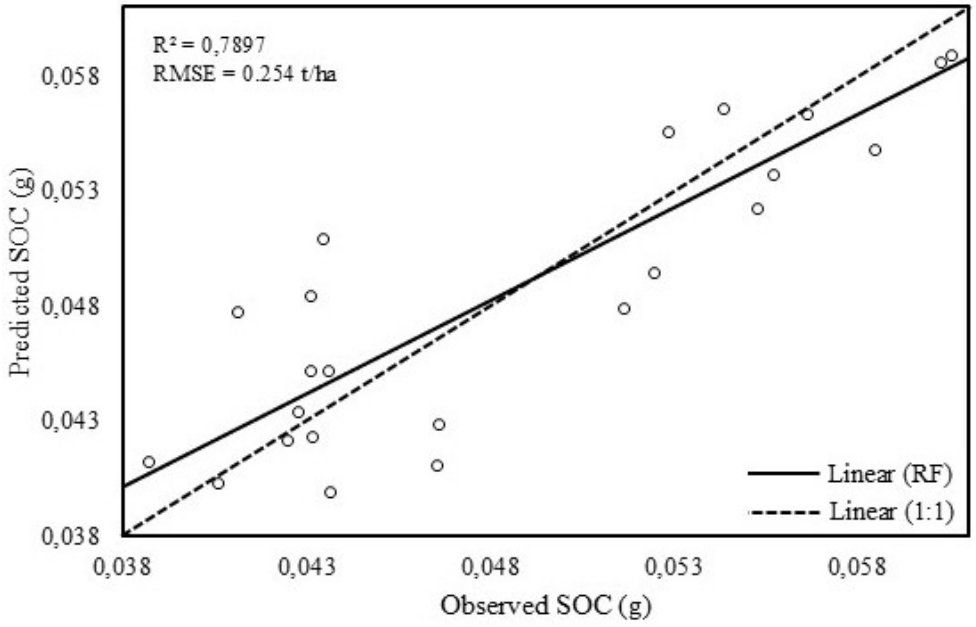

Figure 3.3: Scatterplot of measured verses predicted SOC values at 0-100 cm depth.

3.3.4 Variable Importance Ranking

Figure 3.4 illustrates the importance and contribution of each predictor variable towards the performance of RF model in quantifying SOC distribution at the depth of 0-100 cm. The Figure reveals that VH, Band 11, and Slope were among the top three of the most important variables for predicting SOC in deep soils of a woody encroached grassland. A closer look at Figure 3.3 shows that VH had the highest contribution while NDVI had the smallest importance and contribution toward the performance of the RF model.

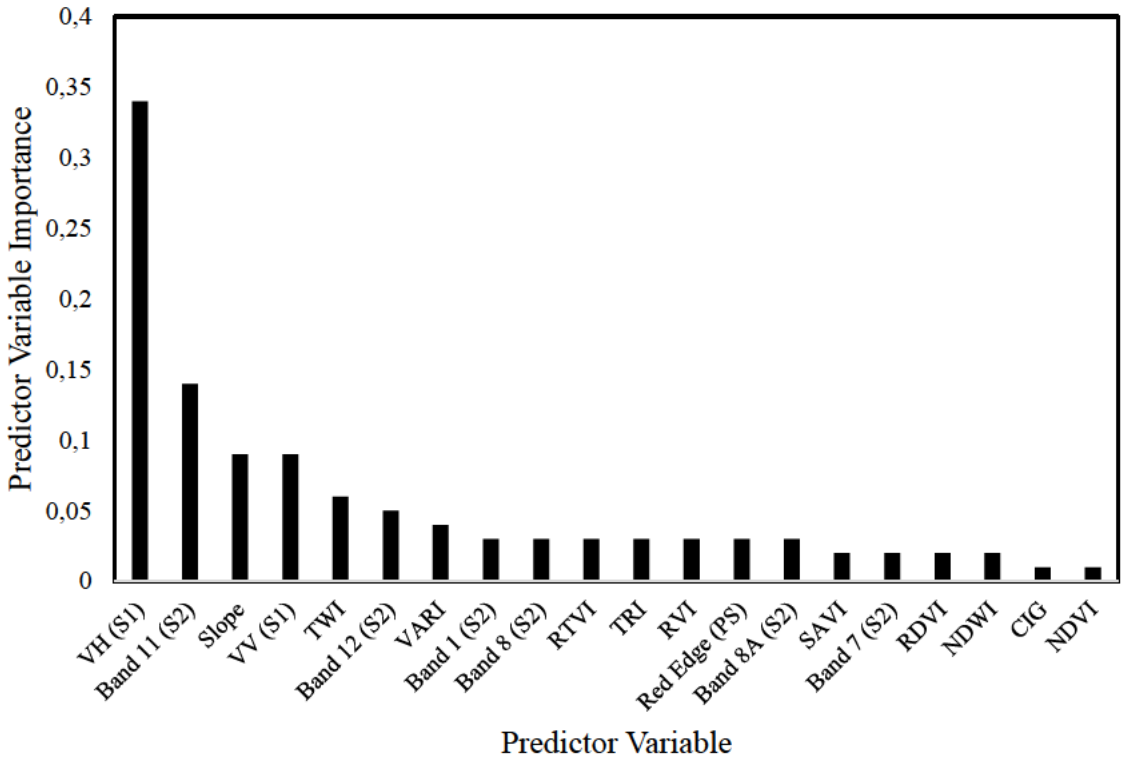

Figure 3.4: Variable importance ranking for SOC across a woody encroached grassland at 0-100 cm.

3.3.5 Spatial Estimation of SOC

Figure 3.5 (a) shows SOC accumulation across a woody encroached grassland, while Figure 3.5 (b) illustrate different landcover classes located at the study site. A closer look at the maps shows that SOC is predominantly located in central and western parts of the grassland for all the soil depths highly invaded by woody vegetation. Furthermore, landscapes located at the southern and northern parts of the grasslands are dominated by low SOC stocks concentration, and these areas are mostly dominated by grasses. The study clearly shows that SOC is concentrated in areas dominated by woody vegetation, compared to areas under the dominance of grasses for all the investigated soil depths. The maps depict that there is minimal variation of SOC distribution for all the investigated soil depths.

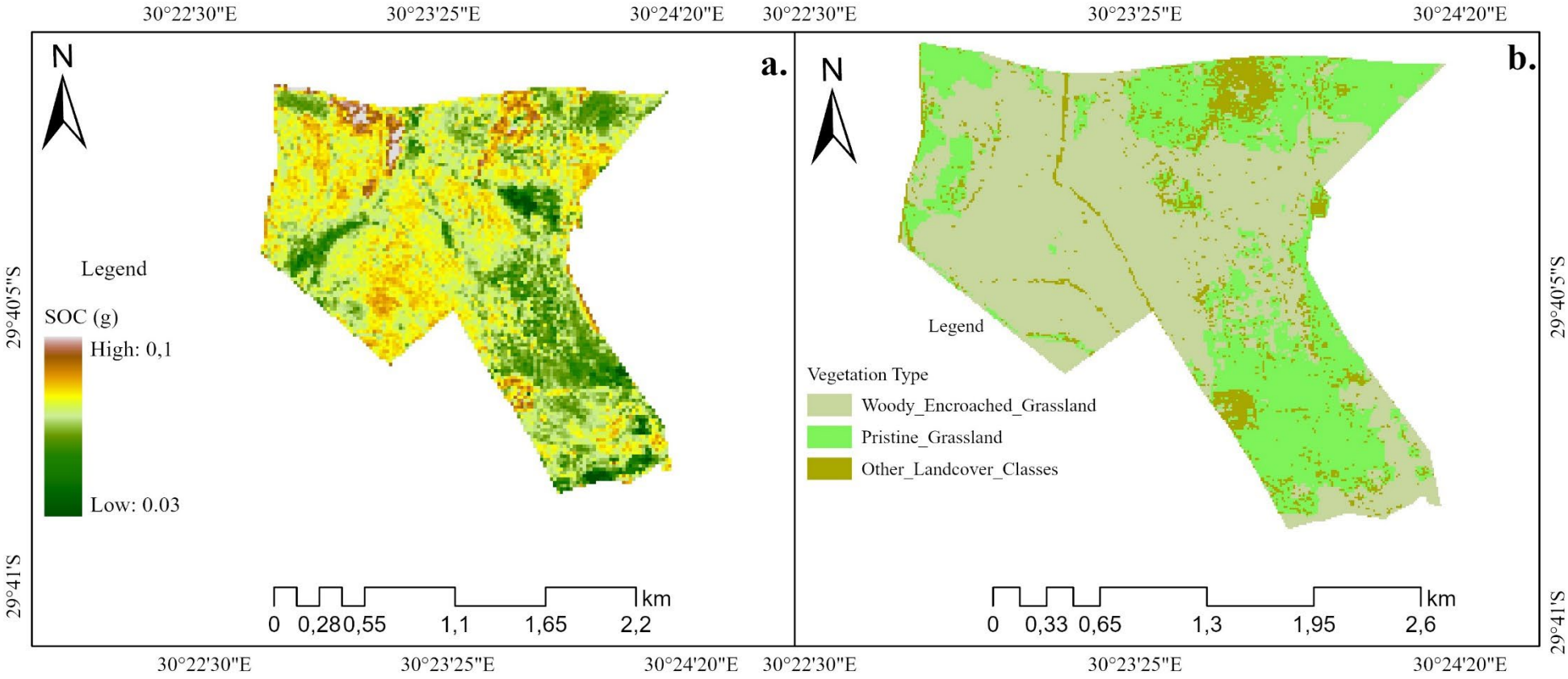

Figure 3.5: (a.) The spatial distribution of SOC for RF model at 0-100 depth. (b.) Landcover classification map of the study area.

3.4 Discussion

This study used RF and remotely sensed data to model SOC distribution at different soil depths of Bisley Nature Reserve affected by woody proliferation. Our statistical analysis revealed that there is noteworthy variation in SOC distribution between the investigated soil depths. Figure 3.2 shows a strong resemblance of SOC stock noted at the depth of 30 cm and low accumulation observed at the depth of 100 cm. This finding is consistent with Zhou et al. (2017) who observed high SOC concentration in the 0-15 cm depth increment, and minimal alteration of SOC in deep soils after woody encroachment. According to Zhou et al. (2018), plant roots and root exudates have pronounced impact on SOC distribution because most organic matter in the soil is derived from the root turnover, while Perry (1989) notes that roots of large trees are well entrenched and tougher than immature trees, preventing root turnover in deeper soils. Notably, the bedrock in study area is near the surface, thus restricting root distribution in deeper soils, reducing SOC sequestration (Angst et al., 2018; Kirfel et al., 2019). Furthermore, in terms of SOC stabilization, the results presented by the current study are a response to woody plants that are relatively young (<100 years)(Kraai et al., 2023). According to Hibbard et al. (2003), it takes more than 200 years for SOC to reach stabilization in deeper soils after woody encroachment. Correspondingly, Zhou et al. (2017) notes that a strong SOC resemblance is expected in deeper soils after at least 400 years of woody invasion. For this reason, most woody trees in the study area have not undergone root decay, thus reducing root turnover at deeper soil depths. Hence, SOC accumulation is more amplified in topsoil compared to deeper soils (Yang et al., 2023). Moreover, deeper roots are associated with endurance and gradual turnover compared to roots closer to the surface.

The RF regression analysis revealed that there is a strong resemblance of SOC distribution in woody encroached landscapes compared to landscapes dominated by pristine grasses (Figure 3.5). This is because woody encroached areas are dominated by significant litter decomposition, which inform SOC accumulation (McGrath & Zhang, 2003; Odebiri, Mutanga, & Odindi, 2022; Zhou et al., 2017). The study also reveals that SOC stocks intensely dominated the central parts of woody vegetated landscapes and declined toward the edges dominated by pure grasses (Figure 3.2 and Figure 3.5). Similarly, the study by (Zhou et al., 2018) established that woody vegetation near the centre is usually older than those closer to the edges, thus explaining the high amount of SOC in the centre of a woody encroached landscape compared to the edges. Previous studies suggest that the establishment of woody vegetation in a grassland

greatly amplifies the distribution of SOC pool, suggesting an interconnected relationship between woody encroachment and SOC stocks (Daryanto et al., 2013; Liu et al., 2011). The translocation of SOC and nutrient particles via fluvial, aeolian, and animal transport processes from grass dominated landscapes to areas dominated by woody encroachment, substantially amplifies nutrient and SOC accumulation in woody invaded landscapes (Liu et al., 2011; Okin et al., 2009).

The results also revealed that VH band had the highest importance and contribution in the performance of the RF model (Figure 3.4). VH polarization is more sensitive to vegetation density and structural changes and is not affected by weather or clouds (Luo et al., 2024). Even though the frequency of backscattered radar signal of S1 sensor is not a direct measure of vegetation, the backscatter detected in VH band is directly linked to the above ground vegetation (Duarte et al., 2024). Therefore, VH cross polarization band provide crucial information about the physiological state and changes of above ground biomass which relate to the accumulation of SOC stocks. The study by Santos et al. (2023) also discovered VH polarization to be the most important variable explaining SOC distribution for RF model since it has the ability to penetrate the ground. Additionally, topographic derivatives, slope and TWI, were among the most significant input residuals for detecting the concentration of SOC at deeper soil depth, across a woody encroached grassland. Similarly, Zhou et al. (2020) observed that topographic variables obtained high accuracy compared to S1 and S2 data when estimating SOC pool. The study by Barbosa et al. (2014) suggest that slope is the key determinant of vegetation density and thus explaining SOC pool. According to Tziachris et al. (2019), toe-slope positions are associated with high SOC concentration, due to high vegetation biomass, compared to mid-slopes, and thus explains the robustness of the slope in explaining SOC distribution.

The integration of topographic variables, S1, PlanetScope and SAR data performed well in modelling the horizontal and vertical accumulation of SOC pool across the study area (Table 3.4). The current study concludes that combining SAR and PlanetScope with topographic factors improves the prediction of SOC within a woody encroached grassland. Moreover, the study indicates that the use of RF regression model, based on combined remotely sensed input variables provide a reliable and effective methodology for modelling SOC in woody encroached ecosystems (R^2 of 0.79 and RMSE of 0.254). The results demonstrated the ability of RF to select the best input variable for producing SOC distribution maps (Figure 3.5). One of the shortcomings of RF regression algorithm is failure to deal with complicated and

heterogeneous data (Odebiri et al., 2020). However, the current study utilized less complicated dataset with few predictor variables, thus explaining the success of RF model in predicting SOC.

Our results have revealed more concentration of SOC in woody encroached landscapes compared to grass-dominated landscapes (Figure 3.5). Previous literature has revealed that SOC stocks promote productivity and stability in grasslands and is crucial for climate change mitigation (Conant, 2010; Ghosh & Mahanta, 2014; Tessema et al., 2020). However, its role on reducing and degrading grasslands has not been well documented, particularly after woody plants encroachment. O. Van Auken (2009) and Wolkovich et al. (2010) note that due to increased soil fertility and SOC enrichment in woody encroached landscapes, the dominance of woody plants is most likely while grasses degrades. Hence, its adverse impact on grasses overrides its positive impacts. Additionally, the study has established more concentration of SOC stocks in central parts of the reserve, hence priority should be given to these landscapes for effective management of the reserve and reduction of woody invasion.

3.5 Conclusion

We used topographic factors, PlanetScope and SAR data to quantify the accumulation of SOC in Bisley Nature Reserve using ML techniques. The study established that integration of spectral information with vegetation indices provide valuable information for monitoring SOC distribution in woody encroached grasslands. RF model performed well in predicting SOC in the study area, e.g. exhibiting R^2 of 79.0 and RMSE of 0.254 t/ha. A strong resemblance of SOC was observed as a depth of 30 cm compared to both 60 cm and 100 cm. Furthermore, SOC was established to be more concentrated in the central parts of the nature reserve and decreased towards the edges. We conclude that freely and accessible SAR and PlanetScope data and topographic factors provide more opportunities to quantify SOC stocks in grasslands. The study is beneficial to reserve managers and policymakers to make informed decisions on managing and conserving the nature reserve. The methodology presented in this study is a cost effective and time efficient procedure of monitoring SOC distribution across woody encroached grasslands. We, however, suggest that future studies explore the combination of bioclimatic variables with remotely sensed data to model SOC stocks in woody encroached landscapes.

Chapter Four: The Synthesis

4.1 Introduction

The primary focus of this study was to apply remote sensing techniques to quantify the spatial distribution of SOC stocks across a woody-encroached grassland. This chapter reviews aims and objectives established in the introduction section (Chapter One) against conclusions. The chapter also contextualize the major conclusions and recommends potential opportunities for future research.

4.2 Objectives Reviewed

4.2.1 First Objective:

Quantifying the influence of woody encroachment on soil organic carbon across Bisley Nature Reserve using PlanetScope data

The shift from a grass- to a woody-dominated ecosystem has an influence on the accumulation of SOC in grassland. The advancement in remote sensing technology promises to revolutionize SOC monitoring particularly in grasslands affected by woody encroachment (Honda & Durigan, 2016; Liu et al., 2011). Therefore, PlanetScope spectral bands and vegetation indices, and DNN model were combined to evaluate the spatial variability of SOC between a woody encroached and pristine landscape of Bisley Nature Reserve. Based on the findings of this study, SOC was noted to be more concentrated in areas dominated by woody vegetation compared to landscapes dominated by pristine grasses. The strong resemblance of SOC in areas dominated by woody encroachment is associated with high litter biomass and decomposition that inform SOC accumulation (Fantappiè et al., 2010). The DNN model and PlanetScope spectral information yielded acceptable accuracies with RMSE of 1.90 t/ha and R^2 value of 0.64. The performance of DNN model can be credited to the unique capability of the model to read complex interrelationships between different environmental variances (Odebiri, Mutanga, & Odindi, 2022). Therefore, the combination of DNN model and PlanetScope data, with improved spatial resolution (3m) and includes red-edge band, provide a huge potential for continuous monitoring of SOC distribution across woody encroached grasslands. Hence, the study provides baseline information that can be utilized by environmental managers to monitor SOC and thus, make informed conservation decisions. Although the study has established that woody encroachment tends to amplify SOC accumulation in grasslands, these results are not conclusive, since SOC was only evaluated at

a depth of 30 cm. This presented an opportunity for further assessments of SOC distribution is deeper soil depths of a woody encroached grassland.

4.2.2 Second Objective:

Modelling of soil organic carbon at different soil depths of a woody encroached grassland using fused remotely sensed data and machine learning

The accumulation of SOC at depths of 0-30 cm, 30-60 cm, and 60-100 cm was estimated using topographic variables, SAR, S2, PlanetScope data and RF model. The results show that there is variation in SOC distribution among the investigated soil depths. According to the findings of the study, among the investigated soil depths, SOC was mostly concentrated at the depth of 0-30 compared to 60 cm and 100 cm. These findings are consistent with the observation made by Zhou et al. (2017) which conclude that SOC is concentrated in topsoil compared to deeper soils. The current study concluded that the combination of SAR, S2, PlanetScope, topographic factors and RF model improves SOC modelling at different soil depths of a woody encroached grassland. Interestingly, RF model performed well in predicting SOC at different soil depths of the study area, i.e. exhibiting R^2 of 79.0 % and RMSE of RMSE of 0.254 t/ha. Therefore, future research can adopt similar technique to quantify SOC distribution or other environmental residuals. This study provides baseline information that can be utilized by Bisley Nature Reserve managers to monitor the spatial distribution of SOC and making informed decisions about the influence of woody encroachment in the reserve.

4.3 Conclusion

The major aim of this study was to quantify the influence of woody encroachment on SOC pool in Bisley Nature Reserve, Pietermaritzburg, South Africa. The study established that landscapes dominated by woody encroachment exhibit high concentration of SOC stocks when compared to landscapes dominated by pure grasses. The spectral bands and vegetation indices of PlanetScope yielded excellent accuracies when predicting SOC stocks across woody encroached and pristine grasslands. Also, DNN model produced acceptable accuracy when quantifying the spatial variability of SOC within woody encroached grasslands at a localized scale. Finally, this study contributes to better understanding of SOC accumulation in grasslands affected by woody encroachment.

The second objective sought to model soil organic carbon at different soil depths of a woody encroached grassland using fused remotely sensed data and machine learning. The study

established that SOC is concentrated at depth of 0-30 cm, when compared to deeper soil depths (< 30 cm). The study established that integration of SAR, S2, PlanetScope, topographic variables and RF model provide valuable information for monitoring SOC accumulation in woody encroached grasslands. The results of this study enhance our comprehension of SOC distribution across woody encroached landscape and improve SOC monitoring and, ultimately, enhancing woody plant invasion management in protected grasslands.

4.4 Recommendations and directions for future research

Regardless the success achieved by this study in quantifying SOC stocks in woody encroached grassland using remote sensing technology, more research is still required to explore the full potential of these cutting-edge approaches in modelling SOC in various areas. Future studies can:

- Evaluate the capability of image fusion at wider spatial extent using ML techniques, particularly at catchment scale.
- Evaluate and invest on high spectral resolution datasets such as hyperspectral remote sensing for more accurate soil organic carbon quantification.
- Evaluate the influence of different management practices such as fires, slashing, and herbicides, on soil organic carbon accumulation in protected grasslands.

Although SOC has been proven to better soil quality and nutrient availability in grasslands, there is a need for future studies to thoroughly investigate trade-offs of having high concentration of SOC after woody encroachment, particularly in protected grasslands and catchment areas.

References

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion*, *76*, 243-297. https://doi.org/https://doi.org/10.1016/j.inffus.2021.05.008

Ahamed, T., Tian, L., Zhang, Y., & Ting, K. (2011). A review of remote sensing methods for biomass feedstock production. *Biomass and bioenergy*, *35*(7), 2455-2469.

Alberti, G., Leronni, V., Piazzi, M., Petrella, F., Mairota, P., Peressotti, A., Piussi, P., Valentini, R., Gristina, L., La Mantia, T., Novara, A., & Rühl, J. (2011). Impact of woody encroachment on soil organic carbon and nitrogen in abandoned agricultural lands along a rainfall gradient in Italy. *Regional Environmental Change*, *11*(4), 917-924. https://doi.org/10.1007/s10113-011-0229-6

Andreatta, D., Gianelle, D., Scotton, M., & Dalponte, M. (2022). Estimating grassland vegetation cover with remote sensing: A comparison between Landsat-8, Sentinel-2 and PlanetScope imagery. *Ecological Indicators*, *141*, 109102.

Angelopoulou, T., Tziolas, N., Balafoutis, A., Zalidis, G., & Bochtis, D. (2019). Remote sensing techniques for soil organic carbon estimation: A review. *Remote Sensing*, *11*(6), 676.

Angst, G., Messinger, J., Greiner, M., Häusler, W., Hertel, D., Kirfel, K., Kögel-Knabner, I., Leuschner, C., Rethemeyer, J., & Mueller, C. W. (2018). Soil organic carbon stocks in topsoil and subsoil controlled by parent material, carbon input in the rhizosphere, and microbial-derived compounds. *Soil Biology and Biochemistry*, *122*, 19-30. https://doi.org/https://doi.org/10.1016/j.soilbio.2018.03.026

Aweto, A. O. (2024). Is woody plant encroachment bad? Benefits of woody plant encroachment—A review. *Landscape Ecology*, *39*(2), 21. https://doi.org/10.1007/s10980-024-01823-1

Bahri, H., Raclot, D., Barbouchi, M., Lagacherie, P., & Annabi, M. (2022). Mapping soil organic carbon stocks in Tunisian topsoils. *Geoderma Regional*, *30*, e00561. https://doi.org/https://doi.org/10.1016/j.geodrs.2022.e00561

Bannari, A., Morin, D., Bonn, F., & Huete, A. (1995). A review of vegetation indices. *Remote sensing reviews*, *13*(1-2), 95-120.

Barati, S., Rayegani, B., Saati, M., Sharifi, A., & Nasri, M. (2011). Comparison the accuracies of different spectral indices for estimation of vegetation cover fraction in sparse

vegetated areas. *The Egyptian Journal of Remote Sensing and Space Science*, *14*(1), 49-56. https://doi.org/https://doi.org/10.1016/j.ejrs.2011.06.001

Barbosa, J. M., Melendez-Pastor, I., Navarro-Pedreño, J., & Bitencourt, M. D. (2014). Remotely sensed biomass over steep slopes: An evaluation among successional stands of the Atlantic Forest, Brazil. *ISPRS Journal of Photogrammetry and Remote Sensing*, *88*, 91-100.

Barger, N. N., Archer, S. R., Campbell, J. L., Huang, C. y., Morton, J. A., & Knapp, A. K. (2011). Woody plant proliferation in North American drylands: a synthesis of impacts on ecosystem carbon balance. *Journal of Geophysical Research: Biogeosciences*, *116*(G4).

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C. R. E., Thompson, T., Lascano, R. J., Li, H., & Moran, M. S. (2000). *Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data* Proc. 5th Int. Conf. Precis Agric,

Bartholomeus, H., Kooistra, L., Stevens, A., van Leeuwen, M., van Wesemael, B., Ben-Dor, E., & Tychon, B. (2011). Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy. *International journal of applied earth observation and geoinformation*, *13*(1), 81-88.

Behnamian, A., Millard, K., Banks, S. N., White, L., Richardson, M., & Pasher, J. (2017). A systematic approach for variable selection with random forests: achieving stable variable importance values. *IEEE Geoscience and Remote Sensing Letters*, *14*(11), 1988-1992.

Belay, L., & Kebede, F. (2010). The impact of woody plants encroachment on soil organic carbon and total nitrogen stocks in Yabello District, Borana Zone, Southern Ethiopia. *J. Drylands*, *3*(2), 234-240.

Bhunia, G. S., Kumar Shit, P., & Pourghasemi, H. R. (2019). Soil organic carbon mapping using remote sensing techniques and multivariate regression model. *Geocarto International*, *34*(2), 215-226.

Blaser, W. J., Shanungu, G. K., Edwards, P. J., & Olde Venterink, H. (2014). Woody encroachment reduces nutrient limitation and promotes soil carbon sequestration. *Ecology and evolution*, *4*(8), 1423-1438.

Broderick, D. E., Frey, K. E., Rogan, J., Alexander, H. D., & Zimov, N. S. (2015). Estimating upper soil horizon carbon stocks in a permafrost watershed of Northeast Siberia by

integrating field measurements with Landsat-5 TM and WorldView-2 satellite data. *GIScience & Remote Sensing*, *52*(2), 131-157.

Browning, D. M., Archer, S. R., Asner, G. P., McClaran, M. P., & Wessman, C. A. (2008). WOODY PLANTS IN GRASSLANDS: POST-ENCROACHMENT STAND DYNAMICS [https://doi.org/10.1890/07-1559.1]. *Ecological applications*, *18*(4), 928-944. https://doi.org/https://doi.org/10.1890/07-1559.1

Bugalho, M. N., & Abreu, J. (2008). The multifunctional role of grasslands. *Sustainable Mediterranean grasslands and their multifunctions*, 25-30.

Bühmann, C., Beukes, D. J., & Turner, D. P. (2006). Clay mineral associations in soils of the Lusikisiki area, Eastern Cape Province, and their agricultural significance. *South African Journal of Plant and Soil*, *23*(2), 78-86. https://doi.org/10.1080/02571862.2006.10634735

Carlier, L., Rotar, I., Vlahova, M., & Vidican, R. (2009). Importance and functions of grasslands. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, *37*(1), 25-30.

Castaldi, F., Chabrillat, S., Don, A., & van Wesemael, B. (2019). Soil organic carbon mapping using LUCAS topsoil database and Sentinel-2 data: An approach to reduce soil moisture and crop residue effects. *Remote Sensing*, *11*(18), 2121.

Chen, X., Zhang, D., Liang, G., Qiu, Q., Liu, J., Zhou, G., Liu, S., Chu, G., & Yan, J. (2016). Effects of precipitation on soil organic carbon fractions in three subtropical forests in southern China. *Journal of Plant Ecology*, *9*(1), 10-19.

Chiti, T., Mihindou, V., Jeffery, K. J., Malhi, Y., De Oliveira, F. L., White, L. J. T., & Valentini, R. (2017). Impact of woody encroachment on soil organic carbon storage in the Lopé National Park, Gabon. *Biotropica*, *49*(1), 9-12. https://doi.org/https://doi.org/10.1111/btp.12369

Cloutis, E. A., Connery, D. R., Major, D. J., & Dover, F. J. (1996). Airborne multi-spectral monitoring of agricultural crop status: effect of time of year, crop type and crop condition parameter. *International Journal of Remote Sensing*, *17*(13), 2579-2601. https://doi.org/10.1080/01431169608949094

Coetsee, C., Gray, E. F., Wakeling, J., Wigley, B. J., & Bond, W. J. (2013). Low gains in ecosystem carbon with woody plant encroachment in a South African savanna. *Journal of Tropical Ecology*, *29*(1), 49-60.

Conant, R. T. (2010). *Challenges and opportunities for carbon sequestration in grassland systems* (Vol. 9). FAO Rome, Italy.

Conant, R. T., & Paustian, K. (2002). Spatial variability of soil organic carbon in grasslands: implications for detecting change at different scales. *Environmental Pollution*, *116*, S127-S135. https://doi.org/https://doi.org/10.1016/S0269-7491(01)00265-2

Daryanto, S., Eldridge, D. J., & Throop, H. L. (2013). Managing semi-arid woodlands for carbon storage: Grazing and shrub effects on above-and belowground carbon. *Agriculture, Ecosystems & Environment*, *169*, 1-11.

Ding, J., & Eldridge, D. J. (2024). Woody encroachment: social–ecological impacts and sustainable management. *Biological Reviews*, *99*(6), 1909-1926. https://doi.org/https://doi.org/10.1111/brv.13104

Duarte, M. L., da Cunha e Silva, D. C., Barbosa, R. L., & Lourenço, R. W. (2024). Modeling of soil organic matter using Sentinel-1 SAR and partial least squares (PLS) regression. *Arabian Journal of Geosciences*, *17*(1), 32.

Eggleston, H., Buendia, L., Miwa, K., Ngara, T., & Tanabe, K. (2006). 2006 IPCC guidelines for national greenhouse gas inventories.

Ehammer, A., Fritsch, S., Conrad, C., Lamers, J., & Dech, S. (2010). Statistical derivation of fPAR and LAI for irrigated cotton and rice in arid Uzbekistan by combining multi-temporal RapidEye data and ground measurements. Remote Sensing for Agriculture, Ecosystems, and Hydrology XII,

Eldridge, D. J., Bowker, M. A., Maestre, F. T., Roger, E., Reynolds, J. F., & Whitford, W. G. (2011). Impacts of shrub encroachment on ecosystem structure and functioning: towards a global synthesis. *Ecology letters*, *14*(7), 709-722.

Eldridge, D. J., & Soliveres, S. (2015). Are shrubs really a sign of declining ecosystem function? Disentangling the myths and truths of woody encroachment in Australia. *Australian Journal of Botany*, *62*(7), 594-608.

Escadafal, R., Belghith, A., & Ben Moussa, H. (1994, 0000-00-00). Indices spectraux pour la dÃ©gradation des milieux naturels en Tunisie aride. 6Ã¨me Symp. Int. Mesures Physiques et Signatures en TÃ©lÃ©dÃ©tection, ISPRS-CNES, Val dâ□□lsÃ¨re, France.

Evans, S. E., Burke, I. C., & Lauenroth, W. K. (2011). Controls on soil organic carbon and nitrogen in Inner Mongolia, China: A cross-continental comparison of temperate grasslands. *Global Biogeochemical Cycles*, *25*(3).

Fantappiè, M., L'Abate, G., & Costantini, E. (2010). Factors influencing soil organic carbon stock variations in Italy during the last three decades. *Land degradation and desertification: assessment, mitigation and remediation*, 435-465.

Francini, S., McRoberts, R. E., Giannetti, F., Mencucci, M., Marchetti, M., Scarascia Mugnozza, G., & Chirici, G. (2020). Near-real time forest change detection using PlanetScope imagery. *European Journal of Remote Sensing*, *53*(1), 233-244. https://doi.org/10.1080/22797254.2020.1806734

Frazier, A. E., & Hemingway, B. L. (2021). A technical review of planet smallsat data: Practical considerations for processing and using planetscope imagery. *Remote Sensing*, *13*(19), 3930.

García-Martínez, H., Flores-Magdaleno, H., Ascencio-Hernández, R., Khalil-Gardezi, A., Tijerina-Chávez, L., Mancilla-Villa, O. R., & Vázquez-Peña, M. A. (2020). Corn Grain Yield Estimation from Vegetation Indices, Canopy Cover, Plant Density, and a Neural Network Using Multispectral and RGB Images Acquired with Unmanned Aerial Vehicles. *Agriculture*, *10*(7), 277. https://www.mdpi.com/2077-0472/10/7/277

Gartzia, M., Alados, C. L., & Pérez-Cabello, F. (2014). Assessment of the effects of biophysical and anthropogenic factors on woody plant encroachment in dense and sparse mountain grasslands based on remote sensing data. *Progress in Physical Geography: Earth and Environment*, *38*(2), 201-217. https://doi.org/10.1177/0309133314524429

Gaston, L., Locke, M., & Zablotowicz, R. (1996). *Sorption and degradation of bentazon in conventional-and no-till Dundee soil* (0047-2425).

Ge, J., & Zou, C. (2013). Impacts of woody plant encroachment on regional climate in the southern Great Plains of the United States [https://doi.org/10.1002/jgrd.50634]. *Journal of Geophysical Research: Atmospheres*, *118*(16), 9093-9104. https://doi.org/https://doi.org/10.1002/jgrd.50634

Gerke, J. (2022a). The central role of soil organic matter in soil fertility and carbon storage. *Soil Systems*, *6*(2), 33.

Gerke, J. (2022b). The Central Role of Soil Organic Matter in Soil Fertility and Carbon Storage. *Soil Systems*, *6*(2).

Geudtner, D., Torres, R., Snoeij, P., Davidson, M., & Rommen, B. (2014, 13-18 July 2014). Sentinel-1 System capabilities and applications. 2014 IEEE Geoscience and Remote Sensing Symposium,

Ghosh, P., & Mahanta, S. (2014). Carbon sequestration in grassland systems. *Range Management and Agroforestry*, *35*(2), 173-181.

Gitelson, A. A., Merzlyak, M., Zur, Y., Stark, R., & Gritz, U. (2001). Non-destructive and remote sensing techniques for estimation of vegetation status.

Grimm, R., Behrens, T., Märker, M., & Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island — Digital soil mapping using Random Forests analysis. *Geoderma*, *146*(1), 102-113. https://doi.org/https://doi.org/10.1016/j.geoderma.2008.05.008

Guo, L., Sun, X., Fu, P., Shi, T., Dang, L., Chen, Y., Linderman, M., Zhang, G., Zhang, Y., Jiang, Q., Zhang, H., & Zeng, C. (2021). Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas. *Geoderma*, *398*, 115118. https://doi.org/https://doi.org/10.1016/j.geoderma.2021.115118

Hair Jr, J., Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Hancock, D. W., & Dougherty, C. T. (2007). Relationships between blue-and red-based vegetation indices and leaf area and yield of alfalfa. *Crop Science*, *47*(6), 2547-2556.

Henrich, V., Krauss, G., Götze, C., & Sandow, C. (2011). The IndexDatabase. https://www.indexdatabase.de/

Hernandez-Clemente, R., Hornero, A., Gonzalez-Dugo, V., Berdugo, M., Quero, J. L., Jimenez, J. C., & Maestre, F. T. (2023). Global monitoring of soil multifunctionality in drylands using satellite imagery and field data. *Remote Sensing in Ecology and Conservation*. https://doi.org/10.1002/rse2.340

Hibbard, K., Schimel, D., Archer, S., Ojima, D., & Parton, W. (2003). Grassland to woodland transitions: integrating changes in landscape structure and biogeochemistry. *Ecological Applications*, *13*(4), 911-926.

Honda, E. A., & Durigan, G. (2016). Woody encroachment and its consequences on hydrological processes in the savannah. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1703), 20150313.

Huete, A., Justice, C., & Van Leeuwen, W. (1999). MODIS vegetation index (MOD13). *Algorithm theoretical basis document*, *3*(213), 295-309.

Huete, A. R., Liu, H. Q., Batchily, K., & van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, *59*(3), 440-451. https://doi.org/10.1016/s0034-4257(96)00112-5

Hunt Jr, E. R., Daughtry, C., Eitel, J. U., & Long, D. S. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agronomy journal*, *103*(4), 1090-1099.

Ismail, R., & Mutanga, O. (2010). A comparison of regression tree ensembles: Predicting Sirex noctilio induced water stress in Pinus patula forests of KwaZulu-Natal, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, *12*, S45-S51.

Jackson, L. E., Potthoff, M., Steenwerth, K. L., O'Geen, A. T., Stromberg, M. R., & Scow, K. M. (2007). 9 Soil Biology and Carbon Sequestration in Grasslands. In S. Mark, C. Jeffrey, & D. A. Carla (Eds.), *California Grasslands* (pp. 107-118). University of California Press. https://doi.org/doi:10.1525/9780520933972-015

Jackson, R. B., Banner, J. L., Jobbágy, E. G., Pockman, W. T., & Wall, D. H. (2002). Ecosystem carbon loss with woody plant invasion of grasslands. *Nature*, *418*(6898), 623-626.

Jaganyi, J. N. U. (1998). *Response of carabid and cicindelid beetles to various types of landscape disturbances*

Kerr, D. D., & Ochsner, T. E. (2020). Soil organic carbon more strongly related to soil moisture than soil temperature in temperate grasslands. *Soil Science Society of America Journal*, *84*(2), 587-596. https://doi.org/https://doi.org/10.1002/saj2.20018

Kgosikoma, O. E., & Mogotsi, K. (2013). Understanding the causes of bush encroachment in Africa: The key to effective management of savanna grasslands. *Tropical grasslands-forrajes tropicales*, *1*(2), 215-219.

Kim, Y., Jackson, T., Bindlish, R., Lee, H., & Hong, S. (2011). Radar vegetation index for estimating the vegetation water content of rice and soybean. *IEEE Geoscience and Remote Sensing Letters*, *9*(4), 564-568.

King, T. S., & Chinchilli, V. M. (2001). Robust estimators of the concordance correlation coefficient. *Journal of biopharmaceutical statistics*, *11*(3), 83-105.

Kirfel, K., Heinze, S., Hertel, D., & Leuschner, C. (2019). Effects of bedrock type and soil chemistry on the fine roots of European beech–A study on the belowground plasticity of trees. *Forest Ecology and Management*, *444*, 256-268.

Konen, M. E., Jacobs, P. M., Burras, C. L., Talaga, B. J., & Mason, J. A. (2002). Equations for predicting soil organic carbon using loss-on-ignition for north central US soils. *Soil Science Society of America Journal*, *66*(6), 1878-1881.

Kooistra, L., Leuven, R., Wehrens, R., Nienhuis, P., & Buydens, L. (2003). A comparison of methods to relate grass reflectance to soil metal contamination. *International Journal of Remote Sensing*, *24*(24), 4995-5010.

Koparan, M., Rekabdarkolaee, H., Sood, K., Westhoff, S., Reese, C., & Malo, D. (2022). Estimating soil organic carbon levels in cultivated soils from satellite image using

parametric and data-driven methods. *International Journal of Remote Sensing*, *43*(9), 3429-3449.

Koparan, M. H. (2019). *Estimating Soil Organic Carbon in Cultivated Soils Using Soil Test Data, Remote Sensing Imagery from Satellites (Landsat 8 and Planetscope), and Web Soil Survey Data*. South Dakota State University.

Koparan, M. H., Rekabdarkolaee, H. M., Sood, K., Westhoff, S. M., Reese, C. L., & Malo, D. D. (2022). Estimating soil organic carbon levels in cultivated soils from satellite image using parametric and data-driven methods. *International Journal of Remote Sensing*, *43*(9), 3429-3449. https://doi.org/10.1080/01431161.2022.2093144

Kraai, U. M., Kraai, M., Tsvuura, Z., Mkhize, N. R., & Tjelele, T. J. (2023). The impacts of Cape porcupines on woody plant mortality. *Austral Ecology*.

Lamichhane, S., Kumar, L., & Wilson, B. (2019). Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma*, *352*, 395-413. https://doi.org/https://doi.org/10.1016/j.geoderma.2019.05.031

Larochelle, H., Bengio, Y., Louradour, J., & Lamblin, P. (2009). Exploring strategies for training deep neural networks. *Journal of machine learning research*, *10*(1).

Li, H., Shen, H., Chen, L., Liu, T., Hu, H., Zhao, X., Zhou, L., Zhang, P., & Fang, J. (2016). Effects of shrub encroachment on soil organic carbon in global grasslands. *Scientific Reports*, *6*(1), 28974.

Li, Z., Zhang, H. K., Roy, D. P., Yan, L., & Huang, H. (2020). Sharpening the Sentinel-2 10 and 20 m Bands to Planetscope-0 3 m Resolution. *Remote Sensing*, *12*(15), 2406.

Liu, F., Wu, X. B., Bai, E., Boutton, T. W., & Archer, S. R. (2011). Quantifying soil organic carbon in complex landscapes: an example of grassland undergoing encroachment of woody plants. *Global change biology*, *17*(2), 1119-1129.

Lorenz, K., & Lal, R. (2005). The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. *Advances in agronomy*, *88*, 35-66.

Loveland, P., & Webb, J. (2003). Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. *Soil and Tillage research*, *70*(1), 1-18.

Lozano-Parra, J., Schnabel, S., Pulido, M., Gómez-Gutiérrez, Á., & Lavado-Contador, F. (2018). Effects of soil moisture and vegetation cover on biomass growth in water-limited environments. *Land Degradation & Development*, *29*(12), 4405-4414.

Luo, C., Zhang, W., Zhang, X., & Liu, H. (2024). Mapping the soil organic matter content in a typical black-soil area using optical data, radar data and environmental covariates. *Soil and Tillage research*, *235*, 105912.

Mahdianpari, M., Jafarzadeh, H., Granger, J. E., Mohammadimanesh, F., Brisco, B., Salehi, B., Homayouni, S., & Weng, Q. (2020). A large-scale change monitoring of wetlands using time series Landsat imagery on Google Earth Engine: a case study in Newfoundland. *GIScience & Remote Sensing*, *57*(8), 1102-1124.

Main, R., Cho, M. A., Mathieu, R., O'Kennedy, M. M., Ramoelo, A., & Koch, S. (2011). An investigation into robust spectral indices for leaf chlorophyll estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 751-761.

Marta, S. (2018). Planet imagery product specifications. *Planet Labs: San Francisco, CA, USA*, *91*.

Matiza, C., Mutanga, O., Odindi, J., & Mngadi, M. (2024). The utility of Planetscope spectral data in quantifying above-ground carbon stock in an urban reforested landscape. *Ecological Informatics*, *80*, 102472. https://doi.org/https://doi.org/10.1016/j.ecoinf.2024.102472

Matongera, T. N., Mutanga, O., Dube, T., & Sibanda, M. (2017). Detection and mapping the spatial distribution of bracken fern weeds using the Landsat 8 OLI new generation sensor. *International Journal of Applied Earth Observation and Geoinformation*, *57*, 93-103.

Mayer, M., Prescott, C. E., Abaker, W. E., Augusto, L., Cécillon, L., Ferreira, G. W., James, J., Jandl, R., Katzensteiner, K., & Laclau, J.-P. (2020). Tamm Review: Influence of forest management activities on soil organic carbon stocks: A knowledge synthesis. *Forest Ecology and Management*, *466*, 118127.

McGrath, D., & Zhang, C. (2003). Spatial distribution of soil organic carbon concentrations in grassland of Ireland. *Applied Geochemistry*, *18*(10), 1629-1639. https://doi.org/https://doi.org/10.1016/S0883-2927(03)00045-3

McKinley, D. C., Norris, M. D., Blair, J. M., & Johnson, L. C. (2008). Altered ecosystem processes as a consequence of Juniperus virginiana L. encroachment into North American tallgrass prairie. *Western North American Juniperus communities: A dynamic vegetation type*, 170-187.

Merino, A. n., Pérez-Batallón, P., & Macías, F. (2004). Responses of soil organic matter and greenhouse gas fluxes to soil management and land use changes in a humid temperate region of southern Europe. *Soil Biology and Biochemistry*, *36*(6), 917-925.

Metternicht, G. (2003). Vegetation indices derived from high-resolution airborne videography for precision crop management. *International Journal of Remote Sensing*, *24*(14), 2855-2877.

Miura, T., Yoshioka, H., Fujiwara, K., & Yamamoto, H. (2008). Inter-comparison of ASTER and MODIS surface reflectance and vegetation index products for synergistic applications to natural resource monitoring. *Sensors*, *8*(4), 2480-2499.

Mngadi, M., Odindi, J., & Mutanga, O. (2021). The Utility of Sentinel-2 Spectral Data in Quantifying Above-Ground Carbon Stock in an Urban Reforested Landscape. *Remote Sensing*, *13*(21), 4281. https://www.mdpi.com/2072-4292/13/21/4281

Moberly, P., & Meyer, J. (1984). Soils: A management factor in sugarcane production in the South African sugar industry. *Proceedings of The South African Sugar Technologists' Association-June*, 193.

Mochi, L. S., Mazía, N., Biganzoli, F., & Aguiar, M. R. (2022). Contrasting effects of grazing on the early stages of woody encroachment in a Neotropical savanna. *Basic and Applied Ecology*, *60*, 13-24. https://doi.org/https://doi.org/10.1016/j.baae.2022.01.007

Montavon, G., Samek, W., & Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1-15. https://doi.org/https://doi.org/10.1016/j.dsp.2017.10.011

Morais, T. G., Jongen, M., Tufik, C., Rodrigues, N. R., Gama, I., Serrano, J., Gonçalves, M. C., Mano, R., Domingos, T., & Teixeira, R. F. (2023). Satellite-based estimation of soil organic carbon in Portuguese grasslands. *Frontiers in Environmental Science*, *11*, 1240106.

Mureva, A., Ward, D., Pillay, T., Chivenge, P., & Cramer, M. (2018). Soil organic carbon increases in semi-arid regions while it decreases in humid regions due to woody-plant encroachment of grasslands in South Africa. *Scientific Reports*, *8*(1), 15506.

Mutanga, O., Rouget, M., Hlanguza, N., & Odindi, J. (2016). Mapping alien and indigenous vegetation in the KwaZulu-Natal Sandstone Sourveld using remotely sensed data. *Bothalia-African Biodiversity & Conservation*, *46*(2), 1-9.

Mzinyane, T., van Aardt, J., & Gebreslasie, M. T. (2015). Soil carbon estimation from eucalyptus grandis using canopy spectra. *South African Journal of Geomatics*, *4*(4), 548-561.

Nabiollahi, K., Eskandari, S., Taghizadeh-Mehrjardi, R., Kerry, R., & Triantafilis, J. (2019). Assessing soil organic carbon stocks under land-use change scenarios using random

forest models. *Carbon Management*, *10*(1), 63-77. https://doi.org/10.1080/17583004.2018.1553434

Nasiri, V., Deljouei, A., Moradi, F., Sadeghi, S. M., & Borz, S. A. (2022). Land Use and Land Cover Mapping Using Sentinel-2, Landsat-8 Satellite Images, and Google Earth Engine: A Comparison of Two Composition Methods. *Remote Sensing*, *14*(9).

Odebiri, O., Mutanga, O., & Odindi, J. (2022). Deep learning-based national scale soil organic carbon mapping with Sentinel-3 data. *Geoderma*, *411*, 115695. https://doi.org/https://doi.org/10.1016/j.geoderma.2022.115695

Odebiri, O., Mutanga, O., Odindi, J., & Naicker, R. (2022). Modelling soil organic carbon stock distribution across different land-uses in South Africa: A remote sensing and deep learning approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, *188*, 351-362. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2022.04.026

Odebiri, O., Mutanga, O., Odindi, J., Peerbhay, K., & Dovey, S. (2020). Predicting soil organic carbon stocks under commercial forest plantations in KwaZulu-Natal province, South Africa using remotely sensed data. *GIScience & Remote Sensing*, *57*(4), 450-463. https://doi.org/10.1080/15481603.2020.1731108

Odebiri, O., Mutanga, O., Odindi, J., Slotow, R., Mafongoya, P., Lottering, R., Naicker, R., Nyasha Matongera, T., & Mngadi, M. (2024). Remote sensing of depth-induced variations in soil organic carbon stocks distribution within different vegetated landscapes. *Catena*, *243*, 108216. https://doi.org/https://doi.org/10.1016/j.catena.2024.108216

Odebiri, O., Odindi, J., & 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.

Okin, G. S., D'Odorico, P., & Archer, S. R. (2009). Impact of feedbacks on Chihuahuan desert grasslands: transience and metastability. *Journal of Geophysical Research: Biogeosciences*, *114*(G1).

Otto, D. (2019). Musina-Makhado Special Economic Zone Development, Vhembe District Municipality, Limpopo Province Soils and Land Capability.

Perry, T. O. (1989). Tree roots: facts and fallacies. *Arnoldia*, *49*(4), 3-29.

Pinheiro, L. F. S., Kansbock, L., Rossatto, D. R., & Kolb, R. M. (2022). Woody plant encroachment constrains regeneration of ground-layer species in a neotropical savanna from seeds. *Austral Ecology*, *47*(3), 674-684.

Pouladi, N., Møller, A. B., Tabatabai, S., & Greve, M. H. (2019). Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. *Geoderma*, *342*, 85-92. https://doi.org/https://doi.org/10.1016/j.geoderma.2019.02.019

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119-126.

Qiu, L., Wei, X., Zhang, X., Cheng, J., Gale, W., Guo, C., & Long, T. (2012). Soil organic carbon losses due to land use change in a semiarid grassland. *Plant and soil*, *355*, 299-309.

Rapinel, S., Clément, B., Magnanon, S., Sellin, V., & Hubert-Moy, L. (2014). Identification and mapping of natural vegetation on a coastal site using a Worldview-2 satellite image. *Journal of Environmental Management*, *144*, 236-246.

Rasse, D. P., Rumpel, C., & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and soil*, *269*(1-2), 341-356.

Ratajczak, Z., Nippert, J. B., & Collins, S. L. (2012). Woody encroachment decreases diversity across North American grasslands and savannas. *Ecology*, *93*(4), 697-703. https://doi.org/https://doi.org/10.1890/11-1199.1

Rossouw, P. (2016). SOIL, AGRICULTURAL POTENTIAL, LAND CAPABILITY AND LAND USE STUDY: IMPACT OF A PROPOSED TAILINGS PIPELINE ON THE FARMS FRISCHGEWAAGD AND MIMOSA, NEAR THE TOWN OF LEDIG, NORTH WEST PROVINCE.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ*, *351*(1), 309.

Roy, D. P., Huang, H., Houborg, R., & Martins, V. S. (2021). A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery. *Remote Sensing of Environment*, *264*, 112586.

Ruiz-Sinoga, J., Martínez-Murillo, J., Gabarrón-Galeote, M., & García-Marín, R. (2011). The effects of soil moisture variability on the vegetation pattern in Mediterranean abandoned fields (Southern Spain). *Catena*, *85*(1), 1-11.

Rumpel, C., & Kögel-Knabner, I. (2011). Deep soil organic matter—a key but poorly understood component of terrestrial C cycle. *Plant and soil*, *338*, 143-158.

Saintilan, N., & Rogers, K. (2015). Woody plant encroachment of grasslands: a comparison of terrestrial and wetland settings. *New Phytologist*, *205*(3), 1062-1070.

Sankaran, M., Ratnam, J., & Hanan, N. P. (2004). Tree–grass coexistence in savannas revisited–insights from an examination of assumptions and mechanisms invoked in existing models. *Ecology letters*, *7*(6), 480-490.

Santos, E. P. d., Moreira, M. C., Fernandes-Filho, E. I., Demattê, J. A. M., Dionizio, E. A., Silva, D. D. d., Cruz, R. R. P., Moura-Bueno, J. M., Santos, U. J. d., & Costa, M. H. (2023). Sentinel-1 Imagery Used for Estimation of Soil Organic Carbon by Dual-Polarization SAR Vegetation Indices. *Remote Sensing*, *15*(23), 5464. https://www.mdpi.com/2072-4292/15/23/5464

Schreiner-McGraw, A. P., Vivoni, E. R., Ajami, H., Sala, O. E., Throop, H. L., & Peters, D. P. (2020). Woody Plant encroachment has a larger impact than climate change on Dryland water budgets. *Scientific Reports*, *10*(1), 8112.

Schulte, E., & Hopkins, B. (1996). Estimation of soil organic matter by weight loss-on-ignition. *Soil organic matter: Analysis and interpretation*, *46*, 21-31.

Scott, R. L., Huxman, T. E., Williams, D. G., & Goodrich, D. C. (2006). Ecohydrological impacts of woody-plant encroachment: Seasonal patterns of water and carbon dioxide exchange within a semiarid riparian environment. *Global change biology*, *12*(2), 311-324.

Shafizadeh-Moghadam, H., Minaei, F., Talebi-khiyavi, H., Xu, T., & Homaee, M. (2022). Synergetic use of multi-temporal Sentinel-1, Sentinel-2, NDVI, and topographic factors for estimating soil organic carbon. *Catena*, *212*, 106077.

Shimizu, K., Ota, T., Mizoue, N., & Saito, H. (2020). Comparison of multi-temporal PlanetScope data with Landsat 8 and Sentinel-2 data for estimating Airborne LiDAR Derived Canopy Height in Temperate forests. *Remote Sensing*, *12*(11), 1876.

Singh, B., Sihag, P., & Singh, K. (2017). Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Modeling Earth Systems and Environment*, *3*, 999-1004.

Smit, I. P., & Prins, H. H. (2015). Predicting the effects of woody encroachment on mammal communities, grazing biomass and fire frequency in African savannas. *PloS one*, *10*(9), e0137857.

Sommer, R., & Bossio, D. (2014). Dynamics and climate change mitigation potential of soil organic carbon sequestration. *Journal of Environmental Management*, *144*, 83-87. https://doi.org/https://doi.org/10.1016/j.jenvman.2014.05.017

Sreenivas, K., Sujatha, G., Sudhir, K., Kiran, D. V., Fyzee, M., Ravisankar, T., & Dadhwal, V. (2014). Spatial assessment of soil organic carbon density through random forests based imputation. *Journal of the Indian Society of Remote Sensing*, *42*, 577-587.

Stevens, N., Lehmann, C. E., Murphy, B. P., & Durigan, G. (2017). Savanna woody encroachment is widespread across three continents. *Global change biology*, *23*(1), 235-244.

Tan, J., Ding, J., Han, L., Ge, X., Wang, X., Wang, J., Wang, R., Qin, S., Zhang, Z., & Li, Y. (2023). Exploring PlanetScope Satellite Capabilities for Soil Salinity Estimation and Mapping in Arid Regions Oases. *Remote Sensing*, *15*(4), 1066. https://www.mdpi.com/2072-4292/15/4/1066

Tessema, B., Sommer, R., Piikki, K., Söderström, M., Namirembe, S., Notenbaert, A., Tamene, L., Nyawira, S., & Paul, B. (2020). Potential for soil organic carbon sequestration in grasslands in East African countries: A review. *Grassland Science*, *66*(3), 135-144.

Thaler, E. A., Larsen, I. J., & Yu, Q. (2019). A new index for remote sensing of soil organic carbon based solely on visible wavelengths. *Soil Science Society of America Journal*, *83*(5), 1443-1450.

Throop, H. L., & Archer, S. R. (2008). Shrub (Prosopis velutina) encroachment in a semidesert grassland: spatial–temporal changes in soil organic carbon and nitrogen pools. *Global change biology*, *14*(10), 2420-2431.

Torres, R., Navas-Traver, I., Bibby, D., Lokas, S., Snoeij, P., Rommen, B., Osborne, S., Ceba-Vega, F., Potin, P., & Geudtner, D. (2017, 8-12 May 2017). Sentinel-1 SAR system and mission. 2017 IEEE Radar Conference (RadarConf),

Tu, C., He, T., Lu, X., Luo, Y., & Smith, P. (2018). Extent to which pH and topographic factors control soil organic carbon level in dry farming cropland soils of the mountainous region of Southwest China. *Catena*, *163*, 204-209.

Tziachris, P., Aschonitis, V., Chatzistathis, T., & Papadopoulou, M. (2019). Assessment of spatial hybrid methods for predicting soil organic matter using DEM derivatives and soil parameters. *Catena*, *174*, 206-216.

Vågen, T.-G., & Winowiecki, L. A. (2013). Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters*, *8*(1), 015011. https://doi.org/10.1088/1748-9326/8/1/015011

Van Auken, O. (2009). Causes and consequences of woody plant encroachment into western North American grasslands. *Journal of Environmental Management*, *90*(10), 2931-2942.

Van Auken, O. W. (2009). Causes and consequences of woody plant encroachment into western North American grasslands. *Journal of Environmental Management*, *90*(10), 2931-2942. https://doi.org/https://doi.org/10.1016/j.jenvman.2009.04.023

Van Huyssteen, C., Rantoa, N., & Du Preez, C. (2021). Organic carbon content in the diagnostic horizons and materials of South African soil forms. *South African Journal of Plant and Soil*, *38*(2), 140-151.

Venter, Z. S., Hawkins, H.-J., Cramer, M. D., & Mills, A. J. (2021). Mapping soil organic carbon stocks and trends with satellite-driven high resolution maps over South Africa. *Science of The Total Environment*, *771*, 145384. https://doi.org/https://doi.org/10.1016/j.scitotenv.2021.145384

Wang, B., Waters, C., Orgill, S., Gray, J., Cowie, A., Clark, A., & Liu, D. L. (2018). High resolution mapping of soil organic carbon stocks using remote sensing variables in the semi-arid rangelands of eastern Australia. *Science of the Total Environment*, *630*, 367-378. https://doi.org/https://doi.org/10.1016/j.scitotenv.2018.02.204

Wang, K., Qi, Y., Guo, W., Zhang, J., & Chang, Q. (2021). Retrieval and mapping of soil organic carbon using Sentinel-2A spectral images from bare cropland in autumn. *Remote Sensing*, *13*(6), 1072.

Wang, W., & Fang, J. (2009). Soil respiration and human effects on global grasslands. *Global and Planetary Change*, *67*(1-2), 20-28.

Wang, X., Han, J., Wang, X., Yao, H., & Zhang, L. (2021). Estimating soil organic matter content using sentinel-2 imagery by machine learning in shanghai. *IEEE Access*, *9*, 78215-78225.

Ward, D., Kirkman, K., & Tsvuura, Z. (2017). An African grassland responds similarly to long-term fertilization to the Park Grass experiment. *PLoS One*, *12*(5), Article e0177208. https://doi.org/10.1371/journal.pone.0177208

Wieczorkowski, J. D., & Lehmann, C. E. R. (2022). Encroachment diminishes herbaceous plant diversity in grassy ecosystems worldwide [https://doi.org/10.1111/gcb.16300]. *Global Change Biology*, *28*(18), 5532-5546. https://doi.org/https://doi.org/10.1111/gcb.16300

Wolkovich, E. M., Lipson, D. A., Virginia, R. A., Cottingham, K. L., & Bolger, D. T. (2010). Grass invasion causes rapid increases in ecosystem carbon and nitrogen storage in a semiarid shrubland. *Global change biology*, *16*(4), 1351-1365.

Wu, J., Wang, D., & Bauer, M. E. (2007). Assessing broadband vegetation indices and QuickBird data in estimating leaf area index of corn and potato canopies. *Field Crops Research*, *102*(1), 33-42. https://doi.org/https://doi.org/10.1016/j.fcr.2007.01.003

Xie, B., Ding, J., Ge, X., Li, X., Han, L., & Wang, Z. (2022). Estimation of soil organic carbon content in the Ebinur Lake wetland, Xinjiang, China, based on multisource remote sensing data and ensemble learning algorithms. *Sensors*, *22*(7), 2685.

Xu, X. B., & Zhai, X. Y. (2023). Mapping Soil Organic Matter Content during the Bare Soil Period by Using Satellite Data and an Improved Deep Learning Network. *SUSTAINABILITY*, *15*(1), Article 323. https://doi.org/10.3390/su15010323

Yang, L., Cai, Y., Zhang, L., Guo, M., Li, A., & Zhou, C. (2021). A deep learning method to predict soil organic carbon content at a regional scale using satellite-based phenology variables. *International Journal of Applied Earth Observation and Geoinformation*, *102*, 102428.

Yang, L., He, X., Shen, F., Zhou, C., Zhu, A.-X., Gao, B., Chen, Z., & Li, M. (2020). Improving prediction of soil organic carbon content in croplands using phenological parameters extracted from NDVI time series data. *Soil and Tillage research*, *196*, 104465.

Yang, R.-M., & Guo, W.-W. (2019). Modelling of soil organic carbon and bulk density in invaded coastal wetlands using Sentinel-1 imagery. *International Journal of Applied Earth Observation and Geoinformation*, *82*, 101906.

Yang, X., Wang, B., Fakher, A., An, S., & Kuzyakov, Y. (2023). Contribution of roots to soil organic carbon: From growth to decomposition experiment. *Catena*, *231*, 107317. https://doi.org/https://doi.org/10.1016/j.catena.2023.107317

Yu, H., Zha, T., Zhang, X., Nie, L., Ma, L., & Pan, Y. (2020). Spatial distribution of soil organic carbon may be predominantly regulated by topography in a small revegetated watershed. *Catena*, *188*, 104459.

Zablotowicz, R., Locke, M., Gaston, L., & Bryson, C. (2000). Interactions of tillage and soil depth on fluometuron degradation in a Dundee silt loam soil. *Soil and Tillage research*, *57*(1-2), 61-68.

Zhang, L., Shao, Z., Liu, J., & Cheng, Q. (2019). Deep learning based retrieval of forest aboveground biomass from combined LiDAR and landsat 8 data. *Remote Sensing*, *11*(12), 1459.

Zhao, Y., & Liu, D. (2022). A robust and adaptive spatial-spectral fusion model for PlanetScope and Sentinel-2 imagery. *GIScience & Remote Sensing*, *59*(1), 520-546.

Zhou, T., Geng, Y., Chen, J., Pan, J., Haase, D., & Lausch, A. (2020). High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. *Science of The Total Environment*, *729*, 138244. https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.138244

Zhou, Y., Boutton, T. W., & Wu, X. B. (2017). Soil carbon response to woody plant encroachment: importance of spatial heterogeneity and deep soil storage. *Journal of Ecology*, *105*(6), 1738-1749. https://doi.org/https://doi.org/10.1111/1365-2745.12770

Zhou, Y., Boutton, T. W., & Wu, X. B. (2018). Woody plant encroachment amplifies spatial heterogeneity of soil phosphorus to considerable depth. *Ecology*, *99*(1), 136-147. https://doi.org/https://doi.org/10.1002/ecy.2051

Žížala, D., Minařík, R., & Zádorová, T. (2019). Soil organic carbon mapping using multispectral remote sensing data: Prediction ability of data with different spatial and spectral resolutions. *Remote Sensing*, *11*(24), 2947.