Grass nutrients estimation as an indicator of rangeland quality using satellite remote sensing

Katlego Kgabo Mashiane

(2018711909)

A thesis submitted to the faculty of Natural and Agricultural Sciences, Department of Geography at the University

of the

Free State, in fulfilment of the academic requirements for the degree of

Doctor of Philosophy in Geography

December 2022

Free State

South Africa

Acknowledgements

First and foremost, I would like to thank God for giving me the strength to carry on; even when it was difficult to do so, I was comforted by the presence of the Almighty throughout the difficulties.

I thank the University of the Free State: Postgraduate School, South African National Parks, and the National Research Foundation in South Africa for funding this PhD study. This is not the support I take for granted; I am, therefore, very grateful for the incentives to conduct this study.

I am thankful for my co-promoter, Prof Abel Ramoelo, who always dedicated his time to answering my questions and guiding me. I am most grateful for the scientific engagements during our field trips to the National Parks (NP) of South Africa: Tankwa NP, Mokala NP, and Golden Gate National NP. I would like to thank my promoter Prof Samuel Adelabu for his support during the fieldwork planning, funding conferences, and contribution to completing my study; I mostly enjoyed the remote sensing seminars organised by Prof Samuel Adelabu on the QwaQwa campus.

I would like to thank the staff at SANParks, especially Mr Ernest Daemane, for warmly welcoming me to the organisation and our discussions on topical scientific issues. Conducting research and working in the Golden Gate Highlands National Park (GGHNP) has been one of the most beautiful experiences of my life. A special thank you to Mr Rodney Makwakwa (Field Assistant), Mpho Mbedzi (SANParks: Scientific Services intern (2019)), Nomcebo Dlamini (Field Ranger), Mpho Mountaung (Field Ranger), and Lucia Mokeona (Environmental Monitor) for assisting me in the field during data collection. A special thank you to the late Ntate Tankiso (Former Senior Ranger: GGHNP), a very supportive, caring, and embracing field ranger; in him, I had a friend, father, and colleague.

I would also like to thank the Department of Geography, Bloemfontein campus staff for always encouraging and supporting me when I needed it the most, especially Dr Rajab Matamanda for his words of wisdom and for instilling courage and, most importantly, giving me hope.

I would like to thank my wife, Palesa, for supporting and encouraging me throughout this PhD study and our son, Kamohelo, for being my pillar of strength. My wife has been and is still is an excellent mother to our son, and I am blessed to have both in my life.

This thesis is approved by:

Prof Samuel Adelabu, **promoter,** University of the Free State

Prof Abel Ramoelo, **co-promoter,** University of Pretoria

Table of Contents

Abstract

South Africa's grasslands are known for their rich biodiversity, which makes for rangelands with expansive landscapes for grazing. Numerous studies have been conducted on how species variety and composition affect the sustainability and productivity of grasslands. The provision of ecosystem functions and services, mainly grazing for both domestic livestock and wild animals, is threatened by rapid biodiversity loss. Hence crucial to develop monitoring systems for biodiversity, especially considering the anticipated impacts of global environmental changes. Nutritionally sufficient grass swards strongly influence the distribution and abundance of grazing animals. Although estimates of crucial ecological and biodiversity indicators can be obtained from *in situ* and remote sensing data, remote sensing offers timely and cost-effective data that can be used for species monitoring.

This study sought to measure essential biodiversity variables in GGHNP related to plant productivity and nutrient availability using *in situ* and remote sensing approaches. In addition, by estimating grass species nitrogen using remote sensing, the research intended to improve carrying capacity models and stocking rate in the park, thus ensuring efficient rangeland management. Assessments of species diversity and richness are pivotal for devising effective conservation strategies. These biodiversity metrics are good indicators of rangeland quality, health, and ecological response to disturbances. Data collected using *in situ* methods showed that species richness are virtually similar across the park under different richness of disturbance regime. However, species richness were higher at landscape richness than at the site richness. Albeit criticised for their use in conservation, species diversity metrics could be useful for measurements of rangeland quality.

Due to its intolerance to harsh environmental circumstances, mountainous grassland vegetation is probably the most susceptible to environmental changes on a worldwide scale. Understanding the factors that influence species distribution in alpine grasslands will be crucial for identifying biodiversity and preserving it. This evident in that an increase in topo-edaphic variables negatively affects species richness, while slope and elevation showed an improvement of species richness. The influence of topography and other accompanying factors on species diversity is highlighted, emphasizing how topography affects species dispersion in mountainous grassland communities.

Understanding the value of environmental conditions are to the geographical distribution of biodiversity has been the focus of the most active ecological research. The selection of appropriate modelling algorithms could be beneficial for gaining insights into biodiversity-environment relations. Non-parametric and parametric modelling frameworks were used to assess these relationships. Topographically controlled edaphic variables continued to be the most significant drivers of species richness and diversity in grassland plant communities in the park, despite higher prediction accuracies being attained using parametric models.

Remote sensing permits rapid and inexpensive recording and assessment of vegetation over short to long-term periods at a local and global scales. This study sought to predict and model species richness and diversity in GGHNP. Near-infrared (NIR) was the most selected spectral interval for predicting species diversity, further ascertaining the efficiency of NIR in vegetation mapping.

Grass species N (grass N) estimation is valueable for rangeland management because it determines their forage quality which has nutritional implications for grazing animals. I used remote sensing data to predict grass N in the park using Sentinel 2 Multi-Spectral Instrument (S2 MSI). The results showed red edge bands as the optimal bands for estimating grass N, which makes S2 MSI superior for modelling grass N throughout grass phenology and among seasons because of its multiple red edge bands.

Grassland monitoring is imperative for both assessing global change impacts and the security of sustainable development goals. Grasslands serve as rangelands that supply forage for domestic and wild animals. Monitoring of forage quality and quantity is crucial for evaluating carrying capacity models and thus ensuring effective rangeland management. Plant species richness and species diversity are key indicators for plant primary productivity in rangelands and many other ecosystem health parameters. Due to the immensity of rangeland landscapes, remote sensing could be the effective technology for determining and keeping track of ecological parameters in grasslands.

Declaration

I, **Katlego K Mashiane**, declare that the thesis that I herewith submit for the Doctoral Degree in **Geography** at the University of the Free State, is my independent work and that I have not previously submitted it for a qualification at another institution of higher education

Chapter 1

Introduction

Global biodiversity at the species level is fast declining as a result of several problems caused by humans, including pollution, habitat loss, and climate change. (Steffen *et al.*, 2004). High levels of biodiversity are pivotal for an ecosystem's resilience, functioning, and provision of services (Oliver *et al.*, 2015). Hence, an extreme loss of biodiversity can have dire ramifications on the health of an ecosystem and its ability to provide services. In grassland utilised for grazing (rangelands), loss of species diversity strongly affects forage quality and nutritive value, thus affecting livestock and wild herbivores production and nutrition (Hector *et al*., 2010, Zarovali, Yiakoulaki and Papanastasis, 2007). These effects may have devastating impacts on rangeland carrying capacity i.e., the maximum number of large herbivores sustained per unit area. Monitoring systems for biodiversity are warranted to gain insights into the loss of grassland biodiversity resulting from unprecedented rates of global change in rangeland quality and ultimately carrying capacity (van Oijen, Bellocchi, and Höglind, 2018).

South African rangelands have high species richness (Siebert, Siebert and Du Toit *et al*., 2011), and consist of a variety of grass communities, that provide important ecosystem services such forage for and habitat to a variety of wildlife and livestock animals. Nonetheless, there is evidence of land-use transformations which are severely affecting the diversity of grass species and communities in South African rangelands (Siebert, Siebert and Du Toit *et.al*., 2011). The rangeland vegetation's ability to harvest vital resources, produce biomass, and recycle vital nutrients may be hampered by these effects on biodiversity in the rangeland (Cardinale *et al.*, 2012). Consequentially, some species may be lost and/or reduced in biomass and/or diversity, which in turn affects the nutrition on offer to grazing animals. Therefore, studies on the predicted dynamic grazing carrying capacity of rangelands are urgently needed in light of climate change, hence it is crucial to take into account the nutritional value of both multiple- and single species of grasses (van Oijen *et al*. 2018).

Rangelands provide crucial ecosystem services such as water quality, soil conservation, wildlife habitat, and recreation (Cingolani, Noy-Meir, and Díaz, 2005; Ferranto *et al.*, 2011). However, rangelands are under tremendous pressure from a wide range of environmental changes, such as population growth, changing land use and cover, and climate change (Scholes and Howden, 2003). Hence, maximising rangeland productivity while preventing land degradation amid global environmental changes is a challenge for range managers (Mccollum *et al.*, 2017). The reasons for this challenge are 1) rangelands are vast, making timely acquisition of spatial information difficult, and 2) the high climatic variability makes vegetation prediction difficult.

The primary driver of rangeland use is the need to predict the production of forage in a ranch. Notwithstanding, the lack of knowledge on how environmental changes will complicate efforts to find suitable management and mitigation measures (Mccollum *et al.*, 2017). For example, uncertain effects of climate variability, which control rangeland productivity, could present a challenge in predicting forage productivity (McCollum et al., 2017). Wessels *et al*. (2004) showed that drought-induced rangeland degradation cause pervasively low forage productivity in the semi-arid northern portions of South Africa.

Rangelands are estimated to cover over two-thirds of the earth's land surface (White, Murray, and Rohweder, 2000) and are home to varying mixtures of grass, forb, and shrub species (West, 1993). As a result, they sustain large populations of both wild and farmed herbivores (White, Murray, and Rohweder, 2000). Although rangelands provide a variety of ecosystem services, forage production for livestock production is the key ecosystem service (Skaggs, 2008). Furthermore, rangelands are of great socio-economic importance as they can provide food security and alleviate poverty (Lund, 2007). Nonetheless, human activities such as pollution, introduction of non-native species directly and indirectly pose threats to rangelands' biodiversity, which in turn their ability to offer their intrinsic ecological functions.

Environmental factors influencing grass species establishment

In grassland ecosystems, three processes of paramount importance influence plant communities; carbon and nitrogen sequestration and distribution, and rainfall-evapotranspiration ratios (McNaughton, 1982). Furthermore, several biotic and abiotic variables, such as grazing, fire, and rainfall patterns, have an impact on the composition and structure of such plant communities (McNaughton, 1983). However, climate change, invasive species, and land use changes will inevitably affect ecological processes, thus influencing grassland plant communities and diversity (O'Connor, 2011). Ultimately, nutrient availability and concentration may be affected, thereby deteriorating or improving the quality of rangeland. Research, therefore, is warranted to understand the nutrients dynamics and status of grass communities in the face of global change.

In grasslands, herbivores are the main agents of accelerating rates of nutrient cycling and altering plant structural and composition elements (McNaughton, 1985; Archibald *et al*., 2005). Additionally, in the temperate region of Northern China, the diversity of plant species is highly associated with the biophysical characteristics of the soil and vice versa (Chen *et al*., 2017; He *et al.*, 2018). Nonetheless, the establishment of grass species in temperate grasslands appears to be a result of competition amongst plant types for below-ground (soil moisture and nutrients) and above-ground (light) resources (Scott *et al*., 2010).

Biodiversity-ecosystem relationship

Due to a variety of environmental changes caused by humans, biodiversity at the species level is quickly declining on a worldwide scale (Steffen *et al.*, 2004). As a result of global change, approximately 50% of the earth's land area has altered, having a significant impact on ecological, climatic, and biological processes (Steffen *et al.*, 2004). Studies show that the severity of the consequences of species loss on ecological processes is equivalent to that of fire, ozone depletion, acidification, high CO_2, drought, climate change, and overgrazing (Hooper *et al*., 2012). Additionally, a decline in plant biodiversity may impact plant growth and nutrient cycling. (Balvanera *et al*., 2006; Cardinale *et al*., 2011). Considering this, Hooper *et al*. (2012) contend that the loss of biodiversity in the 21^{st} century may be one of the primary causes of ecological change.

High levels of biodiversity are necessary for the long-term viability of ecosystem processes and services (Oliver *et al.*, 2015). Depending on its functional qualities, each species in an ecosystem has a role to play in shaping the ecosystem's properties and services; a specific-functional trait phenomenon (Jordano *et al.*, 2013). The resilience and functionalities of ecosystems may also change as a result of species' functional qualities, such as morphological, biochemical, physiological, structural, phenological, or behavioural properties (Suding *et al.*, 2008; Jordano *et al.*, 2013). Therefore, a decline in biodiversity may affect how well ecological groups harvest biologically necessary resources, create biomass, and break down recycled biologically necessary nutrients (Cardinale et al., 2012). The number of species, their identities, evenness within the community, functional characteristics, and their interactions all have an impact on how well an ecosystem functions.

Ecological engineering and stoichiometry, respectively, allow organisms to affect how habitats are physically formed and how elements move through biogeochemical cycles (Gilbert, 2016; Jones *et al.*, 1994). The physiology, appearance, and behaviour of individual organisms at the species level, as well as population and community-level structure and composition, are primarily responsible for changes in an ecosystem (Suding *et al.*, 2008). Plant primary output increased as plant species variety increased in grassland ecosystems (Tilman *et al.*,2012). Nevertheless, research shows the functional modifications brought about shifts in species composition associated with grazing affects soil mineralisation rate (McNaughton *et al.*, 1997; Archibald, 2008).

Forage assessment and determination of rangeland carrying capacity

Large grazing mammals are crucial for increasing carrying capacity in protected areas in Africa by depositing nutrients (McNaughton *et al*., 1997). On the other hand, excessive grazing can slow down mineralisation rates by altering the species composition, cover, and diversity of the vegetation, thus, impacting the quantity and quality of forage (Archibald, 2008; Stock *et al*.,, 2009). Nonetheless, studies have indicated the mutual effect of grazing on grassland biodiversity and productivity (McNaughton, 1985a). The composition of grassland plant species reflects both site conditions and management approaches (Peratoner and Pötsch, 2019). Global change effects can therefore be observed in grassland through ground-based field surveys and remote sensing (Peratoner and Pötsch, 2019).

The distribution and abundance of large grazing mammals in Southern African rangelands are strongly influenced by nutritionally sufficient forages and vegetation dynamics (McNaughton et al.,, 1997). For example, these animals prefer to forage on swards enriched in minerals and this is pivotal for their nutrition and reproduction (McNaughton, 1990). The determination of habitat conditions and usage has important implications on stocking rates and carrying capacity (Ungar, 2019).

The use of remote sensing technology can produce continuous, spatially explicit data that is accessible at a variety of geographical and temporal scales (McCord *et al.*, 2017a). Moreover, the estimation of vegetation is core information derived from remote sensing (McCord *et al.*, 2017a). Hence, remote sensing has many potentials to assess rangeland quality indicators (e.g. nitrogen or protein, phosphorus, etc.) for the determination of rangeland carrying capacity (Hunt and Miyake, 2006; Mirik et al., 2005; Ramoelo et al., 2015, 2012). Grass nutrients and biomass can be monitored using remotely sensed vegetation indices (Ramoelo *et al.*, 2015a).). For instance, using red edge-based vegetation indices can provide more accurate estimations of the nitrogen content of grass (Ramoelo *et al.*, 2015a). Species composition is difficult to identify using remote sensing (Marsett *et al.*, 2006), hence it is important to investigate whether grass species composition and nutrient values can be mapped using *in situ* and satellite remote sensing data.

Problem statement

Many grasslands around the world are known for being hotspots of biodiversity (Blair *et al.*, 2014) which play a critical role in the dynamics and functioning of an ecosystem (Hector *et al.*, 2010). The role of species diversity and composition in the functioning of an ecosystem has been extensively studied (Tilman, 1999a; Tilman *et al.*, 2006; Weisser *et al.*, 2017). For example, biodiversity in grasslands ecosystem influences the productivity Tilman et al., 1996) and stability (Tilman *et al.*, , 2006) of grassland ecosystems. However, the rapid rate of environmental change and the ongoing loss of biodiversity globally endanger the functions and services that ecosystems provide (Oliver *et al.*, 2015). The plant species richness and diversity are good indicators of grassland site conditions and could be used for monitoring the efficiency of conservation interventions. Hence work is needed to determine the role of compositional change due to global change ecosystem resilience and resource pools using both field-based surveys and remote sensing techniques.

South African grasslands are one of the most threatened biomes in world due to a range of natural and man-made issues. Overgrazing and climate change have been shown to affect grassland species in protected areas (Brown *et al.*, 2013b). While in non-protected areas land transformation to commercial and residential use severely lead to biodiversity loss of grassland. In protected areas these threats will likely affect the mammalian herbivores which relies on the forage provided by these grasslands for nutrition (McNaughton *et al.*, 1997). However, certain rangeland with a high meta population have proven to be resilient to the threats (Shoemaker, 2020).

Human-induced environmental changes (global change) are homogenising many terrestrial ecosystems via unprecedented biodiversity loss (Hautier *et al.*, 2015). In grasslands, land-use intensification and climate change are the most important global change drivers contributing to a loss of biodiversity (Allan *et al.*, 2015; van Oijen, *et al.*, 2018), thus a range of ecosystem functions and processes (Tilman *et al.*, 1996). Since plant species vary in their characteristics, their composition has a significant impact on the dynamics and operation of ecosystems (Tilman, 1999b). Global change might result in a decline in local, regional and global species pools which could weaken species' ability to contribute to ecosystem functions and services (Jordano *et al.*, 2013). However, studying how biodiversity affects an ecosystem in the face of climate change is essential since the effects of biodiversity loss on ecosystem and plant community functioning differ (Li *et al.*, 2018).

The plant species diversity in grasslands of protected areas is strongly influenced by management practices such as fire (Myers *et al.*, 2000). However, grazing is the most important driver of vegetation dynamics, species distribution, and forage quality and quantity (Snyman *et al*, 2013). In grasslands utilised for grazing (rangelands), loss of species diversity strongly affects forage quality, especially nutritive value (Zarovali *et al*., 2007; Hector *et al*., 2010). High grazing pressure significantly affects individual grass function. Moreover, there is evidence that environmental changes in grasslands of South Africa may have implications for land degradation-desertification, soil erosion and bush encroachment/plant invasion (Brown and Bezuidenhout, 2020).

Grass communities are highly variable in terms of species composition and function, hence, may respond differently to climate variability, topographical gradient and site management practices i.e. fire and grazing frequencies (Mauchamp *et al.*, 2014). Golden Gates Highlands National park (GGHNP) is a protected mountainous grassland located on the foothills of the Drakensberg (Maluti) and forms part of the mesic highveld grassland with marked variation in geology, elevation, topography, and rainfall (Mucina & Rutherford, 2006). The park comprise the following grasslands units; Eastern Free State sandy grasslands, Basotho montane shrubland, Lesotho highveld basalt grassland and Northern Drakensberg highveld (Mucina & Rutherford, 2006). Therefore, research is necessary to understand how these grass communities use nutrients, especially considering global environmental and climate change, to increase the precision of stocking rates, increase the carrying capacity of protected areas, and, ultimately, improve natural resource management. In addition, there is a need for models for the complex and variable relationship between grassland biodiversity and productivity.

Moreover, GGHNP is grazed by both domesticated livestock and wild animals thus experiences high grazing frequencies and intensity. In addition, although fire is used as a management practice for sustaining grass swards wildfires are often recorded in the park. The information and data on ecosystem parameters derived from this research can be used to plan for monitoring plant species and endangered ecosystems to make conservation decisions. In addition, such studies should be used to inform conservation decisions concerning important plant groups in order to describe the habitat species composition and vegetation structure of plant communities as the dynamics of these elements are important for grassland/rangeland management.

Research objectives:

1. Provide quantitative estimates of species richness and diversity as a measure of rangeland quality across the Golden Gate Highlands National Park and within landscapes of grassland plant communities
2. Determine the relationship between species diversity and a series of environmental factors.
3. Use machine learning modelling approaches to gain insights into the drivers of species diversity and predict the distribution thereof in mountainous grassland communities
4. Explore the relationship between multiple spectral bands and diversity metrics using remote sensing to monitor species diversity in grassland communities.
5. Predict grass species' Nitrogen concentration in mountainous grasslands using multi-spectral remote sensing models.

Study area

The study was conducted at GGHNP, which is in the North-Eastern parts of the Free State province, South Africa. The park covers 32,758.35 ha and lies between 28°27' S - 28°37' S and 28°33' E - 28°42' E. The GGHNP is a mountainous grassland located at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The following soil types were identified in the park: shallow rocky soils (Glenrosa and Mispah), deep soil along drainage lines (Oakleaf), well-developed sand soils e.g., Hutton and Clovelly, as well as Clayey structured soils including Milkwood and Tambakulu (SANParks, 2020). The park is characterized by summer rainfall, temperate summers, and cold winter. The rainfall season stretches from September to April with a mean annual rainfall ranging from 800 mm to 2,000 mm (Kay, Bredenkamp, and Theron, 1993). The park lies between 1,892 m and 2,829 m above sea level and comprises the following grasslands units: Eastern Free State sandy grasslands (Gm 4), Basotho montane shrubland (Gm 5), Lesotho highveld basalt grassland (Gd 8), and Northern Drakensberg highveld (Gd5) (Mucina and Rutherford, 2006).

Chapter 2

Estimating mountainous plant species richness and diversity for monitoring global change in a protected grassland

This chapter is based on:

Mashiane, K.K. Ramoelo, A. Adelabu, S.A. (2022) ‘Estimating mountainous plant species richness and diversity for monitoring global change in a protected grassland park’, *African Journal of Ecology* (**under review after two revisions).**

Abstract

Assessments of species diversity and richness are essential to understand present ecological and biodiversity conditions for effective conservation management strategies. Biodiversity indicators determine rangeland health and response to grazing, fire regimes, and climate change. This research examined species richness, diversity, and composition in a protected mountainous grassland. Two datasets, both collected from a 30m x 30m plot, with similar species composition and cover were combined. One dataset was collected using a 100-step point survey and the other from a series of sixteen plots. A single-factor analysis of variance was used to assess if the mean species richness and diversity of the sites differed across the study area. Species accumulation curves were used to determine the relationship between species richness and the number of sampling units per site. The results from fitting a species-area equation showed that the estimated maximum species richness was slightly greater than the observed species pool in all sites, meaning that the sampling units were not adequate (albeit by small margins) to capture all vascular plant species in the sites. Diversity metrics could thus be used to monitor species change within grassland communities.

Introduction

Biodiversity at the species level is deteriorating rapidly due to a range of human activities resulting in global environmental changes (GEC) (Hooper *et al.*, 2012; Steffen *et al.*, 2004). The main drivers of GEC – increasing atmospheric CO_2 levels and associated climate change, excess deposition of nitrogen, loss and fragmentation of natural habitats, and biotic invasion – may result in extinction and alteration of species distribution (Hooper *et al.*, 2012; Steffen *et al.*, 2004). Human activities have directly altered nearly 50% of terrestrial ecosystems, which has had detrimental effects on biodiversity, nitrogen cycling, soil structure, soil organisms, and climate (Steffen *et al.*, 2004). Studies show that the impact of species loss is comparable to the impact of GEC (Cardinale *et al.*, 2018; Hooper *et al.*, 2012). Plant biodiversity loss may reduce plant production and alter nutrient cycling (Balvanera *et al.*, 2006; Cardinale *et al.*, 2011). Hooper *et al.* (2012) argues that one of the primary drivers of environmental change in the twenty-first century may be biodiversity loss.

High levels of biodiversity are essential for the long-term resilience of ecosystem functions, processes, and services (Oliver *et al.*, 2015). All species in a community influence ecosystem properties and services, however, this is dependent on their respective functional traits (Jordano *et al.*, 2013). Equally, ecosystem changes affect species' functional traits i.e., morphological, biochemical, physiological, structural, phenological, or behavioural characteristics, thus altering ecosystem functions and resilience (Jordano *et al.*, 2013; Suding *et al.*, 2008). Therefore, biodiversity loss may reduce the efficiency of plant communities in capturing biologically essential resources such as water, nutrients, and sunlight, which negatively influence organic nutrient recycling via decomposition (Cardinale *et al.*, 2012). The number and diversity of species, their evenness, functional traits, and interactions within the community all have an impact on how well an ecosystem functions (van Oijen *et al.*, 2018).

Organisms can have an impact on the physical genesis of environments and element concentrations in biogeochemical cycles via ecological engineering and stoichiometry, respectively (Jones *et al.*, 1994; Glibert, 2016). Changes in an ecosystem are attributed to the physiology, morphology, and behaviour of individual organisms at a species level as well as structure, and composition at a population and community level (Suding *et al.*, 2008). In grassland communities, increased plant species richness increases plant productivity and nitrogen input (Tilman *et al*., 2012). Therefore, assessing the levels of species diversity assists in understanding resource use by and distribution of grazing animals. In addition, understanding the distribution of mountainous species diversity is essential for identifying priority areas for conservation (Barros *et al*., 2015). Research shows that high plant diversity often leads to increased plant productivity, nutrient cycling, and ecosystem stability in grassland communities (Tilman and Downing, 1994; Tilman, 1999a). The distribution of large grazing mammals is attributed to the occurrence of nutritionally enriched vegetation species in Serengeti National Park, Tanzania (McNaughton, 1988). Good baseline data that assess biodiversity are essential for monitoring and predicting future biodiversity patterns, which will immensely improve conservation to abate further loss (Cardinale *et al.*, 2018a).

Most losses of biodiversity are seen at a global scale, where both increases and decreases in community diversity are recorded at regional and local scales (Reitalu *et al*., 2012). Global environmental change affects species tolerances and habitats. Therefore, a species with restricted ecological niches is more susceptible to extinction than those with broader ecological tolerance occupying more habitats (Reitalu *et al.*, 2012). Measures of biodiversity parameters are important indicators of ecosystem health in rangelands. As such, Improved datasets explaining patterns of biodiversity in an ecosystem will assist in devising effective conservation strategies (Wardle *et al.*, 2011).Furthermore, quantitative analysis of biodiversity datasets that could explain plant diversity distribution in response to GEC (Wardle, 2016). Measures of biodiversity parameters are important indicators of ecosystem health in rangelands. Species richness and diversity are two very important and suitable measures since they are simple to measure and interpret (Symstad & Jonas, 2011). In addition, useful information about metrics of diversity is derived from species-area equations (Malanson *et al.*, 2020a). Assessment of plant diversity data, especially about species occurrence and abundance, reveals the concentration of dominant species thus enhancing efforts to save biodiversity worldwide (Gaury and Devi, 2017a). Hence, within protected areas, data on species richness and diversity is important for ecological explanations, which enable designs of scientifically defendable conservation strategies (Brown *et al.*, 2013a).

Mountain ecosystems in Southern Africa have high levels of biodiversity and species richness with complex ecosystems (Brown and du Preez, 2020). Golden Gate Highlands National Park (GGHNP) forms part of the Maluti-Drakensberg Mountain range and is a grassland biome that has a diverse range of aesthetically attractive landscapes with high biodiversity value (Kay, Bredenkamp and Theron, 1993) . The park is extremely variable in topography and climate, which results in a complex mosaic of plant communities (Kay *et al*., 1993). Despite being home to diverse species and complex ecosystems, these grasslands are prone to human-induced pressure such as climate change and biological invasions. The lack of data and protection of these biomes worsens the vulnerability to and risk of extinction. Due to the accessibility of GGHNP to the public and local communities, the park's resources are exploited. Overgrazing, thatch harvesting, overstocking, and unplanned fires negatively affect the biodiversity of the park. As part of developing species inventories for long-term monitoring, the South African National Parks collects vegetation data in GGHNP. The data are used for vegetation classification and taxonomic species composition (Kay, Bredenkamp, and Theron, 1993), and land degradation indicators. Unfortunately, data on species diversity and richness for many parts of the grassland biome in Southern Africa is not always available and seldom analysed to determine species diversity and richness, especially in sub-alpine grassland communities of the Drakensberg (Brown and du Preez, 2020). Available information indicates that the grassland biome of Southern Africa contains a mean plant richness of 82 species per 1000m^2(Brown & Bezuidenhout, 2020; Eckhardt *et al*., 1996), and between 9 and 29 species have been recorded in 100m^2 vegetation sampling plots within high altitude grassland of the Eastern Cape and KwaZulu Natal provinces in South Africa (Eckhardt et al., 1996). Thus, the grassland biome is considered the most diverse with alpha diversity in South Africa, second to the fynbos (Brown and Bezuidenhout, 2020). Quantitative estimates of the levels of biodiversity are important for improving our understanding of subsequent changes in ecosystem functioning (Cardinale *et al.*, 2012). This research assessed and explored baseline datasets to test hypotheses about changes in species richness and diversity in grass communities: 1) which sites are habitats to a high number of plant species (herbecious diversity) in GGHNP? 2) Are there any significant variations of species richness and diversity across sites, fire severity and grazing pressures in the park? 3) What is the estimated number of species within each site?

Methodology

Study area

The study was conducted at GGHNP, which is in the North-Eastern parts of the Free State province, South Africa (Figure 1). The park covers 32,758.35 ha and lies between 28°27’ S - 28°37’ S and 28°33’ E - 28°42’ E. The GGHNP is a mountainous grassland located at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The following soil types were identified in the park: shallow rocky soils (Glenrosa and Mispah), deep soil along drainage lines (Oakleaf), well-developed sand soils e.g., Hutton and Clovelly, as well as Clayey structured soils including Milkwood and Tambakulu (SANParks, 2020). The park is characterized by summer rainfall, temperate summers, and cold winter. The rainfall season stretches from September to April with a mean annual rainfall ranging from 800 mm to 2,000 mm (Kay, Bredenkamp, and Theron, 1993). The park lies between 1,892 m and 2,829 m above sea level and comprises the following grasslands units: Eastern Free State sandy grasslands (Gm 4), Basotho montane shrubland (Gm 5), Lesotho highveld basalt grassland (Gd 8), and Northern Drakensberg highveld (Gd5) (Mucina and Rutherford, 2006).

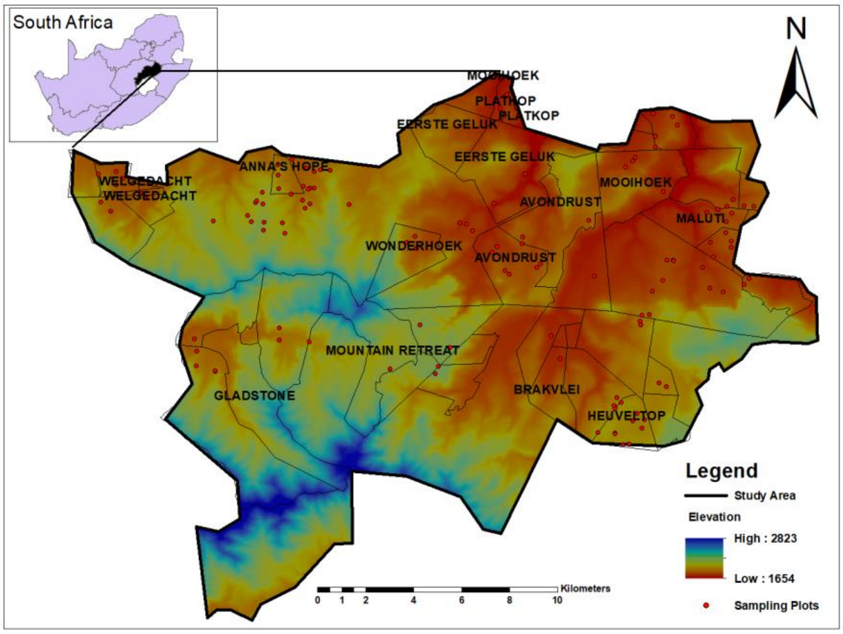

Figure 1: A map of the study area

Data collection

The land type map of GGHNP was used as the first stratification. Sampling sites of homogenous herbaceous units were then located in a randomly stratified manner to acquire baseline species data for rangeland forage species composition. A total of 36 vegetation sampling plots (30m x 30m) ranging between 3 and 5 per site (6 sites with 5 plots and 2 sites with 3 plots) (Figure 2A) were placed randomly within the homogenous grass units. Within each sampling plot, a total of 16 (1 x 1 m) quadrats were placed systematically at every 10 m along four parallel rows (Figure 2). The above-mentioned dataset was merged with another dataset from the same research comprising 12 sites with 106 plots, collected using 100 step points from 4 transects located within a 30m x 30m plot where all species were recorded at every step point (Figure 2B). The joint dataset comprised 142 plots with a total of 13 sites, 7 of which were the same and 6 different (Table 1). The distribution of sampling plots across the land types in GGHNP is given in Figure 3.

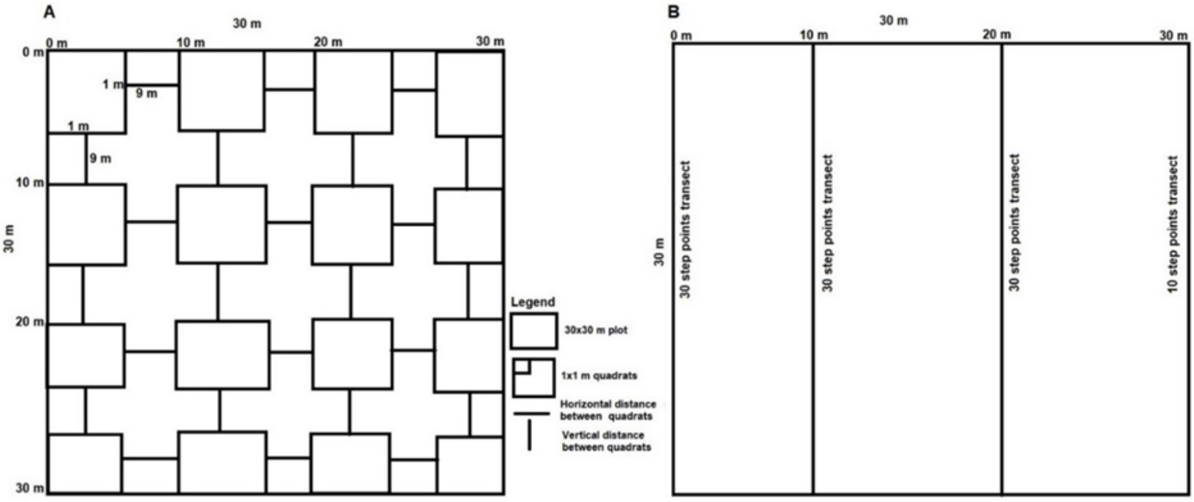

Figure 1: A) Quadrat and B) step-point-based vegetation plot design

In each quadrat, all vascular plant species of the standing vegetation were identified to species level where possible and the visually estimated aerial cover of each species was estimated to the nearest 5% with the quadrat. In each sampling plot where step-point transects were done, Braun-Blanquet cover-abundance values were used as a proxy for species cover.

Table 1: The number of plots at each site from quadrat and step-point vegetation plot surveys

Site name	Quadrat-based plots	Step-point survey plots	Merged dataset
Annas' hope	5	23	28
Avondrust	-	9	9
Basotho	3	18	21
Brakvlei	3	2	5
Eerste Geluk	-	5	5
Gladstone	5	3	8
Heuvel Top	5	16	21
Maluti	5	6	11
Mooihoek	5	5	10
Mountain Retreat	5	-	5
Platkop	-	5	5
Welgedacht	-	8	8
Wonderhoek	-	6	6
Total	36	106	142

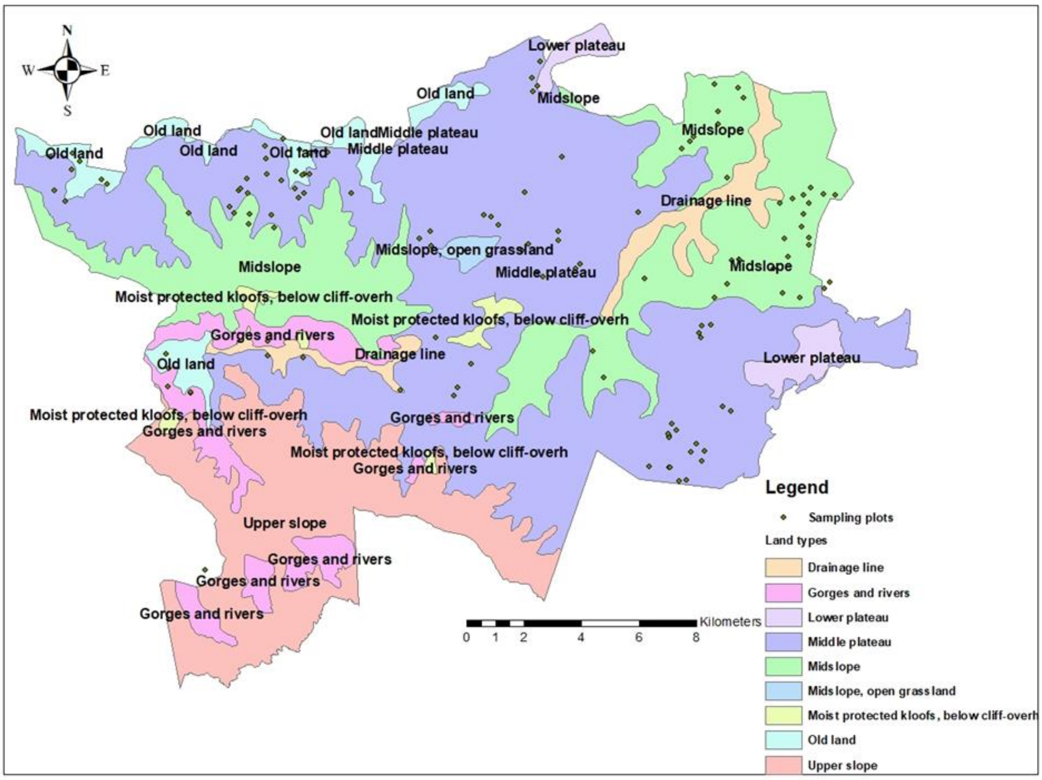

Figure 2: Distribution of sampling plots across the land types in Golden Gate Highlands National Park

The taxonomic composition and percentage cover of both visually estimated quadratic-based and Braun-Blanquet abundance values of each vascular plant species were used to derive species richness and then diversity per plot (Tilman et al., 2006).The values from each plot were averaged to attain mean species richness per site. The variables species richness and diversity were computed using “vegan” and “plyr” R statistical packages (Liang & Seyfried, 2001; Oksanen, 2017), which employed the diversity and apply function for species diversity and richness, respectively.

Total species diversity was calculated using the Shannon-Wiener Index (Eq. 1) where pi is the proportion of the species within the sampling units. Species richness was determined by adding all species from each quadrat and averaging by the number of quadrats in each plot, to obtain the average plot value. Species richness and diversity across selected sites is given in Table 2.

$$H'=-\sum pi*lnpi \qquad (1)$$

Table 2: Data summary of species richness and diversity (mean±standard deviation and coefficient of variability in percentages (CV) and plant species richness across selected sites (n = the number of plots per site)

Site name (n)	Grasses’ richness	Forbs’ richness	Shrubs’ richness	Total Number of species	Total species richness (mean±SD)	CV	Total species diversity (mean±SD)	CV
Welgedacht (8)	17	10	0	27	11,63±2,26	19	1,90±0,31	16
Annas’ hope (28)	41	13	4	58	10,92±3,39	31	1,83±0,35	21
Wonderhoek (10)	19	7	0	26	10,67±1,75	3	1,78±0,36	20
Avondrust (6)	19	9	1	29	9,78±4,12	42	1,76±0,28	31
Mooihoek (10)	17	9	3	29	9,52±3,28	40	1,74±0,38	20
Maluti (11)	23	7	0	39	9,26±3,74	40	1,71±0,24	19
Eerste Geluk (5)	7	10	3	29	8,84±2,64	30	1,67±0,28	14
Brakvlei (5)	9	16	1	26	8,25±3,97	48	1,67±0,50	21
Basotho (21)	23	7	1	31	8,21±3,62	44	1,55±0,49	32
Gladstone (8)	26	23	3	52	7,80±1,73	22	1,48±0,34	15
Heuvel Top (21)	23	7	2	30	7,37±2,95	50	1,45±0,31	23
Platkop (5)	17	8	2	27	5,89±1,43	15	1,16±0,18	36
Mountain Retreat (5)	19	9	0	28	4,28±0,57	13	1,15±0,41	16

Fire data collection

A burn severity map was created for areas affected by fires. The normalized burn ratio (NBR), which utilizes near-infrared (NIR) and shortwave-infrared (SWIR) wavelengths, was used because it highlights burned areas and estimates burn severity (Hawbaker *et al.*, 2020). The NIR and SWIR bands were obtained from Landsat 8 satellite imagery available on the Google Earth Engine. These bands were used to calculate the NBR for pre-and post-fire scenarios. The bands were selected from images of the year before the vegetation survey (2018), before the start of the fire season (January – April), and post-fire season (June – September). Subsequently, an image of delta NBR was determined through the differences between pre-and post-fire NBR (eqs. 2 and 3). Finally, the ΔNBR was classified according to the United States Geological Survey (n.d.) (Table 3).

$$NBR = NIR - SWIR/NIR + SWIR \qquad (2)$$

$$\Delta NBR = NBRprefire - NBRpostfire \qquad (3)$$

Table 3: Burn severity levels (U.S. Geological Survey, n.d.)

Burned severity class	ΔNBR
Unburned	-0.10 to 0.10
Low severity	0.10 to 0.27
Moderate-low severity	0.27 to 0.44
Moderate-high severity	0.44 to 0.66
High severity	>0.66

Grazing data collection

Animal census data of grazers, collected in the park during 2020 were used for the determination of grazing pressure in the park. The data are collected biennially by surveying the area from a helicopter and counting the number of animals including herd sizes found in the park. The hotspot analysis tool Getis-Ord Gi statistics, which identifies statistically significant hotspots and cold spots, was used to find where animals cluster spatially within the park (Getis and Ord, 1992). Subsequently, inverse distance interpolation was used to estimate cell values with the highest and lowest animal concentrations. Estimated values ranging from -1 to 1 were regarded as cold spots and therefore classified as low grazing intensity while 1 to 3 values were regarded as hotspots and classified as high grazing intensity.

Data analysis

The merged dataset was used to compute a single-factor analysis of variance (one-way ANOVA) which was used to test if the mean species richness and diversity of sites, fire severity, and grazing intensity differed across the study area. Subsequently, Turkey Honest Significant Difference was used to test the significance of inter-site variations in richness and diversity. Statistical analyses were performed using R-software and a significant difference was considered at a 95% confidence level ($p<0.05$).

The dataset from the 36 plots comprising 8 sites was used to compute species accumulation curves (SACs) to determine the adequacy of the quadrat-based vegetation sampling plots in capturing local species richness and diversity of sites. A non-linear regression (nls) model for species-area was fitted to the SACs for attaining diversity parameters where the asymptotic maximum number of species was considered the optimal sampling plot for determining species richness. The Lomolino model was the nls: $Asym/(1+slope^{\log(xmid/area)})$ where parameter Asym is the asymptotic maximum number of species, the slope is the maximum slope of increase of richness, and xmid is the area where half of the maximum richness is achieved (Lomolino, 2000). Gamma diversity, which is the highest number of species in each site, was derived from the "saccum" function found in vegan package of R statistical software; (eq. 4) (Oksanen, 2017).

$$S=S_{max}/(1+slope^{\log(xmid/area)}) \tag{4}$$

Results

There is a significant difference in species richness between sites in the park (F=3.55, p<0.05). Similarly, species diversity is also significantly variable across the park (F=2.79, p<0.05) (Table 4). However, species richness and diversity did not differ significantly between burn severity classes ($F_{Richness}$=1.81, $p_{Richness}$=0.13, $D_{iversity}$= 0.45, $p_{Diversity}$=0.77) and grazing pressures ($F_{Diversity}$=0.67, $p_{Diversity=}$0.42, $F_{Richness}$=3.42, $p_{Richness}$=0.06) (Table 4). The Turkey-HSD showed that 61% of sites showed no significant variation, while only 39% of the sites exhibited significant variation in species richness and diversity.

The plots at Annas' hope and Welgedacht showed significantly (p<0.05) higher species richness compared to Mountain Retreat, and Heuvel Top. Sites such as Avondrust, Basotho, Brakvlei, Eerste Geluk, Gladstone, and Platkop showed similar richness values and were not significantly different (p>0.05). Plots at Welgedacht showed a significantly higher species diversity than the rest of the sites (p<0.05). Platkop and Mountain Retreat were significantly lower than the rest of the sites (p<0.05). In general, the remaining sites did not differ in species diversity (p>0.05).

Species diversity was similar between fire severity classes. Species richness did not differ between fire severity classes (p>0.05). However, unburned sites had the highest species richness while sites with moderate to high severity had the lowest species richness. Similarly, both species diversity and richness did not differ between grazing pressures Table 4).

Table 4: Descriptive statistics of species richness and diversity (H') and ANOVA at park-level

Variables of interest	Min	Max	Median	Mean	STDEV	CV	F-VALUE	P-VALUE
Species richness (sites)	2.5	17	9	8.97	3.49	39.9	3.55	0.01
Species diversity (sites)	2.5	17	9	8.97	3.49	24.63	2.79	0.01
Species richness (fire severity)	2.5	17	9	8.97	3.49	39.9	1.81	0.13
Species diversity (fire severity)	2.5	17	9	8.97	3.49	24.63	0.45	0.77
Species richness (grazing pressure)	2.5	17	9	8.97	3.49	39.9	3.42	0.05
Species diversity (grazing pressure)	2.5	17	9	8.97	3.49	24.63	0.67	0.42

The fitted Lomolino model estimated that the asymptotic number of species in a site was less than the gamma diversity (species pool was the total number of species in each site). The slope of the species accumulation curve ranged between 2.07 and 2.35 (Table 5). In addition, the species accumulation curves showed that the sites were approaching the asymptote of the maximum number of species, albeit not reached.

Table 5: The Lomolino model parameters for the fitted species accumulation curves

Site name (n)	Asymptote	Xmid	Slope	Gamma
Annas' hope (5)	56.96	10.49	2.12	49
Basotho (3)	39.94	8.52	2.12	32
Brakvlei (3)	50.1	8.73	2.6	42
Gladstone (5)	54.44	7.89	2.15	47
Heuvel Top (5)	35.72	10.95	2.35	30
Maluti (5)	58.12	10	2.5	51
Mooihoek (5)	42.27	9.71	2.28	36
Mountain Retreat (5)	42.5	15.87	2.07	33

Discussion

The sampled sites in GGHNP are home to many plant species, indicating a high gamma diversity in the grasslands, and are thus an important component of the biome. This is consistent with the high-altitude grasslands of KwaZulu Natal and Eastern Cape provinces where the species ranged between 9 and 29 from 100 m^2 vegetation sampling plots (Eckhardt et al., 1996). Patterns of species richness among plant communities vary through space and time in response to endogenous and exogenous factors (Bagchi et al., 2017). Furthermore, fire and grazing may interactively or non-interactively promote heterogeneity in the composition and diversity of grasslands (Symstad and Jonas, 2011). The results indicate a significant variation of species richness and diversity across the park, even though most sites were similar in this regard. The attributes of the number of species at each site, variation or lack of species richness across sites might be an indication of environmental heterogeneity at different spatial scales (Filibeck et al., 2019) which is pivotal for promoting or maintaining plant richness. Moreover, the results emphasize the importance of spatially explicit data in assessing species richness and diversity within grassland terrestrial landscapes.

Grassland plant communities are highly dynamic due to both large- and small-scale disturbances (Collins, Glenn, and Briggs, 2002). Hence, disturbances such as grazing and fire are pivotal for maintaining species richness in stable grassland (Poschlod *et al.*, 1998). In this study, species richness and diversity did not differ amongst fire severity classes and grazing pressures. Sometimes grassland ecosystems show less response to grazing (Symstad and Jonas, 2011,) and the effect of fire is highly variable but generally neutral to negative (Poschlod *et al.*, 1998). Although species richness and diversity were significantly variable in the GGHNP, most inter-site variations were not significant. Such variations or lack thereof, are mostly attributed to the Intermediate Disturbance Hypothesis, which predicts that diversity will be maximized at intermediate levels of disturbance and lowered by extreme and scant disturbances (Brown et al., 2001). Therefore, species richness is a function of biotic and abiotic factors due to a specific set of species being attributed to a certain condition. Tolerance to environmental changes resulting from disturbances is also pivotal for regulating species richness in ecosystems (Brown et al., 2001). However, in some cases, species richness response to disturbances may be different from that of species diversity where the high number of species in a community does not translate to high species diversity.

The species accumulation curves in this study appear to be in the asymptote and the fitted Lomolino model estimated a slightly higher number of species richness in each site than the actual number recorded. For example, Annas' hope yielded a prediction of fifty-seven species, while 49 species were recorded for the site. Similarly, Miller and Wiegert, (1989) found that datasets of vascular plants have approached an asymptote albeit not necessarily reached in the biogeographic mountainous of Southern Appalachians. In ecosystems that are very difficult to sample, it is important to estimate the total number of species and the sampling effort needed to obtain these estimates (Ugland, Gray, and Ellingsen, 2016). Lomolino, (2000), states that the relationship between species and sampling effort asymptotically approaches the maximum value of the species pool. Williamson et al., (2001) argue that this is only the case for a finite area with finite species.

The species accumulation curves the quadrat-based dataset were similar both in shape and completeness. SACs take account of species identities and the rate at which new species are encountered (Ugland, Gray, and Ellingsen, 2016). The shape of the SACs reflects the evenness, abundance, and distribution of species (Storch, 2016). In addition, asymptotic species richness estimates are not effective when the dataset includes many rare species. Hence, Soberon & Jorge, (1993), highlighted that near the asymptote rare species are likely to be the ones added. This may be the case in grasslands where grasses are the common and predominant taxa. Analysed separately, both rare and dominant species are likely to yield different estimates of species in GGNHP. Estimating the scale i.e., temporal or spatial required to add a given number of species should make it possible to plan rigorously to cut cost, therefore SACs may be useful in GGHNP for monitoring and assessing species response to GEC (Flather, 1996; Malanson et al., 2020a) .

The fitted Lomolino model indicated that the slope of the accumulation curves ranged from 2.07 to 2.35. The slope of the plots is useful for comparing the degree of species turnover between different ecological sites (Filibeck *et al.*, 2019). The slope indicates the rate of increase in diversity and therefore captures the beta-diversity (i.e., variation of identities of species between plots of the observed scale (Malanson *et al*., 2020b). A higher slope of SACs implies a high spatial heterogeneity of vegetation. Therefore, the slopes of the fitted SACs in GGHNP can be used to determine sites with high species turnover, which are essential for setting conservation targets. Hence, monitoring diversity metrics can aid park conservation managers, especially in better response to biodiversity loss.

According to Symstad & Jonas, (2011), "species accumulation curves are often recommended as an ideal indicator of biodiversity because they minimize complications of comparison across studies with varying sampling units". The key distinction of species accumulation curves is that they quantify the rate at which new species are discovered as sampling effort increases (species turnover). In this study, the number of species inhabiting those sampling units increases; rapidly at first, but then more slowly for the larger sampling units, because the SACs were approaching an asymptote. Therefore, within the sites, species only became identical with a higher sampling size as indicated by the species accumulation curve approaching an asymptote. Species richness can either increase, or decrease and become regulated in space and time (Brown *et al*., 2001). This is mainly because environmental changes favour or disfavour some species, altering turnover in species composition (Brown *et al*., 2001). Indeed, the SACs will have more gradual slopes in heterogeneous areas than in homogenous areas. Knowledge of how spatial heterogeneity affects species pools is necessary for predicting species richness and turnover and therefore research on mechanisms controlling local species richness at different spatial scales is warranted.

Conclusion

Ecological consequences of species loss are well studied. Thus, high levels of species richness and/or diversity are pivotal for maintaining many ecosystem functions and vice versa. Monitoring aspects of biodiversity richness and diversity as well as species composition is critical for managing ecosystem functions and services that benefit humanity and determines game-carrying capacity and stocking rates across landscapes. This study showed that species accumulation curves can be used to monitor diversity across and between plant communities in the Park. For average sites, measures of changes in local diversity (alpha diversity) and species turnover (beta diversity) have proved insufficient. This study provides a tool to monitor species change within grassland plant communities, and the information can then be used to manage and improve stocking rates and fire estimation models in the park.

Chapter 3

Plant species richness in mountainous mesic grassland communities: the role of ecological, soil, and topographical variables

This chapter is based on:
Mashiane, K.K. Ramoelo, A. Adelabu, S.A. 'Plant species richness in mountainous mesic grassland communities: the role of ecological, soil, and topographical variables', *Koedoe* (**under review after one revision**).

Abstract

Southern African mountain ranges are characterized by rich and diverse plant species occurring in different habitats with pronounced endemic species. However, globally, biodiversity at the species level is deteriorating rapidly due to environmental changes that leadto habitat degradation and fragmentation. The mountainous grassland communities are likely to be the most vulnerable to rapid ecosystem changes, and there is a need to understand their key vegetation determinants for effective and efficient management. This study determined the relationship between plant species richness and associated environmental drivers in mountainous mesic grasslands communities of Golden Gate Highlands National Park, South Africa. Topographical variables (slope and elevation) were derived from the 30 m resolution Digital Elevation Model (DEM). Soil variables such as bulk density, silt fragments, pH, coarse fragments, soil organic carbon, sand, and nitrogen were acquired from the International Soil Conference and Information Centre and species richness and diversity were derived from species inventory data compiled using a field survey. Species richness was influenced by species diversity, bulk density, the interaction term of elevation - bulk density. Similarly, species composition was mainly related to edaphic factors and elevation. Species diversity is pivotal for maintaining ecosystem integrity and richness, and topography-soil relationships are essential for influencing species richness in mountainous grassland communities. Therefore, soil conservation in the park should be prioritised and degradation avoided to ensure biodiversity conservation.

Introduction

Processes crucial to the ecosystem's functioning are influenced by plant diversity (Cardinale *et al.*, 2011). However, biodiversity at the species level is deteriorating rapidly due to global environmental change drivers such as climate change, pollution, invasive species, and land-use changes, leading to habitat degradation and fragmentation (Hooper *et al.*, 2012). It has been emphasized that declining species richness and, more crucially, concurrent changes in species composition, may have an impact on the essential services that ecosystems supply (Tilman, Reich, and Isbell, 2012). In addition, research found that decreasing plant diversity in grasslands frequently results in plants using fewer soil resources and producing less biomass above ground (Naeem *et al*., 1994; Tilman *et al*., 1996; Hector *et al.,* 1999).

Long-term management techniques have a significant impact on the plant diversity in grasslands, with grazing being the main factor determining vegetation dynamics, species distribution, and biodiversity at the landscape scale (Oliver *et al.*, 2015).Abiotic changes can also lead to biotic change and vice versa. For example, soil fertility leads to changes in forage quality and quantity. In addition, several interrelated environmental factors, including soil properties, moisture regime, and disturbances, affect the species composition of grasslands. Despite interventions such as policy reform and increasing biodiversity coverage by protected areas, most pressures on biodiversity are intensifying (Chanson *et al.*, 2010; Hooper *et al.*, 2012). As a result, the rate of biodiversity loss is accelerating, this could be detrimental to the ecosystem functioning and services at large (Chanson *et al.*, 2010; Oliver *et al.*, 2015; Meyer *et al.*, 2018).

Modelling the distribution of mountainous species diversity is essential for identifying priority areas for conservation (Barros *et al.*, 2015), especially because of the imminent risk of species extinction resulting from pollution, climate change, and habitat loss caused by land conversion (Picó and Van Groenendael, 2007; White *et al.*, 2012). For example, subtropical highland grasslands in Brazil are home to several groups of plants (Iganci *et al.*, 2011), and some taxa are represented by an island-like distribution (Lorenz-Lemke *et al.*, 2010). Furthermore, in Europe, plant diversity is maintained by disturbance regimes such as grazing and soil erosion, which are confined to subalpine environments. This enhances specific-species composition ideal for grazing while restricting the expansion of unwanted species i.e., shrub encroachment (Armas-Herrera *et al.*, 2020a). A better understanding of biodiversity indicators and the factors driving their distribution is crucial for not only conservation interventions but also for modelling future distributions in response to global change (Barros *et al.*, 2015; Kuzemko *et al.*, 2016).

The interaction between biotic and abiotic processes may have important implications for the dynamics of the grazing ecosystem. For example, plant communities are affected compositionally and structurally by a range of biotic and abiotic factors like grazing, fire, and rainfall patterns (McNaughton, 1983). However, climate change, invasive species, and land-use changes will inevitably affect these ecological processes, thus influencing plant communities and diversity (O'Connor, 2008; Liu *et al.*, 2019). Plants within mesic biomes are most susceptible to the pressure of global change because they are extremely stressed and intolerant, especially during drought-related disturbances (Tardella *et al.*, 2016).

Besides evidence supporting a positive relationship between environmental heterogeneity and species richness, conflicting results report negative and monotonous relationships (Yang *et al.*, 2015; Lundholm, 2009). Establishing heterogeneity-diversity relationships (HDRs) is fundamental for developing management strategies and abating biodiversity loss (Filibeck *et al.*, 2019). For example, forage nutrition and spatiotemporal vegetation variations due to geographic rainfall gradients influence the occurrence of large mammals in Serengeti National Park, Tanzania (McNaughton, 1985a, 1988). Soil catena characteristics control species composition in sub-alpine grassland communities of Monte Perdido National Park in Spain (Armas-Herrera *et al.*, 2020b). Although potentially susceptible to global environmental change, the vegetation of temperate mountains in Southern Africa is influenced by ecological, physical, and anthropogenic processes (Brown and du Preez, 2020). However, information and studies on the mountains in Southern Africa are lacking. Therefore, Information on the vegetation and resilience of this ecosystem would be made available by studying and monitoring this area (Brown and du Preez, 2020). To learn more about what affects spatial patterns of biodiversity at certain scales, studies that combine multiple datasets based on ecological features are encouraged (Barros *et al.*, 2015).

Grassland ecosystems rank amongst the most threatened ecosystems in South Africa. Thus, vegetation surveys are imperative to abate the loss of rare and endangered species, through sound conservation plans. In mesic Afromontane grasslands of Southern Africa, vegetation is subjected to intense land-use activities such as grazing by livestock and wildlife (O'Connor, 2008). The Golden Gate Highlands National Park (GGHNP) is a protected conservation area experiencing intense grazing by both livestock and wild animals, which drastically affects grass species composition. The park is extremely variable in topography and climate, which results in a complex mosaic of plant communities (Kay *et al.*, 1993). Most vegetation surveys about environmental drivers were conducted more than two decades ago and did not consider the role of environmental variables in maintaining and driving plant communities. Therefore, this study investigated grass species richness in response to environmental variables. The objectives were to determine: the relationship between species richness and a series of environmental drivers; an optimal model from a set of environmental variables that could explain the variability of the species richness; and the influence of environmental variables on species composition in GGHNP.

Methodology

Study area

The study was carried out at GGHNP, which is in the northeastern region of the South African province: Free State. (Figure 1). The Park is located between 28°27' S and 28°37' S and 28°33' E and 28°42' E, with a total area of 32,758.35 ha. The GGHNP is a mountainous grassland located at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The following soil types were identified in the park: shallow rocky soils (Glenrosa and Mispah), deep soil along drainage lines (Oakleaf), well-developed sand soils (Hutton and Clovelly), as well as Clayey structured soils (Milkwood and Tambakulu) (SANParks, 2020). Rainfall in the summer, mild summers, and cold winters are characteristics of the park. With a typical annual rainfall of 1,800 to 2,000 mm, the rainy season lasts from September to April (Kay, Bredenkamp, and Theron, 1993). The park lies between 1,892 m and 2,829 m above sea level and experiences sporadic thunderstorms; on the other hand, temperatures are expected to exceed 30 degrees Celsius in the summer days.

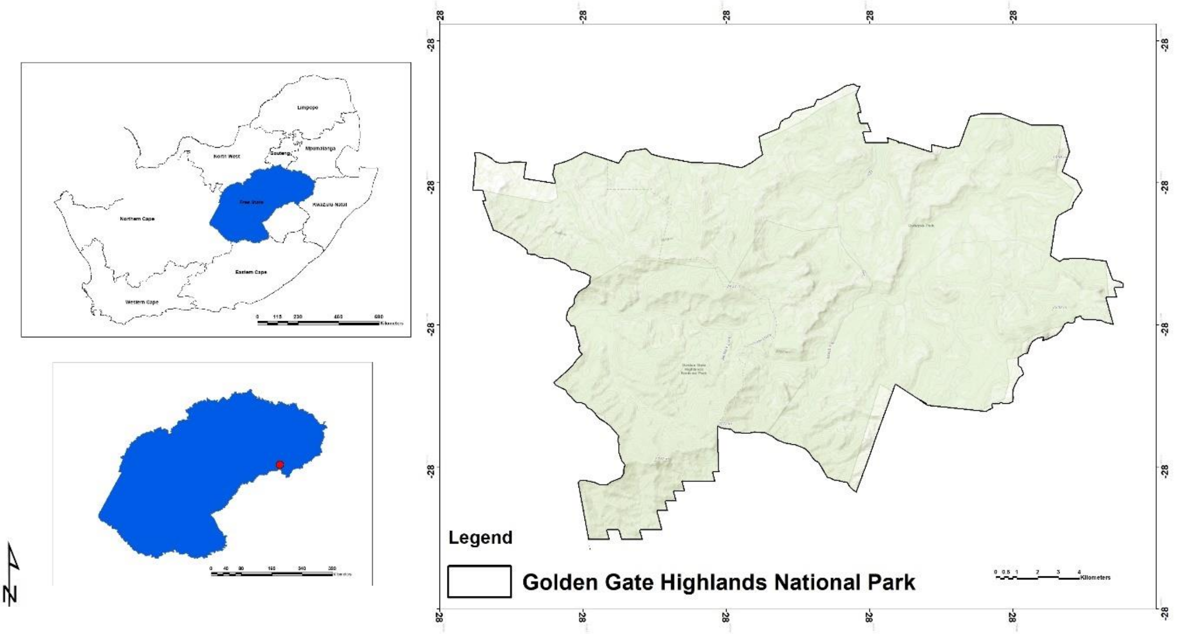

Figure 3: Map of the study area with sampling sites and plots

Data collection

The park was stratified into relatively homogenous physiographic-physiognomic units (Kay *et al.*, 1993). Vegetation sampling was undertaken in a 30m x 30m plot size randomly placed within homogeneous patches of grass communities. Each sampling plot was divided into four transects of 0 – 30 m. Within each 30 m transects, a nested 1m x 1m quadrat was sampled at 10 m intervals. In total 142 plots with 16 quadrats were sampled in the study area. Data were collected in March, April, and May 2019 during the peak growing months for South African montane grasslands. At each quadrat, when possible, every vascular plant species in the standing vegetation was recognized down to the species level, and the 1 m x 1 m grid was divided into 100 cells with a 1 cm^2 square, representing 1% cover to record an aerial cover of each species.

Ancillary data acquisition and preparation

Soil chemical and physical properties such as bulk density (BD), silt fragments (SF), pH, coarse fragments (CF), soil organic carbon (SOC), sand (SD), and nitrogen of topsoil (15 cm) were downloaded from the International Soil Conference and Information Centre (https://soilgrids.org). The elevation data, SRTM DEM at 30 m resolution, were obtained from US Geological Survey's EROS data center (https://earthexplorer.usgs.gov). The slope was derived from DEM using the Spatial Analyst Tool in ArcGIS 10. 4. Fire severity data were acquired from a study conducted in the study area (Adagbasa *et al.*, 2018). By analysing pre- and post-fire season remote sensing photos with burned and unburnt pixels, the study calculated the severity of the fires using the Normalized Burn Ratio (NBR) index (Adagbasa, Adelabu, and Okello, 2018).

To prepare the data for analysis, R Studio statistical software was used to import the raster images after they had been trimmed to the size of the study region, to extract pixel values representing environmental variables in each plot using the "extract" function. The taxonomic composition of the vegetation and its cover were used to derive species richness and diversity for each plot. Total species diversity was calculated using the Shannon-Wiener Index ($H' = -\sum pi*\ln pi$) where pi is the proportion of the species within the sampling unit. The species richness and diversity were computed using "vegan" and "plyr" found R statistical software packages (Oksanen, 2017; Oksanen *et al.*, 2018) by calling the diversity and applying function for species diversity and richness, respectively.

Data analysis

The function "stat. desc" (in pastecs package) in R studio was used to determine descriptive statistics of all the variables in this study. Exploratory analysis i.e., a correlogram was used to test for intercorrelations and the relationships between biodiversity (species richness, and diversity) and environmental variables (soil variables, topographic data, fire severity, and species diversity). Simple multiple linear regression was used to determine the influence of each environmental variable on species richness. To ensure that multicollinearity was low, I calculated variation inflation factors (VIF) and a variable with the highest VIF was removed. This was repeated until the VIF of all variables in the final model was at acceptable levels (VIF<5). Given that the objective of the multivariate GLMs was to identify variables with high direct influence, hierarchical partitioning was employed to evaluate the contribution of each explanatory variable. In order to rank variables according to their independent effects, hierarchical partitioning computes all potential regression subsets and averages model improvements over all hierarchies. This analysis was calculated using the R package ‘hier.part’ R.

Permutational multivariate analysis of variance (PERMANOVA) was used to determine the influence of environmental predictor variables on species composition amongst different sites in the park. Subsequently, fitted environmental variables onto an ordination of species composition to determine variables that significantly affected species composition. Only nine sites were included in this analysis because of ample species composition data of > 5 species per sampling unit. The sites included Annashope, Brakvlei, Mooihoek, Avondrust, Heuweltop, Wilgedactch, Wonderhoek, Maluti, and Basotho. These relationships were displayed using non-metric multidimensional scaling (NMDS).

Results

The maximum number of species recorded in a single plot was 17 while the lowest was 3 and the highest record of species diversity was 2.45 while the lowest was 0.64 in GGHNP (Table1). Of the ecological variables, species richness had the highest coefficient of variation (38.9%) compared to fire severity (25.9%) and species diversity (24.65%) (Table 1). Amongst all edaphic variables, soil organic carbon showed the highest coefficient of variation (31.55%). Slope (80.4%) had the highest coefficient of variation out of the topographic variables when compared to elevation (5.74%) (Table 1).

Table1: Descriptive statistics of species richness, diversity, bulk density, silt fragments, sand, pH, soil organic carbon, coarse fragments, elevation, slope, fire severity and nitrogen for 142 samples

Variables	**Units**	**Minimum**	**Maximum**	**Median**	**Mean**	**Standard deviation**	**Coefficient of variation**	**Standard error**
Species richness	*S*	3.00	17.00	9.00	8.97	3.49	38.91	0.025
Species diversity	H'	0.64	2.45	1.63	1.63	0.40	24.65	0.003
Bulk density	cg/cm^3	1199.00	1454.00	1388.00	1389.57	36.08	2.60	0.254
Silt fragments	g/kg	16.00	26.00	20.00	20.37	1.69	8.30	0.012
Sand	g/kg	44.00	59.00	54.00	54.08	2.56	4.74	0.018
Soil pH	pH$^{\wedge 10}$	56.00	62.00	59.00	59.29	1.47	2.47	0.010
Soil organic carbon	dg/kg	8.00	41.00	13.00	13.09	4.13	31.55	0.029
Coarse fragments	cm^3/dm^3	2.00	19.00	6.50	7.12	2.72	38.16	0.019
Elevation	mabsl	1686.00	2399.00	1842.50	1856.61	106.22	5.72	0.748
Slope	angle$^{\circ}$	0.00	26.73	5.10	6.45	5.21	80.84	0.037
Fire severity	-	2.00	6.00	4.00	3.72	0.96	25.90	0.007
Nitrogen	cg/kg	196.00	334.00	231.00	234.24	19.95	8.52	0.141

Relationship between species richness and environmental variables

There was a strong positive relationship ($r=0.82$, $p < 0.05$) between species richness and diversity (Figure 2). There was a significant negative relationship between species richness, silt fragments, sand, soil organic content, and coarse fragments. Although significant, the other variables had a weak relationship with species richness. Species diversity was not significantly related to any of the topographical and soil variables (see Figure 2).

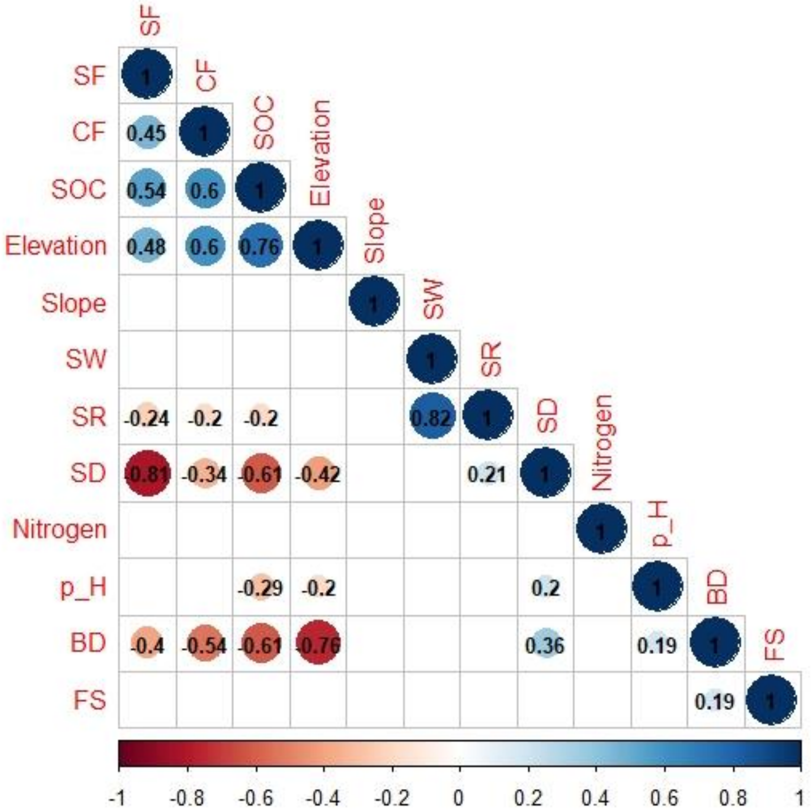

Figure 4: Correlogram of species richness (SR), diversity (SW), bulk density (BD), silt fragments (SF), sand (SD), pH (p_H), soil organic carbon (SOC), coarse fragments (CF), elevation, slope, fire severity (FS) and nitrogen

The multiple linear regression showed that species richness was positively and significantly associated with species diversity within the biotic-based model (Table 2). However, after the removal of species diversity, the model with abiotic variables showed bulk density as the only significant variable associated with species richness albeit negatively. With the topo-edaphic-based model, interaction variables of elevation with bulk density, coarse fragments, and sand fragments contributed significantly to the explanation of variation. Only bulk density was positively associated with species richness.

Table 2: Multiple linear regressions of the species richness measure against biotic, abiotic and topo-edaphic variables (*b* = sign of the regression coefficient, VIF=variable inflation factor)

	b	**t-value**	**VIF**	***b***	**t-value**	**VIF**
	Biotic			Abiotic		
Species diversity	+	16.640*	1.08	n/a	n/a	
Bulk density	-	-0.649	2.78	-	-2.245*	2.68
Silt fragments	-	-1.104	3.64	-	-1.084	3.63
Sand fragments	+	0.005	3.94	-	-0.017	3.95
pH	-	-1.250	1.27	+	0.199	1.26
Soil organic content	-	-0.972	3.75	-	-0.945	3.74
Coarse fragments	-	-0.442	2.02	-	-1.251	2.01
Elevation	+	0.379	3.82	-	-0.455	3.81
Slope	+	0.023	1.04	+	0.589	1.04
Fire severity	-	-2.013	1.17	-	-0.522	1.16
Nitrogen	-	0.348	1.04	-	-0.486	1.04
	Topo-edaphic					
Elevation_Bulk density	+	177.16*	1.319836			
Elevation Silt fragments	-	-13.04	1.870543			
Elevation Coarse fragments	-	-0.61	2.075700			
Elevation Sand fragments	-	-13.63	1.205708			
Slope Bulk density	-	-0.61	1.314365			

Hierarchical partitioning showed species diversity as the variable having the highest independent effects followed by edaphic variables (Figure 3). The abiotic-based model exhibited silt fragments as the variable with the highest explanation of variation in species richness. The interaction term of elevation and bulk density was more pronounced as the variable with the highest direct influence on species richness.

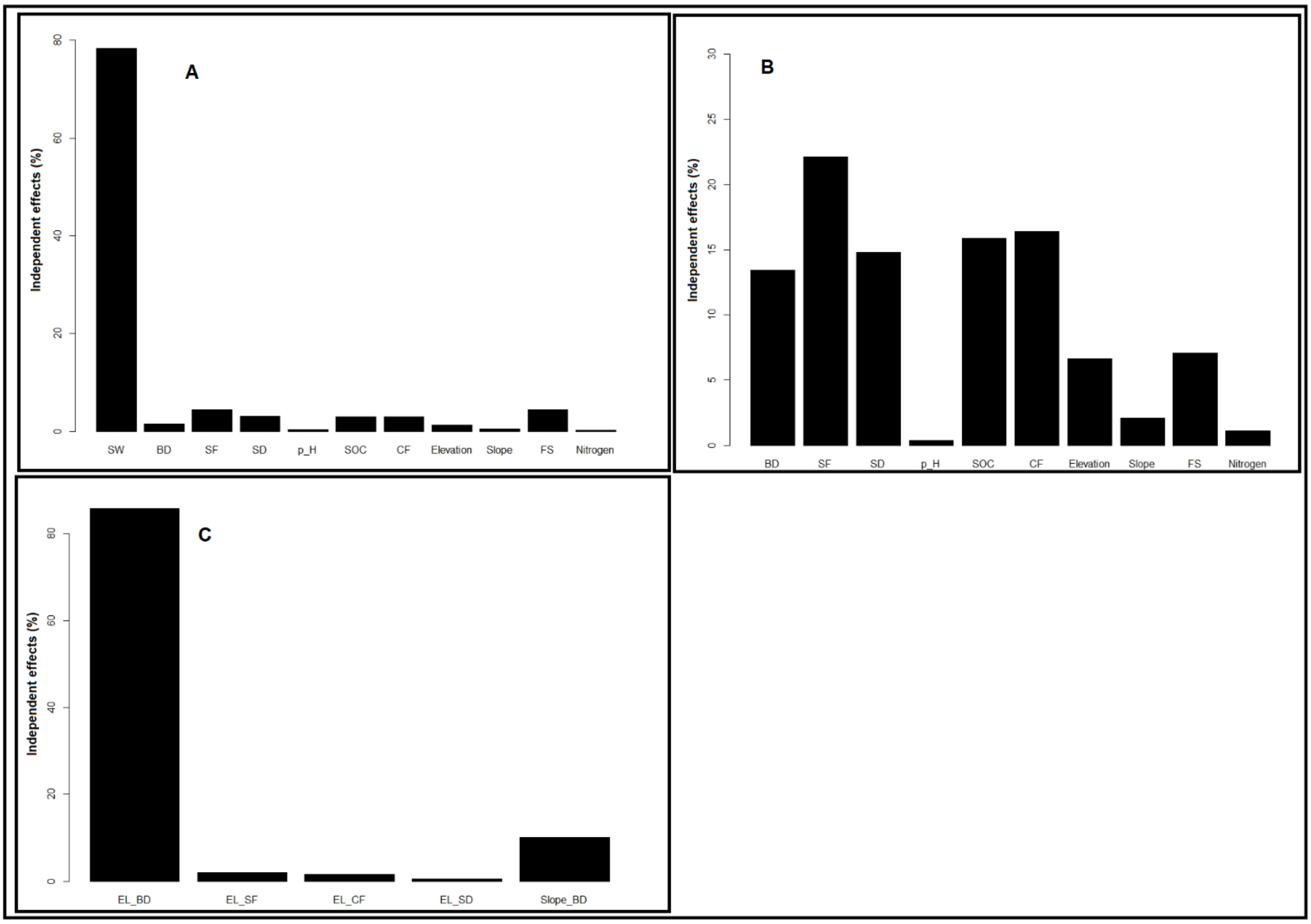

Figure 5: Hierarchical partitioning regression analysis for species richness of (A) biotic-based (B) abiotic-based and (C) topo-edaphic based variables. species diversity (SW) with soil (BD: bulk density; SF: silt fragments; SD: sand fragments; SOC: soil organic content; CF: coarse fragments), pH (p_H), fire severity (FS) and topography (EL: elevation; and Slope)

Species composition was significantly related to the environmental predictors bulk density, silt fragments, sand fragments, pH, and elevation. These variables explained <1% of the variation in species composition. Both sand and pH were identified as the two variables driving species composition at Annashope, however, bulk density was identified as the main driver of species composition at Maluti and Heuwelkop (Figure 4).

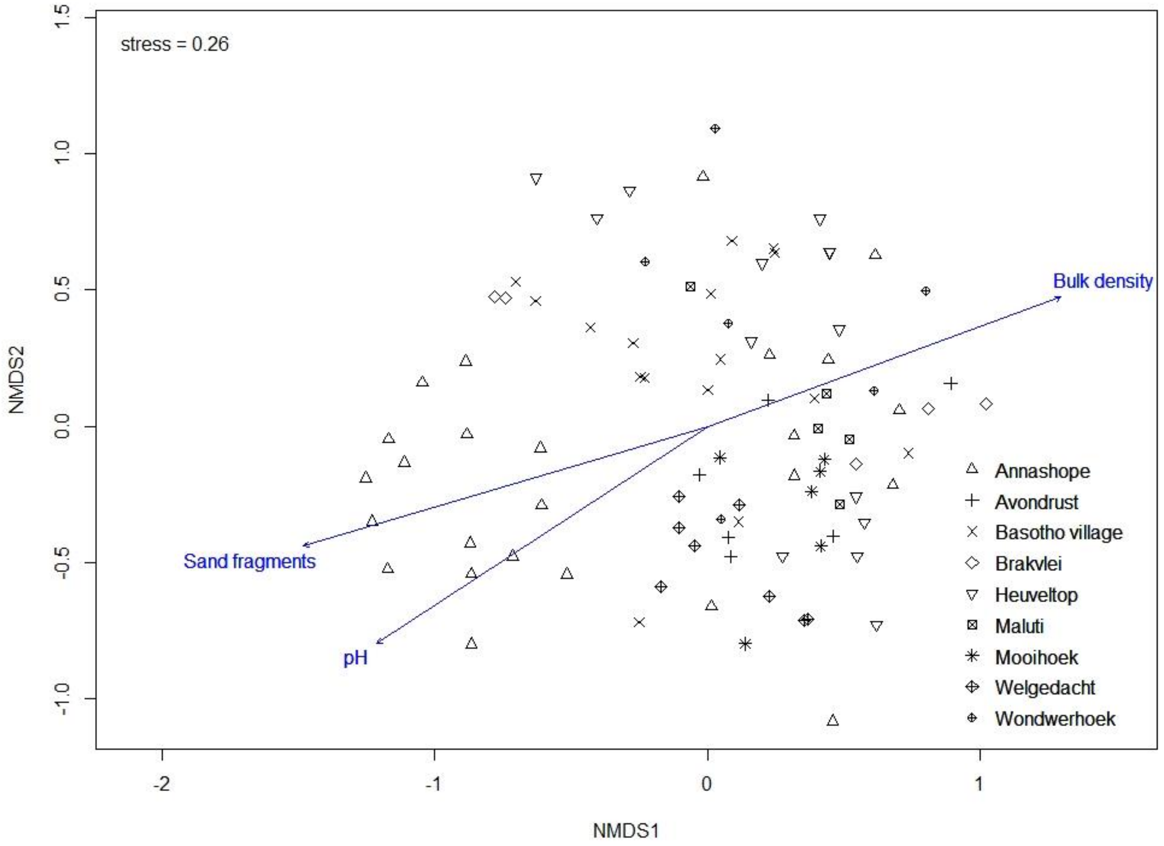

Figure 6: non-metric multidimensional scaling ordination of Golden Gate Highlands National Park, species composition (based on Bray-Curtis dissimilarity scores, after Wisconsin double standardization)

Discussion

In this study, there was no correlation between species richness and any of the topographic variables of slope and elevation and exhibited a high coefficient of variation. Similarly, a metanalysis by Lundholm, (2009) opposes the widely spread assumption that heterogeneity-diversity relationships should be positive by showing many studies that are consistently finding no relationships. Positive heterogeneity-diversity relationships are common but not universal (Lundholm, 2009). In grassland ecosystems, both local and landscape heterogeneity parameters are crucial for determining species richness and interactions (Reitalu *et al.*, 2012; Zulka *et al.*, 2014). The mountain range in GGHNP has extremely variable and unique topography, which could result in a diversity of habitats (Kay *et.al.*, 1993). The variations in mountain topography create micro-environmental conditions that influence habitat heterogeneity and fragmentation (Gaury and Devi, 2017b; Brown and du Preez, 2020). Consequently, these micro-habitats harbour and shelter a diversity of species serving as important biodiversity hotspots, providing ecosystem services such as pollination and controlling biological invasion (Zulka *et al.*, 2014).

A model incorporating species diversity in a set of environmental variables performed well in predicting species richness. In addition, species diversity was strongly related to species richness in Spearman's correlation analysis. The results highlight the importance of incorporating ecological variables in predicting heterogeneity-diversity relationships, which require an understanding of individual species' response to gradients of spatial environmental heterogeneity factors (Lundholm, 2009). Tilman *et.al.* (1996) found that soil minerals and nitrogen utilization in grassland ecosystems increased with increasing species richness and diversity, thus lowering the nutrient leaching of soils. Lundholm (2009) suggests that species diversity allows for compartmentalization and complementary resource use because no single species can fully exploit the conditions of the entire ecosystem, a concept known as niche differentiation (Tilman, 1999b). Therefore, considering species diversity and richness as both independent and dependent variables is useful in understanding their influence on the GGHNP grassland ecosystems' resilience.

Bulk density was the only variable negatively affecting species richness within the abiotic-based model. This is not surprising because the soil in the mountains is affected by topography and thus influences local terrestrial plant diversity patterns (Moeslund *et al.*, 2013). The topographical features can impair or amplify spatial patterns of water loss and accumulation. As such, an essential factor in determining vegetation patterns is the degree to which certain areas of a landscape capture or retain water and nutrient supplies. This finding gives insight into the role of soil drainage in controlling grassland vegetation patterns in the park. In GGHNP, the conservation of good soils can in turn be crucial for maintaining grassland plant diversity.

The results of the topo-edaphic-based model showed that elevation and bulk density jointly affected species richness. This emphasizes the role of topographically controlled edaphic features on landscapes in GGHNP. This could be an indication that compact soil with a higher bulk density can be vital for plant biodiversity in grassland terrestrial plant species. The results indicate that stable landscapes at high elevations harbour niche species diversity. Mid and foot-slope catena are often the recipients of all upslope water and nutrient resources. This causes enrichment that either limits or enhances species richness and diversity, depending on soil quality (McNaughton, 1985b). Steep slopes are prone to soil erosion, a disturbance that could either promote the growth of early-succession species or preclude disturbance-intolerant species in mountainous grassland ecosystems (Filibeck *et al*., 2019; Armas-Herrera et al., 2020a).

Species composition was significantly related to the environmental predictors of soil variables and elevation. The ordination showed that species composition in some sites was driven by sand fragments, pH, or bulk density. This is consistent with the results of Armas-Herrera *et al*.,(2020b), who suggested that grassland vegetation mosaics are maintained by soil-plant feedbacks, especially in subalpine grassland communities where plant co-existence is driven by soil disturbances and inevitable succession. This finding emphasizes the importance of soil drainage in maintaining species diversity in African Mountain vegetation (Brown and du Preez, 2020). A shift in species composition will likely occur in association with increasing soil resources such as water and nutrients. Soil texture as defined by particle size is important for controlling soil-plant relationships, especially seedling germination and establishment. In North America, *Bouteloua gracilis* requires a specific set of environmental conditions for germination and seedling establishment. These conditions are strongly influenced by silt and clay content (Fair, Peters, and Lauenroth, 2001).

Conclusion

This study corroborates earlier studies that drivers of species diversity in grassland communities are multi-dimensional by nature. This is evident by the variations in the importance and direction of multiple environmental drivers about their influence on species richness. To understand community dynamics, especially the diversity-stability relationships, the incorporation of biotic and abiotic factors is beneficial. The notion that diverse habitats, which are influenced by unique and heterogeneous topography in mountainous biomes, enhance different species' composition should be investigated. These habitats could be meta-populations fragmented into islands of similar species composition. The research shows that slope and soil interactions are important influencers of species richness, particularly slope and bulk density. Soil conservation should therefore be a priority, especially in mesic mountainous grasslands where the soil is prone to erosion and nutrient leaching.

Chapter 4

Diversifying modeling techniques to disentangle the complex patterns of species richness and diversity in the protected Afromontane grasslands

This chapter is based on:

Mashiane, K.K. Ramoelo, A. Adelabu, S.A. ‘Diversifying modeling techniques to disentangle the complex patterns of species richness and diversity in the protected Afromontane grasslands’, *Biodiversity and Conservation* (**under review**).

Abstract

Research in ecology has focused on the importance of environmental factors on spatial biodiversity variations and organization. This is important because of scant conservation resources. In this study used stepwise backward selection and recursive feature selection (RFE) to identify a parsimonious model that can predict species richness and diversity metrics in response to three models; biotic, abiotic and topo-edaphic. These results show that both metrics are good predictors of one another, mainly because species diversity is a combination of species richness and abundance, and further highlights the importance of biotic variables in predicting species distribution. Soil texture and its interactions with topographic variables were selected as the most important variables by the two modeling techniques. However, random forest performed worse than multiple linear regression in the prediction of diversity metrics. This study highlights the importance of topographically controlled edaphic factors as drivers of species richness and diversity in mountainous grasslands where topography inherently controls the geomorphic and hydrological characteristics, as a result, ecological processes.

Introduction

Many environmental factors affect plant species richness and diversity in grassland communities (Auestad *et al*., 2008). Habitat fragmentation and loss are causing major declines in grassland biodiversity and consequently, ecosystem services are dwindling (Zulka *et al.*, 2014). The most crucial factors affecting grassland biodiversity are ecological, topography, and management intensity (Orlandi *et al.*, 2016). However, in the Afromontane grasslands, these features are many and site-specific, making it difficult for efficient conservation management. Conservation resources are often scarce, and knowledge about which environmental drivers promote species diversity and richness in mountainous protected areas could allow for prioritization of hotspots and allocation of conservation efforts (Olea *et al*., 2010). For example, nutrient-poor grasslands in Central Europe are identified as one of the biodiversity hotspots in the region and therefore highly valued by conservationists (Becker and Brändel, 2007).

Vegetation response to environmental factors should be studied at specific yet multiple scales (recitals 2009). Variations in grassland vegetation are a result of complex gradients of soil moisture and element concentrations locally, and climatic variables regionally (Auestad *et al*., 2008). Research on vegetation-environment relationships is scanty for most parts of the Afromontane grasslands, especially locally, hence, as a result, conservation efforts may be haphazard and on an ad-hoc basis. For example, some edaphic variables, such as soil carbon content increased species richness while heavy metals and the C/N ratio were unrelated and decreased species richness, respectively (Becker and Brändel, 2007). Grassland landscapes can be better understood when good predictors of species richness are used (Zulka *et al.*, 2014). Predictors can be used to comprehend the complex patterns of species richness, especially in mountainous grasslands, which are vulnerable to global environmental changes.

In alpine grassland, environmental factors, such as climate, topography and soil are pivotal for maintaining plant diversity. For example, altitude and slope influence species composition. Inherently, topography in mountainous areas determines temperature, elevation, and hydrology; therefore, is an important determinant of vascular plant diversity (Moeslund *et al.*, 2013). In addition, topography in mountainous habitats facilitates geomorphological processes such as erosion and thus soil fertility along slope gradients as well as aspects affecting ecological processes. Ultimately, this interplay of environmental factors influences vegetation variations in mountainous areas and can be used to identify biodiversity hotspots for direct conservation protocols in protected areas (Lee and Chun, 2016). Mountain ranges in Southern Africa vary in topography and aspect, thereby providing suitable habitats for various plants and animals (Brown and du Preez, 2020). The role of topography needs to be investigated, especially in habitats where topographic interplays are poorly understood (Moeslund *et al.*, 2013).

The most active research in ecology has been to understand the importance of environmental factors on spatial biodiversity variations and organization. Species distribution models, which use a combination of species occurrence or abundance and environmental aspects, can aid in gaining insight into vegetation distribution across landscapes and allow for extrapolation in space and time. Furthermore, key modeling steps such as gathering data, selection of relevant predictors, and appropriate modeling algorithms can influence the robustness and realism of the model (Elith and Leathwick, 2009). Modeling algorithms such as parametric generalized linear models and non-parametric random forest (RF) can be used to assess the uncertainty of modeling algorithms (Bittner *et al.*, 2011). In protected areas of developing countries, such as Golden Gate Highlands National Park in South Africa, conservation resources are limited and ought to be used sparingly. Therefore, good models that can predict the potential distribution of vascular plant diversity, which is linked to nutrient-rich forage, can give initial insights into animal distribution and forage preference. Subsequently, precise carrying capacity models can be developed. This research aimed to gain insights into the main drivers of vascular plant diversity in the mountainous grassland of Golden Gate Highlands National Park by two different modelling techniques.

Methodology

Study area

The study was conducted at Golden Gate Highlands National Park, which is in the North-Eastern part of the Free State province, South Africa (Figure 1). The park is 32,758.35 ha in size and is located between 28°27' S and 28°37' S and 28°33' E and 28°42' E latitudes. The park is in mountainous grasslands at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The following soil types are identified in the park: shallow rocky soils (Glenrosa and Mispah), deep soil along drainage lines (Oakleaf), well-developed sand soils (Hutton and Clovelly), and clayey structured soils (Milkwood and Tambakulu) (SANParks, 2020). Rainfall in the summer, mild summers, and cold winters define the park. The rainy season lasts from September to April, and the average annual rainfall is between 800 and 2,000 mm (Kay, Bredenkamp, and Theron, 1993). The park lies between 1,892 m and 2,829 m above sea level and hence, comprises the following grasslands units: Eastern Free State Sandy grasslands (Gm 4), Basotho Montane Shrubland (Gm 5), Lesotho Highveld Basalt Grassland (Gd 8), and Northern Drakensberg Highveld (Gd5) (Mucina and Rutherford, 2006).

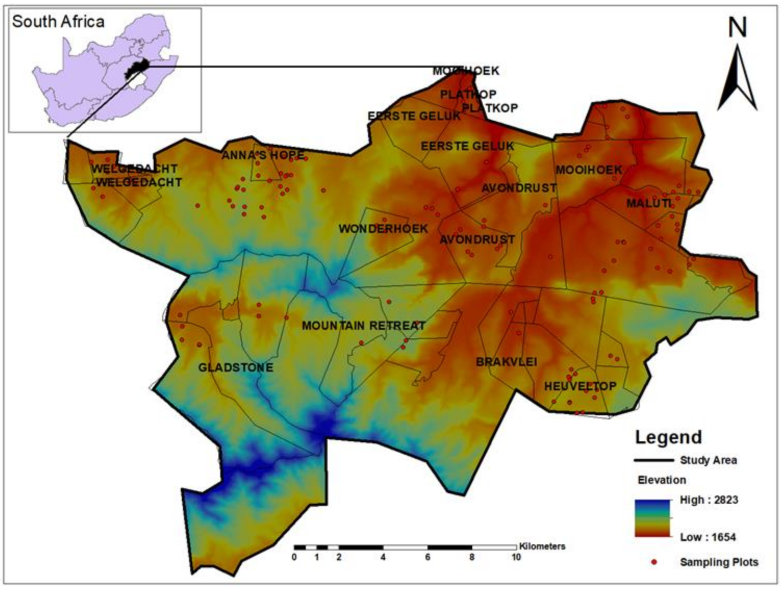

Figure 1: A map of the study area

Data collection

Field data collection

The park was stratified into relatively homogenous physiographic-physiognomic units (Kay, Bredenkamp, and Theron, 1993). Vegetation sampling was undertaken in a 30m × 30m plot size randomly placed within homogeneous patches of grass communities (Figure 2). Four 30 m transect were placed horizontally at every 10 m interval within plot, the plot was generated in using the “generate random points” tool in ArcMap 10.7. In each plot, a 1m × 1m quadrat size was placed at every 10 m interval along the transect for vascular plant species identification and ariel visual cover estimations. In total, 142 plots with 16 quadrats were sampled in the study area. Data were collected in March, April, and May 2019 which are the rainy growing months for South African montane grasslands, this area receives late rains for summer, thus prolonged vegetation growing season, in addition to xeromorphic characteristics of plant species. Wherever possible, every vascular plant species of the standing vegetation in each quadrat was identified down to the species level, and the 1 m quadrat was gridded into 100 cells with a surface area of 1 cm2, or 1% cover, to record the aerial cover of each species Figure 3.

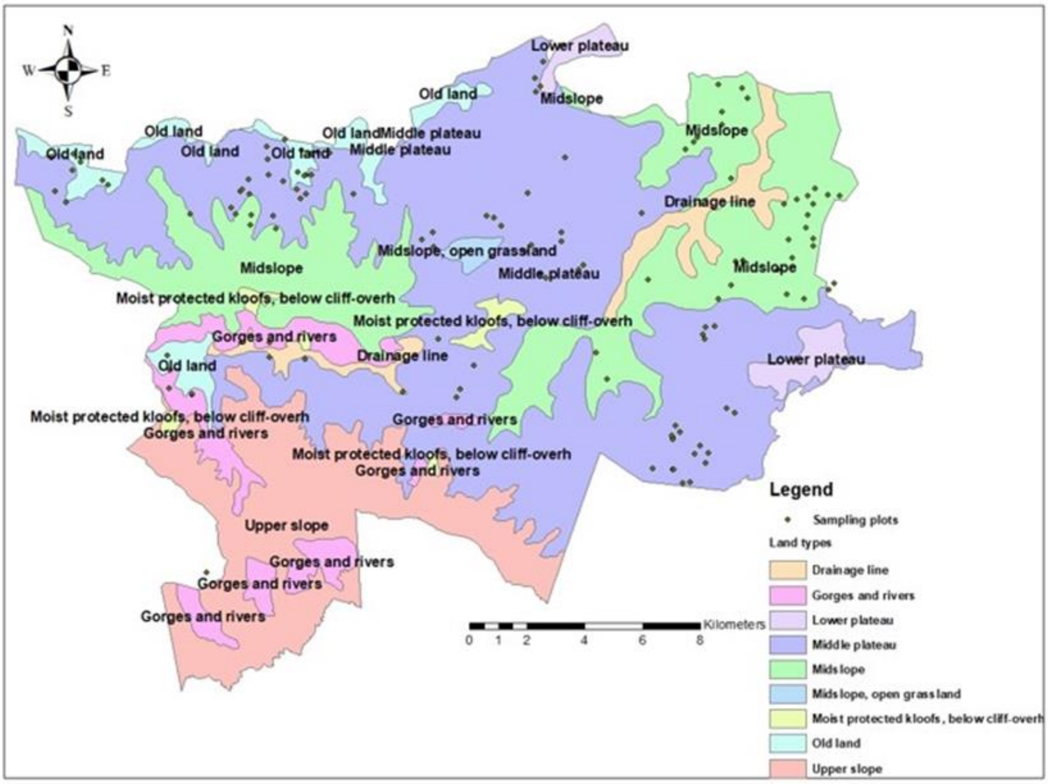

Figure 2: The distribution of different vegetation types including sampling location

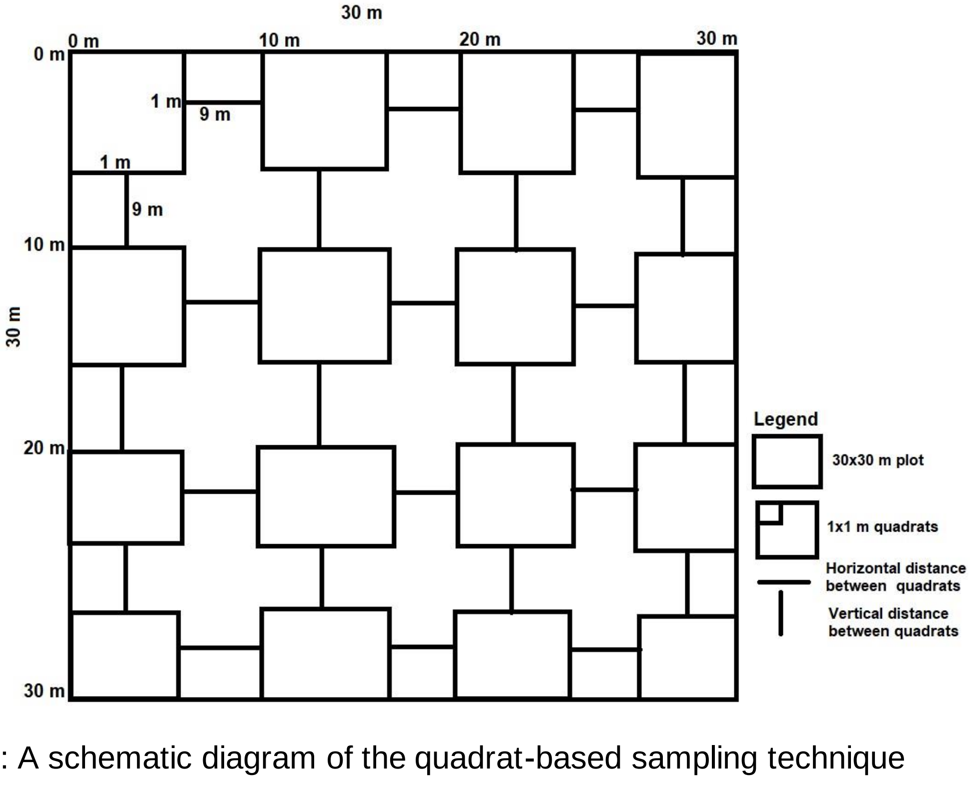

Figure 3: A schematic diagram of the quadrat-based sampling technique

Environmental variables and fire severity

A variety of environmental predictor variables from different data sets were used to measure and model their influence on species richness and diversity. For soil, both chemical and physical variables were used; bulk density (BD), silt fragments (SF), pH, coarse fragments (CF), soil organic carbon (SOC), sand (SD), and nitrogen of topsoil (15 cm) were downloaded from International Soil Reference and information Center (https://soilgrids.org). The park comprises of shallow rocky soils, while field sampling is encouraged in underrepresented areas (Hengl *et al.*, 2017), physical soil sampling may be damaging in the park. Using this grided soil data with relatively good accuracies (Hengl *et al.*, 2017) is non-destructive, therefore, ideal for a conservation area. The elevation data, i.e., SRTM DEM at 30 m resolution was obtained from US Geological Survey's EROS data center (https://earthexplorer.usgs.gov). The slope was derived from DEM using the Spatial Analyst Tool in ArcGIS 10. 4. Topographic and edaphic variables were multiplied to determine the interaction terms of topo-edaphic variables. Fire severity data were acquired from a study conducted in the area (Adagbasa *et al*., 2018). By comparing remote sensing photos from before and after a fire season that contained burned and unburnt pixels, the study calculated the intensity of the fire using the Normalized Burn Ratio (NBR) index (Adagbasa *et al*., 2018).

Diversity metrics

Utilizing the Shannon diversity index H'=-pi*lnpi, where pi is the proportion (species cover), species diversity was calculated for each species in the quadratwas computed using the vegan package (Oksanen *et al.*, 2012) in R studio (R core Team, 2022). Furthermore, total species richness was measured by counting all vascular plants species that were recorded for each quadrat. Subsequently, all the quadrat values were averaged to attain the plot level value of each diversity metric.

Data analysis

Simple multiple linear regression

Simple multiple linear regression were used to examine the relationship between diversity metrics (species diversity and richness) and a set of three models of explanatory environmental variables. The biotic model included diversity metrics as predictor variables amongst abiotic-based variables. The abiotic model included bulk density, silt fragments, sand, pH, soil organic content, coarse fragments, elevation, slope, fire severity, and soil nitrogen. The topo-edaphic model included interactions between topographic and soil variables. The diversity metrics (species diversity and richness) were used as both response and predictor variables, because of the high colinearity between the two. Subsequently, a backward selection was used to eliminate redundant variables and select the optimal variables explaining species richness and diversity from the set of environmental variables based on the lowest Akaike's Information Criterion. Statistical analyses were performed using R statistical software (R Core Team, 2022).

Random forest

A set of biotic, abiotic, and topo-edaphic variables was used as input into the nonparametric RF method to predict species richness and diversity. Random forest is a highly recommended method for ecologists and remote sensing scientists and was developed to improve classification and regression trees (CART), by making use of a large set of regression trees (Ramoelo *et al.*, 2015a). The 'random forest' package available in R statistical software was used to analyze the data. The optimization of the number of variables required to predict species richness and diversity was determined using a random feature selection (RFE) based on leave-one-out cross-validation and the root mean square error (RMSE). The RFE is a simple backward selection algorithm implemented via the 'caret' package in R statistical software (R Core Team, 2022).

Model performance

The statistical measure of model precision and robustness, the r-squared (R^2) and root mean square error (RMSE), was determined to test the performance of the modeling algorithms using the selected stepwise and RFE variables in predicting the diversity metrics. The parameters were used to assess the strength of the relationship between the observed and predicted species richness and diversity.

Results

The species and richness diversity metrics had a coefficient of variation of 25% and 39%, respectively (Table 1). All edaphic variables had a very low coefficient of variation (<10%), except for coarse fragments (38%) and soil organic content (32%). For topographic variables, the slope had the highest (80%) and elevation the lowest (0.06%) coefficient of variation. Fire severity had a coefficient variability of 26%. All the topo-edaphic variables exhibited a substantially high coefficient of variation (>82%).

Table 1: Descriptive statistics of biotic, abiotic, and topo-edaphic variables in grassland communities of Golden Gate Highland National Park

Variables	units	**Min**	**Max**	**Range**	**Mean**	**Variance**	**STDEV**	**CV (%)**
Biotic								
Species richness	*S*	2.50	17.00	14.50	8.97	12.18	3.49	0.39
Species diversity	H'	0.64	2.45	1.82	1.63	0.16	0.40	0.25
Abiotic								
Bulk density	cg/cm^3	1199.00	1454.00	255.00	1389.57	1301.99	36.08	0.03
Silt fragments	g/kg	16.00	26.00	10.00	20.37	2.86	1.69	0.08
Sand	g/kg	44.00	59.00	15.00	54.08	6.57	2.56	0.05
pH	pH$^{\wedge 10}$	56.00	62.00	6.00	59.29	2.15	1.47	0.02
Soil organic content	dg/kg	8.00	41.00	33.00	13.09	17.06	4.13	0.32
Coarse fragments	cm^3/dm^3	2.00	19.00	17.00	7.12	7.38	2.72	0.38
Elevation	masl	1686.00	2399.00	713.00	1856.61	11281.87	106.22	0.06
Slope	angle°	0.00	26.73	26.73	6.45	27.17	5.21	0.81
Fire severity	-	2.00	6.00	4.00	3.72	0.93	0.96	0.26
Nitrogen	cg/kg	196.00	334.00	138.00	234.24	398.17	19.95	0.09
Topo-edaphic								
Elevation_bulk density		4455.00	32147.00	27692.00	16604.98	41079757.52	6409.35	0.39
Elevation _silt fragments		28880.00	62374.00	33494.00	37910.51	22278807.63	4720.04	0.12

Elevation _coarse fragments	3372.00	40783.00	37411.00	13391.62	33430633.81	5781.92	0.43
Elevation _pH	98439.00	139142.00	40703.00	110045.37	40048610.86	6328.40	0.06
Elevation _sand	82614.00	112690.00	30076.00	100301.60	31684077.83	5628.86	0.06
Slope_ bulk density	0.00	352.65	352.65	58.91	3295.60	57.41	0.97
Slope_silt fragments	0.00	614.70	614.70	131.02	11465.87	107.08	0.82
Slope_coarse fragments	0.00	237.07	237.07	45.85	1696.60	41.19	0.90
Slope_pH	0.00	1576.84	1576.84	382.78	97720.28	312.60	0.82
Slope_sand	0.00	1449.44	1449.44	350.12	81511.77	285.50	0.82

STDEV: standard deviation; CV: coefficient of variation. Shapiro-Wilk normality was computed in R statistical software for all variables and confirmed that all variables were normally distributed ($p<0.05$).

The optimal environmental variables explaining species richness and diversity in grassland communities in Golden Gate Highland National Park are given in Table 2. The step-wise backward selection retained species diversity, silt fragments, and fire severity as the optimal variables explaining species richness. However, species diversity had a significantly positive (7.04, $p<0.05$), while silt fragments and fire severity had a significantly ($p<0.05$) negative relationship (-0.24 and -0.41, respectively). Bulk density (-0.02) and silt fragments (-0.03) were identified as the two optimal abiotic variables with a significantly ($p<0.05$) negative relationship with species richness. Similarly, in the topo-edaphic model, the stepwise selected elevation, soil bulk density, silt fragments, sand, and pH had a significantly negative ($p<0.05$) albeit minute relationship with species richness. Species diversity was optimally explained by species richness (0.10), soil pH (0.02), and fire severity (0.04). Coarse fragments had a significantly ($p<0.05$) negative relationship (-0.03) with species diversity amongst the selected abiotic variables. However, two topographic variables were identified as optimally explaining species diversity, namely elevation interaction with soil bulk density and sand, but with very small estimate coefficients.

Table 2: The optimal environmental variables explaining species richness and diversity in grassland communities in Golden Gate Highland National Park

Variables	Estimate	Standard error	T-value	p-value
Species richness				
Biotic model				
Species diversity	7.042	0.40	17.57	0.00
Silt fragments	-0.24	0.09	-2.53	0.01
Fire severity	-0.41	0.16	-2.47	0.01
Abiotic model				
Bulk density	-0.02	0.02	-2.62	0.00
Silt fragments	-0.03	0.20	-1.86	0.00
Soil organic content	-0.14	0.09	-1.45	0.14
Coarse fragments	-0.20	0.13	-1.49	0.14
Topo-edaphic model				
Elevation_bulk density	-0.01	0.00	201.88	0.00
Elevation_silt fragments	-0.01	0.00	-8.16	0.00
Elevation_pH	-0.01	0.00	-2.06	0.00
Elevation_sand	-0.01	0.00	-8.36	0.00
Species Diversity				
Biotic model				
Species richness	0.10	0.00	17.77	0.00
pH	0.02	0.00	1.64	0.10
Fire severity	0.04	0.00	2.18	0.03
Abiotic model				
Bulk density	0.00	0.00	-1.74	0.09
pH	0.04	0.02	-1.63	0.09
Coarse fragments	-0.03	0.01	-2.34	0.02
Topo-edaphic model				
Elevation_bulk density	0.00	0.00	17.11	0.00
Elevation_sand	0.00	0.00	-2.25	0.03

Measures of model precision and robustness (R^2 and RMSE) to test the performance of the modelling algorithms in predicting species richness and diversity in grassland communities of Golden Gate Highland National Park are given in Table 3. The top 5 variables selected within the biotic model by RFE were species diversity (RMSE =

2.22, $R^2 = 0.60$), silt fragments (RMSE = 2.06, $R^2 = 0.65$), elevation (RMSE = 2.12, $R^2 = 0.65$), sand (RMSE = 2.14, $R^2 = 0.65$) and coarse fragments (RMSE = 2.06, $R^2 = 0.67$). Within the aboitic model, elevation (RMSE = 3.27, $R^2 = 0.12$), sand (RMSE = 3.36, $R^2 = 0.09$), silt fragments (RMSE = 3.26, $R^2 = 0.13$), soil organic content (RMSE = 3.26, $R^2 = 0.14$), and bulk density (RMSE = 3.79, $R^2 = 0.02$) were the top 5 variables explaining species richness. An interaction between elevation and bulk density (RMSE = 0.56, $R^2 = 0.97$) was identified as the top variable within the topo-edaphic model to explain species richness. Species richness (RMSE = 0.24, $R^2 = 0.64$) and elevation (RMSE = 0.26, $R^2 = 0.59$) were selected as the top two variables explaining species diversity within the biotic model. However, within the aboitic based model, sand (RMSE = 0.45, $R^2 = 0.09$), elevation (RMSE = 0.38, $R^2 = 0.12$), soil organic content (RMSE = 0.38, $R^2 = 0.09$), silt fragments (RMSE = 0.41, $R^2 = 0.03$), and bulk density (RMSE = 0.45, $R^2 = 0.37$) were the top five most important variables. Lastly, elevation-bulk density (RMSE = 0.28, $R^2 = 0.54$) and slope-bulk density (RMSE = 0.26, $R^2 = 0.59$) interactions were the two most important variables within the topo-edaphic model.

Table 3: Measures of model precision and robustness (R^2 and the root means square error (RMSE)) to test the performance of the modeling algorithms in predicting the species richness and diversity in grassland communities of Golden Gate Highland National Park

Model variables	**RMSE**	**R^2**	**Mean absolute error**	**RMSE**	**R^2**	**Mean absolute error**
	Species richness			**Species diversity**		
Biotic						
Species diversity	2.24	0.60	1.82	0.24	0.64	0.19
Bulk density	2.11	0.63	1.66	0.25	0.60	0.20
Silt fragments	2.04	0.67	1.66	0.26	0.60	0.21
Sand	2.13	0.64	1.75	0.26	0.60	0.21
pH	2.22	0.63	1.83	0.27	0.57	0.22
Soil organic content	2.06	0.66	1.69	0.25	0.61	0.20
Coarse fragments	2.09	0.66	1.71	0.25	0.62	0.20
Elevation	2.13	0.65	1.74	0.26	0.59	0.22
Slope	2.06	0.66	1.67	0.25	0.61	0.21
Fire severity	2.11	0.66	1.72	0.25	0.61	0.21
Nitrogen	2.14	0.65	1.78	0.26	0.61	0.21
Abiotic						
Bulk density	3.79	0.03	3.05	0.42	0.01	0.34
Silt fragments	3.53	0.04	2.94	0.41	0.04	0.33
Sand	3.39	0.08	2.80	0.39	0.09	0.32
pH	3.32	0.10	2.73	0.39	0.09	0.31
Soil organic content	3.25	0.13	2.68	0.38	0.10	0.31
Coarse fragments	3.24	0.14	2.66	0.38	0.12	0.31

Elevation	3.31	0.11	2.72	0.38	0.11	0.31
Slope	3.35	0.09	2.77	0.38	0.11	0.31
Fire severity	3.35	0.09	2.77	0.39	0.08	0.32
Nitrogen	3.35	0.08	2.77	0.39	0.07	0.32
Topo-edaphic						
Elevation_bulk density	0.56	0.97	0.46	0.28	0.54	0.22
Elevation_silt fragments	0.76	0.96	0.58	0.26	0.59	0.21
Elevation_coarse fragments	0.92	0.95	0.69	0.26	0.59	0.21
Elevation_pH	1.13	0.93	0.81	0.26	0.57	0.21
Elevation_sand	1.32	0.91	0.92	0.27	0.56	0.22
Slope_bulk density	0.92	0.95	0.63	0.26	0.59	0.21
Slope_silt fragments	1.06	0.94	0.71	0.26	0.58	0.21
Slope_coarse fragments	1.13	0.93	0.78	0.26	0.58	0.22
Slope_pH	0.93	0.95	0.63	0.26	0.58	0.21
Slope_sand	1.03	0.94	0.72	0.26	0.59	021

The multiple linear regression models (Figure 4) performed better than the RF models (Figure 5). The multiple linear regression biotic model explained 72% of the variation in species richness with an RMSE of 1.84, while abiotic and topo-edaphic models explained 12% (RMSE = 3.06) and 99% (RMSE = 0.11) respectively. However, the models' performances were reduced when species diversity was the response variable. The biotic model explained 71% variation (RMSE = 0.12). On the other hand, abiotic and topo-edaphic models explained 0.8% and 69% variation, respectively. The

RF biotic model explained 62% of the variation in species richness with the abiotic model explaining 0.03 (RMSE = 3.46) and the topo-edaphic model 91% (RMSE = 1.20). When species diversity was the response variable, the biotic model explained 59% of the variation, abiotic model -0.2% (RMSE = 0.45) and topo-edaphic 64% (RMSE = 0.26).

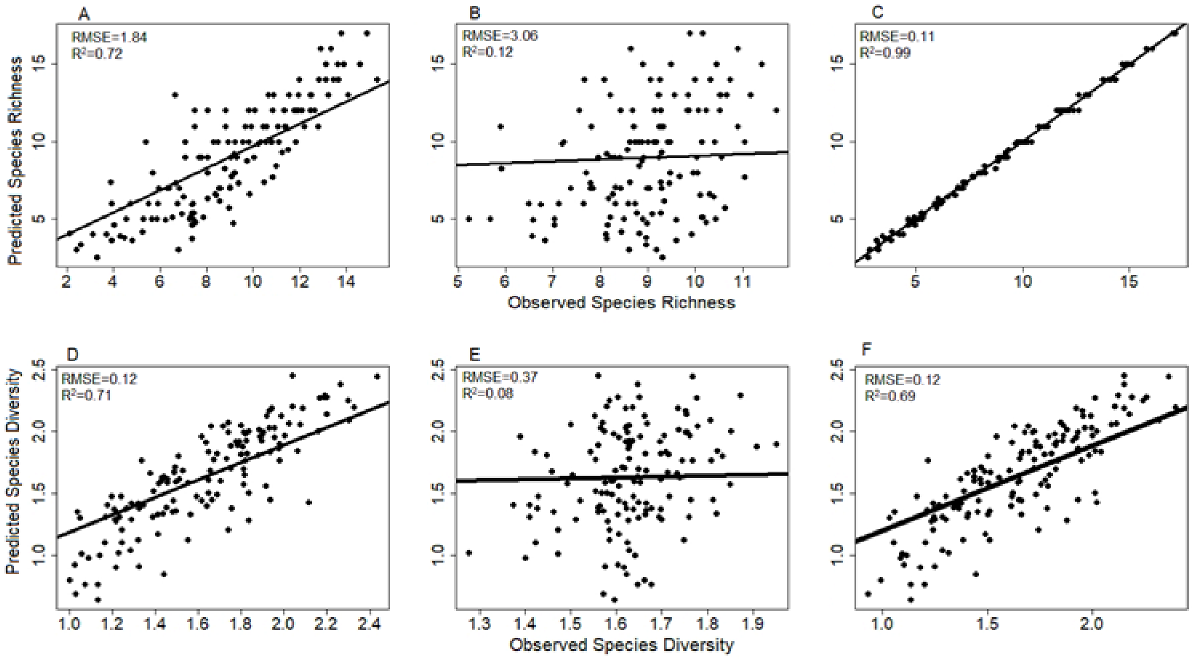

Figure 4: Multiple linear regression species richness (A-C) and diversity (D-F) estimation performance of biotic (A and D), abiotic (B and E), and topo-edaphic (C and F) based models

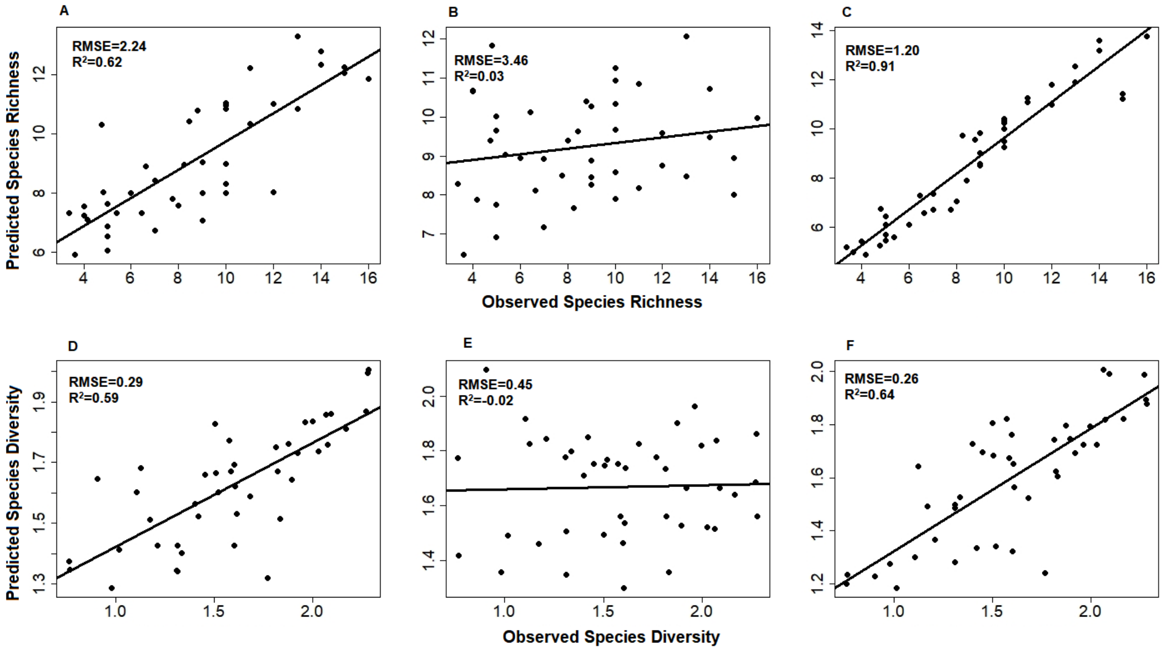

Figure 5: Random Forest regression species richness (A-C) and diversity (D-F) estimation performance of biotic (A and D), abiotic (B and E), and topo-edaphic (C and F) based models

Discussion

Soil edaphic variables exhibited a very low coefficient of variation (<10%), while slope showed a very high variation of 80%, second to all the topo-edaphic variables with a variation of >82%. The low variation of the soil edaphic dataset in Golden Gate Highland National Park can be attributed to unvarying land-use activities both current (protected area with grazing activities) and historic (farming: crop cultivation) because soil stoichiometry differs amongst and is influenced by land-use activities. For example, soil total carbon (C) and nutrients varied among different grassland types and land use in Alpine Qinghai-Tibetan Plateau (Han *et al.*, 2019). Furthermore, a mean slope of 6.4% that ranged from completely flat (0.0%) to gentle (0.26%) was observed, despite the high variation in the dataset, which may have been because of the difficulty of sampling steep slopes >45 degrees.

The diversity metrics served as both response and predictor variables. This gives insight into forces controlling species richness and diversity in response to biotic and abiotic factors (Tilman, 1993). The high collinearity between the two metrics was inevitable, because species richness comprises the number of species, and species richness and abundance are considered by the Shannon-Wiener diversity index. Despite having similar traits, the diversity metrics may differ in their sensitivity to biotic and abiotic factors. For example, species richness may be sensitive to rapid changes in key rare species, while species diversity may react to the changes of abundances dominant species. Thus, Symstad and Jonas (2011) suggested that the response of richness and diversity to drivers reflect different changes in the plant communities.

This research showed a significantly negative relationship between species richness and silt fragments as well as fire severity. pH and fire severity, however, influenced species diversity via the stepwise algorithm in the biotic model. Edaphic factors are related to geomorphic heterogeneity, especially slope stability can have an effect on species richness at specific scales in mountainous areas (Malanson *et al.*, 2020b). Moreover, fire severity is one of the variables selected by the RF as explaining negatively species richness. High fire severity will cause more physical/structural damage to the grassland vegetation community, hence vegetation seldom fully recovers to pre-fire conditions (Adagbasa *et al.*, 2020). Generally, however, variations in fire regimes can serve as a stimulant for seedling establishment and emergence (Olea *et al.*, 2010), while high fire severity, in particular, can obliterate existing species in grassland plant ecosystems. . Interestingly, elevation and edaphic interaction variables were identified as optimal for influencing species richness and diversity. This emphasizes the role of topography in influencing local plant diversity. For instance, elevation may limit spatial seed distribution, especially by animals that find it difficult to ascend steep slopes (Moeslund *et al.*, 2013). The RF algorithm similarly identified silt fragments, elevation, and other soil texture variables as important for explaining species richness. This also shows that topographically controlled edaphic factors such as silt fragments can be drivers of species richness in mountainous areas (Filibeck *et al.*, 2019). It is, therefore, important to prioritize soil conservation in protected areas to improve plant species diversity, especially in humid mountainous areas, which are susceptible to erosion and nutrient leaching.

Species richness and diversity were affected differently by biotic, abiotic, and topo-edaphic models. As such, species richness was better predicted by all the multiple linear regression algorithms when compared to species diversity. The RF performed worse in predicting diversity metrics, because of the high prediction error, i.e., RMSE. The results concur with studies that question the predictive power of RF in estimating vegetation parameters(Filibeck *et al.*, 2019; Kosicki, 2020), especially because it is a non-parametric technique that does not make assumptions about statistical distribution (Ramoelo *et al.*, 2015a). However, RF is recommended for ecological use, because many variables can be tested, are robust to multicollinearity, and are independent of any assumption, in comparison, stepwise multiple regression is prone to multicollinearity and overfitting; while machine learning methods are prone to different specific conditions in an ecosystem and lack analytical assumptions, which influence their predictive power.

Conclusion

Conservation of natural resources warrants active research to gain insight into the distribution of species occurrence, abundance, and composition. Especially, because conservation resources are dwindling with heightened threats to biodiversity. This research can, therefore, provide a framework for identifying environmental drivers of species richness through the selection of parsimonious models. The two metrics were useful rangeland indicators. When species distribution is well understood they can be used to model carrying capacity because they are linked to plant productivity and nutrition. Furthermore, the importance of topographically controlled edaphic grasslands as drivers of species richness and diversity in mountainous grasslands is highlighted. Topography inherently controls the geomorphic and ecological processes. Fire severity is one of the most concerning predictors of species diversity to park managers in the park, which is susceptible to sporadic anthropogenic fire, drivers of fire severity ought to be investigated for abating species loss as a result of high fire severity. The findings show that diversifying model techniques can help to disentangle the complex patterns of species richness

Chapter 5

Predicting Species Richness and Diversity Using Satellite Remote Sensing and Random Forest Machine Learning Algorithm

This chapter is based on:

Mashiane, K.K. Ramoelo, A. Adelabu, S.A. ‘Using machine learning random forest to examine correlations between and predict key biodiversity by remote sensing factors’, *Geomatics* (**submitted**)

Abstract

Compiling species inventory in mountainous grasslands has been challenging owing to the amount of time and cost needed to conduct field surveys. Since remote sensing techniques are affordable and enable rapid coverage of huge areas, they may be useful for monitoring crucial biodiversity variables. This study sought to investigate the relationship between multispectral remote sensing data from both Landsat 8 and Sentinel 2, and species richness and diversity in mountainous and protected grasslands. This study found weak relationships between remote sensing vegetation indices and the diversity metrics, but significant relationships were found between diversity metrics, red-edge and visible bands in Sentinel 2 and Landsat 8, respectively . Moreover, using machine learning random forest, the multispectral datasets exhibited strong predictive powers. In this investigation, for both sensors, near-infrared (NIR) section of the electromagnetic spectrum seemed to be the most selected band to explain species diversity in mountainous grasslands, also vegetation indices such as EVI emerged as important. This finding further ascertains the efficiency of using NIR in vegetation mapping. The results are also indicative of the limitation associated with using broadband remote sensing sensors to estimate species diversity on a smaller scale. This research affirms that the NIR as most effective at estimating species richness and diversity in mountainous grasslands.

Introduction

The maintenance of ecosystem services and functioning, which ultimately promote human well-being, depends on biodiversity (Oliver *et al.*, 2015). In grassland ecosystems, plant diversity contributes to multiple ecosystem services at local and landscape scales, which makes conservation efforts within and among ecological communities imperative (Hautier *et al.*, 2018). However, monitoring biodiversity at multiple scales is a common conservation challenge in protected areas (Ferreira *et al.*, 2011). Different approaches to monitoring vegetation i.e., *in situ* species and remote sensing approaches are complex with pros and cons; hence a hybrid or coupling of approaches is necessary (Lausch *et al.*, 2018). Remote sensing approaches could be beneficial for monitoring essential biodiversity variables because it is cost-effective and allows for coverage of large areas over a short period in contrast to *in situ* methods (Lausch *et al.*, 2018). Consistent monitoring of vegetation forms the basis of wildlife conservation, which is essential as changes in the environment are initially observed in the vegetation (Brown *et al.*, 2013b).

Despite compelling evidence showing a link between biodiversity loss and ecosystem health, the matter is still contentious (Cardinale *et al.*, 2012). The loss of biodiversity as a major conservation issue has received varied scientific reporting and views worldwide (Cardinale *et al.*, 2018b). For example, according to a new analysis of time-series data, species richness is constant on average but is declining in some places and rising in others (Cardinale *et al.*, 2018b). Species diversity may be changing at varying spatial scales, and biodiversity may not be declining at local spatial scales across the globe (Cardinale *et al.*, 2007, 2018a). By observing species variety at various scales, it is feasible to compute the rates of species turnover and explain the homogeneity of the world's biota (Cardinale *et al.*, 2018b).

Plant species richness and diversity are key ecosystem indicators because the consequences of ecosystem disturbance can be explicitly observed through these diversity metrics (Symstad and Jonas, 2011). They also describe ecosystem health, stability, and resilience and can, therefore, be used for monitoring plant species (Lausch *et al.*, 2018), especially in conservation areas. Hence, the Group of Earth Biodiversity Observation Network identifies taxonomic diversity as one of the key biodiversity variables (Pereira *et al.*, 2013). Vegetation characteristics could be measured by remote sensing techniques, which are suitable for discriminating species turnover and floristic composition (Lausch *et al.*, 2018). Discriminating species using remote sensing is dependent on multiple biological and physical factors, appropriate data (Lausch *et al.*, 201,8), and modeling algorithms (Richter *et al.*, 2016). However, remote sensing approaches could be beneficial for monitoring essential biodiversity variables because it is cost-effective and allows for coverage of large areas over a short time (Rocchini *et al.*, 2016). Studies showed that when species relative abundances were taken into consideration, the association between species alpha diversity and remotely sensed spectral heterogeneity became stronger . This improved the capability of local species diversity estimations, especially while using spectral information in addition to the commonly used spectral indices (Rocchini *et al*. 2007a).

The use of multispectral sensors have limitations on properties such as species identification, for example, they are inferior compared to hyperspectral data in spectral information (Lyon and Huete, 2016). As such, alpha diversity is commonly predicted and mapped based on the spectral variation hypothesis (SVH) which starts with a heterogeneity map from a satellite sensor image correlated with field sampling data (Rocchini et al., 2007b; Rocchini *et al.*, 2016). In grassland studies, the spectral properties of grass species may be difficult to detect because of the similarity in the taxa, especially due to broadband remote sensing's inability to identify small differences in green vegetation. Furthermore, weak and moderate relationships between species richness, diversity, and spectral vegetation indices that are derived from remote sensing sensors (Rocchini *et al*., 2007b). This seemingly universal observation may be because VIs are poorly related to structural properties. For example, a study by excluded NDVI from the predictive analysis because of its weak correlation with species' structural features. However, the relationship between remotely sensed data and species diversity is scale-dependent (Rocchini *et al*, 2004). Specifically, Sentinel 2 has shown enormous potential for vegetation mapping because it the only multispectral sensor that can acquire images with 13 spectral bands (Torresani *et al.*, 2019) including Red-Edge bands (Xulu *et al*., 2021), in addition, its free and open data policy may be beneficial for nature conservation in developing and under-resourced countries (Torresani *et al.*, 2019).

Despite the availability of remote sensing data and applications, their use has yet to be widely incorporated by managers and researchers working in biodiversity monitoring (Reddy, 2021). Many studies have been conducted to map species diversity using remote sensing datasets and study designs, resulting in varying outcomes and accuracies (Schmidtlein and Fassnacht, 2017). Coarser spatial resolution datasets perform better when estimating species diversity than finer spatial resolution (Gessner *et al.*, 2015). Albeit, increased spectral resolution is deemed beneficial for improving the estimation of species diversity; however, hyperspectral remote sensing data is cumbersome, and most of its bands are redundant (Thenkabail *et al.*, 2004). Methodological advances such as machine learning algorithms may present significant opportunities for utilising multispectral datasets with high accuracies. The accuracy difference between hyperspectral and multispectral bands could be much higher for grass species (Gessner *et al.*, 2015)

The loss of biodiversity is now more prevalent for many biomes across the globe due to overexploitation and land use transformations; thus, monitoring and modelling biodiversity local scale via remote sensing can abate this crisis (Cardinale *et al.*, 2018b). Monitoring patterns of species diversity over time is essential for decision-making in conservation, especially in the context of unprecedented global change. Remote sensing is one of the most cost-effective approaches to identifying biodiversity hotspots and predicting changes in good time. This study aims to explore the relationship between multiple spectral bands and diversity metrics using remote sensing to monitor biodiversity in grassland communities.

Methodology

Study Area

The study was carried out in the north-eastern Free State, South Africa (Figure 1). The park is 32,758.35 ha in size and is located between 28°27' S and 28°37' S and 28°33' E and 28°42' E. The park is in mountainous grasslands at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The soil types in the park include shallow rocky soils (Glenrosa and Mispah), deep soil along drainage lines (Oakleaf), well-developed sand soils (Hutton and Clovelly), and clayey structured soils (Milkwood and Tambakulu) (SANParks, 2020). Rainfall in the summer, mild summers, and cold winters define the park. The rainy season lasts from September to April, and the average annual rainfall is between 800 and 2,000 mm (Kay *et al*., 1993). The park lies between 1,892 m and 2,829 m above sea level and comprises the grasslands units: Eastern Free State sandy grasslands (Gm 4), Basotho montane shrubland (Gm 5), Lesotho highveld basalt grassland (Gd 8), and Northern Drakensberg highveld (Gd5) (Mucina and Rutherford, 2006).

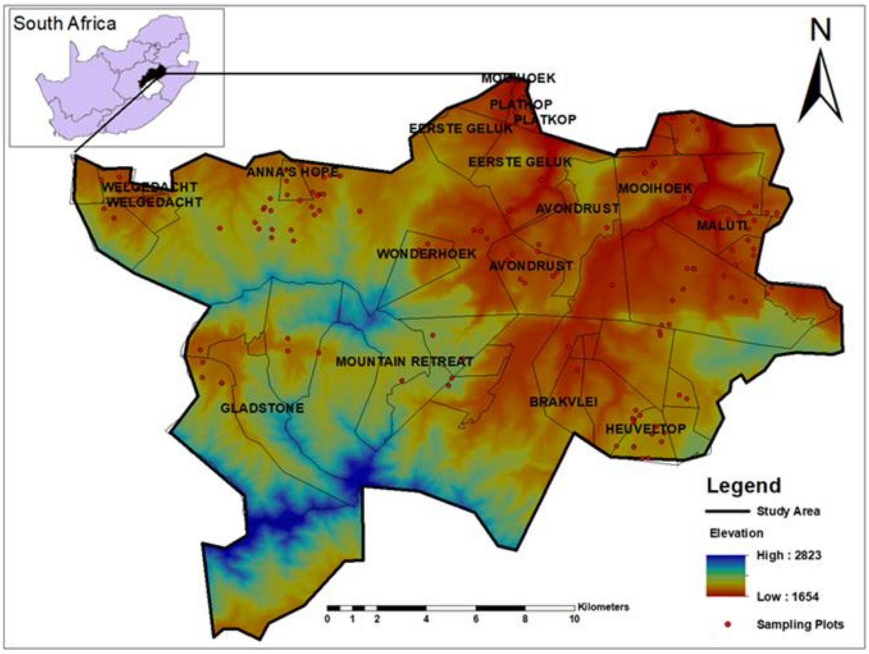

Figure 1: A map of the study area

Data collection

The land type map of GGHNP was used as the first level of stratification. Sampling sites of relatively homogenous grass units were in a randomly stratified manner. A total of 36 vegetation sampling plots (30m x 30m) ranging between 3 and 5 per site (6 sites with 5 plots and 2 sites with 3 plots) were placed randomly within the homogenous grass units. Within each vegetation sampling plot, a total of 16 (1m x 1m) quadrats were placed systematically at every 10 m along four parallel rows. In another 107 plots from 12 sites (plots ranging between 2 and 23 plots), 100 step points transect from 4 transects located within a 30m x 30m plot was used to record all species at every step. The dataset of the 36 plots was merged with that of 107 plots. The combined dataset comprised 142 plots with 13 sites.

The taxonomic composition of the vegetation and its cover were used to derive species richness and diversity per plot. The values from each plot were averaged to attain mean species richness per site. The variables, species richness, and diversity were computed using “vegan” and “plyr” in R studio statistical packages (Oksanen, 2017; Oksanen *et al.*, 2018), which employed the diversity and apply function for species diversity and richness, respectively. Species diversity was calculated using the Shannon-Wiener Index (Eq. 1) where pi is the proportion of the species within the sampling units. Species richness was determined by adding all species from each quadrat and averaged by the number of quadrats in each plot, to obtain the average plot value

Remote sensing data collection

Satellite images from Sentinel-2 and Landsat 8 datasets were extracted and processed from the JavaScript code editor Google Earth Engine (GEE). All the images in this research were near cloud-free. The mean spectral value of images, which were filtered using monthly dates from January to April and the average of January to March the monthly variables were for assessing phenology’s effects on species diversity’s predictiveness. From the extracted spectral images of both sensors, vegetation indices (Normalised Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Simple Ratio (SR), and Enhanced Vegetation Index (EVI) in Eqs. (1:4)) were calculated within GEE.

$$NDVI = NIR - R/NIR+R \qquad 1$$

$$SAVI = ((NIR - R) / (NIR + R + L)) * (1 + L)) \qquad 2$$

$$SR = NIR/Red \qquad 3$$

$$EVI = G*((NIR-R)/(NIR+RED*R-C2*B+L) \qquad 4$$

All the spectral images and indices were imported into RStudio, and the GPS coordinates of each sampling plot were used to obtain values. This was achieved using the extract function from the raster library in RStudio (Table 1).

Table 1. Spectral configuration information of Sentinel-2 and Landsat 8 sensors.

Band Number	Band Name	Wavelengths	Band Abbreviation	Description
		Landsat 8 Collection 2 Level 2 -Surface Reflectance		
Band 1	SR_B1	0.435–0451 µm		Ultra-blue
Band 2	SR_B2	0.452–0512 µm		Blue
Band 3	SR_B3	0.533–0.590 µm		Green
Band 4	SR_B4	0.636–0.673 µm		Red
Band 5	SR_B5	0.851–0.879 µm	NIR	Near Infrared
Band 6	SR_B6	1.566–1.651 µm	SWIR	Shortwave infrared)
Band 7	SR_B7	2.107–2.294 µm		Shortwave infrared)
		Sentinel-2 MSI (S2A/S2B)		
Band 1	B1	443.9nm/442.3 nm		Aerosols
Band 2	B2	496.6nm/492.1 nm		Blue
Band 3	B3	560nm/559 nm		Green
Band 4	B4	664.5nm/665 nm		Red
Band 5	B5	703.9nm/703.8 nm		Red Edge 1
Band 6	B6	740.2nm/739.1 nm		Red Edge 2
Band 7	B7	782.5nm/779.7 nm		Red Edge 3
Band 8	B8	835.1nm/833 nm	NIR	Near Infrared
Band 9	B8A	864.8nm/864 nm		Red Edge 4
Band 10	B9	945nm/943.2 nm		Water vapor
Band 11	B11	1613.7nm/1610.4 nm		Shortwave infrared 1
Band 12	B12	2202.4nm/2185 nm		Shortwave infrared 2

Data analysis

A correlogram depicting Pearson correlations between species richness and diversity, S2 MSI bands, and vegetation indices including the averages of January-March, was computed to show relationships between explanatory and response variables. A machine learning approach: random forest enhances regression and classification trees by fusing a large number of decision trees. In this study, this technique was used to select the optimal variables that can be used to estimate species richness and diversity. Subsequently, a set of selected optimal variables were used as input into the RF model to predict species richness and diversity.

The RF model's *ntree*, which in this study was set to 500, had three parameters optimized; the numbers of regression bootstrap sample values are used to base trees on observations. The *mtry*, or the number of predictors to be tested at each node, for this investigation, a default value of four *mtry* was employed based on the square root of the number of predictors(Xulu *et al*., 2021).

To validate the performance of the random forest regression model, the datasets were split into 70% and 30% for training and testing respectively. A training dataset was used to develop an RF model that could estimate the response variables, while the test dataset was used to validate the final model. A one-to-one relationship between the measure and predicted species diversity and richness was fitted and the coefficient of determination (R^2) and root mean square error was used to assess the predictive performance.

Results

Figure 2 illustrates a correlogram of the relationship between Sentinel-2 MSI bands, indices, and diversity metrics (i.e., species richness and diversity). This relationship yielded a poor coefficient of determination. Similarly, the relationship between Landsat 8 spectral bands, indices, and diversity metrics yield was very weak (Figures 2 and 3).

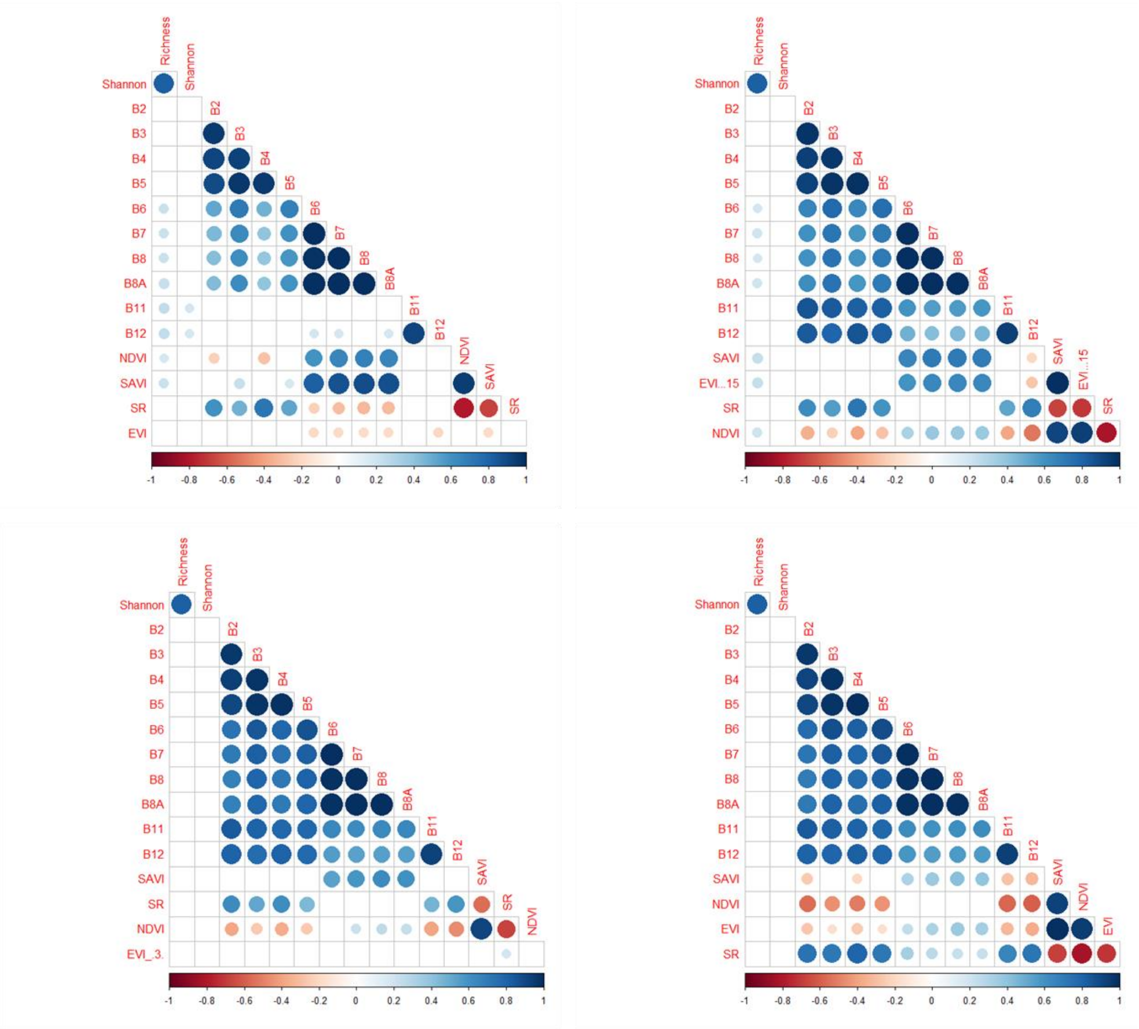

Figure 2. The correlation coefficients between species richness (SR) and diversity (SW) and vegetation were calculated from all possible combinations of Sentinel 2 Image in January (top-left), February (top-right), March (bottom-left), and January-March (bottom-right) ((blank = insignificant)

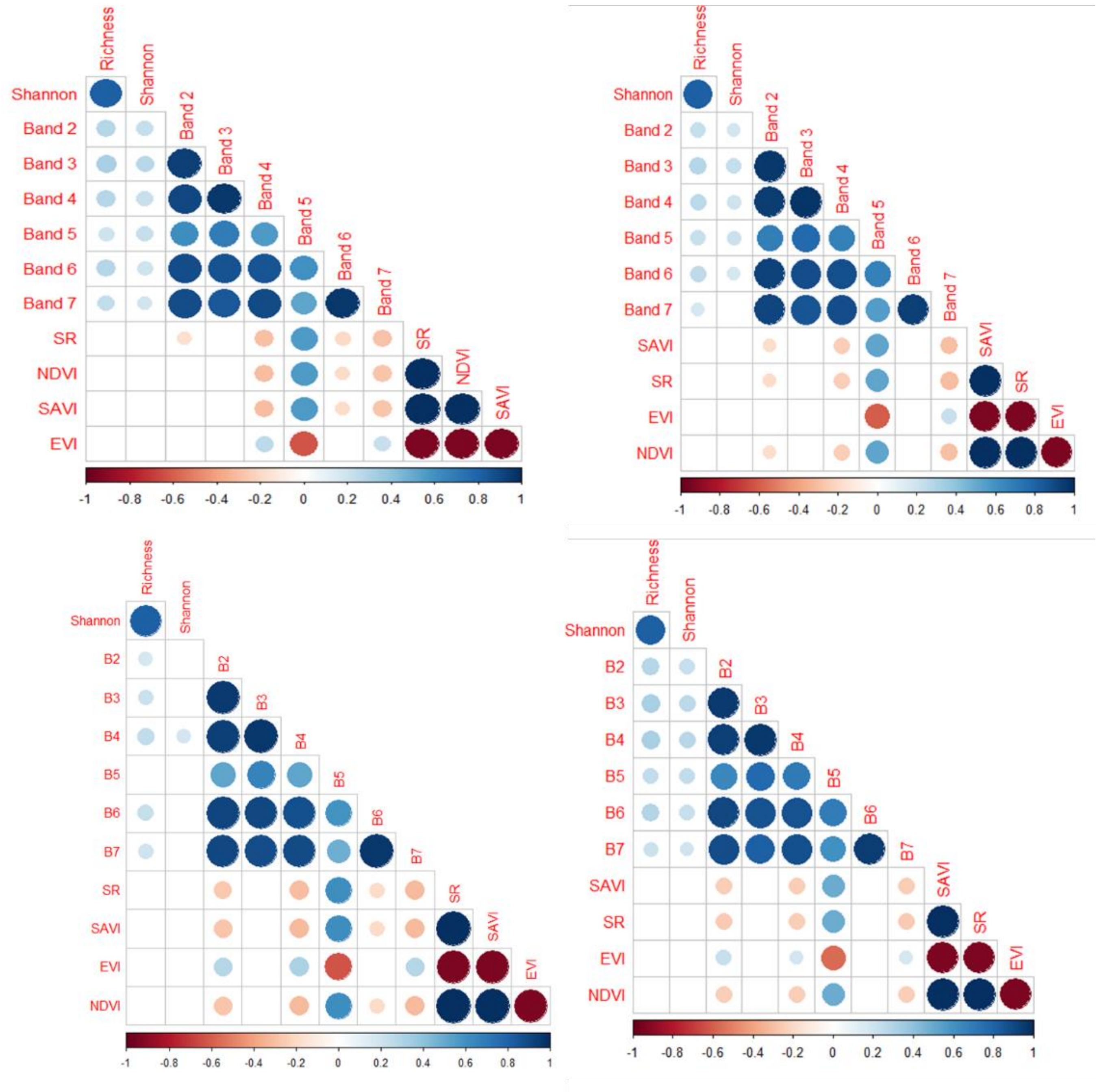

Figure 3. The correlation coefficients between species richness (SR) and diversity (SW) and vegetation were calculated from all possible combinations of Landsat 8 Image in January (top-left), February (top-right), March (bottom-left), and January-March (bottom-right) (blank=insignificant).

Random feature selection

The significant Landsat 8 variables (Table 2) optimally explaining species richness were EVI for January (RMSE=3.621, R^2=0.044), February (RMSE=3.493, R^2=0.080), and NIR for March (RMSE=4.126, R^2=0.008) and January–March (RMSE=3.935, R^2=0.006). Moreover, the selected Sentinel-2 variables explaining species richness were SR for January (RMSE=3.928, R^2=0.009), NIR for February (RMSE=3.872, R^2=0.001), Red edge for March (RMSE=3.711, R^2=0.001), and NIR for January–March (RMSE=3.657, R^2=0.014). The Landsat 8 variables explaining species diversity were EVI for January, and February, SAVI for March, and NIR for January–March. The Sentinel-2 variables for explaining species diversity were Red-Edge 1 for January, February, and January–March, and SR for March. In January, the selected band which was the most significant was EVI.

Table 2. Optimal Landsat 8 and Sentinel-2 variables explaining species richness and diversity

Landsat 8									
Species richness					Species diversity				
Selected Bands	RMSE	R squared	MAE		Month	RMSE	R squared	MAE	
January					Jan				
NIR	3.386	0.04008	3.063		NIR	0.4485	0.005913	0.3631	
SR	3.631	0.04269	3.015		SR	0.4367	0.001187	0.3547	
EVI	3.621	0.04362	3.019	*	EVI	0.4281	0.008509	0.349	*
February					Feb				
NIR	3.462	0.001955	3.228		NIR	0.4735	0.07716	0.3788	
SAVI	3.511	0.053272	2.949		SR	0.4666	0.06849	0.3768	
EVI	3.493	0.080437	2.859	*		0.4632	0.06849	0.3753	*
Mar					Mar				
NIR	4.126	0.008531	3.456	*	NIR	0.465	0.03137	0.3872	
SAVI	4.109	0.027952	3.445		SR	0.4521	0.0136	0.3735	*
EVI	4.087	0.028821	3.424		SAVI	0.458	0.03173	0.3782	
Jan-Mar					Jan-Mar				
NIR	3.935	0.006241	3.287	*	NIR	0.4244	0.0192	0.3519	*
SR	4.002	0.006201	3.445		SR	0.4343	0.01147	0.3563	
NDVI	4.01	0.011824	3.457		SAVI	0.4274	0.01579	0.3543	
Sentinel-2									
January					January				
Red-Edge 1	3.928	0.009438	3.34		Red-Edge 1	0.4366	0.000819	0.3656	*
Red-Edge 4	3.936	0.026493	3.369		Red-Edge 4	0.447	0.00168	0.3679	
SR	3.873	0.010987	3.319	*	SR	0.4404	0.0000055	0.3645	
February					February				
Red-Edge 1	3.872	0.001696	3.251	*	Red-Edge 1	0.4411	0.0000271	0.3464	*
Red-Edge 4	3.968	0.02039	3.374		Red-Edge 4	0.4504	0.0010034	0.3587	
SR	3.983	0.0293	3.43		SR	0.4428	0.0002043	0.3515	
March					March				
Red-Edge 1	3.733	0.001088	3.137		Red-Edge 1	0.4303	0.007374	0.3534	
Red-Edge 4	3.711	0.001685	3.074	*	Red-Edge 4	0.4286	0.006517	0.3502	
SWIR	3.77	0.000864	3.135		SR	0.4231	0.011342	0.3438	*
Jan-Mar					Jan-Mar				
Red-Edge 1	3.657	0.014519	3.076	*	Red-Edge 1	0.4325	0.0028	0.3593	*
Red-Edge 4	3.677	0.003288	3.079		Red-Edge 4	0.4392	0.00000309	0.3671	
SR	3.76	0.001014	3.146		SR	0.4366	0.000513	0.3631	

*=p<0.05

The random forest model of Landsat-8 variables explained 87% of the species diversity in January, 89% in February, 90% in March, and 79 % in January -March duration. While for species richness they explained 90% in January, 88% in February and 87% in March, and 81% in January-March. On the other hand, the variation of species richness and diversity explained by Sentinel-2 variables ranged between 82 and 91% (Table 3). The test exhibited similar predictive performance compared to the training dataset for most of the months.

Table 3. Random forest regression for predicted species richness and diversity, and remote sensing datasets.

	Species richness				Species diversity			
	Landsat 8		Sentinel-2- MSI		Landsat 8		Sentinel-2 - MSI	
	Training	Test	Training	Test	Training	Test	Training	Test
Number of plots	99	43	99	43	99	43	99	43
January								
R^2	0.90	0.88	0.90	0.80	0.87	0.86	0.85	0.86
RMSE	1.477	1.558	1.614	1.78	0.182	0.197	0.201	0.201
February								
R^2	0.88	0.87	0.84	0.92	0.89	0.90	0.89	0.89
RMSE	1.626	1.722	0.178	1.646	0.189	0.178	0.152	0.161
March								
R^2	0.87	0.88	0.91	0.91	0.90	0.88	0.91	0.91
RMSE	1.704	1.871	1.544	1.539	0.186	0.209	0.167	0.175
January-March								
R^2	0.81	0.88	0.88	0.85	0.79	0.85	0.82	0.83
RMSE	1.712	2.008	1.788	1.746	0.203	0.219	0.195	0.188

Discussion

The response of species richness and diversity often reflects diverse outcomes, which shows their varying sensitivity to external factors in an ecosystem and measuring tools. However, for these diversity metrics to be ideal, grassland indicator estimation methods ought to be improved and augmented (Symstad and Jonas, 2011). Machine learning algorithms provide an opportunity for exploring automated environmental monitoring using these diversity metrics and remote sensing data. Our research used the random forest approach using species richness and diversity plot measurements corresponding to remote sensing bands of Landsat 8, Sentinel-2, and vegetation indices derived thereof. The random forest feature selection of EVI, NIR, and SAVI remote sensing variables explained species richness and diversity, although with very low variations. This is not surprising given the moderate predictive performance of random forest machine learning algorithms on datasets with plant species richness (Adjorlolo and Botha, 2015). Rocchini, Ricotta and Chiarucci (2007) confirmed PCA as an effective tool for compressing multispectral data without losing information. Furthermore, the spectral band selected in this study was consistent with the findings of Rocchini *et al*., (2007), where near-infrared explained 41% of the variation in species richness.

The random forest model in this study managed to predict species richness and diversity with relatively high accuracies. This shows that using different model techniques other than the ones based on SVH can improve the predictive power of satellite remote sensing (SRS) variables (Rocchini *et al*., 2007a). Machine learning SRS models are important and imperative for species monitoring especially with the present biodiversity crisis concomitant with global environmental changes (Rocchini *et al.*, 2016). This research shows that a machine learning algorithm can improve the predictions of plant alpha diversity compared to the commonly used SVH, providing an alternative for species mapping. Moreover, the study shows that the machine learning models used for biodiversity mapping may not require satellite sensors with high spatial and spectral resolution.

Both multispectral (presented in this study) and hyperspectral data present opportunities and limitations in understanding relationships between spectral indices and species richness. Hyperspectral data are cumbersome with high multicollinearity. Although this research sought to overcome these challenges using ML random forest, the variables from the remote sensing data did not improve the predictability of species diversity. Indeed, random forest exhibited the least predictive accuracy when compared to other machine learning algorithms in investigating species (Richter *et al.*, 2016). This research ascertains that most of the variation in species richness and diversity could be explained by remote sensing variables and remote sensing predictors produced in this study could be used to identify hotspots. However, remote sensing models for estimating plant diversity variables must still be refined to confirm this assertion. These models are dependent on the characteristics of remote sensors, statistical algorithms, and the scale of the investigation. As such, high-resolution data at spectral and temporal levels from remote sensing data may also be explored for these estimations (Rocchini *et al.*, 2007).

A wide range of relationships between satellite-based VIs and vegetation characteristics have been established, with low to moderate predictive performance (Haboudane, 2004). Our research corroborates studies that propose using VIs in addition to spectral bands to improve the predictiveness of species diversity metrics (Rocchini et al., 2016). This is mainly because when used alone, VIs yield poor predictive performance and correlations, despite the enhanced spatial resolution of a satellite sensor; hence, our study elucidates the use of machine learning and relevant spectral bands, and VIs improves the estimation of species richness and diversity. Even so, the differences in predictive accuracies between Sentinel-2 and Landsat 8 sensor was not substantial; this is not surprising since the two sensors are not too distinct spectrally and spatially. This finding aligns with studies that postulate multispectral sensors with relatively moderate to high spatial resolution could be good candidates for biodiversity mapping (Rocchini et al., 2016).

The relationship between spectral bands, indices and species diversity indices somewhat depends on the correct measure of species diversity metric. However, our study demonstrated little to no effect in using different diversity metrics; opposing studies suggest that using the Shannon-Wiener index improves predictive performance three-fold (Oldeland et al., 2010). Nonetheless, the prediction accuracies were highest in both January (richness) and March (diversity) for Landsat and March (richness and diversity for Sentinel. The estimation accuracies increased with deteriorating phenology for Sentinel, rendering it beneficial for mapping grass species diversity in senescence. This is because of its strategically positioned spectral bands, especially the inclusion of red-edge bands, making it helpful in studying vegetation characteristics (Thenkabail et al., 2004).

This research explored satellite remote sensing as a primary tool for identifying biodiversity hotspots in South Africa's mountainous grasslands and predicting changes. The random forest (RF) remote sensing model predicted species richness and diversity with relatively high accuracy. These models present an opportunity for plant species monitoring using remote sensing, which has always been associated with many challenges concerning species diversity monitoring (Rocchini et al., 2016). Remote sensing is a cost-effective and less labour-intensive tool for biodiversity management, and its development is imperative for monitoring the inevitable consequences of global environmental change. Previously, species richness at local scales was studied using the spectral variation hypotheses (SVH). However, the random forest models in this study provide better estimates of plant species richness than the proposed SVH (Rocchini et al., 2016, 2018). In contrast to SVH, the RF predictive models do not require remote sensing sensors with high spatial and spectral resolution. SVH starts with a heterogeneity map correlated with field sampling data for estimation models.

Conclusion

Due to minute variations in spectral signature measures, it may be challenging to distinguish between species in tropical environments when attempting to build correlations between remote sensing data and species.. Remote sensing approaches have always been limited by the type of sensor and modeling algorithms. Near-infrared (NIR) as the selected spectral band for predicting species richness and diversity circumvents these limitations. This selection also augments NIR as the spectral band that can allow for species discrimination as they are related to species traits, especially chlorophyll which can also be measured using NIR-based vegetation indices. This method deviates from the traditional use of the spectral variation hypothesis which is based on the hypothesis that each species or group of species can be identified by their spectral signature. However, remote sensing models derived from SVH to estimate species diversity could improve by testing the spectral variability of NIR with species. Furthermore, Remote sensing images with high spectral and spatial resolution, special unmanned aerial vehicles, could improve biodiversity mapping in complex environments., especially in GGHNP where field surveys in the mountainous terrain are difficult.

Chapter 6

Grass species nutrient estimation using satelite remote sensing in protected mountainous grasslands

This chapter is based on:

Mashiane, K.K. Ramoelo, A. Adelabu, S.A. ‘Grass species nutrient estimation using satellite remote sensing in protected mountainous grasslands’, *Geocarto International* (**submitted**).

Abstract

The role of biodiversity to improve the primary productivity within terrestrial ecosystems is well documented. Each species in an ecosystem has a role in the overall productivity of an ecosystem. Grass species nitrogen (grass N) estimation is an important task in rangelands because of the forage quality which has nutritional implications on the grazing animals. This study aimed to use a random forest machine learning algorithm to identify the optimal variables that can accurately predict grass N from three different datasets derived from Sentinel 2 Multispectral instrument (S2 MSI) i.e., bands only, vegetation indices of optimal variable, and combined bands and vegetation indices. The selected optimal variables were used as input in the random forest model to predict grass N. The red edge bands from the S2 MSI were selected as the most important bands for optimally predicting grass N. This research ascertains the role of the red edge position in investigating foliar N and changes thereof. The prediction of grass N improved slightly in the vegetation indices model compared to the bands model and the highest prediction was achieved in a model combining the two. This ascertains that the inclusion of red-edge vegetation indices improves the estimation of foliar N. The multiple red-edge bands of S2 MSI provide an opportunity for the estimation of biochemical concentrations of grass N across their phenology and during specific periods of their life cycle with improved accuracies. S2 MSI remains the ideal remote sensing tool for estimating grass N because of its strategically positioned red-edge bands, which are highly correlated with chlorophyll content in plants.

Introduction

The effect of biodiversity loss in accelerating changes in ecosystem processes is now evident. Thus, globally, biodiversity loss is regarded as a major driver of ecosystem change. Moreover, the sustainability and productivity of the earth's ecosystem are largely dependent on biodiversity conditions. In rangelands plant communities, there is an indication that species loss will reduce plant production and alter decomposition thereby affecting the carrying capacity of grazing animals (Tilman *et al*., 2012). These impacts on rangeland biodiversity may impair the efficiency of rangeland vegetation in capturing essential resources, producing biomass, and recycling essential nutrients (Cardinale *et al.*, 2012). Consequentially, some species may be lost and/or reduced in biomass and/or diversity, affecting grazing animals' nutrition. Therefore, it is imperative to consider the nutritional value of multi- and single species cover of grasses, as research on predictive dynamic grazing carrying capacity of rangelands becomes pressing in the face of global change.

The distribution of grazing mammals in grasslands is largely attributed to the occurrence of nutritionally enriched vegetation species (McNaughton, 1988). Ascertaining the relationships between species diversity and nutrient levels will improve the determination of stocking rates in conservation, especially in the advent of ecosystem changes. Remote sensing data is constantly used in conjunction with field data at multiple scales to estimate rangeland indicators (McCord *et al.*, 2017b). Infact, modelling approaches using data from different satellite imagery were used to predict vegetation and foliar cover as indicators of rangelands (McCord *et al.*, 2017a), further cementing the role of remote sensing data in producing spatially explicit and continuous surface estimates of rangeland indicators. Ramoelo *et al.*, (2015) showed that vegetation indices derived from high spatial resolution remote sensing data explained leaf nitrogen in Southern African rangeland. The performance of taxa-specific nitrogen estimation is useful for rangeland monitoring, such as being used for the identification of grazing lawns and the distribution of grazing mammals and as a result, improved carrying capacity models (McNaughton, 1988).

In African protected areas, nutrient deposition by large grazing mammals is pivotal for enhancing carrying capacity (McNaughton, 1985, 1990; McNaughton *et al*., 1997). Conversely, excessive grazing can reduce mineralization rates via changes in vegetation dynamics, i.e., species composition, cover, and diversity. Thus, affecting forage quality and quantity (Stock *et al*., 2010). Nonetheless, studies have indicated the mutual effect of grazing on grassland biodiversity and productivity (McNaughton, 1985a). The distribution and abundance of large grazing mammals in Southern African rangeland are strongly influenced by nutritionally sufficient forages and vegetation dynamics (McNaughton *et al.,* 1997). For example, these animals prefer to forage on swards enriched in minerals, which is pivotal for their nutrition and reproduction (McNaughton, 1990). The determination of habitat conditions, usage (Ungar, 2019), and mineral nutrients (McNaughton, 1990) has important implications on stocking rates and carrying capacity

Estimations of foliar nitrogen have the potential to provide insight into animal feeding patterns and distribution (Mutanga and Skidmore, 2007; Chabalala *et al.*, 2020). Remote sensing sensors with red edge bands can benefit the accuracy of mapping vegetation biochemical concentrations (Mutanga and Skidmore, 2007). Sentinel-2 MSI (S2 MSI) has become a highly sought-after tool for vegetation mapping because of free access to datasets and its superior spectral quality (Chabalala *et al.*, 2020; Xulu*et al*., 2021), especially the red edge band which is better suited for chlorophyll estimations (Mutanga and Skidmore, 2007). The launch of Sentinel-2 with the inclusion of red edge present an opportunity to validate Nitrogen models of coarser remote sensing sensor against high resolution sensors (Ramoelo *et al.*, 2015a). Nonetheless, using machine learning, both Landsat 8 and Sentinel-2 achieved comparable superior accuracies in forest mapping (Xulu *et al*., 2021). The objectives of this study were to 1) determine optimal variables for the predictions of grass N using a dataset derived from S2 MSI, 2) measure the contribution of each variable to the prediction model, and 3) test the predictive performance of random forest remote sensing models in predicting grass N.

Methodology

Study area

The study was conducted in the Golden Gate Highlands National Park (GHNP) in the north-eastern Free State province, South Africa (Figure 1). The park is 32,758.35 ha in size and is located between 28°27' S and 28°37' S and 28°33' E and 28°42' E (SANParks, 2020). The park is in mountainous grasslands at the foothills of the Drakensberg and forms part of the mesic highveld grassland with marked variation in geology, topography, and rainfall. The soil types in the park include shallow rocky soils (Glenrosa and Mizpah), deep soil along drainage lines (Oakleaf), well-developed sand soils (Hutton and Clovelly), and clayey structured soils (Milkwood and Tambakulu) (SANParks, 2020). The park is characterized by summer rainfall, temperate summers, and cold winters. The rainfall season stretches from September to April with a mean annual rainfall ranging from 800 mm to 2,000 mm (Kay *et al.*, 1993). The park lies between 1,892 m and 2,829 m above sea level and comprises the grasslands units: Eastern Free State sandy grasslands (Gm 4), Basotho montane shrubland (Gm 5), Lesotho highveld basalt grassland (Gd 8), and Northern Drakensberg highveld (Gd5) (Mucina and Rutherford, 2006).

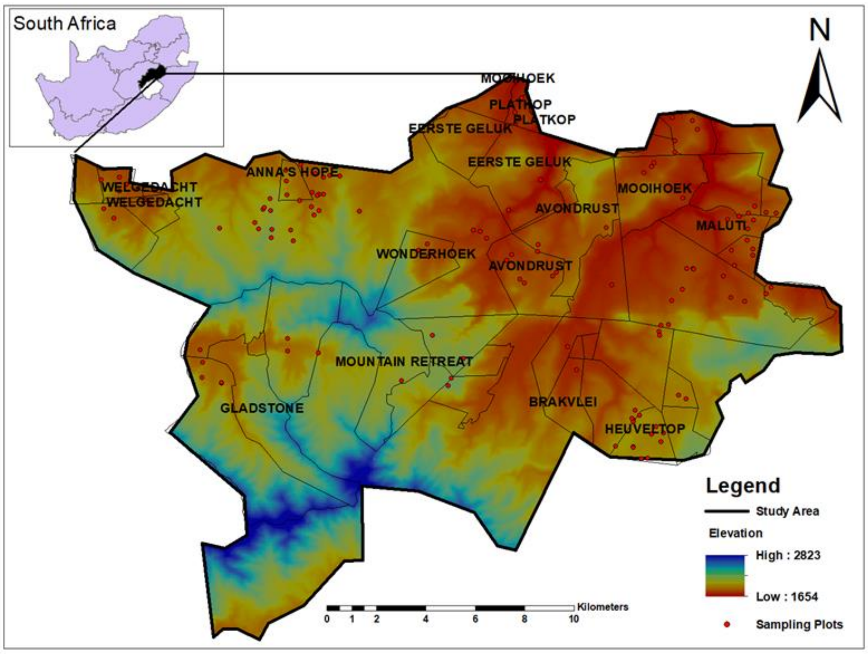

Figure 1: A map of the study area

Data collection

Satellite images were downloaded and processed from the Javas Script code editor Google Earth Engine (GEE). Sentinel satellites are ideal for vegetation monitoring as it provides high-resolution images with a global 5-day revisit frequency. Most importantly Sentinel 2 has spectral bands comparable to those satellites with high spatial resolution. The addition of red edge image values of the image, for this study, the band's images of S2 MSI were filtered using monitoring dates average of January and March

The land type of map of Golden gate was used as the first level of stratification, sampling sites of relatively homogenous grasses were located randomly. A 30m × 30m plot of homogenous grasses was defined within which, 16 subplots of 1m × 1m were systematically placed in each species cover and composition was collected, and the dominant species identified based on the highest cover in all or most of the subplots. Subsequently, grass samples on the dominant grass species were collected from each of the plots. The sample was dried (80°C in 24 hours) and thereafter taken to the laboratory for LECO chemical analysis (Sweeney and Rexroad, 1987) to determine each species' grass Nitrogen (grass N).

Data Analysis

The spectral reflectance of Sentinel-2 band images (Table 1) for the average January-March period was extracted corresponding to each sampling GPS point with grass N value. A random forest (RF) modelling algorithm based on three modelling scenarios (bands only, vegetation indices, and combination bands and vegetation indices) were used to predict the grass N concentrations from the remote sensing spectral bands and the vegetation indices. Random forest recursive feature selection was used to determine which variable from S2 MSI's spectral bands can optimally predict N. The selected band from the random forest feature selection was used to compute vegetation indices (Table 2) in GEE. In total, 18 predictor variables were used for data analysis in this study, 10 remote sensing bands, and 8 vegetation indices were chosen for their strong correlation with foliar N. Subsequently, the optimal vegetation indices and combination of bands and vegetation indices that can predict N were identified using recursive feature elimination (RFE) in the “caret” package of RStudio. Thus, the optimal variables from the selected RFE were then used as input into the RF models to predict grass N under the three modelling scenarios implemented in R statistical programming language using a random forest package (R Core Team, 2020).

Table 3: Sentinel 2 – MSI Datasets

Band Number	Band Name	Wavelength	Description
Band 1	B1	443.9nm/442.3nm	Aerosols
Band 2	B2	496.6nm/492.1nm	Blue
Band 3	B3	560nm/559nm	Green
Band 4	B4	664.5nm/665nm	Red
Band 5	B5	703.9nm/703.8nm	Red Edge 1
Band 6	B6	740.2nm/739.1nm	Red Edge 2
Band 7	B7	782.5nm/779.7nm	Red Edge 3
Band 8	B8	835.1nm/833nm	Near Infrared
Band 9	B8A	864.8nm/864nm	Red Edge 4
Band 10	B9	945nm/943.2nm	Water vapor
Band 11	B11	1613.7nm/1610.4nm	Shortwave infrared 1
Band 12	B12	2202.4nm/2185nm	Shortwave infrared 2

Table 2: Vegetation indices used in this study

Index	Used formulae	Reference
MCARIR4	((NIR-Red edge 4)-0.2*(NIR-Red edge 4)) * (NIR/Red edge 4)	(Daughtry, 2000)
MSAVIR4	0.5*(2*NIR+1-SQRT((2*NIR+1)-8(NIR-Red edge 4)))	(Qi *et al.*, 1994)
NDVIR4	(NIR-Red edge 4)/(NIR+Red edge 4)	(Rouse *et al.*, 1974)
OSAVIR4	(1+0.6)*(NIR-Red edge 4)/(NIR-Red edge 4+0.16)	(Rondeaux, Steven and Baret, 1996)
RDVIR4	(NIE-Red edge 4)/SQRT(NIR+Red edge 4)	(Daughtry, 2000)
SAVIR4	2.5*NIR-Red edge 4)/((NIR+Red edge 4)+2)	(Huete, 1997)
SR4	NIR/Red edge 4	(Jordan, 1969)
TCARIR4	3*(NIR-Red edge 4) - 0.2*(NIR-Rededge4)*(NIR/Red edge 4))	(Haboudane, 2004)
TVIR4	0.5*(120*(NIR-Red edge 4)-200*(NIR-Red edge 4))	(Broge and Leblanc, 2001)

MCARI = modified chlorophyll absorption ratio index, MSAVI = modified soil adjusted vegetation index, TVI = triangular vegetation index, NDVI = normalized difference vegetation, OSAVI = soil adjusted vegetation index, RDVI = renormalized difference vegetation index, SAVI = soil adjusted vegetation index, SR = simple ratio, TCARI = transformed chlorophyll absorption ratio, TVI = triangular vegetation index.

Results

Figure 2 illustrates a correlogram of correlation coefficients between remote sensing bands, red-edge 4 vegetation indices, and grass N. The relationship between the vegetation indices and remote sensing variables yielded a poor coefficient of determination and was non-significant, excluding TVIR4 (Figure 2).

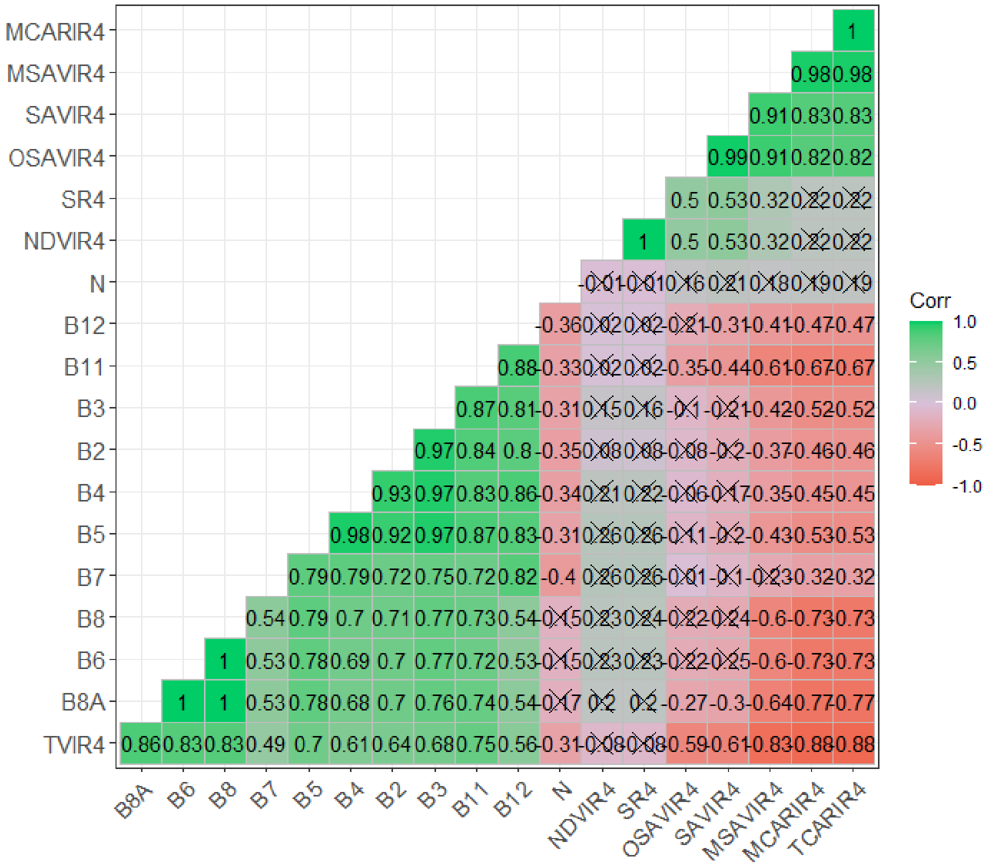

Figure 2: Correlogram of red edge-based vegetation indices, grass N, and remote sensing variables (the crossed-out cells show their insignificance).

Table 3 shows the descriptive statistics of the measured grass N in the park. The grass N concentration averaged 0.89 and with a variability of 29%. Grass N ranged between 0.45 and 1.71% throughout the park (Table *3*).

Table 3: Descriptive statistics: grass Nitrogen (%).

Datasets	*N*	Min	Max	Mean	SD	CV (%)
Training	35	0.58	1.6	0.86	0.22	25
Test	15	0.44	1.71	0.93	0.34	36
All combined	50	0.45	1.71	0.89	0.25	29

The red and red-edge bands were among the selected bands optimally explaining grass N in the bands-. In the bands-only dataset, red-edge 4 (B8A) was the select band variable for optimally explaining grass N (RMSE=0.22, R^2=0.47), along with SWIR (B11) and Red band (B4). TVIR4 and NDVIR4 were the selected variables optimally explaining grass in the red-edge-based vegetation indices dataset. Similarly, TVIR4 was also the selected band optimally explaining grass N in the dataset combining bands and vegetation indices, in addition, the selected variables were also red (B4) red-edged 4 and SR4 (Table 4).

Table 4: The optimal variable(s) selected for estimation of grass N.

Selected variables	R-Squared	RMSE	MAE	$p<0.05$
Bands only				
Red	0.4223	0.2272	0.1727	
Red Edge 4	0.4794	0.2154	0.1672	*
SWIR	0.4274	0.2249	0.1721	
Red-Edge 4 Indices				
NDVIR4	0.00025	0.3184	0.2347	
TVIR4	0.0000482	0.3156	0.2284	*
Band + Vegetation Indices				
Red	0.02986	0.2655	0.205	
Red-Edge 4	0.10306	0.2473	0.1879	
SR4	0.14878	0.2396	0.1842	
TVIR4	0.15517	0.2387	0.1813	*

The variable that contributed the most towards explaining grass N was red-edge 4 (B8A) according to the random forest variable of importance score (Figure 3), in the bands-only dataset. In the vegetation indices-only dataset, TCARIR4 and MCARIR4 were the top two variables with the highest score of variable importance. All red-edge bands featured as contributing the most towards predicting grass N in a dataset combining bands and vegetation indices, however, the latter contributed the least and red-edge 1, 2, and 4 were the top three variables with the highest score of a variable of importance (Figure 3).

The random forest model explained 80 % of the grass N in the bands-only model. For the vegetation indices model, a prediction of 81% was achieved. The highest prediction was achieved in the model combining S2 and vegetation indices (Figure 4).

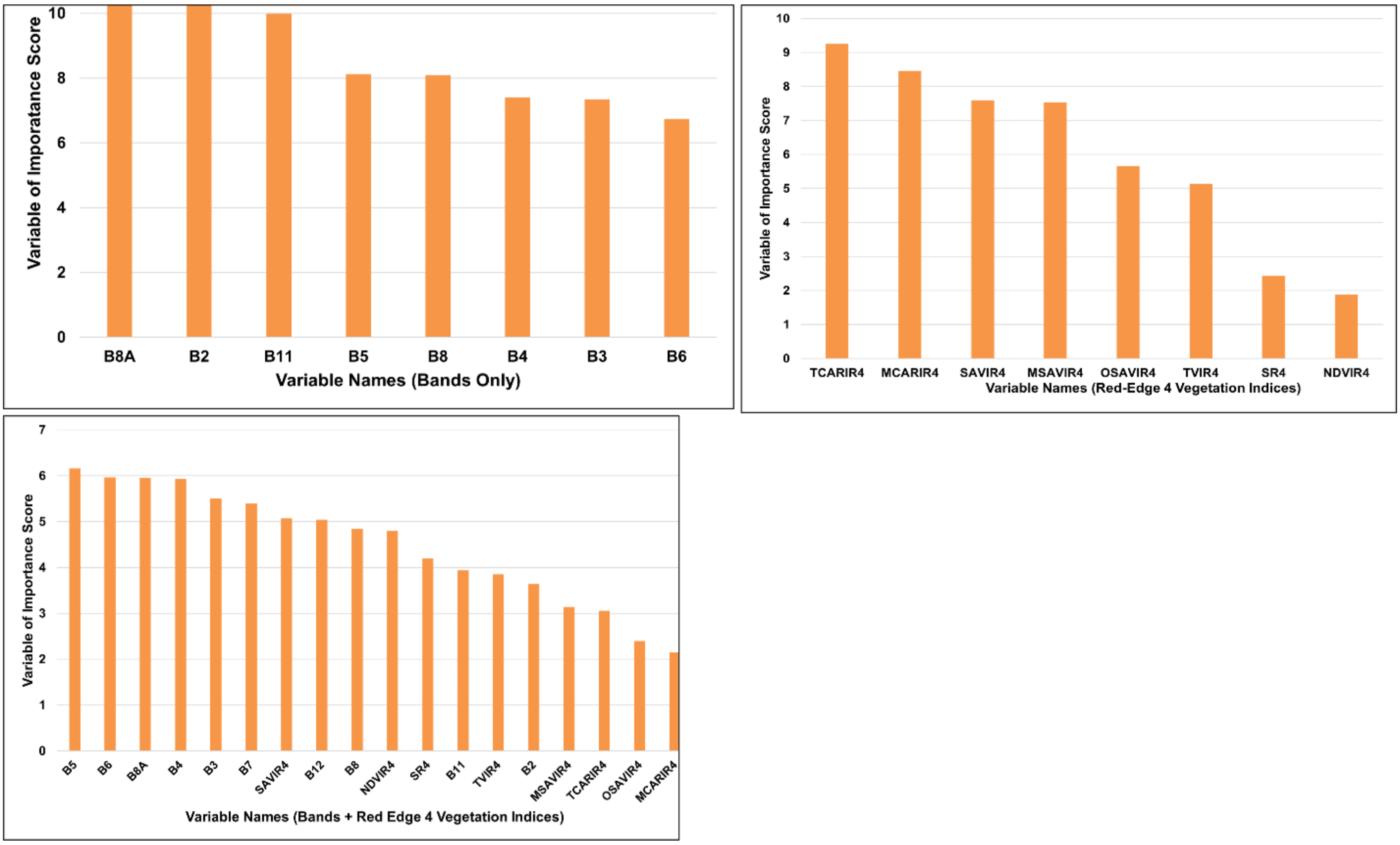

Figure 3: Measuring variables of importance for estimating grass N; top left (Bands), top right (Vegetation indices), and bottom left (a combination of bands and vegetation indices. The higher the variable of importance score the more important the variable is for estimating grass N.

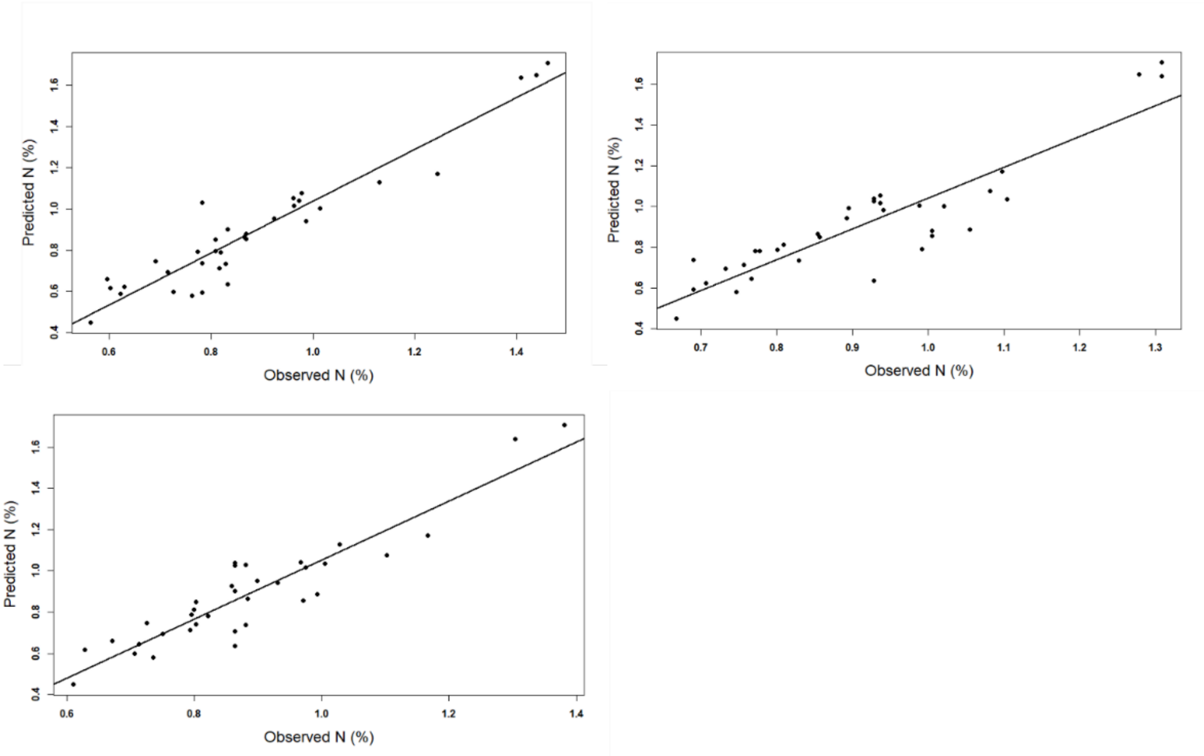

Figure 4: Grass N estimation performance of grass for various modeling scenarios: top left: Bands (80 %), top right: Vegetation indices (81 %), and bottom left (85%).

Discussion

In this study, grass N averaged between 0.45 and 0.71. Similarly, research shows that grass N is affected by environmental conditions and seasonality (Ramoelo *et al.*, 2015; 2018), but it is also higher during the plant's peak productivity (Ramoelo *et al.*, 2015). The correlation coefficients between grass N and remote sensing bands were moderate and significant, except for the vegetation indices which were not significant. Indeed, studies show that leaf N correlates more with spectral bands, especially those in the red edge position (Cho *et al.*, 2007; Ramoelo *et al.*, 2015a). This may be because microwave technologies provide information on the structural properties of plants, but multispectral optical remote sensing bands provide additional estimations of vegetation characteristics (Xie *et al.*, 2008). This study showed a poor correlation between vegetation indices derived from broad-based sensors and leaf N, attributed to signal saturation (Cho *et al.*, 2007), phenology, and seasonality (Ramoelo *et al.*, 2012b). Leaf N and biomass have been demonstrated to be effectively estimated by red edge-based vegetation indices using both hyperspectral (Cho *et al.*, 2007) and multispectral remote sensing (Ramoelo *et al.*, 2015; 2018).

Sentinel -2 multispectral sensor has become a very useful remote sensing tool because it is freely-available and includes the red edge bands (Xulu *et al.*, 2021) which are related to chlorophyll in plants (Cho and Skidmore, 2006). It is not perplexing, therefore, that the red and red edge bands are among the most selected bands in this study because they are related to chlorophyll which is related to leaf N (Cho and Skidmore, 2006). This research shows that Sentinel-2 can provide important information about vegetation spectra with results comparable to commercial hyperspectral sensors. This study shows that models of predicting leaf N of grasses using higher resolution sensors could be calibrated and used for data from coarser sensors as they provide similar information on vegetation characteristics. Furthermore, this research is indicative that rangeland monitoring with sensors that have rangeland capability is possible (Cho and Skidmore, 2006). Indeed, Sentinel-2 performed considerably well predicting grass N compared to Rapid Eye, which has a higher spatial resolution sensor (Chabalala *et al.*, 2020).

Notably, the red edge 4 was the band selected for optimally estimating grass N in this study, this is indicative of the vegetation stress observed during the sampling period as it averages grass phenology. The relationship between red-edge position (REP) and foliar N depends on Nitrogen and chlorophyll (Cho and Skidmore, 2006). Thus, shifts in the red edge position are a good indicator of changes in foliar N and water stress, because changes in the REP is mainly attributed to the chlorophyll content which peaks and decreases during the wet and dry season respectively (Clevers *et al.*, 2002; Cho and Skidmore, 2006). This study shows that insights into vegetation senescence are possible using Sentinel-2. This has a major implication on rangeland management because this is where plants lose their primary productivity rate and in turn affect grazing animals and, as a result, stocking rates and carrying capacity. Hence, Ramoelo *et al.*, (2015) confirmed that leaf N can be estimated in the dry season with good accuracy, using hyperspectral data. This study shows that the multispectral sensors with the REP can achieve similar estimates.

The prediction of grass N improved slightly in the vegetation indices model compared to the bands model and the highest prediction was achieved in a model combining the two. This ascertains that the inclusion of red-edge vegetation indices improves the estimation of foliar N (Ramoelo *et al.*, 2012b). Ramoelo *et al.*, (2015b) found that the prediction accuracies of grass N by remote sensing variables decrease as they senesce. However, the multiple red-edge bands of Sentinel 2 MSI provide an opportunity for the estimation of biochemical concentrations of grass N across their phenology with improved accuracies. Notwithstanding, the TVIR4 bands was the band selected for optimally predicting grass N in this study, however, TCARIR4 and MCARIR4 had the highest variable of importance. In univariate modelling techniques, the above vegetation indices performed poorly in the prediction of foliar N (Ramoelo *et al.*, 2012b). The significance of using machine learning approaches to explore the range of vegetation indicators obtained from several red-edge bands, particularly of S2 MSI, is highlighted by this study.

Conclusion

This study sought to investigate the N concentration of grass species across the landscape The study was limited in that the grass samples were not collected monthly and averaged. However, grass N was better predicted by a combination of remote sensing data from the multispectral Sentinel 2 and vegetation indices. The S2 MSI remains the ideal remote sensing tool for the estimation of foliar N as it incorporates the strategically positioned red edge bands. Furthermore, red edge-based vegetation indices have been reported to provide better estimates in other studies, compared to traditional vegetation indices. Further research is still pending on utilizing the remote sensing variables of S2 MSI during the different seasons in the mountainous grasslands, especially by incorporating topographic variables into the models in such environments.

Chapter 7

Synthesis and Conclusions

The loss of biodiversity due to global environmental changes is fast becoming evident and will result in further accelerated changes in the ecosystem processes. For instance, local species extinction may limit ecosystem variety and spatial uniformity of plant life (Hautier *et al.*, 2018). There is broad consensus that greater diversity of species at small sizes can enhance ecosystem functioning and benefits (Cardinale *et al.*, 2012). This is mainly because different species contribute to different ecosystem functioning. This study showed (Chapter 2) the variation of plant species diversity at local and landscape scales, hence, this information contributes to the understanding of ecosystem functioning and services in mountainous grasslands.

This research focused on the premise that species composition and diversity can influence ecological function and services, including the movement of organisms and resources in grasslands. The primary aim of chapters 1, 2, and 3 was to identify and explain plant species diversity in mountainous grasslands to contribute to the knowledge of grazing patterns and most importantly refinement of carrying capacity models. Knowledge about species within and among plant communities is imperative for preserving ecosystem functioning and conservation of biodiversity (Cardinale *et al.*, 2012). As such, one of the most important ecosystem services in grasslands, especially GGHNP, is grazing for both domestic and wild animals. The occurrence and distribution of animals are attributed to the occurrence of nutritionally enriched grass species (McNaughton, 1988). This study provided insights into the role of diversity in the distribution of grazing levels.

Grasslands are renowned for having a high diversity of plant species at small geographical scales (Franzén and Eriksson, 2008). In the South African grasslands of the Drakensberg, it is estimated that between 9 and 29 species have been observed at 100 m^2 plots (Brown and Bezuidenhout, 2020). In this research, a much higher gamma diversity of between 30 and 51 was observed per site, this shows that GGHNP is an important part of the grassland biome in South Africa. The pattern of high species richness and diversity is explained by many factors such as fire and grazing in grasslands, however, the lack of nutrient-enriched soil could result in lower productivity and enhance competition which could lead to spatial homogenization by precluding other species (Franzén and Eriksson, 2008). This research shows that species pool may be used to identify plant species richness patterns, the species pool theory shows that species richness at a larger scale is dependent on species richness at a lower scale. The postulations of the metapopulations theories within the mountainous grasslands, especially in topographic conservation areas ought to be studied, patterns of species composition across topographic gradients would serve as an important contribution to the understanding of the mountainous grassland systems.

This study forms an important base for developing a framework for biodiversity monitoring and conservation thereof. Measurement of changes in local and landscape diversity is essential for nature conservation, especially for an average location where there is scant biodiversity data (Cardinale *et al.*, 2018b). This is important for grassland conservation areas because biodiversity loss resulting from global environmental change will likely affect rangeland quality and animal nutrition (van *et al.*, 2018). Biodiversity plays an important role in grasslands, however, the role of grassland biodiversity on grassland productivity and animal nutrition still must be investigated, for a different site in GGHNP. Data and information derived from this research are intended to form the bases on which plant species diversity models can be used to explain animal nutrition and thus, refining carrying capacity models for rangelands in conservation areas.

The drivers of plant species richness and diversity.

There is compelling evidence that shows topography is an important driver of vascular plant species diversity and can preclude and/or enhance the existence of some species (Moeslund *et al.*, 2013). This study highlights the importance of topographically controlled factors in influencing plant species diversity at small scales. This is known as microtopography, which is important for understanding the productivity and geomorphology of ecological landscapes (Moeslund *et al.*, 2013; Lundholm, 2009). Soil texture and drainage are important and play an important role in the establishment of seedlings in grassland communities. The variety of plant species in the mountainous grasslands of the Drakensberg is often ascribed to unique ecological conditions created by the complex topographic landscapes (Brown and du Preez, 2020). I recommend further investigation on the role of topography and its impact on species composition and changes thereof along topographic gradients in mountainous areas, this allows for topography to be used as a proxy for species diversity (Lee and Chun, 2016) and can also contribute the understanding of animals foraging patterns and quality.

The machine learning algorithm for modelling species richness and diversity in this study is indicative that taking advantage of novel modelling techniques can augment research about the mechanisms underlaying the relationship between topography and local vegetation patterns. This research sought to eliminate the challenges of assessing species distribution models' uncertainty and provide a framework for selecting relevant modelling variables (Elith and Leathwick, 2009). In areas where conservation resources are scarce, selecting relevant predictors of species richness and diversity is important for setting precise conservation priorities (Olea *et al.*, 2010). Most of the modelling techniques for predicting species diversity suffer from multicollinearity, this research shows that the use of non-parametric machine learning techniques can circumvent this challenge.

The remote sensing of plant species diversity and Nitrogen in grasslands

This research explored satellite remote sensing as a primary tool for identifying biodiversity hotspots in mountainous grasslands of South Africa and predicting changes thereof. Random forest (RF) remote sensing model predicted species richness and diversity with relatively high accuracies. These models present an opportunity for plant species monitoring using remote sensing which has always been associated with many challenges when it comes to species diversity monitoring (Rocchini *et al.*, 2016). Remote sensing is a cost-effective and less labour-intensive tool for biodiversity management and its development is imperative for monitoring the inevitable consequences of global environmental change. Previously, species richness at local scales was studied using the spectral variation hypotheses (SVH) (Rocchini *et al.*, 2007b), however, the random forest models in this study provide better estimates of plant species richness than the proposed SVH. In contrast to SVH, the RF models do not require remote sensing sensors with high spatial and spectral resolution. SVH starts with a heterogeneity map that is correlated with field sampling data for estimation models.

Remote sensing is a superior approach for identifying biodiversity hotspots and monitoring thereof, in contrast to *in situ* approaches, because it allows for complete coverage of the earth's surface and it can record short to long-term data (Lausch *et al.*, 2018). However, the remote sensing approaches for estimating vegetation diversity are always limited by the characteristics of the sensor and the algorithm utilised for modelling. Near infrared portion of the electromagnetic spectrum was selected as the most important variable for mapping species diversity, in addition to vegetation. This augments the use of both NIR and vegetation indices in predicting species richness as they are associated with species traits, especially chlorophyll (Rocchini *et al.*, 2016). Although hyperspectral techniques are seen as superior for discriminating species, this research shows that NIR and its vegetation indices can be used to improve the delineation of species in mountainous grasslands. Nonetheless, the use of the hyper, multi and ultra-spectral sensors in predicting species diversity ought to be compared in their performance against one another at various spatial scales. Vegetation indices derived from NIR ought to be further investigated to be used as a proxy for species richness at all spatial scales, especially at the site-specific level.

This research deviates from the traditional use of SVH which is reliant on the characteristics of the plant taxa such as their chemical, biochemical, biophysical, and physiological traits. The assumption with SVH is that each species or group thereof can be identified based on its spectral signature (Rocchini, Ricotta, and Chiarucci, 2007b). However, the spectral variation hypothesis is dependent on a combination of high spatial and spectral resolution sensors which have been lacking for validation of remote sensing of species diversity. For this reason, vegetation indices and NIR remain paramount for studying species diversity (Rocchini, Ricotta, and Chiarucci, 2007b), these remote sensing variables do not necessarily require high spatial and spectral remote sensing sensors to be efficient. However, a possibility exists for aerial and airborne sensors, especially unmanned automated vehicles also known as drones or UAVs, to improve the relationship between species diversity and spectral heterogeneity in mountainous grassland with difficult terrain for vegetation sampling. Research of these relationships at local, regional and even global scales is still outstanding and thus warranted. The selection of an appropriate remote sensing sensor may be beneficial for gaining insights into these relationships of remote sensing of plant species richness.

In African grasslands, the distribution and feeding patterns of grazing animals have largely been related to nutrients enriched vegetation (McNaughton, 1988), which are further controlled by the diversity of species in plant communities (David Tilman, Wedin, and Knops, 1996). Remote sensing has been shown as an important tool for the estimation of grass N concentration for vast landscapes such as rangelands (Ramoelo *et al.*, 2015b). This research is mainly reliant on the use of narrow bands, especially red-edge which is affected by leaf biochemical characteristics (Mutanga and Skidmore, 2007). In fact, in this study, the red-edge bands of Sentinel-2 were selected as the most optimal variable for the estimation of grass N. The vegetation indices derived thereof, performed well when used for modelling grass N. This show that the combination of machine learning techniques and Sentinel -2 multispectral instruments can be used to estimate grass N with a similar precision and accuracy as the hyperspectral sensor with a high spatial resolution (Ramoelo *et al.*, 2015b).

Conclusions

This study sought to provide a framework for the estimation and prediction of species richness, diversity, and nutrients for vegetation monitoring in GGHNP to allow for accurate carrying capacity models of wildlife, especially in the event of inevitable global environmental change. Furthermore, the relationships between species quality and quantity ought to still be investigated, and whether either can be used as a proxy for the other. Nonetheless, models created in this study will be used to map the spatial distribution of species and their abundances. These maps will be used to determine stocking rates for wildlife management. The main conclusions are based on the following findings from the different objectives addressed in this study:

1. Data collected using *in situ* methods showed that species levels are virtually similar across the park under different levels of disturbance regimes. However, species-level were higher at landscape than at site levels. Albeit criticised for their use in conservation, species diversity metrics could be useful for measurements of rangeland quality.
2. Topographical controlled factors played a major role in influencing species diversity, thus, highlighting the role of topography in influencing species in mountainous grassland communities.
3. Non- and parametric modelling frameworks were used to assess these relationships. Although higher prediction accuracies were achieved using

parametric models, topographically controlled edaphic factors remained the most important drivers of species richness and diversity.

4. This study sought to predict and model species richness and diversity in GGHNP using satellite remote sensing. Near-infrared was the most selected spectral band for predicting species diversity during the grasses phenology, and further, ascertaining the efficiency of NIR in vegetation mapping.
5. We used remote sensing data to predict Nitrogen in the park using Sentinel 2 Multi-Spectral Instrument (S2 MSI). The results show red edge bands as the optimal bands for estimating grass N. As such, S2 MSI present an opportunity for modelling grass N through their phenology and among seasons because of its multiple red edge bands

References

Adagbasa, E.G., Adelabu, S.A. and Okello, T.W. (2020) ‘Development of post-fire vegetation response-ability model in grassland mountainous ecosystem using GIS and remote sensing’, *ISPRS Journal of Photogrammetry and Remote Sensing*, 164(September 2019), pp. 173–183. doi:10.1016/j.isprsjprs.2020.04.006.

Adagbasa, G.E., Adelabu, S.A. and Okello, T.W. (2018) ‘Spatio-temporal assessment of fire severity in a protected and mountainous ecosystem’, *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July(December), pp. 6572–6575. doi:10.1109/IGARSS.2018.8518268.

Adjorlolo, C. and Botha, J.O. (2015) ‘Integration of remote sensing and conventional models for modeling grazing/browsing capacity in southern African savannas’, *Journal of Applied Remote Sensing*, 9(1), p. 096041. doi:10.1117/1.JRS.9.096041.

Allan, E. *et al.* (2015) ‘Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition’, *Ecology Letters*. Edited by J. Knops, 18(8), pp. 834–843. doi:10.1111/ele.12469.

Archibald, S. (2008) ‘African Grazing Lawns—How Fire, Rainfall, and Grazer Numbers Interact to Affect Grass Community States’, *Journal of Wildlife Management*, 72(2), pp. 492–501. doi:10.2193/2007-045.

Armas-Herrera, C.M. *et al.* (2020a) ‘Plant-topsoil relationships underlying subalpine grassland patchiness’, *Science of The Total Environment*, 712, p. 134483. doi:10.1016/j.scitotenv.2019.134483.

Armas-Herrera, C.M. *et al.* (2020b) ‘Plant-topsoil relationships underlying subalpine grassland patchiness’, *Science of the Total Environment*, 712, p. 134483. doi:10.1016/j.scitotenv.2019.134483.

Auestad, I., Rydgren, K. and Økland, R.H. (2008) ‘Scale-dependence of vegetation-environment relationships in semi-natural grasslands’, *Journal of Vegetation Science*, 19(1), pp. 139–148. doi:10.3170/2007-8-18344.

Barros, M.J.F. *et al.* (2015) ‘Environmental drivers of diversity in Subtropical Highland Grasslands’, *Perspectives in Plant Ecology, Evolution and Systematics*, 17(5), pp. 360–368. doi:10.1016/j.ppees.2015.08.001.

Becker, T. and Brändel, M. (2007) ‘Vegetation-environment relationships in a heavy metal-dry grassland complex’, *Folia Geobotanica*, 42(1), pp. 11–28. doi:10.1007/BF02835100.

Bittner, T. *et al.* (2011) ‘Comparing modelling approaches at two levels of biological organisation - Climate change impacts on selected Natura 2000 habitats’, *Journal of Vegetation Science*, 22(4), pp. 699–710. doi:10.1111/j.1654-1103.2011.01266.x.

Blair, J., Nippert, J. and Briggs, J. (2014) *Ecology and the Environment*. Edited by R.K. Monson. New York, NY: Springer New York. doi:10.1007/978-1-4614-7501-9.

Broge, N.. and Leblanc, E. (2001) ‘Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density’, *Remote Sensing of Environment*, 76(2), pp. 156–172. doi:10.1016/S0034-4257(00)00197-8.

Brown, J.H. *et al.* (2001) ‘Regulation of diversity: maintenance of species richness in changing environments’, *Oecologia*, 126(3), pp. 321–332. doi:10.1007/s004420000536.

Brown, L.R. *et al.* (2013a) ‘Guidelines for phytosociological classifications and descriptions of vegetation in southern Africa’, *Koedoe*, 55(1), pp. 1–10. doi:10.4102/koedoe.v55i1.1103.

Brown, L.R. *et al.* (2013b) ‘Guidelines for phytosociological classifications and descriptions of vegetation in southern Africa’, *Koedoe*, 55(1). doi:10.4102/koedoe.v55i1.1103.

Brown, L.R. and Bezuidenhout, H. (2020) ‘Grassland Vegetation of Southern Africa’, in *Encyclopedia of the World’s Biomes*. Elsevier, pp. 814–826. doi:10.1016/B978-0-12-409548-9.11960-8.

Brown, L.R. and du Preez, J. (2020) ‘Alpine Vegetation of Temperate Mountains of Southern Africa’, in *Encyclopedia of the World’s Biomes*. Elsevier, pp. 395–404. doi:10.1016/B978-0-12-409548-9.11892-5.

Cardinale, B.J. *et al.* (2007) ‘Impacts of plant diversity on biomass production increase through time because of species complementarity’, *Proceedings of the*

National Academy of Sciences, 104(46), pp. 18123–18128. doi:10.1073/pnas.0709069104.

Cardinale, B.J. *et al.* (2011) ‘The functional role of producer diversity in ecosystems’, *American Journal of Botany*, 98(3), pp. 572–592. doi:10.3732/ajb.1000364.

Cardinale, Bradley J *et al.* (2012) ‘Biodiversity loss and its impact on humanity.’, *Nature*, 486(7401), pp. 59–67. doi:10.1038/nature11148.

Cardinale, Bradley J. *et al.* (2012) ‘Biodiversity loss and its impact on humanity’, *Nature*, 486(7401), pp. 59–67. doi:10.1038/nature11148.

Cardinale, B.J. *et al.* (2018a) ‘Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends’, *Biological Conservation*, 219(December 2017), pp. 175–183. doi:10.1016/j.biocon.2017.12.021.

Cardinale, B.J. *et al.* (2018b) ‘Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends’, *Biological Conservation*, 219(January), pp. 175–183. doi:10.1016/j.biocon.2017.12.021.

Chabalala, Y. *et al.* (2020) ‘Exploiting the capabilities of Sentinel-2 and RapidEye for predicting grass nitrogen across different grass communities in a protected area’, *Applied Geomatics*, 12(4), pp. 379–395. doi:10.1007/s12518-020-00305-8.

Chanson, J. *et al.* (2010) ‘Global Biodiversity: Indicators of Recent Declines’, *Science*, 328(May), pp. 1164–1169.

Cho, M.A. *et al.* (2007) ‘Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression’, *International Journal of Applied Earth Observation and Geoinformation*, 9(4), pp. 414–424. doi:10.1016/j.jag.2007.02.001.

Cho, M.A. and Skidmore, A.K. (2006) ‘A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method’, *Remote Sensing of Environment*, 101(2), pp. 181–193. doi:10.1016/j.rse.2005.12.011.

Cingolani, A.M., Noy-meir, I. and Díaz, S. (2005) ‘Grazing effects on rangeland diversity: A sysnthesis of contemporaty models’, *Ecological Applications*, 15(2), pp. 757–773. doi:10.1890/03-5272.

Clevers, J.G.P.W. *et al.* (2002) ‘Derivation of the red edge index using the MERIS standard band setting’, *International Journal of Remote Sensing*, 23(16), pp. 3169–3184. doi:10.1080/01431160110104647.

Collins, S.L., Glenn, S.M. and Briggs, J.M. (2002) ‘Effect of local and regional processes on plant species richness in tallgrass prairie’, *Oikos*, 99(3), pp. 571–579. doi:10.1034/j.1600-0706.2002.12112.x.

Daughtry, C. (2000) ‘Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance’, *Remote Sensing of Environment*, 74(2), pp. 229–239. doi:10.1016/S0034-4257(00)00113-9.

Eckhardt, H.C., van Rooyen, N. and Bredenkamp, G.J. (1996) ‘Plant communities and species richness of the Agrostis lachnantha–Eragrostis plana Wetlands of northern KwaZulu-Natal’, *South African Journal of Botany*, 62(6), pp. 306–315. doi:10.1016/S0254-6299(15)30670-0.

Elith, J. and Leathwick, J.R. (2009) ‘Species Distribution Models: Ecological Explanation and Prediction Across Space and Time’, *Annual Review of Ecology, Evolution, and Systematics*, 40(1), pp. 677–697. doi:10.1146/annurev.ecolsys.110308.120159.

Fair, J.L., Peters, D.P.C. and Lauenroth, W.K. (2001) ‘Response of individual Bouteloua gracilis (Gramineae) plants and tillers to small disturbances’, *American Midland Naturalist*, 145(1), pp. 147–158. doi:https://doi.org/10.1674/0003-0031(2001)145[0147:ROIBGG]2.0.CO;2.

Ferranto, S. *et al.* (2011) ‘Forest and rangeland owners value land for natural amenities and as financial investment’, *California Agriculture*, 65(4), pp. 184–191. doi:10.3733/ca.v065n04p184.

Ferreira, S. *et al.* (2011) ‘From numbers to ecosystems and biodiversity: A mechanistic approach to monitoring’, *Koedoe*, 53(2), pp. 1–12. doi:10.4102/koedoe.v53i2.998.

Filibeck, G. *et al.* (2019) ‘Exploring the drivers of vascular plant richness at very fine spatial scale in sub-Mediterranean limestone grasslands (Central Apennines, Italy)’, *Biodiversity and Conservation*, 28(10), pp. 2701–2725. doi:10.1007/s10531-019-

01788-7.

Flather, C.H. (1996) ‘Fitting species-accumulation functions and assessing regional land use impacts on avian diversity’, *Journal of Biogeography*, 23(2), pp. 155–168. doi:10.1046/j.1365-2699.1996.00980.x.

Franzén, D. and Eriksson, O. (2008) ‘Small-scale patterns of species richness in Swedish semi-natural grasslands: the effects of community species pools’, *Ecography*, 24(5), pp. 505–510. doi:10.1111/j.1600-0587.2001.tb00485.x.

Gaury, P.K. and Devi, R. (2017a) ‘Plant Species Composition and Diversity at the Aravalli Mountain Range in Haryana, India’, *Journal of Biodiversity*, 8(1), pp. 34–43. doi:10.1080/09766901.2017.1336306.

Gaury, P.K. and Devi, R. (2017b) ‘Plant Species Composition and Diversity at the Aravalli Mountain Range in Haryana, India’, *Journal of Biodiversity*, 8(1), pp. 34–43. doi:10.1080/09766901.2017.1336306.

Gessner, U. *et al.* (2015) ‘Multi-sensor mapping of West African land cover using MODIS, ASAR and TanDEM-X/TerraSAR-X data’, *Remote Sensing of Environment*, 164, pp. 282–297. doi:10.1016/j.rse.2015.03.029.

Getis, A. and Ord, J.K. (1992) ‘The Analysis of Spatial Association by Use of Distance Statistics’, *Geographical Analysis*, 24(3), pp. 189–206. doi:10.1111/j.1538-4632.1992.tb00261.x.

Glibert, P.M. (2016) ‘Ecological stoichiometry’, in *Encyclopedia of Earth Sciences Series*, pp. 228–231. doi:10.1007/978-94-017-8801-4_15.

Haboudane, D. (2004) ‘Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture’, *Remote Sensing of Environment*, 90(3), pp. 337–352. doi:10.1016/j.rse.2003.12.013.

Han, Y. *et al.* (2019) ‘Response of soil nutrients and stoichiometry to elevated nitrogen deposition in alpine grassland on the Qinghai-Tibetan Plateau’, *Geoderma*, 343(September 2018), pp. 263–268. doi:10.1016/j.geoderma.2018.12.050.

Hautier, Y. *et al.* (2018) ‘Local loss and spatial homogenization of plant diversity reduce ecosystem multifunctionality’, *Nature Ecology and Evolution*, 2(1), pp. 50–56.

doi:10.1038/s41559-017-0395-0.

Hawbaker, T.J. *et al.* (2020) ‘The Landsat Burned Area algorithm and products for the conterminous United States’, *Remote Sensing of Environment*, 244(April), p. 111801. doi:10.1016/j.rse.2020.111801.

Hector, A. *et al.* (2010) ‘General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding’, *Ecology*, 91(8), pp. 2213–2220. doi:10.1890/09-1162.1.

Hengl, T. *et al.* (2017) ‘SoilGrids250m: Global gridded soil information based on machine learning’, *PLOS ONE*. Edited by B. Bond-Lamberty, 12(2), p. e0169748. doi:10.1371/journal.pone.0169748.

Hooper, D.U. *et al.* (2012) ‘A global synthesis reveals biodiversity loss as a major driver of ecosystem change’, *Nature*, 486(7401), pp. 105–108. doi:10.1038/nature11118.

Huete, A. (1997) ‘A comparison of vegetation indices over a global set of TM images for EOS-MODIS’, *Remote Sensing of Environment*, 59(3), pp. 440–451. doi:10.1016/S0034-4257(96)00112-5.

Hunt, E.R. and Miyake, B.A. (2006) ‘Comparison of Stocking Rates From Remote Sensing and Geospatial Data. Rangeland Ecology & Management’, *Rangeland Ecology & Management*, 59(1), pp. 11–18.

Iganci, J.R. V. *et al.* (2011) ‘Campos de Cima da Serra: the Brazilian Subtropical Highland Grasslands show an unexpected level of plant endemism’, *Botanical Journal of the Linnean Society*, 167(4), pp. 378–393. doi:10.1111/j.1095-8339.2011.01182.x.

Jones, C.G., Lawton, J.H. and Shachak, M. (1994) ‘Organisms as Ecosystem Engineers’, *Oikos*, 69(3), p. 373. doi:10.2307/3545850.

Jordan, C.F. (1969) ‘Derivation of Leaf-Area Index from Quality of Light on the Forest Floor’, *Ecology*, 50(4), pp. 663–666. doi:10.2307/1936256.

Jordano, P. *et al.* (2013) ‘Functional traits, the phylogeny of function, and ecosystem service vulnerability’, *Ecology and Evolution*, 3(9), pp. 2958–2975. doi:10.1002/ece3.601.

Kay, C., Bredenkamp, G.J. and Theron, G.K. (1993) ‘The plant communities of the Golden Gate Highlands National Park in the north-eastern Orange Free State’, *South African Journal of Botany*, 59(4), pp. 442–449. doi:10.1016/S0254-6299(16)30717-7.

Kosicki, J.Z. (2020) ‘Generalised Additive Models and Random Forest Approach as effective methods for predictive species density and functional species richness’, *Environmental and Ecological Statistics*, 27(2), pp. 273–292. doi:10.1007/s10651-020-00445-5.

Kuzemko, A.A. *et al.* (2016) ‘Patterns and drivers of phytodiversity in steppe grasslands of Central Podolia (Ukraine)’, *Biodiversity and Conservation*, 25(12), pp. 2233–2250. doi:10.1007/s10531-016-1060-7.

Lausch, A. *et al.* (2018) ‘Understanding and assessing vegetation health by in situ species and remote-sensing approaches’, *Methods in Ecology and Evolution*, 9(8), pp. 1799–1809. doi:10.1111/2041-210X.13025.

Lee, C.-B. and Chun, J.-H. (2016) ‘Environmental Drivers of Patterns of Plant Diversity Along a Wide Environmental Gradient in Korean Temperate Forests’, *Forests*, 7(12), p. 19. doi:10.3390/f7010019.

Li, Wenjin *et al.* (2018) ‘Effect of loss of plant functional group and simulated nitrogen deposition on subalpine ecosystem properties on the Tibetan Plateau’, *Science of the Total Environment*, 631–632, pp. 289–297. doi:10.1016/j.scitotenv.2018.02.287.

Liu, Y. *et al.* (2019) ‘Assessing the effects of climate variation and human activities on grassland degradation and restoration across the globe’, *Ecological Indicators*, 106(July), p. 105504. doi:10.1016/j.ecolind.2019.105504.

Lorenz-Lemke, A.P. *et al.* (2010) ‘Diversification of plant species in a subtropical region of eastern South American highlands: A phylogeographic perspective on native Petunia (Solanaceae)’, *Molecular Ecology* [Preprint]. doi:10.1111/j.1365-294X.2010.04871.x.

Lund, H.G. (2007) ‘Accounting for the World ’ s Rangelands’, *Rangelands*, 29(1), pp. 3–10.

Lundholm, J.T. (2009) ‘Plant species diversity and environmental heterogeneity: Spatial scale and competing hypotheses’, *Journal of Vegetation Science*, 20(3), pp.

377–391. doi:10.1111/j.1654-1103.2009.05577.x.

Lyon, J.G. and Huete, A. (2016) *Hyperspectral Remote Sensing of Vegetation*. Edited by P.S. Thenkabail and J.G. Lyon. CRC Press. doi:10.1201/b11222.

Malanson, G.P. *et al.* (2020a) ‘Alpine plant community diversity in species–area relations at fine scale’, *Arctic, Antarctic, and Alpine Research*, 52(1), pp. 41–46. doi:10.1080/15230430.2019.1698894.

Malanson, G.P. *et al.* (2020b) ‘Alpine plant community diversity in species–area relations at fine scale’, *Arctic, Antarctic, and Alpine Research*, 52(1), pp. 41–46. doi:10.1080/15230430.2019.1698894.

Marsett, R.C. *et al.* (2006) ‘Remote Sensing for Grassland Management in the Arid Southwest Remote Sensing for Grassland Management in the Arid Southwest’, 59(5), pp. 530–540.

Mauchamp, L. *et al.* (2014) ‘Impact of management type and intensity on multiple facets of grassland biodiversity in the French Jura Mountains’, *Applied Vegetation Science*, 17(4), pp. 645–657. doi:10.1111/avsc.12116.

Mccollum, D.W. *et al.* (2017) ‘Climate change effects on rangelands and rangeland management: affirming the need for monitoring’, *Ecosystem Health and Sustainability*, 3(3), p. e01264. doi:10.1002/ehs2.1264.

McCord, S.E. *et al.* (2017a) ‘Integrating Remotely Sensed Imagery and Existing Multiscale Field Data to Derive Rangeland Indicators: Application of Bayesian Additive Regression Trees’, *Rangeland Ecology and Management*, 70(5), pp. 644–655. doi:10.1016/j.rama.2017.02.004.

McCord, S.E. *et al.* (2017b) ‘Integrating Remotely Sensed Imagery and Existing Multiscale Field Data to Derive Rangeland Indicators: Application of Bayesian Additive Regression Trees’, *Rangeland Ecology & Management*, 70(5), pp. 644–655. doi:10.1016/j.rama.2017.02.004.

McNaughton, S.J. (1985a) ‘Ecology of a Grazing Ecosystem: The Serengeti’, *Ecological Monographs*, 55(3), pp. 259–294. doi:10.2307/1942578.

McNaughton, S.J. (1985b) ‘Ecology of a Grazing Ecosystem: The Serengeti’, *Ecological Monographs*, 55(3), pp. 259–294. doi:10.2307/1942578.

McNaughton, S.J. (1988) ‘Mineral nutrition and spatial concentrations of African ungulates’, *Nature*, 334(6180), pp. 343–345. doi:10.1038/334343a0.

McNaughton, S.J. (1990) ‘Mineral nutrition and seasonal movements of African migratory ungulates’, *Nature*, 345(6276), pp. 613–615. doi:10.1038/345613a0.

McNaughton, S.J., Banyikwa, F.F. and McNaughton, M.M. (1997) ‘Promotion of the Cycling of Diet-Enhancing Nutrients by African Grazers’, *Science*, 278(5344), pp. 1798–1800. doi:10.1126/science.278.5344.1798.

Meyer, S.T. *et al.* (2018) ‘Biodiversity-multifunctionality relationships depend on identity and number of measured functions’, *Nature Ecology and Evolution*, 2(1), pp. 44–49. doi:10.1038/s41559-017-0391-4.

Moeslund, J.E. *et al.* (2013) ‘Topography as a driver of local terrestrial vascular plant diversity patterns’, *Nordic Journal of Botany*, 31(2), pp. 129–144. doi:10.1111/j.1756-1051.2013.00082.x.

Mucina (2006) ‘The vegetation of South Africa, Lesotho and Swaziland. South African National Biodiversity Institute’, (December 2015).

Mucina, L. and Rutherford, M.C. (2006) ‘The vegetation of South Africa, Lesotho and Swaziland.’, *Strelitzia 19*, (December), pp. 1–30. Available at: http://ebooks.cambridge.org/ref/id/CBO9781107415324A009.

Mutanga, O. and Skidmore, A.K. (2007) ‘Red edge shift and biochemical content in grass canopies’, *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1), pp. 34–42. doi:10.1016/j.isprsjprs.2007.02.001.

Myers, N. *et al.* (2000) ‘Biodiversity hotspots for conservation priorities’, *Nature*, 403: 853(February), pp. 853–858.

O’Connor, T. (2008) ‘Influence of land use on phytomass accumulation in Highland Sourveld grassland in the southern Drakensberg, South Africa’, *African Journal of Range & Forage Science*, 25(1), pp. 17–27. doi:10.2989/AJRFS.2008.25.1.3.381.

van Oijen, M., Bellocchi, G. and Höglind, M. (2018) ‘Effects of Climate Change on Grassland Biodiversity and Productivity: The Need for a Diversity of Models’, *Agronomy*, 8(2), p. 14. doi:10.3390/agronomy8020014.

Oksanen, A.J. *et al.* (2012) *Community Ecology Package*, *Comprehensive R Archive Network*. Available at: http://mirror.bjtu.edu.cn/cran/web/packages/vegan/vegan.pdf.

Olea, P.P., Mateo-Tomás, P. and de Frutos, Á. (2010) ‘Estimating and Modelling Bias of the Hierarchical Partitioning Public-Domain Software: Implications in Environmental Management and Conservation’, *PLoS ONE*. Edited by S. Plaistow, 5(7), p. e11698. doi:10.1371/journal.pone.0011698.

Oliver, T.H. *et al.* (2015) ‘Biodiversity and Resilience of Ecosystem Functions’, *Trends in Ecology & Evolution*, 30(11), pp. 673–684. doi:10.1016/j.tree.2015.08.009.

Orlandi, S. *et al.* (2016) ‘Environmental and land use determinants of grassland patch diversity in the western and eastern Alps under agro-pastoral abandonment’, *Biodiversity and Conservation*, 25(2), pp. 275–293. doi:10.1007/s10531-016-1046-5.

Peratoner, G. and Pötsch, E.M. (2019) ‘Methods to describe the botanical composition of vegetation in grassland research Methoden zur Beschreibung des Pflanzenbestandes in Grünland-Feldversuchen’, 70(1), pp. 1–18.

Pereira, H.M. *et al.* (2013) ‘Essential Biodiversity Variables’, *Science*, 339(6117), pp. 277–278. doi:10.1126/science.1229931.

Picó, F.X. and Van Groenendael, J. (2007) ‘Large-scale plant conservation in European semi-natural grasslands: A population genetic perspective’, *Diversity and Distributions*, 13(6), pp. 920–926. doi:10.1111/j.1472-4642.2007.00349.x.

Poschlod, P. *et al.* (1998) ‘Plant species richness in calcareous grasslands as affected by dispersability in space and time’, *Applied Vegetation Science*, 1(1), pp. 75–91. doi:10.2307/1479087.

Qi, J. *et al.* (1994) ‘A modified soil adjusted vegetation index’, *Remote Sensing of Environment*, 48(2), pp. 119–126. doi:10.1016/0034-4257(94)90134-1.

Ramoelo, A. *et al.* (2012a) ‘Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor’, *International Journal of Applied Earth Observation and Geoinformation*, 19(1), pp. 151–162. doi:10.1016/j.jag.2012.05.009.

Ramoelo, A. *et al.* (2012b) ‘Regional estimation of savanna grass nitrogen using the red-edge band of the spaceborne rapideye sensor’, *International Journal of Applied*

Earth Observation and Geoinformation, 19(1), pp. 151–162. doi:10.1016/j.jag.2012.05.009.

Ramoelo, A. *et al.* (2015a) ‘Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data’, *International Journal of Applied Earth Observation and Geoinformation*, 43, pp. 43–54. doi:10.1016/j.jag.2014.12.010.

Ramoelo, A. *et al.* (2015b) ‘Monitoring grass nutrients and biomass as indicators of rangeland quality and quantity using random forest modelling and WorldView-2 data’, *International Journal of Applied Earth Observation and Geoinformation*, 43, pp. 43–54. doi:10.1016/j.jag.2014.12.010.

Reddy, C.S. (2021) ‘Remote sensing of biodiversity: what to measure and monitor from space to species?’, *Biodiversity and Conservation*, 30(10), pp. 2617–2631. doi:10.1007/s10531-021-02216-5.

Reitalu, T. *et al.* (2012) ‘Responses of grassland species richness to local and landscape factors depend on spatial scale and habitat specialization’, *Journal of Vegetation Science*. Edited by J. Fridley, 23(1), pp. 41–51. doi:10.1111/j.1654-1103.2011.01334.x.

Richter, R. *et al.* (2016) ‘The use of airborne hyperspectral data for tree species classification in a species-rich Central European forest area’, *International Journal of Applied Earth Observation and Geoinformation*, 52, pp. 464–474. doi:10.1016/j.jag.2016.07.018.

Rocchini, D. *et al.* (2016) ‘Satellite remote sensing to monitor species diversity: potential and pitfalls’, *Remote Sensing in Ecology and Conservation*. Edited by A. Skidmore and A. Chauvenet, 2(1), pp. 25–36. doi:10.1002/rse2.9.

Rocchini, D., Chiarucci, A. and Loiselle, S.A. (2004) ‘Testing the spectral variation hypothesis by using satellite multispectral images’, *Acta Oecologica*, 26(2), pp. 117–120. doi:10.1016/j.actao.2004.03.008.

Rocchini, D., Ricotta, C. and Chiarucci, A. (2007a) ‘Using satellite imagery to assess plant species richness: The role of multispectral systems’, *Applied Vegetation Science*, 10(3), pp. 325–331. doi:10.1111/j.1654-109X.2007.tb00431.x.

Rocchini, D., Ricotta, C. and Chiarucci, A. (2007b) ‘Using satellite imagery to assess plant species richness: The role of multispectral systems’, *Applied Vegetation Science*, 10(3), pp. 325–331. doi:10.1111/j.1654-109X.2007.tb00431.x.

Rondeaux, G., Steven, M. and Baret, F. (1996) ‘Optimization of soil-adjusted vegetation indices’, *Remote Sensing of Environment*, 55(2), pp. 95–107. doi:10.1016/0034-4257(95)00186-7.

Rouse, J.W. *et al.* (1974) *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation*, *Final Report, RSC 1978-4, Texas A & M University, College Station, Texas*.

SANParks (2020) *Golden Gate Highlands National Park Park Management Plan*, *South African National Parks*.

Schmidtlein, S. and Fassnacht, F.E. (2017) ‘The spectral variability hypothesis does not hold across landscapes’, *Remote Sensing of Environment*, 192, pp. 114–125. doi:10.1016/j.rse.2017.01.036.

Scholes, B. and Howden, M. (2003) ‘Management for sustainable use — global climate change and rangelands’, *African Journal of Range & Forage Science*, 20(2), pp. 153–156. doi:10.2989/10220110309485810.

Shoemaker, L.G. (2020) ‘Intergrating the underlying structure of stochasticity into community ecology’, *Concepts and Synthesis*, 101(2), pp. 1–17. doi:10.1002/ecy.2922.

Siebert, S.J., Siebert, F. and Du Toit, M.J. (2011) ‘The extended occurrence of Maputaland Woody Grassland further south in KwaZulu-Natal, South Africa’, *Bothalia*, 41(2), pp. 341–350. doi:10.4102/abc.v41i2.77.

Skaggs, R. (2008) ‘Ecosystem Services and Western U . S . Rangelands’, *American Agricultural Economics Association*, pp. 37–41.

Snyman, H.A., Ingram, L.J. and Kirkman, K.P. (2013) ‘Themeda triandra: a keystone grass species’, *African Journal of Range and Forage Science*, 30(3), pp. 99–125. doi:10.2989/10220119.2013.831375.

Soberon, J. and Jorge, L. (1993) ‘The Use of Species Accumulation Functions for the Prediction of Species Richness’, *Conservation Biology*, 7(3), pp. 480–488.

doi:10.1046/j.1523-1739.1993.07030480.x.

Steffen, W. *et al.* (2004) *Global Change and the Earth System: A Planet Under Pressure.*

Stock, W.D., Bond, W.J. and van de Vijver, C.A.D.M. (2010) ‘Herbivore and nutrient control of lawn and bunch grass distributions in a southern African savanna’, *Plant Ecology*, 206(1), pp. 15–27. doi:10.1007/s11258-009-9621-4.

Sweeney, R.A. and Rexroad, P.R. (1987) ‘Comparison of LECO FP-228 “nitrogen determinator” with AOAC copper catalyst Kjeldahl method for crude protein.’, *Journal - Association of Official Analytical Chemists*, 70(6), pp. 1028–1030. doi:10.1093/jaoac/70.6.1028.

Symstad, A.J. and Jonas, J.L. (2011) ‘Incorporating Biodiversity Into Rangeland Health: Plant Species Richness and Diversity in Great Plains Grasslands’, *Rangeland Ecology & Management*, 64(6), pp. 555–572. doi:10.2111/REM-D-10-00136.1.

Tardella, F.M. *et al.* (2016) ‘Environmental gradients and grassland trait variation: Insight into the effects of climate change’, *Acta Oecologica*, 76, pp. 47–60. doi:10.1016/j.actao.2016.08.002.

Thenkabail, P.S. *et al.* (2004) ‘Accuracy assessments of hyperspectral waveband performance for vegetation analysis applications’, *Remote Sensing of Environment*, 91(3–4), pp. 354–376. doi:10.1016/j.rse.2004.03.013.

Tilman, D. (1993) ‘Species Richness of Experimental Productivity Gradients: How Important is Colonization Limitation?’, *Ecology*, 74(8), pp. 2179–2191. doi:10.2307/1939572.

Tilman, D. (1999a) ‘The ecological consequences of changes in biodiversity: A search for general principles’, in *Ecology*.

Tilman, D. (1999b) ‘The ecological consequences of changes in biodiversity: A search for general principles’, *Ecology*, 80(5), pp. 1455–1474. doi:https://doi.org/10.1890/0012-9658(1999)080[1455:TECOCI]2.0.CO;2.

Tilman, D. and Downing, J.A. (1994) ‘Biodiversity and stability in grasslands’, *Nature*, 367(6461), pp. 363–365. doi:10.1038/367363a0.

Tilman, D., Reich, Peter B and Isbell, F. (2012) ‘Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory’, *Proceedings of the National Academy of Sciences*, 109(26), pp. 10394–10397. doi:10.1073/pnas.1208240109.

Tilman, D., Reich, Peter B. and Isbell, F. (2012) ‘Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory’, *Proceedings of the National Academy of Sciences*, 109(26), pp. 10394–10397. doi:10.1073/pnas.1208240109.

Tilman, D., Reich, P.B. and Knops, J.M.H. (2006) ‘Biodiversity and ecosystem stability in a decade-long grassland experiment’, *Nature*, 441(7093), pp. 629–632. doi:10.1038/nature04742.

Tilman, D., Wedin, D. and Knops, J. (1996) ‘Productivity and sustainability in grassland ecosystems’, *Nature*, 379(3), pp. 718–720.

Tilman, David, Wedin, D. and Knops, J. (1996) ‘Productivity and sustainability influenced by biodiversity in grassland ecosystems’, *Nature*, 379(6567), pp. 718–720. doi:10.1038/379718a0.

Torresani, M. *et al.* (2019) ‘Estimating tree species diversity from space in an alpine conifer forest: The Rao’s Q diversity index meets the spectral variation hypothesis’, *Ecological Informatics*, 52(February), pp. 26–34. doi:10.1016/j.ecoinf.2019.04.001.

Ugland, K.I., Gray, J.S. and Ellingsen, K.E. (2016) ‘The Species-Accumulation Curve and Estimation of Species Richness Published by : British Ecological Society Linked references are available on JSTOR for this article : The species-accumulation curve and estimation of species richness’, 72(5), pp. 888–897.

Ungar, E.D. (2019) ‘Perspectives on the concept of rangeland carrying capacity, and their exploration by means of Noy-Meir’s two-function model’, *Agricultural Systems*, 173(March), pp. 403–413. doi:10.1016/j.agsy.2019.03.023.

Wardle, D.A. *et al.* (2011) ‘Terrestrial Ecosystem Responses to Species Gains and Losses’, *Science*, 332, pp. 1273–1277. doi:10.1210/jcem-10-10-1361.

Wardle, D.A. (2016) ‘Do experiments exploring plant diversity-ecosystem functioning relationships inform how biodiversity loss impacts natural ecosystems?’, *Journal of*

Vegetation Science. Edited by M. Palmer, 27(3), pp. 646–653. doi:10.1111/jvs.12399.

Weisser, W.W. *et al.* (2017) ‘Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: Patterns, mechanisms, and open questions’, *Basic and Applied Ecology*, 23, pp. 1–73. doi:10.1016/j.baae.2017.06.002.

West, N.E. (1993) ‘Biodiversity of Rangelands’, *Journal of Range Management*, 46(1), pp. 2–13.

White, R., Murray, S. and Rohweder, M. (2000) ‘Research Watch: Forest ecosystems’, *Environmental Science & Technology*, 34(11), pp. 274A-274A. doi:10.1021/es0032881.

White, S.R. *et al.* (2012) ‘Climate change experiments in temperate grasslands: synthesis and future directions’, *Biology Letters*, 8(4), pp. 484–487. doi:10.1098/rsbl.2011.0956.

Williamson, M., Gaston, K.J. and Lonsdale, W.M. (2001) ‘The species-area relationship does not have an asymptote!’, *Journal of Biogeography*, 28(7), pp. 827–830. doi:10.1046/j.1365-2699.2001.00603.x.

Xie, Y., Sha, Z. and Yu, M. (2008) ‘Remote sensing imagery in vegetation mapping: a review’, *Journal of Plant Ecology*, 1(1), pp. 9–23. doi:10.1093/jpe/rtm005.

Xulu, S., Mbatha, N. and Peerbhay, K. (2021) ‘Burned area mapping over the southern cape forestry region, South Africa using sentinel data within gee cloud platform’, *ISPRS International Journal of Geo-Information*, 10(8). doi:10.3390/ijgi10080511.

Yang, Z. *et al.* (2015) ‘The effect of environmental heterogeneity on species richness depends on community position along the environmental gradient’, *Scientific Reports*, 5(1), p. 15723. doi:10.1038/srep15723.

Zarovali, M.P., Yiakoulaki, M.D. and Papanastasis, V.P. (2007) ‘Effects of shrub encroachment on herbage production and nutritive value in semi-arid Mediterranean grasslands’, *Grass and Forage Science*, 62(3), pp. 355–363. doi:10.1111/j.1365-2494.2007.00590.x.

Zulka, K.P. *et al.* (2014) ‘Species richness in dry grassland patches of eastern

Austria: A multi-taxon study on the role of local, landscape and habitat quality variables', *Agriculture, Ecosystems & Environment*, 182, pp. 25–36. doi:10.1016/j.agee.2013.11.016.