The importance of hydrologic refugia for the diversity of the Cape Flora

Supervisor: Dr Jasper Slingsby

Co-supervisor: Professor Tony Verboom

Abstract

Refugia are sites with more stable climate and hydrology than the surrounding area. There is growing recognition of the need to conserve refugia due to their role in harboring rare and endemic plants that might otherwise be prone to extinction due to global change. These sites are considered crucial for preserving species richness in the Cape Floristic Region (CFR) of South Africa, an area renowned for its exceptional plant species diversity and endemism, which is increasingly threatened by the impacts of global change. Despite the importance of refugia in the CFR, very little is known about their distribution. To fill this knowledge gap, I identify the distribution and character of refugia (mesic or xeric) by mapping total species richness (TSR), the distribution of widespread species (WS), narrow-range endemics (NREs), and wetland dependent narrow-range endemics (WD-NREs) and explore the environmental correlates of these distributions. I make use of Quarter degree square grid cells as a measuring unit, which are larger (~20 km across), than the scale at which microrefugia typically occur, however, the broad environmental heterogeneity within these units provides a meaningful basis for identifying potential refugial areas. From this study I find that the spatial distributions of TSR, WS, NREs and WD-NREs exhibited topographic and longitudinal gradients, with more species found in the southwestern CFR and in the mountains. Differences were observed in their correlates, where productivity, environmental stability explained TSR and WS, whereas environmental heterogeneity and environmental stability explained NREs and WD-NREs. Hydrological stability (such as groundwater-fed wetlands and seeps) and climatic stability (climatic stability index) have been documented to be indicators of refugia. Where hydrologic refugia provide broad-scale stable conditions for endemics and were found to play a significant role in maintaining persistence species in the CFR. These refugial sites face significant threats from groundwater abstraction, invasive species, and global change, highlighting the need for detailed fine-scale mapping and conservation of microrefugia within the broader refugial areas. Protecting these sites is essential for conserving biodiversity and ecosystem function in the CFR.

Declaration

1. I know that plagiarism is a serious form of academic dishonesty.

2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.

3. Where I have used the words of others, I have indicated this by the use of quotation marks.

4. I have referenced all quotations and other ideas borrowed from others.

5. I have not and shall not allow others to plagiarize my work.

Signature: Signed by candidate

February 2025

Acknowledgements

First and foremost, I would like to express my gratitude to the Building Research Active Academic Staff (BRAAS) Grant and the South African Environmental Observation Network (SAEON) for their generous financial support. Their funding played a crucial role in enabling me to complete this degree.

I am profoundly grateful to my supervisor, Dr. Jasper Slingsby, for his unwavering support, expert guidance, and invaluable insights throughout this thesis. His patience, constructive feedback, and deep GIS and statistical expertise in R were instrumental in shaping this research. I truly appreciate his dedication to both my academic and personal growth, and I feel incredibly fortunate to have had his mentorship.

I would also like to extend my sincere appreciation to Professor Tony Verboom for his continuous encouragement and thoughtful suggestions. His vast knowledge of the Cape Floristic Region has greatly enriched my research, and his tough yet supportive approach has challenged me to think critically. His experience and expertise have been invaluable in this thesis.

A special thanks to my fellow postgraduate and postdoctoral colleagues in the Department of Biological Sciences: Dr. Seth Musker, Dr. Anabelle Cardoso, Dr. Justin Van Blerk, Simcelile Chenge, Hana Petersen, and Ruan van Mazijk. Their technical support, camaraderie, and feedback and encouragement made this journey far more manageable. I am also grateful for the sense of community in the Plant Ecology Lab and for their help in navigating the deep end of this process.

I am deeply indebted to my mother for her unwavering love, support, and belief in me. Her encouragement has been a constant source of strength, reminding me that I am capable even in the toughest moments. A heartfelt thank you to my girlfriend, Nompumelelo, whose boundless love, patience, and encouragement carried me through long nights and moments of self-doubt. She has been my confidant, my anchor, and my greatest source of stability, and I feel incredibly lucky to share this journey with such a thoughtful and inspiring partner.

Lastly, I would like to thank my friends from back home and my soccer teammates, who provided me with a space to unwind, recharge, and find balance throughout this process. Their support and companionship have been invaluable, and I am truly grateful.

Table of Contents

Table of Figures

List of Tables

List of abbreviations

CAR	Conditional autoregressive model
CFR	Cape Floristic Region
CHELSA	Climatologies at high resolution for the Earth's land surface
CloudFQ	Cloud cover frequency
CSI	Climatic stability index
DIC	Deviance information criterion
IUCN	International Union for Conservation of Nature
GIS	Geographic information system
GLM	Generalised Linear Model
LM	Lagrange's Multiplier
LMPL	Log marginal predictive likelihood
MAP	Mean annual precipitation,
MAT	Mean annual temperature,
MCMC	Markov chain Monte Carlo
MDQ	Mean precipitation in the driest quarter,
MIH	More-individuals hypothesis
NFEPA	National Freshwater Ecosystem Priority Areas Map
NRE	Narrow-range endemics
P	Phosphorus
PET	Potential evapotranspiration,
POSA	Plants of Southern Africa
QDS	Quarter-degree squares
SAC	Spatial autocorrelation
SANBI	South African National Biodiversity Institute
SANPC	South African National Plant Checklist
SARMA	Spatial autoregressive moving average

SR	Species richness
TSR	Total species richness
TSPP	Threatened Species Programme
WD-NREs	Wetland-dependent narrow range endemics
Wetland%	Wetland percentage cover
WS	Widespread Species

Introduction

The term "refugium" has been used increasingly in the scientific literature to refer to an area which provides buffering capacity against environmental fluctuations and is resilient to disturbances (Bennet and Provan, 2008; Harrison and Noss, 2017). A growing body of work suggests that various forms of refugium contribute to preserving biodiversity (Harrison and Noss, 2017). Moreover, there is reason to believe that sites which have historically been climatically stable, and so functioned as refugia in the past, will continue to buffer species against future climate fluctuations by providing cooler and/or wetter habitats (Loarie *et al*., 2008; Keppel *et al*., 2017; McLaughlin *et al.,* 2017). As such, refugia have been identified as priority sites for biodiversity conservation in the face of a changing climate (Noss, 2001; Taberlet and Cheddadi, 2002; Loarie *et al*., 2008). Identifying and mapping refugia is challenging, because they are at times not detectable at the scale of most readily available datasets relating to biodiversity, climate and other environmental variables, resulting in exaggerated predictions of extinction risk (Randin *et al.,* 2009; Austin and Van Niel, 2011; Franklin *et al.,* 2013). However, one expectation is that species' persistence through past climate change should result in historical refugia accumulating greater numbers of species, especially narrow range endemics (NREs; Harrison and Noss 2017). While this notion holds at global extents and for broader scale "macrorefugia" (Sandel *et al.,* 2011; Cai *et al.,* 2023), it is less well explored within smaller regions and for finer scale "microrefugia" (<1km)– the scale at which most conservation planning and action takes place (Rull, 2009; Mee & Moore, 2014).

A range of methods have been used to identify refugia using climate or biological data, or both (Keppel *et al*., 2012; Morelli *et al*., 2016). Ultimately, the best approach depends on the scale and resolution of the analysis and the data available. Global analyses typically identify regional "macrorefugia" (Harrison and Noss 2017) based on long-term climatic stability (Jansson, 2003; Gómez and Lunt, 2007; Sandel *et al*., 2011). This approach is less useful at finer scales where "microrefugia", with favourable local environmental features that allow small populations to survive in otherwise unfavourable regional environmental conditions (Rull, 2009), may be maintained by microtopographic and hydrological landscape features (Morelli *et al*., 2016; McLaughlin *et al*., 2017; Ackerly *et al*., 2020; Cartwright *et al*., 2020). In such cases, useful indicators of refugia include areas of greater genetic diversity (Carnaval *et al.,* 2009). or the presence of relict populations of species that are typical of cooler regions (Tang *et al.,* 2018). While endemism is often used as an indicator in global analyses and studies of microrefugia (Jansson, 2003; Sandel *et al.,* 2011), it appears to be less commonly used in regional studies of microrefugia, possibly because it is a less useful indicator in regions that have been subject to glaciation and/or have few NRE species, or because we rarely have suitable data at fine enough spatial scales.

NREs are defined as species which occupy restricted geographic ranges because of niche specialisation, limited dispersal ability, and/or reduced adaptive capacity (Picker and Samways, 1996; Chichorro *et al.,* 2019; Manes *et al.,* 2021). NREs have become a point of interest for conservation due to their rarity and influence on the global extinction rate, as their limited geographic range and specialised niche requirements may make them especially prone to extinction (Linder, 1995; Myers *et al.,* 2000). Thomas *et al.,* (2004) estimated that more than one million species, mainly those with narrow ranges, could become extinct by 2050 due to global climate change. These predictions are heightened for water-limited regions (e.g. Mediterranean climate regions) where climate change is expected to shift the water-balance towards more xeric conditions (Flint and Flint, 2012; Dobrowski *et al.,* 2013). However, Thomas *et al.,* (2004) modelled species range shifts using data for macroclimates, which would not necessarily account for the buffering effect of microrefugia, so some NREs may not be as vulnerable as suggested. Ascertaining where NREs are supported by microrefugia or not is critical for the management of the microenvironments and species in question.

Identifying microrefugia and their importance for maintaining biodiversity is challenged by the spatial scale of available data and the conflation of micro- and macro- scale processes. While the distribution of NREs may be a useful proxy for refugia, it is difficult to tease apart the influence of microenvironments (microrefugia) relative to broader macroclimatic stability (macrorefugia). NREs usually inhabit areas that have remained stable over long periods of time, but this stability may be due to macroclimate, microenvironments, or both (Işik, 2011; Harrison and Noss, 2017; McLaughlin *et al.,* 2017). In general, refugia have been linked to modern-day hotspots of both genetic and species diversity (Carnaval *et al.,* 2009), with climatic stability and environmental heterogeneity implicated as contributors of high concentrations of paleo-endemics and neo-endemics in regions such as south-western Australia (Hopper and Gioia, 2004), the North American Coastal Plain (Sorrie and Weakley, 2001) and the Cape region of South Africa (Linder, 2008). While climatic stability is generally inferred from coarse macroclimatic data, it can be affected at fine scales by topographically mediated microclimates and other microenvironments that may support hydrological refugia (McLaughlin et al. 2017). Similarly, environmental heterogeneity may act to buffer species against changing climate by reducing climate velocities (Loarie et al. 2009) but may also support more microenvironments with potential to act as microrefugia. Untangling the roles of microenvironments versus stable macroclimate in the creation and maintenance of global biodiversity hotspots may be key to their conservation, as the management options for mitigation and adaptation differ between the two scales.

In the absence of fine-scale data it is impossible to map and examine fine-scale environments that may act as microrefugia, but there are ways to infer the potential importance of microrefugia from course-scale data. Firstly, the distribution of total species richness (TSR) and NRE richness may show differences in their response to environmental drivers that reveal information about their respective genesis and maintenance. These differences may be more pronounced if one examines the subset of non-endemic widespread species (WS) richness independently. This approach is somewhat marred in that the three sets will always be related to some degree since TSR = NRE richness + WS richness, and the split between NRE and WS is somewhat subjective. Secondly, one may learn about the role of specific microenvironments in supporting microrefugia by exploring patterns of NREs with known habitat associations. For instance, when an environment is consistently wetter and cooler than the regional climate and allows species to survive and thrive over long periods of time, we characterize those sites as hydrological refugia (McLaughlin *et al*, 2017; Cartwright and Johnson, 2018; Cartwright *et al*., 2020). By mapping and analysing the richness patterns of narrow-range endemic species that depend on habitats such as wetlands relative to TSR or WS richness, there is the potential to reveal the distribution and determinants of microrefugia (Ashcroft, 2010).

Using patterns of NRE richness across an entire flora to identify and map microrefugia is not new (Tribsch and Schönswetter 2003; Weber *et al.,* 2014). However, this approach comes with challenges of untangling the individual drivers of TSR and NRE richness, because NRE richness and TSR often respond to the same drivers and/or they depend on each other, although the direction of this dependency is not always clear (Eriksson, 1993; Ricklefs and Schluter, 1993). The relationship between NRE richness and TSR is often examined by regressing the former against the latter (Cornell and Lawton, 1992, Cornell, 1993; Ricklefs and Schluter, 1993). At low values of TSR, NRE richness may increase as TSR increases. However, as TSR continues to increase, NRE richness may either keep increasing, level off and reach an asymptote, or become independent of further increases in TSR (Cornell and Lawton 1992, Cornell, 1993). The relationship between NRE richness and TSR is important to understand, because it may inform the extent to which NREs contribute to TSR (Cornell and Lawton, 1992; Cornell, 1993). It becomes challenging to distinguish the drivers of TSR and NRE richness, particularly in biodiversity hotspots that have been identified as potential refugia sites, because their distributions may be highly correlated (Cornell and Lawton, 1992). Understanding the factors that differentiate and shape the distribution of NRE richness and TSR in biodiversity hotspots requires a closer examination of classic species richness drivers. It is within these contexts—where environmental *stability* (Dynesius and Jansson, 2000), *productivity* (O'Brien, 1993), and environmental *heterogeneity* (Kreft and Jetz, 2007 can also explain NRE richness patterns—and its links to refugia and its persistence (Harrison and Noss, 2017).

Environmental stability

Long term stability of climate and hydrology are known to influence the persistence and distribution of TSR and NRE richness (Dynesius and Jansson, 2000; Jansson, 2003; Jetz *et al.,* 2004; Cartwright and Johnson, 2018; McLaughlin *et al.,* 2017). Climatic fluctuations from the Pleistocene (i.e., last 3 Myr) had a detrimental effect on biodiversity, contributing to widespread extinctions of flora and fauna (Hughes *et al.,* 2013). However, regions that experienced climatic stability during that period were less affected, leading to the accumulation of TSR, paleoendemics and neoendemics (Keppel *et al.,* 2012; Harrison and Noss, 2017). Similarly, hydrologically stable habitats (like seepages and bogs) also facilitate the persistence and distribution of TSR and NREs especially in regions experiencing increasing aridification (Morelli *et al.,* 2016; McLaughlin *et al.,* 2017). These habitats maintain stable all year-round hydrological flow which provides cooler and wetter conditions that enable species to survive effects of climate change, essentially decoupling the habitat from regional climate fluctuations (McLaughlin *et al.,* 2017; Cartwright and Johnson, 2018). Therefore, climatically and hydrologically stable environments are expected to support the persistence of species with narrow ranges, despite otherwise unfavourable surrounding environments (Rull, 2009; Sandel *et al.,* 2011; Keppel *et al.,* 2012; MoruetaHolme *et al.,* 2013; Hannah *et al.,* 2014; Feng *et al.,* 2016).

Numerous studies have highlighted the use of GIS climate grids (Hijmans *et al.,* 2005), and remote sensing tools for delimiting sites of hydrologic stability (Cartwright and Johnson, 2018). These tools can be used along with models to make inference on species distribution, endemism and refugia (Harrison and Noss, 2017). However, the risk of using GIS and remote sensing tools is that they often fail to capture the climate-forcing factors generated by habitat features that decouple regional and local conditions (Grotch and MacCracken, 1991; Ashcroft, 2010; Dobrowski, 2011). While, GIS and remote-sensing data are valuable, they may not fully capture the intricate, localized climate-forcing factors that are essential for identifying true microrefugia (Franklin et al., 2013; Keppel *et al.,* 2017). Therefore, integrating detailed field studies and local floristic data is necessary to complement broader scale climate models and improve our understanding of how environmental stability potentially shapes biodiversity and refugia (Cartwright and Johnson, 2018; Linder, 2019; Colville *et al.,* 2020).

Environmental heterogeneity

Another potential driver of NRE richness and species richness is environmental heterogeneity. This is the spatial or temporal variation in environmental conditions such as soil composition, microclimatic conditions, and other physical and biological factors within a given area (Ricklefs, 1977). A physically heterogeneous region directly influences species richness by providing a variety of ecological niches that different species can exploit (Kreft and Jetz, 2007; Laliberte *et al.,* 2014; Stein *et al.,* 2014). This creates opportunities for niche partitioning, reduce competition, enhances evolutionary processes like speciation, increases habitat complexity (Rosenzweig, 1995; Antonelli and Sanmartín, 2011), and promotes both rare and endemic species (Stebbins and Major, 1965). Heterogeneous regions are more likely to buffer species from adverse environmental conditions and periods of climate change, which in turn promotes persistence (Kallimanis et al., 2010).

A spatially heterogeneous region has the potential to function as a refugium (Jansson, 2003; Dobrowski, 2011). However, delimiting refugia and the extent to which environmental heterogeneity contributes to these stable sites can be challenging, unless the refugia are strongly associated with distinct environmental features, such as mountains and deep gorges (Médail and Diadema, 2009). These sites may enable NREs to survive periods of climatic oscillation while steep environmental gradients reduce the distance required to disperse into more favourable environments (Dobrowski, 2011). This buffering effect of heterogeneous landscapes on climate can be quantified using a metric termed climate change velocity (Loarie et al. 2009). Climate change velocity is lowest where stable climates coincide with mountain ranges, this is important to consider, because this is often used as an approach to track species distribution shifts linking them to climatic variables and topography (Loarie *et al.,* 2008; 2009). There are other important aspects of environmental heterogeneity that are not climate related (Harrison and Noss, 2017). Heterogeneity in environments is often measured using a ruggedness index, most often applied to terrain/topography, but which can be applied to other variables such as climate or soils, along with measures of the range and roughness of the environmental variables (Cramer and Verboom, 2017).

Environmental productivity

Environmental productivity sometimes referred to as the "water and energy hypothesis" can potentially explain patterns of TSR and endemism (O'Brien, 1993; 1998). For example, habitat productivity is sometimes used to explain the latitudinal TSR gradients observed globally, where the number of species in an area is associated with the ambient available environmental energy (Gaston, 2000). However, the relationship between environmental productivity and TSR is dependent on the spatial scale at which this relationship is studied. For instance, at a relatively local scale, there is a marked tendency for a general hump-shaped relationship between TSR and available energy, with TSR increasing from low to moderate levels of energy and then declining again towards high levels of productivity when a sufficient range of energy values are recorded (Rosenzweig and Abramsky, 1993; Rosenzweig, 1995; Rosenzweig and Sandlin, 1997). The best correlates of productivity for plants are measures of both heat and water (such as net primary productivity, mean annual temperature and potential evapotranspiration; Wright, 1983; Currie *et al.,* 2004).

There are several potential mechanisms that link TSR and endemism patterns to productivity. A possible mechanism is the more-individuals hypothesis (MIH; Currie *et al.,* 2004), which states that higher energy availability promotes a higher total number of individuals in a community, which consequently increases species richness by allowing for a greater number of species with viable populations (Srivastava and Lawton, 1998). The MIH is also known as competitive exclusion theory (Grime, 1973), suggests that TSR within a community follows a U-shaped pattern in relation to productivity. At lower levels of productivity, there is insufficient energy to support a high number of species, limiting richness, while, at higher productivity levels, competition intensifies, which can suppress TSR, as dominant species outcompete others (Currie *et al.,* 2004). Another possible mechanism that may drive TSR and NREs through productivity is the rate-of-evolution theory (Rohde, 1992). This theory states that evolution occurs more rapidly in more productive environments and that this leads to greater rates of speciation and species accumulation (Rohde, 1992; Losos and Schulter, 2000; Hubbell, 2001; Gillooly et al. 2005). All two theories agree at high productivity NREs can potentially maintain larger populations (Mittelbach et al., 2001) and be more resistant to extinction and indirectly drive genetic isolation between local populations (Rohde, 1992; Willig et al., 2003).The Cape Floristic Region of South Africa

The Cape Floristic Region (CFR) of South Africa has exceptional angiosperm species diversity with 9400 species in an area of 90,000 km^2 and 68.8% endemism at the species level (Manning and Goldblatt, 2012). While high plant species richness is typical of Mediterranean-type ecosystems (Cowling *et al.,* 2015), the floristic richness and endemism of the CFR is exceptional among these, being equal to or higher than comparable areas of neotropical ecosystems (Latimer *et al.,* 2005; Rundel *et al.,* 2016). The high endemism of the CFR flora is associated mostly with fynbos, which is the dominant and most species-rich vegetation in the region (Cowling *et al.,* 1992), covering half of its geographic area and accounting for more than 80% of its species (Goldblatt, 1978; Linder, 2003).

The distribution of diversity and endemism in the CFR is not spatially homogeneous (Cowling and Holmes, 1992; Cowling *et al.,* 1997; Cowling *et al.,* 2017). There are well documented patterns of richness at a regional scale, the most notable of these being a concentration of species in the west with a decline toward the east, this being attributed to a longitudinal gradient in climatic stability (Levyns, 1964; Oliver *et al.,* 1983; Cowling *et al.,* 2017). There is also a topographic pattern, with species richness being greater at high elevations, these areas being cooler, moister, and more climatically buffered (Cowling and Holmes, 1992; Cowling *et al.,* 1997). These longitudinal and elevational diversity gradients observed are heavily influenced by rainfall as well as seasonality (Deacon *et al.,* 1992). Cowling *et al.,* (1992, 1997), showed that western (winter-rainfall) landscapes (i.e. west of 21°E), had more than double the species richness of eastern (non-seasonal rainfall) landscapes across all area sizes. Linder (1991) showed that the diversity in quarter-degree squares (QDS; 634–671 km^2) was significantly positively correlated with elevation and range in rainfall, both measures of environmental heterogeneity. Linder (2019) also showed that range size variation in Restionaceae species was largely associated with conditions of climatic stability. These studies highlight how heterogeneity as well as climatic stability has driven TSR as well as NRE richness patterns in the CFR (Van Mazijk *et al.,* 2021; Wuëst *et al.,* 2019; Linder, 2019). However, like Mediterranean ecosystems elsewhere, the CFR is vulnerable to the current climate change occurring in the Anthropocene.

Climate data from the last century have shown that the CFR has become significantly warmer (Midgley *et al.,* 2003; Midgley *et al.,* 2005; New *et al.,* 2006) and is likely to depart from its historical limits in the coming century (Mora, 2013). Slingsby *et al.,* (2017) showed a climate-driven decline in diversity and selection for species that are tolerant of hotter temperatures. Thus identifying, mapping, monitoring, and forecasting species composition and ecosystem function in specialised habitats like the CFR's fynbos region is critical for conserving the diverse Cape flora (Cowling and Lombard, 2002; Linder, 2019). Of particular importance are refugial sites which, while functioning as important sites of species persistence in the past, are not immune to the threat of global change, with moist microenvironments that may act as hydrological refugia threatened by extreme regional aridification, invasive species and water abstraction by humans (Slingsby *et al.,* 2021; Boon *et al.,* 2023). Assessing the importance and location of these potential refugia, and how they may change over time, will help us highlight specific locations for conservation action.

Study aims

The goals of this study are twofold. First, I seek to identify areas that may have historically served as refugia (macro and micro) and may continue to do so in the future. To achieve this, I will create distribution maps of TSR, widespread (WS), NRE and wetland-dependent NRE (WD-NRE) richness. I anticipate that the results will align with previous studies that have shown there to be higher species richness and endemism in the southwestern Cape (Cowling and Holmes, 1992; Cowling *et al.,* 1997; Cowling and Lombard, 2002; Cowling *et al.,* 2017; Linder, 2019).

Second, I evaluate the degree of covariation of WS, NRE and WD-NRE richness with TSR to see if the factors determining variation in their distribution are similar or different among the groups.

Finally, to understand the factors that determine the richness of the different species sets and to what extent these factors suggest a role of macro- or microrefugia, I test their relative dependence on variables relating to environmental stability, environmental heterogeneity, and productivity. Environmental stability has been documented to indicate patterns of refugia (Harrison and Noss, 2017), thus we expect NREs and WD-NREs patterns in the CFR to be driven by climatically and hydrologically stable sites, while we expect TSR and WS be driven by environmental heterogeneity and climatic stability. To explore these relationships, I fit a Bayesian hierarchical model to assess the correlations between environmental variables and TSR, WS, NRE and WD-NRE richness, while accounting for spatial autocorrelation.

Methods

Species data

Data assembly and preparation

A list of NRE vascular species in the CFR was assembled from various data sources. The primary data source was the Plants of Southern Africa (POSA) online database, maintained by the South African National Biodiversity Institute (SANBI). This database encompasses all vascular plant species, derived from vegetation sampling and monitoring programs, as well as herbarium collections. It contains 25163 unique species across 7440 unique quarter-degree squares (QDS) in southern Africa. Only species with accurate and precise (<10 000 m precision) localities were selected within each QDS using the *filter()* function in the tidyverse package (v2.0.0; Wickham *et al.,* 2019). The unit of sampling for this study was the QDS, which varies in size and is roughly 27.4 km long and 27.4 km wide (QDS area roughly 634–671 km^2) in South African latitudes. This sampling unit was used to ensure a consistent and standardised representation of spatial distribution.

Although microrefugia are typically understood to occur at much finer spatial scales (<1 km), the use of QDS units in this study was necessitated by the resolution of available species distribution data (e.g., POSA data). While environmental variables were available at a finer 1 km resolution, they were aggregated to match the spatial resolution of the species data. This scale mismatch is acknowledged as a possible limitation; however, there is evidence that plant diversity patterns at broader spatial scales can still reflect underlying fine-scale ecological processes (Huston, 1999). Furthermore, each QDS has the potential to capture considerable range of environmental conditions, and this heterogeneity itself is a strong predictor of species and endemic richness (Stebbins and Major, 1965). Therefore, despite their coarser resolution, QDS units may still be useful for identifying areas that exhibit characteristics consistent with microrefugia.

To improve the completeness and representativeness of species richness data, two datasets were integrated. The first, the POSA dataset, does not include all plant species—particularly threatened species —which were censored to prevent risks such as poaching or overharvesting. It is also important to note that most NREs are threatened and are species of concern, therefore, most NREs would not be found under the POSA database, hence the inclusion of the threatened species dataset.

To address this gap, the addition records from SANBI Threatened Species Programme (TSPP)were added. This data set, which was compiled by numerous research institutions and conservation non-governmental organisations, comprised of 3239 unique species from the Western, Eastern and Northern Cape listed as endangered, vulnerable, near threatened or critically endangered according to the International Union for Conservation of Nature (IUCN) Red List Categories or Critically Rare according to the SANBI Threatened Species Programme - http://redlist.sanbi.org/redcat.php. The TSPP data was also cleaned to include only species with accurate and precise (<10 000 m precision) localities using the *filter()* function in the tidyverse package (v2.0.0; Wickham *et al.,* 2019), to ensure that the TSPP data mirrored the cleaning done on the POSA data and the two datasets had the same resolution. The POSA and TSPP databases were consolidated using the *anti_join()* function from the dplyr package (v1.0-16; Wickham *et al.,* 2023), to return a full list of species within the CFR without duplicating records that may have been in both the POSA and TSPP datasets. Finally, the data were cropped with a mask of the CFR drawn in Google Earth™ as outlined in Manning and Goldblatt (2012), to ensure the inclusion of vegetation types exclusive to the CFR. This was done by overlapping the full species data and the CFR mask using the *st_intersection()* from the sf package (v1.0-16; Pebesma, 2018), where the output was a list of all species in the fynbos biome of the CFR. To complete the cropping of these data an additional fynbos lithology layer raster file from the Rutherford *et al.,* (2006) study was used to crop the data to only have fynbos vegetation, with the *st_crop()* function in the terra package (v1.7-78; Hijmans *et al.,* 2024). The rationale here was that the Manning and Goldblatt (2012) CFR layer did not fully cover the area, so it was necessary to add the Rutherford *et al.,* (2006) fynbos layer to crop it further to bring the extents into alignment. Finally, the South African National Plant Checklist (SANPC) 40504 unique species (including Mosses, Liverworts and Hornworts) was used to qualify species as invasive or non-invasive. Invasive species in the SANPC dataset (1715 unique species), were matched to the consolidated full TSPP and POSA species list in the CFR region so that invasive plants could be removed from the latter using the *anti_join()* function.

Response variables

Four response variables were compiled, namely TSR, WS, NREs and WD-NREs richness. To define NREs I used the South African Red List definition of range-restrictedness (Extent of occurrence <500 km^2), which identified NREs as species whose ranges do not exceed four species within a QDS (< /= 4). From these data, a set of NREs were identified and the number of such species occurring in each QDS was determined using the *summarise()* function.

To classify NREs as wetland-associated the Sieben *et al.*, (2021) data was used. This data included a list of facultative (generalists) and obligate (specialised) wetland species in Southern Africa (4124 unique species). A total of 1607 species from the Sieben *et al.* (2021) data was not classified as either facultative or obligate wetland species and removed from the dataset. Species were classified as obligate wetland species (994 unique species) if they were exclusively found in wet environments. Species that were collected in both wet and dry habitats were considered facultative wetland species (1466 unique species). The comprehensive species list, derived by Sieben *et al.,* (2021), was compiled from two distinct sources to ensure its completeness. The initial source involved herbarium data, where descriptors indicative of wetland habitat (e.g., 'wetland', 'moist', 'wet') were sought on the herbarium labels. This method may have overlooked some facultative (generalist) wetland species that predominantly occur outside wetlands but still exhibit substantial coverage within wetland extents (Sieben *et al.,* 2017). To mitigate this limitation, Sieben *et al.,* (2021) incorporated a vegetation database from diverse wetland habitats across the country. This dual-source approach aimed to provide a more comprehensive and robust list of wetland dependent plant species (Sieben *et al.,* 2021).The Sieben et al., (2021) data were compared to the full list of NREs found in the CFR using the *left_join()* function in the dplyr package (v1.0-16; Wickham *et al.,* 2023), to determine NREs that are wetland-dependent. The TSR of the CFR was retrieved from Van Mazijk (2021), while the WS were the difference between TSR and NREs. Ultimately, four response variables were considered for spatial mapping and analysis: TSR, WS, NRE and WD-NRE richness. Here each QDS (a total of 192), directly corresponded to a count of the respective response variable within the CFR. Finally, spatial heatmaps were constructed using the ggplot2 package (v3.5.1; Wickham *et al.,* 2016), to give an indication of how each response variable is distributed per QDS. Scatterplots were also generated to understand the association between all response variables.

Environmental predictor variables

The covariates were split into three sets representing environmental stability, environmental heterogeneity and productivity (Table 1). There were six productivity variables, describing the availability of energy (mean annual temperature, MAT; potential evapotranspiration, PET), water (mean annual precipitation, MAP; mean precipitation in the driest quarter, MDQ), and soil fertility (pH; extractable soil [P]). These variables were chosen from WorldClim (Fick and Hijmans, 2017), Climatologies at high resolution for the Earth's land surface (Karger *et al.,* 2017; CHELSA) and Cramer *et al* (2019). There were three environmental stability variables, including the climatic stability index from HerrandoMoraira *et al.,* (2022; CSI), cloud cover frequency (CloudFq) from Wilson & Jetz (2016) and wetland percentage cover (Wetland %) from the National Freshwater Ecosystem Priority Areas (NFEPA) map 5 from Deventer *et al.,* (2020). These variables were selected to explain the stability of the environment from a historical, climatic and hydrological perspective. There were eight environmental heterogeneity variables, described by the range (the maximum and minimum mean differences of the focal pixel, averaged out within a QDS) and the roughness (the difference of the aggregated mean value between the focal pixel and its eight neighboring pixels). The “range” and “roughness” variables are known to align with environmental heterogeneity, because they account for the variability in pixels for each of the environmental variables.

Table 1: A list of all covariates used for the analysis along with their definition, description source, citation, reasoning, resolution and units.

Covariates	Definition	Description	Source & citation	Reasoning	Resolution and units	Covariate type
Average Climatic stability index (CSI)	This metric was used to assess the stability of climate conditions over periods ranging from 21000 BP (Before Present) to 2100 CE (Common Era).	CSI raster values were extracted using the *extract()* function where 36 pixels within each QDS were aggregated and then averaged to find the mean CSI within each QDS.	The data was from Herrando-Moraira *et al.,* (2022) and was stored in an online Figshare drive	This metric provided a measure of historical variability of the climate during the Pleistocene period, shaping today's climate and biodiversity.	This raster was at 1 km resolution and the variable is an index without units.	Environmental stability

NFEPA Map 5 (Wetland%)	This metric was a percentage cover of water bodies (rivers, seeps, flood plains and depressions) within each QDS. They represent groundwater systems.	The wetland percentage shapefile was intersected with the NRE QDS data, and the percentage of intersection was calculated using *st_area()* to give the wetland percentage per QDS.	NFEPA map 5 was derived from Van Deventer *et al.,* (2020) and was stored in an online repository managed by SANBI.	This metric provided a proxy measure for groundwater or hydrology, which has the potential to become refugia hydrological.	This geospatial vector was at 1: 5000 scale and presented as a percentage cover per QDS	Environmental stability

Average summer Cloud frequency (CloudFq)	This metric was the mean remote sensing-derived cloud cover frequency for the summer months from December to February. This is from 2000 - 2014	Mean cloud fq rasters were stacked (Dec, Jan, Feb) and averaged, and values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean cloud fq within each QDS.	Cloud frequency rasters were from Wilson & Jetz (2016) and were stored in an online repository known as EarthEnv	Provided a proxy measure for potential moisture input and reduced evapotranspiration due to shading. This was a measure of spatial cooling within a QDS	This raster was at 1 km resolution, while the variable was a percentage frequency per QDS	Environmental stability

Average Mean Annual Temperature (MAT)	This metric was the average maximum and minimum temperatures of the hottest and the coldest months of the year from 19702000.	MAT raster values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean MAT within each QDS	Average MAT raster data were from Fick and Hijmans, (2017), while the weather station data were stored in [Worldclim]	This metric provided a proxy for the available energy within a QDS.	This raster was at 1 km resolution, while the variable units were in Celsius (°C)	Productivity

Average Mean annual precipitation (MAP)	This metric was the average total amount of precipitation recorded during a given year from 1970-2000.	MAP raster values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean MAP within each QDS.	Average MAP raster data were from Fick and Hijmans, (2017), while the weather station data were stored in Worldclim	Provided a proxy for water available within a QDS.	This raster was at 1 km resolution, while the variable units were in millimetres (mm)	Productivity

Mean evapotranspiration (PET)	This metric was the average value of water that would evaporate and transpire from plants, soil, water bodies, and snow from 1979 to 2018.	PET raster values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean PET within each QDS	Average MAP raster data were from Karger *et al.,* (2017), while the weather station data were stored in [CHELSA]	Provided a proxy for average water loss within a QDS to the atmosphere.	This raster was at 1 km resolution, while the variable units were in millimetres (mm)	Productivity

Average soil extractable Phosphorus (P)	This metric was a measure of P in the soil that can be readily dissolved by plants.	Average P raster values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean P within each QDS.	Average P raster data was from Cramer *et al.,* (2019), these data were stored in an online repository in the Supplementary information	This represents the mean concentration of phosphorus in the soil which is a limiting resource in soil nutrients.	This raster was at 1 km resolution, while the variable units were in mg/kg.	Productivity

Average soil fertility (pH)	This metric was a measure of the acidity or alkalinity of the soil, which is a key factor influencing the availability of nutrients to plants.	pH raster values were extracted using the *extract()* function where 900 pixels within each QDS were aggregated and then averaged to find the mean pH within each QDS	Average pH raster data was from Cramer *et al.,* (2019), these data were stored in an online repository in the supplementary information	Fynbos ranges from acidic sandstone to limestone. Therefore, our modelling efforts need to be able to accommodate this variability in soil fertility in fynbos.	This raster was at 1 km resolution, while the variable units were in pH scale.	Productivity

Range in Mean Annual Temperature	This metric was the difference between the highest and lowest MAT pixel values within each QDS.	Range MAT values were extracted from the MAT raster using the *extract()* function where 900 pixels within each QDS were aggregated, then ranged to indicate heterogeneity of the pixels MAT within each QDS.	Range MAT raster data was from Fick and Hijmans, (2017), these data were stored in an online repository in [Worldclim]()	This represented the spatial variability in the highest MAP pixel and the lowest MAP pixel within each QDS.	This raster was at 1 km resolution, while the variable units were in Celsius (°C)	Environmental heterogeneity

Roughness dry season rainfall	This metric was the difference in MDQ of the focal pixel and eight neighbouring pixels within a QDS.	Roughness in MDQ values was obtained by using the *terrain()* function, these are absolute differences found between one focal pixel and eight others, then *extract()* function was used to get the values where 900 pixels within each QDS were aggregated, then averaged to find the roughness of the MDQ within each QDS.	Roughness in MDQ raster data was from Fick and Hijmans, (2017), these data were stored in an online repository in [Worldclim]()	This represented the spatial variability of all the pixels relative to each of its adjacent pixels for values of MDQ within a QDS.	This raster was at 1 km resolution, while the variable units were in millimeters (mm)	Environmental heterogeneity

Roughness Potential evapotranspiration (PET)	This metric was the difference in PET of the focal pixel and eight neighbouring pixels.	Roughness in PET values was obtained by using the *terrain()* function, these are the absolute differences found between one focal cell and eight others, then *extract()* function was used to get the values where 900 pixels within each QDS were aggregated, then averaged to find the roughness of the MDQ within each QDS.	Roughness in PET raster data was from Fick and Hijmans, (2017), these data were stored in an online repository in [Worldclim]	This represented the spatial variability of all the pixels relative to each of its adjacent pixels for values of PET within a QDS.	This raster was at 1 km resolution, while the variable units were in millimeters (mm)	Environmental heterogeneity

Analysis

Model fitting

A regression approach was used to explore the relationship between the four response variables (TSR, WS, NREs and WD-NREs) and the complete set of covariates (full model) and subsets by covariate type (stability, productivity and heterogeneity). To develop a robust model, assumptions and biases were sequentially tested and accounted for. These included collinearity among the covariates, the assumption of equal mean and variance (for the use of a Poisson distribution), and spatial autocorrelation. I began with pairwise correlation comparison between all environmental covariates using the *cor()* function. Then proceeded to removed covariates which were correlated with $|r| > 0.7$. To conclude I fit a Bayesian hierarchical model with spatially correlated random effects for each of the response variables. A Generalised Linear Model (GLM) with a Poisson distribution was used since the response variables are count data (Nelder and Wedderburn, 1972). The Poisson GLM structure and function was as follows:

$$Y \sim \text{Poisson}\ (u_{i,}\ )$$

$$\text{Var}\ (Y_i) = \mu_i$$

$$\mu_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon \qquad (1)$$

This model aimed to examine the relationship between the response variables (μ_i) and explanatory variables ($\beta_0 + \beta_1 x_1 \ldots$). For the Poisson GLM, we assume $Y \sim$ Poisson (u_i,) where the mean of the *i*th observation is allowed to vary as a function of the covariates for that observation. Var (Y_i) represents the variance of the response for the *i*th observation. Once all Poisson GLMs were fitted using the *glm()*function*,* a Likelihood Ratio Test was conducted, using the *anova()* function, to test whether the count data exhibited more (overdispersion) or less variability (underdispersion) than would be expected under a Poisson distribution. Since the data showed more variability (overdispersion) than expected under a Poisson distribution, it was necessary to fit a Quasi-Poisson GLM, with the use of the *glm()* function, which relaxes the assumption of equidispersion (equal mean and variance) by adding an extra parameter (θ) to scale the variance (Ver Hoef and Boveng, 2007).

The Quasi-Poisson GLM structure and function was as follows:

$$Y \sim \text{Quasi-Poisson } (u_{i,}\ )$$

$$\text{Var } (Y_i) = \theta \cdot \mu_i$$

$$\log(\mu_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n i + \epsilon \quad (2)$$

Similarly, with Quasi-Poisson GLM, we assume $Y \sim$ Quasi-Poisson (u_i,) where the mean for the *i*th observation is allowed to vary as a function of the covariates for that observation. However, the relationship between the mean response (μ_i) and the predictors ($\beta_0 + \beta_1 x_1 \ldots$) was considered non-linear. To accommodate this non-linearity, a log-link function ($\log(\mu_i)$) was incorporated. This function effectively relates the mean parameter (μ_i) with the explanatory variables ($\beta_0 + \beta_1 x_1 \ldots$), creating an exponential relationship and allowing for better estimation of the parameter coefficients. This model becomes a Quasi-Poisson GLM when combined with the presence of the theta parameter (θ) that allows for adjusting the variance to accommodate overdispersion. However, the method of fitting a Quasi-Poisson GLM with a response to environmental variables was not appropriate, because all response variables and their residuals showed spatial dependency.

Testing for spatial autocorrelation

Moran's I was used to assess the extent of spatial autocorrelation (SAC) of the response variables and their residuals in the Quasi-Poisson GLM (Anselin, 1988, 2002). The *moran.test()* function from the spdep package (v1.3-5; Bivand *et al.,* 2005), was used to assess SAC, specifically to determine whether values of the respective response variables and their residuals were clustered or dispersed across the CFR. To perform the test, two main inputs were used: first, a numeric vector containing the values of the response of each interest for each QDS, and second, a list of spatial weights, which defined the relationships between QDS units and how strongly they are connected. The spatial weights help the function understand the spatial structure of the data, allowing it to test if nearby areas tend to have similar or dissimilar values. If significant SAC is found, it suggests that the variable of interest exhibits spatial patterns that might warrant further investigation. The data was further investigated using the Lagrange's Multiplier (LM) diagnostics with the *lm.LMtests()* function

(Anselin et al., 1996). This diagnostic test was used to determine the effect of five separate spatial terms on the spatial dependency exhibited by the response and its residuals. The five types of spatial terms assessed were: a spatial error (LMerr), a spatial lag (LMlag), a more robust LMerr (RLMerr), a more robust LMlag (RLMlag) or a spatial autoregressive moving average (SARMA; Bivand *et al.,* 2008). The LM tested against the null hypothesis of no SAC, where a lower p-value indicated stronger evidence for SAC.

Modelling spatial dependence

The observed response variables (TSR, WS, NRE richness, and WD-NRE richness) and their residuals, all exhibited SAC. Failing to deal with SAC in responses and residuals — is problematic, because the model coefficients have inflated standard errors and wider confidence intervals. The most common remedy used to account for SAC in the response and their residuals is to augment a linear predictor with a set of spatially autocorrelated random effects (Besag *et al.,* 1991). This was done by computing a Bayesian hierarchical model, which is a statistical model that incorporates multiple levels of random variables, where parameters at one level are treated as random variables whose distributions are influenced by parameters at a higher level and estimated in a Bayesian framework (Fei-Fei and Perona, 2005).

Our random effects were represented with a conditional autoregressive model (CAR, Besag *et al.,* 1991), which induces SAC through the adjacency structure of the QDS. The CAR prior used in this study was developed by Leroux *et al.,* (2000). This Leroux prior, forced the random effects to exhibit a single global level of SAC, ranging from independence through to strong spatial smoothness. This model was fitted using *S.CARleroux()* from the CARBAYES package (v6.1.1; Lee *et al.,* 2018). This function made use of the Markov chain Monte Carlo (MCMC) algorithm to sample the posterior distribution of the Bayesian model. In this study, we used three chains (chains = 3), that run concurrently, to assess that the chains converged to the same distribution. The MCMC sampler was used to sample 1 000000 samples (n.sample = 1 000000), of which the first 500 000 were discarded as burn-in. To reduce autocorrelation (Besag *et al.,* 1991), the post-burn-in samples were further rarefied by sampling every 500th sample (thin = 500), to give a total of 1000 samples. Model convergence was assessed by plotting the model's β samples of each covariate for one thousand iterations (N = 1000), to see whether the 3 MCMC chains converged (Fig. A2). Once the model has converged, the individual chains of the MCMC sampler should produce samples that look like bell curves (normal distributions). These distributions represent the posterior distribution of the covariates (β).

The model structure and function of the conditional autoregressive Bayesian hierarchical (CAR BYH) model is as follows:

$$Y \sim \text{Poisson}\ (\mu_{i,}\ )$$

$$\beta \sim N\ (\mu_\beta, \Sigma_\beta)$$

$$\tau^2 \sim \text{Inverse-Gamma}\ (\alpha_\tau, \beta_\tau)$$

$$\phi_i|\phi_{-i}\tau^2 \sim N\left(\frac{\sum nj=1\ \omega ij}{\sum nj=1\ \omega ij}, \frac{\tau 2}{\sum nj=1\ \omega ij}\right) \quad (3)$$

Here, the parameters (β_i) were distributed normally ($\beta \sim N\ (\mu_\beta, \Sigma_\beta)$), with the sum of weights ($\sum^n_{j=1} \omega_{ij}$) and defined the relationship between a unique focal QDS and its neighbours as well as its prior mean (μ_β) as a weighted average. The weights (ω_{ij}) defined the spatial relationship between the QDS area and its neighbours; the prior mean was a weighted average of the neighbouring QDS. The prior means sampled by the MCMC algorithm had their variance (τ^2) follow an inverse-gamma distribution with shape parameter (α_τ) and scale parameter (β_τ). The distribution of the variance of the prior means was important to consider because it guides the model towards more realistic estimates of variance while accounting for the spatial random effect (ϕ_i). Ultimately, our BYH model ($\phi_i|\phi_{-i}\tau^2$) was normal distribution for the ϕ_i given all other spatial random effects (ϕ_{-i}) and the τ^2. The μ_β of this conditional distribution is a weighted average of the neighbouring random effects, and the variance is inversely proportional to the sum of the weights. This structure enforces spatial smoothness, implying that ϕ_i is more similar to its neighbours' values if the weights are higher. This helped in drawing more accurate and meaningful inferences about the response in the CFR.

Model performance and plotting

For each response variable, the best models were selected using the deviance information criterion (DIC) or the change in the deviance information criterion (ΔDIC) and the log marginal predictive likelihood (LMPL). The DIC selected the most appropriate model (with the lowest value) that best explains of the response variable given the observed patterns, while ΔDIC was used to compare model performance between models with the same response variable, where a ΔDIC of two units or more was considered better than the alternative models. The LMPL was used to quantify how well a model predicts new data, considering the uncertainty in the model parameters, where higher values of LMPL indicated that the model was better at predicting new data, while lower values suggest poorer predictive performance.

Reproducibility

The entire study is fully reproducible, as both the data and R code used in the analysis are publicly available. These resources have been securely stored in a GitHub repository (https://doi.org/10.5281/zenodo.14608156) to ensure easy access for researchers and students. The repository includes well-documented and clearly annotated R scripts, which provide step-by-step instructions for reproducing all analyses presented in the thesis. Additionally, raw and processed data files are available for download, with the exception of the TSP data, which has been made exclusive due to the species' vulnerability to poaching. The reproducibility of this study enables others to verify the results and build upon the work for further research or application. By making these resources publicly available and reproducible, the study fosters academic collaboration, knowledge sharing, and transparency.

Results

Patterns of species and endemism richness

The complete dataset, which consisted of 9219 vascular plant species within the CFR, revealed notable spatial variation in TSR. When exploring this variation, the Cape Town region (QDS 3318CD), Cape Peninsula (QDS 3418BB), and the Hottentots-Holland Mountain range (QDS 3318DD) emerged as the most species-rich areas, each with over 2000 species recorded per QDS. (Fig. 1a). Similarly, 4231 WS were found within the CFR, with most of these species also found in the southwestern Cape region (Fig. 1b). Our study further identified 4988 NREs in the CFR (Fig. 1c) distributed across 186 QDS. Similar to TSR and WS, we found a high concentration of NRE species in the Cape Peninsula, the Cape Town region and the southwestern region, where more than 500 species were found per QDS. In stark contrast, only a single NRE species was recorded in the fynbos of the Namaqualand region (QDS 3118AA). Using data from Sieben *et al.*, (2021), we identified 763 unique facultative and 397 unique obligate WD-NRE species, producing a total of 1160 unique WD-NRE species in the CFR (Fig. 1d). As with the distribution of overall NRE richness, WD-NRE richness was greatest in the southwestern part of the CFR, especially on the Cape Peninsula. Overall, the distribution patterns of TSR, WS, NRE richness, and WD-NRE richness exhibit similar spatial patterns of longitudinal (highest in the east) and topographic (highest in the mountains) gradients.

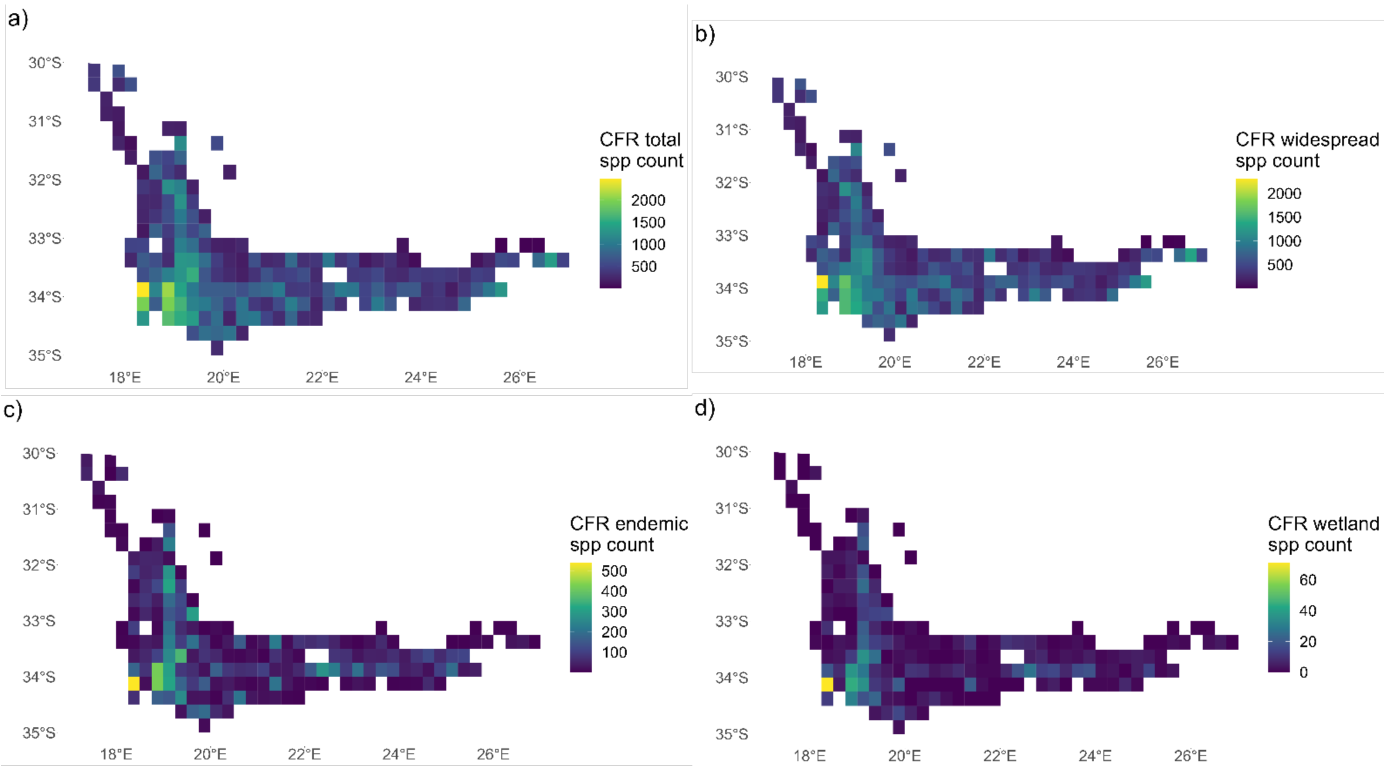

Figure 1: Spatial heat maps at a QDS-scale for the CFR of, a) the total vascular plant species richness (TSR), b) widespread species (WS), c) narrow-range endemics (NREs), and d) wetland-dependent narrow-range endemics (WD NREs), all at a spatial resolution of 1 km. The map projection used is WGS84.

Relationship among species sets

WS, NRE, and WD-NRE richness are observed to have a strong positive relationship with TSR ($\rho > 0.67$; Fig. 2). We also observe that TSR explains less than half of the variability in WS, NRE, and WD-NRE richness ($R^2 < 0.40$). A one unit increase in TSR leads to a 0.84 unit increase in the number of WS (an almost 1:1 ratio; Fig. 2a). Additionally, a one unit increase in the TSR leads a 0.16 (Fig. 2b) and 0.01 (Fig. 2c) unit increase in NRE and WDNRE species respectively. The strong relationship observed between NREs, and TSR makes it difficult to untangle the drivers of NREs and TSR.

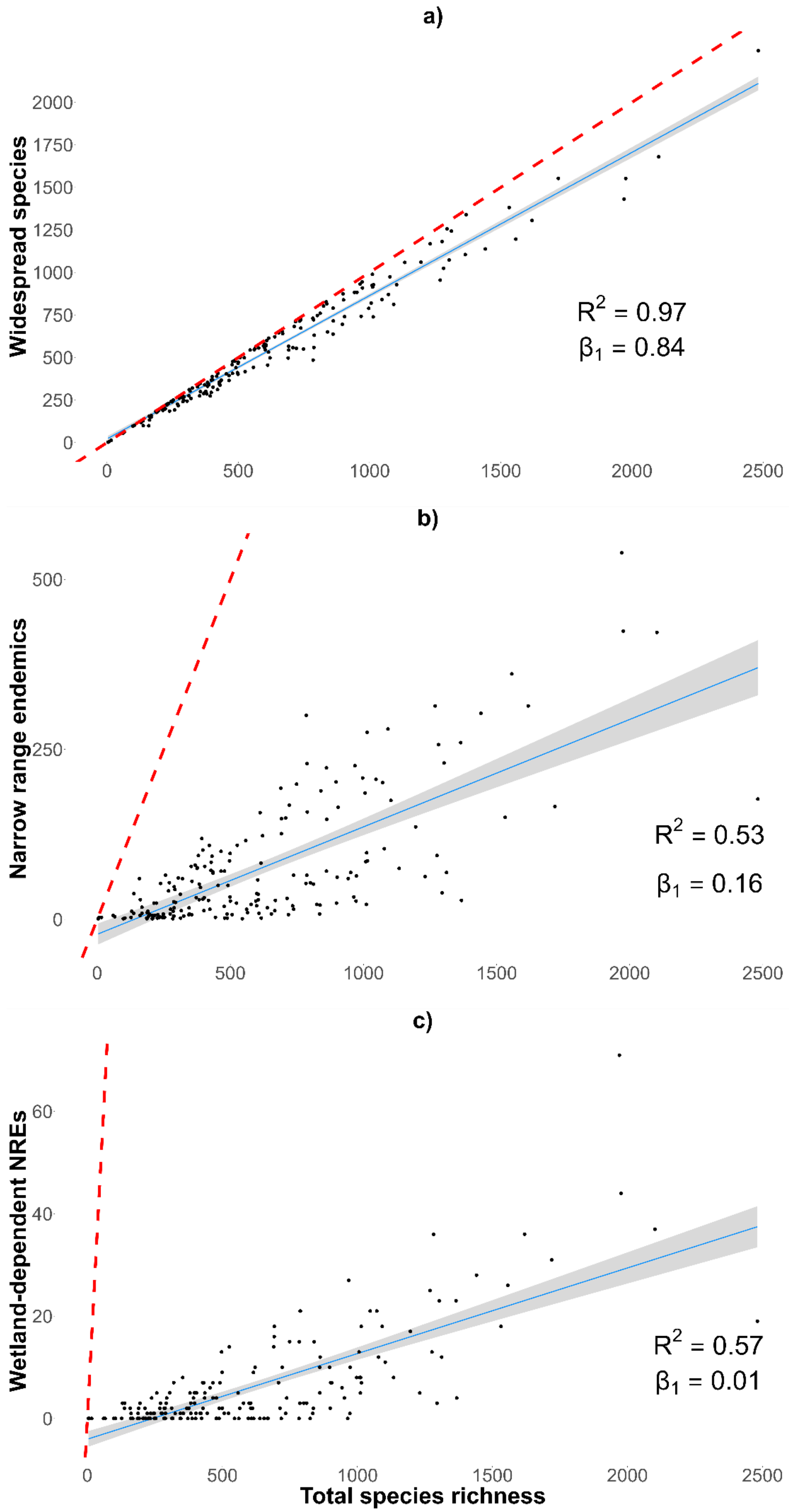
a)
Widespread species
2000
1750
1500
1250
1000
750
500
250
0
0
500
1000
1500
2000
2500
$R^2 = 0.97$
$\beta_1 = 0.84$
b)
Narrow range endemics
500
250
0
0
500
1000
1500
2000
2500
$R^2 = 0.53$
$\beta_1 = 0.16$
c)
Wetland-dependent NREs
60
40
20
0
0
500
1000
1500
2000
2500
Total species richness
$R^2 = 0.57$
$\beta_1 = 0.01$

Figure 2: The scatter plots of, a) WS vs TSR, b) NREs vs TSR, and c) WD-NRE vs TSR. These plots also depict the adjusted R^2, confidence interval, as well as the regression line, the x = y dashed line, and the slope (β), denoting the rate of change between the two variables.

Spatial variation in environmental covariates

Figure 3 shows that both CSI and Cloud Fq exhibit a longitudinal gradient, with CSI being highest in the southwestern CFR (> -0.09) and lowest in the east, and Cloud Fq highest in the east (> 50%) and lowest in the west. Additionally, wetland% is highest in the southwestern region of the CFR, particularly on the Agulhas Plain (>16%). MAT and pH are low in the mountains and higher in the lowlands (> 17 °C; > 6), while PET is highest in the northwestern cape (1500 > $kg.m^{-2}. year^{-1}$). Range in MAT is higher (> 8°C) in the Cape Fold Mountains and decreases across areas without this elevation. Roughness in MDQ increases in the west (> 0.3 mm) and decreases in the east. Similarly, the roughness in PET is higher on the west (> 6 $kg\ m^{-2}\ year^{-1}$) and lower on the east

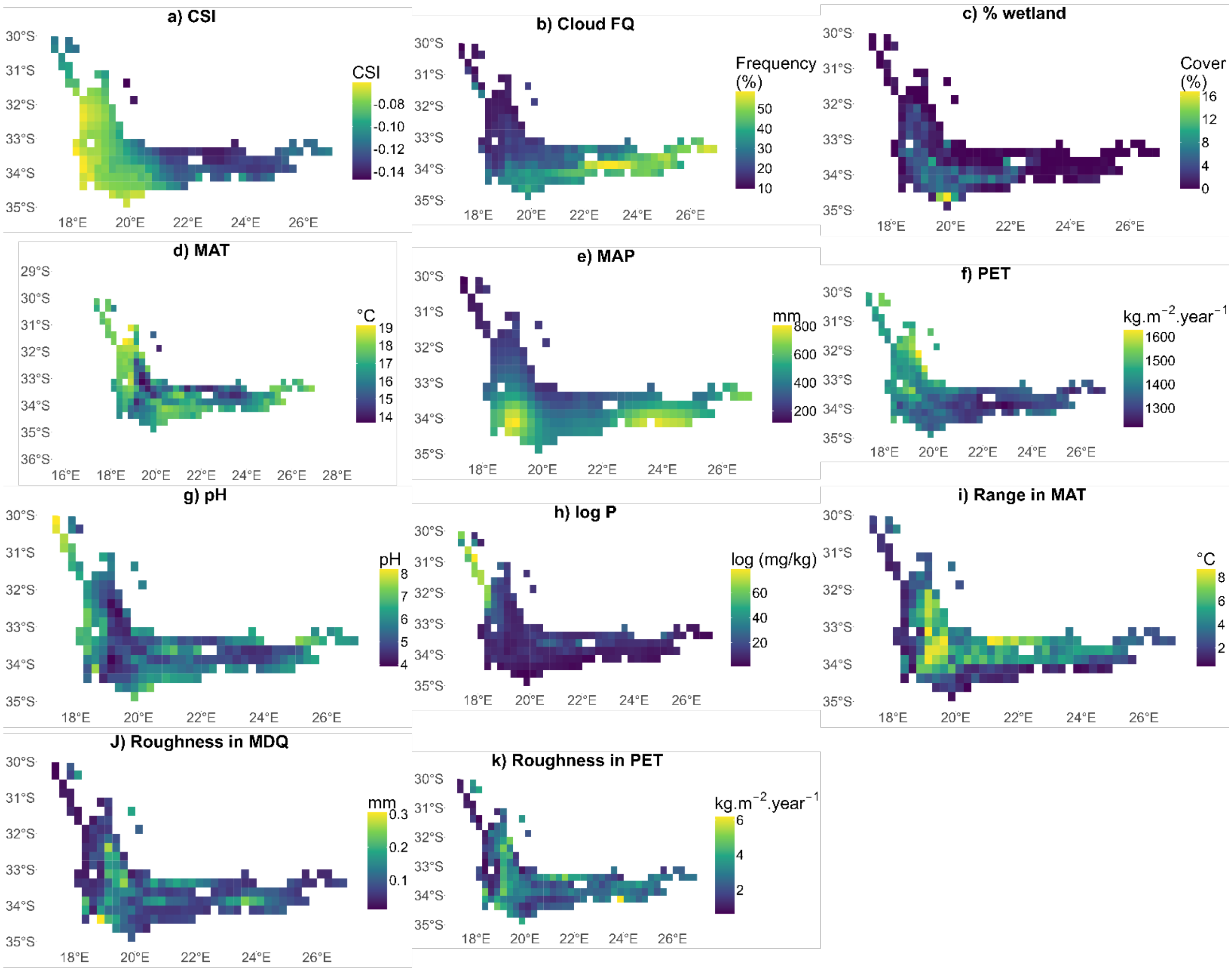

Figure 3: Maps of a) climatic stability index (CSI), b) cloud cover frequency (Cloud Fq), c) percentage wetland cover (wetland%), d) mean annual temperature (MAT), e) mean annual precipitation (MAP), f) mean potential evapotranspiration (PET), g) mean pH, h) logged mean soil Phosphorus (log P), i) range in mean annual temperature (Range in MAT), j) the roughness in the mean temperature of driest quarter (Roughness in MDQ), and k) the roughness in the potential evapotranspiration (Roughness in PET) all at a spatial resolution of 1 km.

Model performance

A Bayesian model performance table was the starting point of our analysis to test the correlates of TSR, WS, NREs and WD-NREs against environmental stability, heterogeneity and productivity. The best model was used in our CAR BYH model, where significant covariates signal a strong association with the respective response variables. Below (Tab. 2) provided outputs of a model comparison of all response variables. For each model, the deviance information criterion (DIC), delta DIC (ΔDIC) and log marginal predictive likelihood (LMPL) are shown.

Table 2: The table below is for model performance for various model types, including environmental stability, productivity, environmental heterogeneity, and the full model. The best models (ΔDIC < 2) are italicised according

to a low DIC, ΔDIC and higher LMPL values, while models marked with an asterisk (*) failed model diagnostics, indicating that the three Markov Chain Monte Carlo (MCMC) chains did not converge during sampling from the prior distributions.

Model	DIC	ΔDIC	LMPL	Response
Stability	*1900.808*	*0.00*	*-1009.28*	*TSR*
Productivity	1902.149	1.34	-1006.21	TSR
Heterogeneity	1901.731	0.92	-1014.00	TSR
Full	1901.473	0.66	-1016.49	TSR
Stability	1878.133	0.56	-1009.34	WS
Productivity	1878.396	0.82	-1004.88	WS
Heterogeneity	*1877.571*	*0.00*	*-996.2*	*WS*
Full	1877.857	0.28	-1005.08	WS
Stability	1389.415	8.07	-765.08	NREs
Productivity	1494.859	113.51	-2257.07	NREs

Heterogeneity	1386.415	5.07	-753.68	NREs
Full	*1381.347*	*0.00*	*-751.3*	*NREs*
Stability	821.8767	16.63	-461.43	WD-NREs
Productivity*	3455.447	2650.20	-67628.55	WD-NREs
Heterogeneity	813.515	8.27	-455.51	WD-NREs
Full	*805.246*	*0.00*	*-440.69*	*WD-NREs*

The full models (which included all the covariates) best explained NRE and WD-NRE richness, therefore these models were deemed suitable for explaining their spatial richness patterns and distribution in the CFR. For TSR and WS, environmental stability and heterogeneity respectively best described their spatial richness patterns and distribution in the CFR, however the full model for both response variables do not lag far behind (ΔDIC < 2). Therefore, the full model (encompassing call covariates) was chosen as the suitable model to be used to fit the fit the CAR BYH model for all response variables because it displayed the same level of support for across all response variables (ΔDIC < 2). The productivity model for WD-NREs is marked "*", because the three Markov chains did not converge, resulting in unreliable estimates (Fig. A5).

Bayesian model outputs

The full model for each response variable were plotted, allowing for easy comparison of similarities and differences in the effects of covariates across responses. For environmental stability, we observe a significant positive effect for wetland across all response variables. CSI was positive for all response variables, but only significant for endemic richness (NRE and WD-NRE). While Wetland% is also positive for all response variables but proving to be important for TSR, WS and NREs, while WD-NRE being nearly significant. The productivity covariates showed varying effects, with no effect for P, a significant negative effect of pH on endemics as well as a positive effect for MAP across all response variables. MAT showed weak positive effects for TSR and WS while negative effects were observed for endemic richness. PET had little effect on endemic richness and significant negative effects for TSR and WS. Lastly, two heterogeneity covariates, roughness in MDQ and roughness in PET had positive effects across all response variables, with range in MAT displaying significant positive effects for all response variables except for WS.

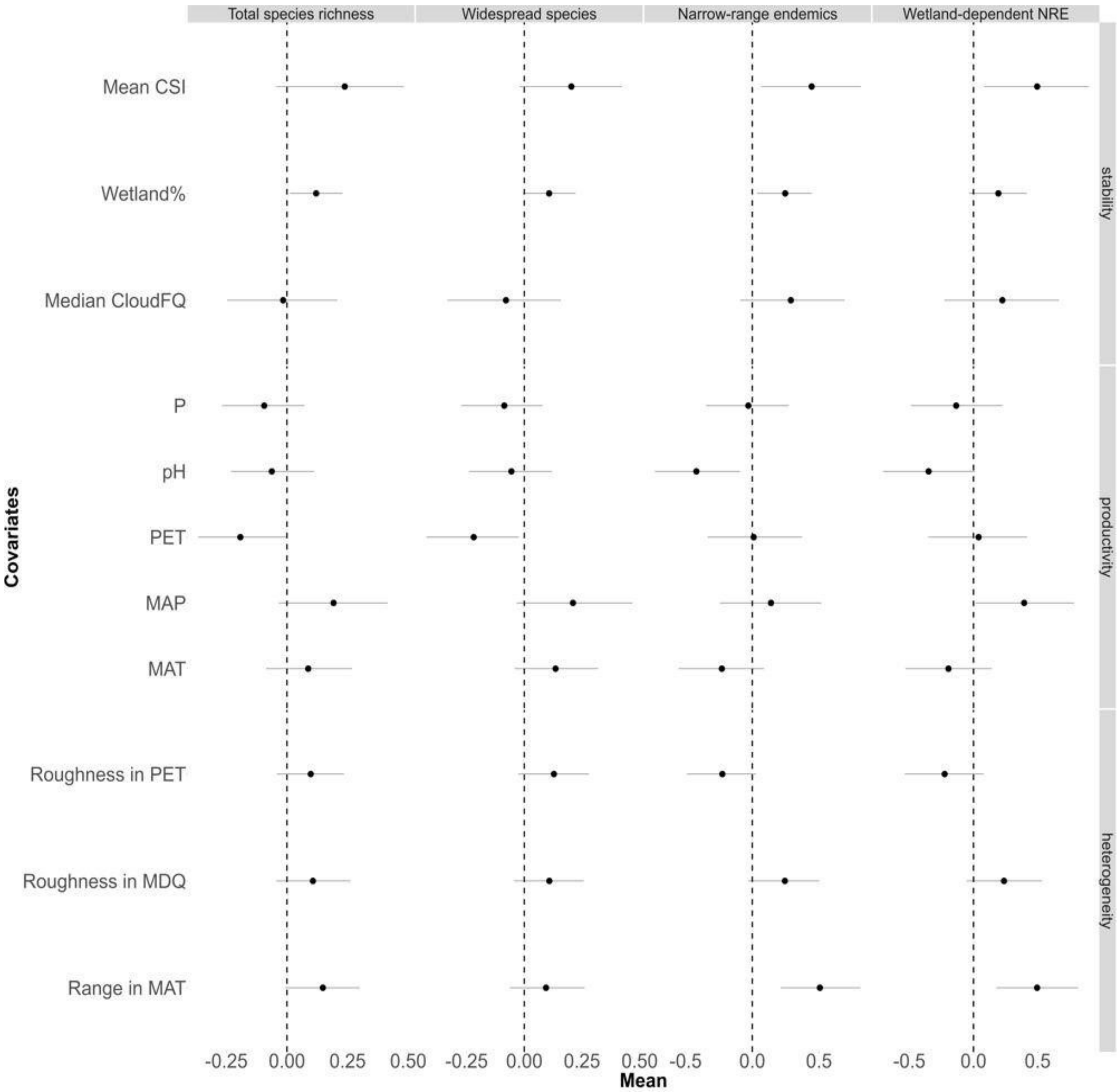

Figure 4: The Bayesian hierarchical conditional model posterior means, 95% confidence intervals and all 11 covariates compared to four response variables: TSR, WS, NREs and WD-NREs. The posterior means represent the covariates' effects on the respective response variables.

Discussion

Understanding the character and distribution of refugia and the relative role of micro and macrorefugia on species richness and endemism is crucial, especially in biodiversity hotspots like the CFR of South Africa that are threatened by global change. Our study revealed distinct longitudinal and topographic gradients for all response variables studied—TSR, WS, NREs, and WD-NREs (Fig. 1). Higher richness is observed in the mountainous southwestern region of the CFR, particularly in the Cape Town area, the Cape Peninsula, and the Hottentots-Holland Mountain ranges. Our analysis highlighted the importance of environmental stability in driving patterns of endemic richness, particularly for NREs and WD-NREs. Wetland% was found to be important in explaining endemic richness patterns, whereas CSI emerged as a strong determinant of TSR, WS, and NRE richness. These findings underscore the critical role of stable hydrologic and climatic microsites in species persistence, potentially acting as microrefugia. We refer to these as microsites, because they are relatively smaller, therefore support smaller refugial populations and harboured less diverse biotic communities than microrefugia (Mee and Moore, 2014). In contrast, environmental heterogeneity, the range in MAT, was important in explaining NRE and WD-NRE richness patterns in the CFR. While pH and PET, moderately explained NRE and WDNRE richness patterns as well as TSR and WS patterns respectively (Fig. 4).

TSR, WS, NRE and WD-NRE richness show some similarity in their spatial distribution, which was expected under our hypothesis. These richness patterns align with previous studies by Cowling and colleagues (Cowling and Lombard 2002; Cowling *et al.,* 2017).
The western CFR has been wetter during the last glacial period than it is today (Cartwright *et al.,* 1999; Parkington *et al.,* 2000). This explains the relatively higher TSR, WS richness and prevalence of endemics in the west (Levyns' Law; Cowling *et al.*, 2017). Beyond the east-west diversity gradient, a study on Restionaceae by Linder (2019) identified the southwestern mountains of the CFR as a "central cradle of diversification" that functioned as a source area for lineages colonizing the northwestern and eastern mountains. Linder (2019) also noted that NRE Restionaceae tend to associate with high elevations in the Kogelberg, western Langeberg and the Klein Swartberg. Our fundings corroborate Linder

(2019) as well as previous work which found exceptional species richness and endemism in the western CFR and southwestern mountains (Oliver *et al.,* 1983; Cowling *et al.,* 2017). However, >40% of the variability in NRE richness is not explained by TSR (Fig. 2b), suggesting that some of the factors influencing the genesis and persistence of NREs differ from those driving TSR. The opposite is true for WS richness, The variability in its relationship with TSR is less than 4% and exhibits a 1:1 relationship (Fig. 2a), which means the relationship between TSR and WS counts is nearly linear and strongly correlated.

The distribution and correlates of TSR and WS richness are very similar, with comparable directions and magnitude of the effects, but there are clear differences relative to NRE and WD-NRE richness. Environmental stability (+Wetland%) and environmental heterogeneity (+Range in MAT) are both important for explaining TSR patterns, while productivity (-PET) is important for explaining both TSR and WS richness. These effects are consistent with earlier work by Cowling and Holmes, (1992); Cowling, (1997) and Linder, (2019) who all identified higher elevations, especially in the southwestern CFR, as important centres of hydrological stability. Therefore, hydrologically stable sites (with high wetland%), with strong topography (reflected by range in MAT) and cool, moist conditions (low PET) favour high TSR and WS richness. For NRE and WD-NRE richness, environmental heterogeneity (+Range in MAT) and environmental stability (+CSI; +Wetland%) were the best predictors, with low productivity (-pH) in the form of soil substrate acidity being important for the persistence of endemics species (Fig 4). These patterns make sense because at a finescale, environmental heterogeneity is an important prerequisite for speciation, leading to the emergence of neo-endemic species (Verboom *et al.,* 2015; Harrison and Noss, 2017), while both stability and heterogeneity offer NRE species protection against extinction (Harrison and Noss, 2017; Jansson, 2003).

It is evident that environmental stability, heterogeneity and productivity play an important role in both richness and endemism. We observe that climatic stability is important for explaining NRE and WD-NRE richness in the CFR, we suggest that this is linked to the CFR climate during the Pleistocene (the last 1.8 million years) being relatively stable (Dynesius and Jansson 2000; Jansson and Dynesius 2002), this stability has buffered rare and endemic species and has been a key determinant of the region’s exceptional floristic richness (Cowling *et al.,* 1996, 2015; Latimer *et al.,* 2005; Kreft and Jetz 2008; Sniderman *et al.,* 2012). We also observe that hydrological stability plays an important role in TSR, WS, and NRE richness. We suggest that the southwestern mountains, which harbour a high number of NRE species and TSR, contain microsites with perennial moisture (such as seepages, bogs, and streams). These sites are often associated with climatic stability from the Pleistocene. (Linder, 2019). These microsites, along with the broader climatic stability of the region from the Pleistocene, function as both microrefugia and macrorefugia, contributing to species persistence. However, it's important to acknowledge the role precipitation (MAP), particularly winter rainfall, plays in recharging wetland reservoirs in drier periods, by infiltrating deep soils and mountainous catchments. The nearly significant effect of this productivity covariate (MAP) emphasises that these mesic environments are not entirely decoupled from the regional climate (Jobbagy et al., 2011; McLaughlin *et al.,* 2017).

Productivity (or lack thereof) also plays an important role in the persistence of endemic species in the CFR. For instance, soil fertility (pH) is key to explaining the patterns of NRE and WD-NRE richness, with higher richness linked to nutrient-poor acidic soils. These patterns may be linked to acidic soils being commonly found in mountainous quartzite areas, which are characterized by greater climatic heterogeneity, which drives speciation (Cowling *et al.,* 2009; Britton and Verboom, 2014, Verboom *et al.,* 2015, Linder, 2019). The strong association pH has with NREs and WD-NREs is linked to acidic limestone, which tends to have larger ranges due to less topographic complexity and isolation of climates (Willis *et al.,* 1996). These quartzite soils are derived from Ordovician sandstones of the Table Mountain Group (Cape Supergroup), dating back approximately 400 million years (Young *et al.,* 2004). They form part of the Table Mountain Group Aquifer and support many groundwater-fed seeps and bogs, here water is predominantly stored in the quartzites which serve as important hydrological refugia (Blake *et al.,* 2010).

However, while our study highlights the importance of these environmental drivers on TSR, WS, NREs, and WD-NRE richness, our observations are limited by levels of uncertainty associated with the environmental predictions of NRE and WD-NRE richness. We suspect that the wider confidence intervals observed for NRE and WD-

NRE richness may result from the omission of key explanatory variables, as well as a scale mismatch between our spatial unit of analysis (QDS), the resolution of the environmental data, and the finer spatial scale at which refugia—particularly microrefugia—are expected to operate. Missing such variables in our models likely increases the uncertainty of our estimates, as the absence of crucial environmental factors can lead to less precise predictions and wider confidence intervals.

The omission of key ecological variables known to influence species richness in the fynbos region—such as fire regimes (Magadzire et al., 2019)—may have limited the explanatory power of our environmental predictors for species richness and endemism. Fire plays a crucial ecological role in the CFR's fynbos vegetation, promoting species persistence and post-disturbance regeneration (Kruger and Bigalke, 1984; Magadzire et al., 2019). Much like hydrologic refugia, fire acts as a dynamic driver of biodiversity, potentially functioning as a form of disturbance-mediated refugia by creating niche opportunities that support coexistence and diversification (Cowling and McDonald, 1995; Procheş, 2006). However, fire regimes are increasingly altered by climate change, invasive species, and direct human activity, which may challenge reseeding species that require longer fire-free intervals for successful recruitment (Cowling et al., 2015). The absence of fire regime data from our model likely contributes to increased uncertainty in parameter estimates, as reflected in the wider credible intervals observed for NRE and WD-NRE richness (Fig. 4).

Another limitation of this study is the potential mismatch in spatial scale between the measuring units (QDS) and the fine-scale nature of microrefugia — where localized features that may not be fully captured at the resolution of our analysis. Nevertheless, the observed correlation TSR and finer-scale richness patterns suggests that meaningful ecological variation is still being detected at the QDS level. Furthermore, the considerable environmental heterogeneity within individual grid cells likely reflects microrefugia-like conditions, reinforcing the ecological relevance of our findings for identifying broader refugial patterns. Finally, issues related to data quality and sampling bias must also be acknowledged. Species richness records are often spatially biased toward more accessible or frequently surveyed regions (Reddy and Davalos, 2003;

Newbold, 2010). If this sampling bias aligns with environmental gradients, it can introduce systematic bias into species–environment relationships and potentially skew model outcomes. These spatial dependencies were addressed using a random spatial effect term, which helped account for spatial bias and improve the robustness of our estimates.

Conclusion

Ultimately, TSR, WS, NRE, and WD-NRE richness show a longitudinal and topographic gradient, with greater species richness in the southwestern CFR and at higher elevations. Environmental stability, through CSI and Wetland%, drives high richness by buffering species in both microrefugia (e.g., seepages and bogs) and macrorefugia (larger, climatically stable topographic features), offering localized and broader species buffering over time (Cowling and Lombard 2002; Cowling *et al.,* 2017).
While productivity (in the form of PET and pH) is important for TSR, WS, NRE, and WD-NRE richness, high persistence is driven by poor soil nutrient quality, often associated with acidic soils from the Table Mountain Group Aquifer. These soils support groundwater-fed seeps and bogs that provide moist conditions in the Anthropocene (Willis *et al.,* 1996; Blake, *et al.,* 2010). Environmental heterogeneity (in the form of Range in MAT) is only important for NREs and WD-NREs, driving high persistence of endemics possibly by facilitating speciation and coexistence through a strong topographic effect.

Our findings support Linder (2019), who suggested that some parts of the southwestern mountains (the Kogelberg region and western Langeberg) are potential refugia sites, due to the stability they exhibit. In our case, hydrological stability coupled with climatic stability has created these wetter conditions in the southwestern region since the last glacial period (Parkington et al., 2000). However, it's important to note that stable sites like these are not entirely decoupled from the regional climate, as rainfall typically plays an important role in the replenishment of hydrological refugia (Cowling and Lombard, 2002; Linder, 2019). As global temperatures rise and habitats shift, these putative refugial sites may serve as critical buffer zones that enable the persistence of genetic diversity and evolutionary potential (Harrison and Noss, 2017). Thus, identifying and protecting these sites is essential for biodiversity and ecosystem function (Keppel et al., 2015). However, in the CFR, many hydrological refugia face threats not only from groundwater abstraction but also from invasive species. Invasive plants, alter the hydrology and fire regimes of these refugial sites, which may reduce their capacity to support native species (Gallardo *et al.,* 2017). Recent proposals to abstract water from the Kogelberg and Hottentots Holland Mountains present a potentially serious biodiversity threat, especially to species associated with mesic environments (Slingsby et al., 2021). Therefore, there is a need to critically analyse the consequences of groundwater abstraction and invasive species in the context of a rapidly changing climate in NRE-rich sites in the CFR. There is a pressing need for more detailed finescale spatial analyses of refugia, particularly in mesic regions with high species richness and endemism, such as the Cape Peninsula, to guide efforts to conserve and manage these protectors of the Cape's unique floristic biodiversity.

References

Ackerly, D.D., Kling, M.M., Clark, M.L., Papper, P., Oldfather, M.F., Flint, A.L. and Flint, L.E., 2020. Topoclimates, refugia, and biotic responses to climate change. *Frontiers in Ecology and the Environment*, *18*(5), pp.288-297.

Altwegg, R., West, A., Gillson, L. and Midgley, G.F., 2014. Impacts of climate change in the Greater Cape Floristic Region. *Fynbos: Ecology, evolution, and conservation of a megadiverse region*, pp.299-320.

Anderson, S., 1994. Area and endemism. *The quarterly review of biology*, *69*(4), pp.451-471.

Anselin, L., 1988. Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. *Geographical analysis*, *20*(1), pp.1-17.

Anselin, L., 2002. Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, *27*(3), pp.247-267.

Antonelli, A. and Sanmartín, I., 2011. Why are there so many plant species in the Neotropics? *Taxon*, *60*(2), pp.403-414

Ashcroft, M.B., 2010. Identifying refugia from climate change. *Journal of biogeography*, *37*(8), pp.1407-1413.

Ashcroft, M.B., Gollan, J.R., Warton, D.I. and Ramp, D., 2012. A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Global change biology*, *18*(6), pp.1866-1879.

Austin, M.P. and Van Niel, K.P., 2011. Improving species distribution models for climate change studies: variable selection and scale. *Journal of biogeography*, *38*(1), pp.1-8.

Axelius, B., 1991. Areas of distribution and areas of endemism. *Cladistics*, *7*(2), pp.197-199.

Baldwin, B.G., 2005. Origin of the serpentine-endemic herb Layla discoidea from the widespread L. landulosa (Compositae). *Evolution*, *59*(11), pp.2473-2479.

Bendix, J., Wiley Jr, J.J. and Commons, M.G., 2017. Intermediate disturbance and patterns of species richness. *Physical Geography*, *38*(5), pp.393-403.

Bennett, K.D. and Provan, J., 2008. What do we mean by 'refugia'? *Quaternary Science Reviews*, *27*(27-28), pp.2449-2455.

Besag, J., York, J. and Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, *43*, pp.1-20.

Bivand, R., Bernat, A., Carvalho, M., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Lewin-Koh, N., Ma, J. and Millo, G., 2005. The spdep package. *Comprehensive R Archive Network, Version*, pp.05-83.

Bivand, R., 2008. Spatial Dependence: Weighting Schemes, Statistics and Models [R package spdep version 0.5-92].

Blake, D., Mlisa, A. and Hartnady, C., 2010. Large scale quantification of aquifer storage and volumes from the Peninsula and Skurweberg Formations in the southwestern Cape. *Water SA*, *36*(2), pp.177-184.

Born, J., Linder, H.P. and Desmet, P., 2007. The greater cape floristic region. *Journal of Biogeography*, *34*(1), pp.147-162.

Boon, J.S., Keith, S.A., Exton, D.A. and Field, R., 2023. The role of refuges in biological invasions: A systematic review. *Global Ecology and Biogeography*, *32*(8), pp.12441271.

Bramwell, D. and Caujapé-Castells, J. eds., 2011. *The biology of island floras*. Cambridge University Press.

Britton, M.N., Hedderson, T.A. and Verboom, G.A., 2014. Topography as a driver of cryptic speciation in the high-elevation cape sedge Tetraria triangularis (Boeck.) CB Clarke (Cyperaceae: Schoeneae). *Molecular phylogenetics and Evolution*, *77*, pp.96-109.

Bruchmann, I. and Hobohm, C., 2013. Factors that create and increase endemism. In *Endemism in vascular plants* (pp. 51-68). Dordrecht: Springer Netherlands.

Cai, L., Kreft, H., Taylor, A., Schrader, J., Dawson, W., Essl, F., Van Kleunen, M., Pergl, J., Pyšek, P., Winter, M. and Weigelt, P., 2023. Climatic stability and geological history shape global centers of neo-and paleoendemism in seed plants. *Proceedings of the National Academy of Sciences*, *120*(30), p.e2300981120.

Carnaval, A.C., Hickerson, M.J., Haddad, C.F., Rodrigues, M.T. and Moritz, C., 2009. Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. *science*, *323*(5915), pp.785-789.

Carrión, J.S., Yll, E.I., Walker, M.J., Legaz, A.J., Chaín, C. and López, A., 2003. Glacial refugia of temperate, Mediterranean and Ibero-North African flora in south-eastern Spain: new evidence from cave pollen at two Neanderthal man sites. *Global Ecology and Biogeography*, *12*(2), pp.119-129.

Cartwright, J. and Johnson, H.M., 2018. Springs as hydrologic refugia in a changing climate? A remote-sensing approach. *Ecosphere*, *9*(3), p.e02155.

Cartwright, J.M., Dwire, K.A., Freed, Z., Hammer, S.J., McLaughlin, B., Misztal, L.W., Schenk, E.R., Spence, J.R., Springer, A.E. and Stevens, L.E., 2020. Oases of the future? Springs as potential hydrologic refugia in drying climates. *Frontiers in Ecology and the Environment*, *18*(5), pp.245-253.

Chichorro, F., Juslén, A. and Cardoso, P., 2019. A review of the relation between species traits and extinction risk. *Biological Conservation*, 237, pp.220-229.

Corlett, R.T. and Tomlinson, K.W., 2020. Climate change and edaphic specialists: irresistible force meets an immovable object? *Trends in Ecology & Evolution*, *35*(4), pp.367-376.

Cornell, H.V., 1985. Species assemblages of cynipid gall wasps are not saturated. *The American Naturalist*, *126*(4), pp.565-569.

Cornell, H.V. and Lawton, J.H., 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. *Journal of animal ecology*, pp.1-12.

Cornell, H.V., 1993. Unsaturated patterns in species assemblages: the role of regional processes in setting local species richness. *Species diversity in ecological communities. University of Chicago Press, Chicago, Illinois, USA*, pp.243-252.

Cowling, R.M., 1983. Diversity relations in Cape shrublands and other vegetation in the southeastern Cape, South Africa. *Vegetatio*, *54*(2), pp.103-127.

Cowling, R.M. and Gxaba, T., 1990. Effects of a fynbos overstorey shrub on understorey community structure: implications for the maintenance of community-wide species richness. *South African Journal of Ecology*, *1*, pp.1-7.

Cowling, R.M. and Holmes, P.M., 1992. Endemism and speciation in a lowland flora from the Cape Floristic Region. *Biological Journal of the Linnean Society*, *47*(4), pp.367383.

Cowling, R.M., Holmes, P.M. and Rebelo, A.G., 1992. Plant diversity and endemism. *The ecology of fynbos: nutrients, fire and diversity*, *4*, pp.62-112.

Cowling, R.M., Rundel, P.W., Lamont, B.B., Arroyo, M.K. and Arianoutsou, M., 1996. Plant diversity in Mediterranean-climate regions. *Trends in Ecology & Evolution*, *11*(9), pp.362-366.

Cowling, R.M., Richardson, D.M. and Pierce, S.M. eds., 1997. *Vegetation of southern Africa*. Cambridge University Press.

Cowling, R.M. and McDonald, D.J., 1998. Local endemism and plant conservation in the Cape Floristic Region. *Landscape disturbance and biodiversity in Mediterranean-type ecosystems*, pp.171-188.

Cowling, R.M. and Lombard, A.T., 2002. Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. *Diversity and Distributions*, *8*(3), pp.163-179.

Cowling, R.M., Potts, A.J., Bradshaw, P.L., Colville, J., Arianoutsou, M., Ferrier, S., Forest, F., Fyllas, N.M., Hopper, S.D., Ojeda, F. and Procheş, Ş., 2015. Variation in plant diversity in mediterranean-climate ecosystems: The role of climatic and topographical stability. *Journal of Biogeography*, *42*(3), pp.552-564.

Cowling, R.M., Logie, C., Brady, J., Middleton, M. and Grobler, B.A., 2019. Taxonomic, biological and geographical traits of species in a coastal dune flora in the southeastern Cape Floristic Region: regional and global comparisons. *PeerJ*, *7*, p.e7336.

Clark, V.R., Barker, N.P. and Mucina, L., 2009. The Sneeuberg: a new centre of floristic endemism on the Great Escarpment, South Africa. *South African Journal of Botany*, *75*(2), pp.196-238.

Cramer, M.D. and Verboom, G.A., 2017. Measures of biologically relevant environmental heterogeneity improve prediction of regional plant species richness. *Journal of Biogeography*, *44*(3), pp.579-591.

Cramer, M.D., Wootton, L.M., van Mazijk, R. and Verboom, G.A., 2019. New regionally modelled soil layers improve prediction of vegetation type relative to that based on global soil models. *Diversity and Distributions*, *25*(11), pp.1736-1750.

Crimmins, S.M., Dobrowski, S.Z., Greenberg, J.A., Abatzoglou, J.T. and Mynsberge, A.R., 2011. Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. *Science*, *331*(6015), pp.324-327.

Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guegan, J.F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O'Brien, E. and Turner, J.R.G., 2004. Ideas and perspectives: Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology Letters*, *7*(12), pp.1121-1134.

Dahl, E., 1946. On different types of unglaciated areas during the ice ages and their significance to phytogeography. *The New Phytologist*, *45*(2), pp.225-242.

Dawson, T.E. and Pate, J.S., 1996. Seasonal water uptake and movement in root systems of Australian phraeatophytic plants of dimorphic root morphology: a stable isotope investigation. *Oecologia*, *107*, pp.13-20.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C. and Mace, G.M., 2011. Beyond predictions: *biodiversity conservation in a changing climate. science*, 332(6025), pp.53-58.

De Candolle, A.P., 1820. *Essai élémentaire de géographie botanique*. éditeur non identifié.

De Frenne, P., Rodríguez-Sánchez, F., Coomes, D.A., Baeten, L., Verstraeten, G., Vellend, M., Bernhardt-Römermann, M., Brown, C.D., Brunet, J., Cornelis, J. and Decocq, G.M., 2013. Microclimate moderates plant responses to macroclimate warming. *Proceedings of the National Academy of Sciences*, *110*(46), pp.1856118565.

Dietl, G.P. and Flessa, K.W., 2011. Conservation paleobiology: putting the dead to work. *Trends in Ecology & Evolution*, *26*(1), pp.30-37.

Dobrowski, S.Z., 2011. A climatic basis for microrefugia: the influence of terrain on climate. *Global change biology*, *17*(2), pp.1022-1035.

Dobrowski, S.Z., Abatzoglou, J., Swanson, A.K., Greenberg, J.A., Mynsberge, A.R., Holden, Z.A. and Schwartz, M.K., 2013. The climate velocity of the contiguous U nited S tates during the 20th century. *Global change biology*, *19*(1), pp.241-251.

Duan, H., Li, Y., Xu, Y., Zhou, S., Liu, J., Tissue, D.T. and Liu, J., 2019. Contrasting drought sensitivity and post-drought resilience among three co-occurring tree species in subtropical China. *Agricultural and Forest Meteorology*, *272*, pp.55-68.

Dynesius, M. and Jansson, R., 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proceedings of the National Academy of Sciences*, *97*(16), pp.9115-9120.

Engelbrecht, F.A., McGregor, J.L. and Engelbrecht, C.J., 2009. Dynamics of the Conformal-Cubic Atmospheric Model projected climate-change signal over southern Africa. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *29*(7), pp.1013-1033.

Eriksson, O., 1993. The species-pool hypothesis and plant community diversity. *Oikos*, pp.371-374.

Fei-Fei, L. and Perona, P., 2005, June. A bayesian hierarchical model for learning natural scene categories. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 2, pp. 524-531).

Feng, G., Mao, L., Sandel, B., Swenson, N.G. and Svenning, J.C., 2016. High plant endemism in China is partially linked to reduced glacial-interglacial climate change. *Journal of Biogeography*, *43*(1), pp.145-154.

Fick, S.E. and Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, *37*(12), pp.43024315.

Flint, L.E. and Flint, A.L., 2012. Downscaling future climate scenarios to fine scales for hydrologic and ecological modelling and analysis. *Ecological Processes*, *1*, pp.1-15.

Franklin, J., Davis, F.W., Ikegami, M., Syphard, A.D., Flint, L.E., Flint, A.L. and Hannah, L., 2013. Modelling plant species distributions under future climates: how fine scale do climate projections need to be? *Global change biology*, *19*(2), pp.473-483

Frenzel, B., 1973. B. Foreland of the Alps: 2. Some Remarks on the Pleistocene Vegetation. *E&G Quaternary Science Journal*, *23*(1), pp.281-292.

Fridley, J.D., 2009. Downscaling climate over complex terrain: high finescale (< 1000 m) spatial variation of near-ground temperatures in a montane forested landscape (Great Smoky Mountains). *Journal of Applied Meteorology and Climatology*, *48*(5), pp.1033-1049.

Gallardo, B., Aldridge, D.C., González-Moreno, P., Pergl, J., Pizarro, M., Pyšek, P., Thuiller, W., Yesson, C. and Vilà, M., 2017. Protected areas offer refuge from invasive species spreading under climate change. *Global change biology*, *23*(12), pp.5331-5343.

García, M.B., Miranda, H., Pizarro, M., Font, X., Roquet, C. and González-Sampériz, P., 2022. Habitats hold an evolutionary signal of past climatic refugia. *Biodiversity and Conservation*, *31*(5-6), pp.1665-1688.

Gaston, K.J., 2000. Global patterns in biodiversity. *Nature*, *405*(6783), pp.220-227.

Gillman, L.N. and Wright, S.D., 2006. The influence of productivity on the species richness of plants: a critical assessment. *Ecology*, *87*(5), pp.1234-1243.

Gillooly, J.F., Allen, A.P., West, G.B. and Brown, J.H., 2005. The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proceedings of the National Academy of Sciences*, *102*(1), pp.140-145.

Goldblati, P., 1978. An analysis of the flora of southern Africa: its characteristics, relationships, and origins. *Annals of the Missouri Botanical Garden*, pp.369-436.

Goldblatt, P., 1997. Floristic diversity in the Cape flora of South Africa. *Biodiversity & Conservation*, *6*, pp.359-377.

Goldblatt, P. and Manning, J.C., 2002. Plant diversity of the Cape region of southern Africa. *Annals of the Missouri Botanical Garden*, pp.281-302

Goldie, X., Gillman, L., Crisp, M. and Wright, S., 2010. Evolutionary speed limited by water in arid Australia. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1694), pp.2645-2653.

Gómez, A. and Lunt, D.H., 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. *Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity*, pp.155-188.

Grime, J.P., 1973. Competitive exclusion in herbaceous vegetation. *Nature*, *242*(5396), pp.344-347.

Grime, J.P., 1979, November. Primary strategies in plants. In *Transactions of the Botanical Society of Edinburgh* (Vol. 43, No. 2, pp. 151-160).

Grotch, S.L. and MacCracken, M.C., 1991. The use of general circulation models to predict regional climatic change. *Journal of climate*, *4*(3), pp.286-303.

Guerrina, M., Dagnino, D., Minuto, L., Médail, F. and Casazza, G., 2024. Unveiling the hypotheses of endemic richness: A study case in the Southwestern Alps. *Perspectives in Plant Ecology, Evolution and Systematics*, *63*, p.125792.

Hannah, L., Flint, L., Syphard, A.D., Moritz, M.A., Buckley, L.B. and McCullough, I.M., 2014. Fine-grain modeling of species' response to climate change: holdouts, steppingstones, and microrefugia. *Trends in ecology & evolution*, *29*(7), pp.390-397.

Harold, A.S. and Mooi, R.D., 1994. Areas of endemism: definition and recognition criteria. *Systematic biology*, *43*(2), pp.261-266.

Harrison, S. and Noss, R., 2017. Endemism hotspots are linked to stable climatic refugia. *Annals of Botany*, 119(2), pp.207-214.

Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M. and Porter, E.E., 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, *84*(12), pp.3105-3117.

Hendrych, R., 1982. Material and notes about the geography of the highly stenotopic to monotopic endemic species of the European flora. *Acta Universitatis CarolinaeBiologica*, *3*(4), pp.335-372.

Herrando-Moraira, S., Nualart, N., Galbany-Casals, M., Garcia-Jacas, N., Ohashi, H., Matsui, T., Susanna, A., Tang, C.Q. and López-Pujol, J., 2022. Climate Stability Index maps, a global high-resolution cartography of climate stability from Pliocene to 2100. *Scientific data*, *9*(1), p.48.

Heusser, C.J., 1955. Pollen profiles from the Queen Charlotte Islands, British Columbia. *Canadian Journal of Botany*, *33*(5), pp.429-449.

Hewitson, B.C. and Crane, R.G., 2006. Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *26*(10), pp.1315-1337.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), pp.907913.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very highresolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *25*(15), pp.1965-1978.

Hijmans, R.J., Bivand, R., Forner, K., Ooms, J., Pebesma, E. and Sumner, M.D., 2024. terra: Spatial Data Analysis; 2022. *URL https://CRAN. R-project. org/package= terra. R package version*, pp.1-6.

Hobohm, C. ed., 2014. Endemism in vascular plants (No. 14760). Dordrecht: Springer Netherlands.

Holt, R.D., 1984. Spatial heterogeneity, indirect interactions, and the coexistence of prey species. *The American Naturalist*, *124*(3), pp.377-406.

Hopper, S.D. and Gioia, P., 2004. The southwest Australian floristic region: evolution and conservation of a global hot spot of biodiversity. *Annu. Rev. Ecol. Evol. Syst.*, *35*(1), pp.623-650.

Hubbell, S.P., 2005. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional ecology*, *19*(1), pp.166-172.

Hughes, P.D., Gibbard, P.L. and Ehlers, J., 2013. Timing of glaciation during the last glacial cycle: evaluating the concept of a global 'Last Glacial Maximum'(LGM). *EarthScience Reviews*, *125*, pp.171-198.

Hutchinson, G.E., 1959. Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist*, *93*(870), pp.145-159.

Huston, M.A., 1994. *Biological diversity: the coexistence of species*. Cambridge University Press.

Huston, M.A., 1999. Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. *Oikos*, pp.393-401.

Igu, N.I., 2017. Species diversity and structure of an intact freshwater swamp forest in the Niger Delta. *Open Journal of Forestry*, *7*(02), p.242.

I.P.C.C., 2014. Climate change 2014 synthesis report. *IPCC: Geneva, Switzerland*, pp.10591072.

Işik, K., 2011. Rare and endemic species: why are they prone to extinction? *Turkish Journal of Botany*, *35*(4), pp.411-417.

Jansson, R., 2003. Past climatic changes explain global patterns in endemism. Proceedings of the Royal Society of London. *Series B: Biological Sciences*, 270(1515), pp.583-590.

Jetz, W., Rahbek, C. and Colwell, R.K., 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters*, 7(12), pp.11801191.

Jobbágy, E.G., Nosetto, M.D., Villagra, P.E. and Jackson, R.B., 2011. Water subsidies from mountains to deserts: Their role in sustaining groundwater-fed oases in a sandy landscape. *Ecological Applications*, *21*(3), pp.678-694.

Kallimanis, A.S., Bergmeier, E., Panitsa, M., Georghiou, K., Delipetrou, P. and Dimopoulos, P., 2010. Biogeographical determinants for total and endemic species richness in a continental archipelago. *Biodiversity and Conservation*, *19*, pp.1225-1235.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. and Kessler, M., 2017. Climatologies at high resolution for the earth's land surface areas. *Scientific data*, *4*(1), pp.1-20.

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yates, C.J., Byrne, M., Mucina, L., Schut, A.G., Hopper, S.D. and Franklin, S.E., 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global ecology and biogeography*, *21*(4), pp.393-404.

Keppel, G., Mokany, K., Wardell-Johnson, G.W., Phillips, B.L., Welbergen, J.A. and Reside, A.E., 2015. The capacity of refugia for conservation planning under climate change. *Frontiers in Ecology and the Environment*, *13*(2), pp.106-112.

Keppel, G., Robinson, T.P., Wardell-Johnson, G.W., Yates, C.J., Van Niel, K.P., Byrne, M. and Schut, A.G., 2017. A low-altitude mountain range as an important refugium for two narrow endemics in the Southwest Australian Floristic Region biodiversity hotspot. *Annals of Botany*, *119*(2), pp.289-300.

Kiedrzyński, M., Zielińska, K.M., Kiedrzyńska, E. and Rewicz, A., 2017. Refugial debate: On small sites according to their function and capacity. *Evolutionary Ecology*, *31*, pp.815-827.

Kraft, N.J., Baldwin, B.G. and Ackerly, D.D., 2010. Range size, taxon age and hotspots of neoendemism in the California flora. *Diversity and Distributions*, *16*(3), pp.403-413.

Kreft, H. and Jetz, W., 2007. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, *104*(14), pp.5925-5930.

Kreft, H., Jetz, W., Mutke, J., Kier, G. and Barthlott, W., 2008. Global diversity of island floras from a macroecological perspective. *Ecology letters*, *11*(2), pp.116-127.

Kruger, F.J. and Bigalke, R.C., 1984. Fire in fynbos. In *Ecological effects of fire in South African ecosystems* (pp. 67-114). Berlin, Heidelberg: Springer Berlin Heidelberg.

Laliberté, E., Zemunik, G. and Turner, B.L., 2014. Environmental filtering explains variation in plant diversity along resource gradients. *science*, *345*(6204), pp.1602-1605.

Latimer, A.M., Silander Jr, J.A. and Cowling, R.M., 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. *Science*, *309*(5741), pp.17221725.

Lee, D., Rushworth, A. and Napier, G., 2018. Spatio-temporal areal unit modelling in R with conditional autoregressive priors using the CARBayesST package. *Journal of Statistical Software*, *84*(9), pp.1-39.

Lenoir, J., Gégout, J.C., Guisan, A., Vittoz, P., Wohlgemuth, T., Zimmermann, N.E., Dullinger, S., Pauli, H., Willner, W. and Svenning, J.C., 2010. Going against the flow: potential mechanisms for unexpected downslope range shifts in a warming climate. *Ecography*, *33*(2), pp.295-303.

Leroux, B.G., Lei, X. and Breslow, N., 2000. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In *Statistical models in epidemiology, the environment, and clinical trials* (pp. 179-191). Springer New York.

Levyns, M.R., 1964. Migrations and origin of the Cape flora. *Transactions of the Royal Society of South Africa*, *37*(2), pp.85-107.

Linder, H.P., 1991. Environmental correlates of patterns of species richness in the southwestern Cape Province of South Africa. *Journal of Biogeography*, pp.509-518.

Linder, H.P., 1995. Setting conservation priorities: the importance of endemism and phylogeny in the southern African orchid genus Herschelia. *Conservation Biology*, *9*(3), pp.585-595.

Linder, H.P., 2003. The radiation of the Cape flora, southern Africa. *Biological Reviews*, *78*(4), pp.597-638.

Linder, H.P., 2005. Evolution of diversity: the Cape flora. *Trends in plant science*, *10*(11), pp.536-541.

Linder, H.P., 2008. Plant species radiations: where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), pp.3097-3105.

Linder, H.P., 2019. Rare species, Restionaceae, and the Cape flora. *Journal of Biogeography*, *46*(12), pp.2637-2650.

Loarie, S.R., Carter, B.E., Hayhoe, K., McMahon, S., Moe, R., Knight, C.A. and Ackerly, D.D., 2008. Climate change and the future of California's endemic flora. *PloS one*, 3(6), p.e2502.

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. and Ackerly, D.D., 2009. The velocity of climate change. *Nature*, *462*(7276), pp.1052-1055.

Losos, J.B. and Schluter, D., 2000. Analysis of an evolutionary species–area relationship. *Nature*, *408*(6814), pp.847-850

MacKellar, N., New, M. and Jack, C., 2014. Observed and modelled trends in rainfall and temperature for South Africa: 1960-2010. *South African Journal of Science*, *110*(7-8), pp.1-13.

Magadzire, N., De Klerk, H.M., Esler, K.J. and Slingsby, J.A., 2019. Fire and life history affect the distribution of plant species in a biodiversity hotspot. *Diversity and Distributions*, *25*(7), pp.1012-1023.

Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., Olsen, J., Perez, K.E., Stam, W., Väinölä, R. and Viard, F., 2008. Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology*, *89*(sp11), pp.S108-S122.

Manes, S., Costello, M.J., Beckett, H., Debnath, A., Devenish-Nelson, E., Grey, K.A., Jenkins, R., Khan, T.M., Kiessling, W., Krause, C. and Maharaj, S.S., 2021. Endemism increases species' climate change risk in areas of global biodiversity importance—Biological *Conservation*, *257*, p.109070.

Manning, J. and Goldblatt, P., 2012. Plants of the greater cape floristic region. 1: the Core Cape flora (pp. xiv-853).

McDonald, D.J. and Morley, M., 1988. A checklist of the flowering plants and ferns of Swartboschkloof, Jonkershoek, Cape Province. *Bothalia*, *18*(2), pp.261-270.

McLaughlin, B.C., Ackerly, D.D., Klos, P.Z., Natali, J., Dawson, T.E. and Thompson, S.E., 2017. Hydrologic refugia, plants, and climate change. *Global change biology*, 23(8), pp.2941-2961.

McMahon, S.M., Harrison, S.P., Armbruster, W.S., Bartlein, P.J., Beale, C.M., Edwards, M.E., Kattge, J., Midgley, G., Morin, X. and Prentice, I.C., 2011. Improving assessment and modelling of climate change impacts on global terrestrial biodiversity. *Trends in Ecology & Evolution*, 26(5), pp.249-259.

Mee, J.A. and Moore, J.S., 2014. The ecological and evolutionary implications of microrefugia. *Journal of Biogeography*, *41*(5), pp.837-841.

Médail, F. and Diadema, K., 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, *36*(7), pp.1333-1345.

Mee, J.A. and Moore, J.S., 2014. The ecological and evolutionary implications of microrefugia. *Journal of Biogeography*, *41*(5), pp.837-841.

Mercier, J.L. and Wilmé, L., 2013. The Eco-Geo-Clim model: explaining Madagascar's endemism. *Madagascar Conservation & Development*, *8*(2), pp.63-68.

Midgley, G.F., Hannah, L., Millar, D., Thuiller, W. and Booth, A., 2003. Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic Region. *Biological conservation*, *112*(1-2), pp.87-97.

Midgley, G.F., Chapman, R.A., Hewitson, B., Johnston, P., De Wit, M., Ziervogel, G., Mukheibir, P., Van Niekerk, L., Tadross, M., Van Wilgen, B.W. and Kgope, B., 2005. A status quo, vulnerability and adaptation assessment of the physical and socioeconomic effects of climate change in the Western Cape.

Miller, G.A.D., Suzuki, N., Ciftci-Yilmaz, S.U.L.T.A.N. and Mittler, R.O.N., 2010. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant, cell & environment*, *33*(4), pp.453-467.

Mishler, B.D., Knerr, N., González-Orozco, C.E., Thornhill, A.H., Laffan, S.W. and Miller, J.T., 2014. Phylogenetic measures of biodiversity and neo-and paleo-endemism in Australian Acacia. *Nature Communications*, *5*(1), p.4473.

Mittelbach, G.G., Steiner, C.F., Scheiner, S.M., Gross, K.L., Reynolds, H.L., Waide, R.B., Willig, M.R., Dodson, S.I. and Gough, L., 2001. What is the observed relationship between species richness and productivity? *Ecology*, *82*(9), pp.2381-2396.

Mora, C., Frazier, A.G., Longman, R.J., Dacks, R.S., Walton, M.M., Tong, E.J., Sanchez, J.J., Kaiser, L.R., Stender, Y.O., Anderson, J.M. and Ambrosino, C.M., 2013. The projected timing of climate departure from recent variability. *Nature*, *502*(7470), pp.183-187.

Morelli, T.L., Daly, C., Dobrowski, S.Z., Dulen, D.M., Ebersole, J.L., Jackson, S.T., Lundquist, J.D., Millar, C.I., Maher, S.P., Monahan, W.B. and Nydick, K.R., 2016. Managing climate change refugia for climate adaptation. *PloS one*, *11*(8), p.e0159909.

Morrone, J.J., 1994. On the identification of areas of endemism. *Systematic biology*, *43*(3), pp.438-441.

Morueta-Holme, N., Enquist, B.J., McGill, B.J., Boyle, B., Jørgensen, P.M., Ott, J.E., Peet, R.K., Šímová, I., Sloat, L.L., Thiers, B. and Violle, C., 2013. Habitat area and climate stability determine geographical variation in plant species' range sizes. *Ecology Letters,* 16(12), pp.1446-1454.

Mosblech, N.A.S., Bush, M.B. and van Woesik, R., 2011. On metapopulations and microrefugia: palaeoecological insights. *Journal of Biogeography*, *38*(3), pp.419-429.

Mummenhoff, K., Al-Shehbaz, I.A., Bakker, F.T., Linder, H.P. and Mühlhausen, A., 2005. Phylogeny, morphological evolution, and speciation of endemic Brassicaceae genera in the Cape flora of southern Africa. *Annals of the Missouri Botanical Garden*, pp.400-424.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A. and Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), pp.853-858.

New, M., Hewitson, B., Stephenson, D.B., Tsiga, A., Kruger, A., Manhique, A., Gomez, B., Coelho, C.A., Masisi, D.N., Kululanga, E. and Mbambalala, E., 2006. Evidence of trends in daily climate extremes over southern and west Africa. *Journal of Geophysical Research: Atmospheres*, *111*(D14).

Newbold, T., 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in physical geography*, *34*(1), pp.3-22.

Nelder, J.A. and Wedderburn, R.W., 1972. Generalized linear models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *135*(3), pp.370-384.

Noss, R.F., 2001. Beyond Kyoto: Forest management in a time of rapid climate change. *Conservation biology*, 15(3), pp.578-590.

Nilsson, C., 1983. Frequency distributions of vascular plants in the littoral vegetation along two rivers in northern Sweden. *Journal of Biogeography*, pp.351-369.

O'Brien, E.M., 1993. Climatic gradients in woody plant species richness: towards an explanation based on an analysis of southern Africa's woody flora. *Journal of biogeography*, pp.181-198.

O'Brien, E., 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. *Journal of Biogeography*, *25*(2), pp.379-398.

Ohlemüller, R., Anderson, B.J., Araujo, M.B., Butchart, S.H., Kudrna, O., Ridgely, R.S. and Thomas, C.D., 2008. The coincidence of climatic and species rarity: high risk to small-range species from climate change. *Biology letters*, *4*(5), pp.568-572.

Oliver, E.G.H., Linder, H.P. and Rourke, J.P., 1983. Geographical distribution of present-day Cape taxa and their phytogeographical significance. *Bothalia*, *14*(3/4), pp.427-440.

Ollis, D., Snaddon, K., Job, N. and Mbona, N., 2013. *Classification system for wetlands and other aquatic ecosystems in South Africa*. South African National Biodiversity Institute.

Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L. and Hofreiter, M., 2004. Genetic analyses from ancient DNA. *Annu. Rev. Genet.*, *38*, pp.645-679.

Park, D.S., Willis, C.G., Xi, Z., Kartesz, J.T., Davis, C.C. and Worthington, S., 2020. Machine learning predicts large scale declines in native plant phylogenetic diversity. *New Phytologist*, *227*(5), pp.1544-1556.

Parkington, J., Cartwright, C., Cowling, RM, Baxter, A. and Meadows, M., 2000. Palaeovegetation at the last glacial maximum in the western Cape, South Africa: wood charcoal and pollen evidence from Elands Bay Cave. *South African Journal of Science*, *96*(11-12), pp.543-546.

Pavlik, B.M., 1995. Species diversity in ecological communities: historical and geographical perspectives.

Pebesma, E.J., 2018. Simple features for R: standardized support for spatial vector data. *R J.*, *10*(1), p.439.

Picker, M.D. and Samways, M.J., 1996. Faunal diversity and endemicity of the Cape Peninsula, South Africa—a first assessment. *Biodiversity & Conservation*, *5*, pp.591606.

Post-Leon, A.C., Dryak, M., Zhu, E., De Guzman, M.E., Salladay, R., Moritz, M.A., Parkinson, A.M.L. and Ramirez, A.R., 2022. Integration of landscape-level remote sensing and tree-level ecophysiology reveals drought refugia for a rare endemic, bigcone Douglas-fir. *Frontiers in Forests and Global Change*, *5*, p.946728.

Procheş, Ş., Cowling, R.M., Goldblatt, P., Manning, J.C. and Snijman, D.A., 2006. An overview of the Cape geophytes. *Biological Journal of the Linnean Society*, *87*(1), pp.27-43.

Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B., Vittoz, P., Thuiller, W. and Guisan, A., 2009. Climate change and plant distribution: local models predict high-elevation persistence. *Global change biology*, *15*(6), pp.15571569.

Reddy, S. and Dávalos, L.M., 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, *30*(11), pp.1719-1727.

Rensch, B., 1954. Recent problems of the theory of descent. Transspecific evolution. *Recent problems of the theory of descent. Transspecific evolution.*, (2nd ed).

Ricklefs, R.E., 1977. Environmental heterogeneity and plant species diversity: a hypothesis. *The American Naturalist*, *111*(978), pp.376-381.

Ricklefs, R.E., 1987. Community diversity: relative roles of local and regional processes. *Science*, 235(4785), pp.167-171.

Ricklefs, R.E. and Schluter, D., 1993. Species diversity: regional and historical influences. *Species diversity in ecological communities*, *1*, pp.350-363.

Rohde, K., 1992. Latitudinal gradients in species diversity: the search for the primary cause. *Oikos*, pp.514-527.

Rosenzweig, M.L. and Abramsky, Z., 1993. How are diversity and productivity related? In *Species diversity in ecological communities* (pp. 52-65). University of Chicago Press.

Rosenzweig, M.L., 1995. *Species Diversity in Space and Time*. Cambridge: University Press. 436 p.

Rosenzweig, M.L. and Sandlin, E.A., 1997. Species diversity and latitudes: listening to area's signal. *Oikos*, pp.172-176.

Rouget, M., Richardson, D.M. and Cowling, R.M., 2003. The current configuration of protected areas in the Cape Floristic Region, South Africa—reservation bias and representation of biodiversity patterns and processes. *Biological conservation*, *112*(1-2), pp.129-145.

Rouget, M., Robertson, M.P., Wilson, J.R., Hui, C., Essl, F., Renteria, J.L. and Richardson, D.M., 2016. Invasion debt–quantifying future biological invasions. *Diversity and Distributions*, *22*(4), pp.445-456.

Ruhi, A., Catford, J.A., Cross, W.F., Escoriza, D. and Olden, J.D., 2019. Understanding the nexus between hydrological alteration and biological invasions. In *Multiple stressors in river ecosystems* (pp. 45-64). Elsevier.

Rull, V., 2009. Microrefugia. *Journal of Biogeography*, *36*(3), pp.481-484.

Rundel, P.W., Arroyo, M.T., Cowling, R.M., Keeley, J.E., Lamont, B.B. and Vargas, P., 2016. Mediterranean biomes: evolution of their vegetation, floras, and climate. *Annual Review of Ecology, Evolution, and Systematics*, *47*(1), pp.383-407.

Rutherford, M.C., Mucina, L. and Powrie, L.W., 2006. Biomes and bioregions of southern Africa. *The vegetation of South Africa, Lesotho and Swaziland*, *19*, pp.30-51.

Rutherford, M.C., Powrie, L.W., Husted, L.B. and Turner, R.C., 2011. Early post-fire plant succession in Peninsula Sandstone Fynbos: The first three years after disturbance. *South African journal of botany*, *77*(3), pp.665-674.

Sandel, B., Arge, L., Dalsgaard, B., Davies, R.G., Gaston, K.J., Sutherland, W.J. and Svenning, J.C., 2011. The influence of Late Quaternary climate-change velocity on species endemism. *Science*, *334*(6056), pp.660-664.

Saxon, E., Baker, B., Hargrove, W., Hoffman, F. and Zganjar, C., 2005. Mapping environments at risk under different global climate change scenarios. *Ecology Letters*, *8*(1), pp.53-60.

Schmidt-Lebuhn, A.N., Knerr, N.J., Miller, J.T. and Mishler, B.D., 2015. Phylogenetic diversity and endemism of Australian daisies (Asteraceae). *Journal of Biogeography*, *42*(6), pp.1114-1122.

Schmitt, T. and Varga, Z., 2012. Extra-Mediterranean refugia: the rule and not the exception? *Frontiers in Zoology*, *9*, pp.1-12.

Selwood, K.E. and Zimmer, H.C., 2020. Refuges for biodiversity conservation: A review of the evidence. *Biological Conservation*, *245*, p.108502.

Sieben, E.J.J., Kotze, D.C., Job, N.M. and Muasya, A.M., 2017. The sclerophyllous wetlands on quartzite substrates in South Africa: Floristic description, classification and explanatory environmental factors. *South African Journal of Botany*, *113*, pp.54-61.

Sieben, E.J.J., Glen, R.P., Van Deventer, H. and Dayaram, A., 2021. The contribution of wetland flora to regional floristic diversity across a wide range of climatic conditions in southern Africa. *Biodiversity and Conservation*, *30*(3), pp.575-596.

Slingsby, J.A., Merow, C., Aiello-Lammens, M., Allsopp, N., Hall, S., Kilroy Mollmann, H., Turner, R., Wilson, A.M. and Silander Jr, J.A., 2017. Intensifying postfire weather and biological invasion drive species loss in a Mediterranean-type biodiversity hotspot. *Proceedings of the National Academy of Sciences*, *114*(18), pp.4697-4702.

Slingsby, J.A., Moncrieff, G.R., Rogers, A.J. and February, E.C., 2020. Altered ignition catchments threaten a hyperdiverse fire-dependent ecosystem. *Global Change Biology*, *26*(2), pp.616-628.

Slingsby, J.A., de Buys, A., Simmers, A.D., Prinsloo, E., Forsyth, G.G., Glenday, J. and Allsopp, N., 2021. Jonkershoek: Africa's oldest catchment experiment-80 years and counting. *Hydrological Processes*, *35*(4), p.e14101.

Sorrie, B.A. and Weakley, A.S., 2001. Coastal plain vascular plant endemics: phytogeographic patterns. *Castanea*, pp.50-82.

Srivastava, D.S. and Lawton, J.H., 1998. Why more productive sites have more species: an experimental test of theory using tree-hole communities. *The American Naturalist*, *152*(4), pp.510-529.

Stebbins, G.L. and Major, J., 1965. Endemism and speciation in the California flora. *Ecological Monographs*, *35*(1), pp.2-35

Stein, A., Gerstner, K. and Kreft, H., 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology letters*, *17*(7), pp.866-880.

Štubňová, E., Hodálová, I., Kučera, J., Mártonfiová, L., Svitok, M. and Slovák, M., 2017. Karyological patterns in the European endemic genus Soldanella L.: Absolute genome size variation uncorrelated with cytotype chromosome numbers. *American Journal of Botany*, *104*(8), pp.1241-1253.

Smyčka, J., Roquet, C., Renaud, J., Thuiller, W., Zimmermann, N.E. and Lavergne, S., 2017. Disentangling drivers of plant endemism and diversification in the European Alps–A phylogenetic and spatially explicit approach. *Perspectives in Plant Ecology, Evolution and Systematics*, *28*, pp.19-27.

Taberlet, P. and Cheddadi, R., 2002. Quaternary refugia and persistence of biodiversity. *Science*, *297*(5589), pp.2009-2010.

Tadross, M., Jack, C. and Hewitson, B., 2005. On RCM-based projections of change in southern African summer climate. *Geophysical Research Letters*, *32*(23).

Tang, C.Q., Matsui, T., Ohashi, H., Dong, Y.F., Momohara, A., Herrando-Moraira, S., Qian, S., Yang, Y., Ohsawa, M., Luu, H.T. and Grote, P.J., 2018. Identifying long-term stable refugia for relict plant species in East Asia. *Nature Communications*, *9*(1), p.4488.

Team, R.D.C., 2010. R: A language and environment for statistical computing. *(No Title)*.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F., De Siqueira, M.F., Grainger, A., Hannah, L. and Hughes, L., 2004. Extinction risk from climate change. *Nature*, *427*(6970), pp.145-148.

Thwaites, R.N. and Cowling, R.M., 1988. Soil-vegetation relationships on the Agulhas plain, South Africa. *Catena*, *15*(3-4), pp.333-345.

Tribsch, A. and Schönswetter, P., 2003. Patterns of endemism and comparative phylogeography confirm palaeo-environmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon*, *52*(3), pp.477-497.

Tzedakis, P.C., Emerson, B.C. and Hewitt, G.M., 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in ecology & evolution*, *28*(12), pp.696-704.

VanDerWal, J., Murphy, H.T., Kutt, A.S., Perkins, G.C., Bateman, B.L., Perry, J.J. and Reside, A.E., 2013. Focus on poleward shifts in species' distribution underestimates the fingerprint of climate change. *Nature climate change*, *3*(3), pp.239-243.

Van Deventer, H., Van Niekerk, L., Adams, J., Dinala, M.K., Gangat, R., Lamberth, S.J., Lötter, M., Mbona, N., MacKay, F., Nel, J.L. and Ramjukadh, C.L., 2020. National Wetland Map 5: An improved spatial extent and representation of inland aquatic and estuarine ecosystems in South Africa. *Water SA*, *46*(1), pp.66-79.

Van Mazijk, R., Cramer, M.D. and Verboom, G.A., 2021. Environmental heterogeneity explains contrasting plant species richness between the South African Cape and southwestern Australia. *Journal of Biogeography*, *48*(8), pp.1875-1888.

Verboom, G.A., Bergh, N.G., Haiden, S.A., Hoffmann, V. and Britton, M.N., 2015. Topography as a driver of diversification in the C ape Floristic Region of South Africa. *New Phytologist*, *207*(2), pp.368-376.

Ver Hoef, J.M. and Boveng, P.L., 2007. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology*, *88*(11), pp.2766-2772.

Weber, L.C., VanDerWal, J., Schmidt, S., McDonald, W.J. and Shoo, L.P., 2014. Patterns of rain forest plant endemism in subtropical Australia relate to stable mesic refugia and species dispersal limitations. *Journal of Biogeography*, *41*(2), pp.222-238.

Webb, T.J. and Gaston, K.J., 2003. On the heritability of geographic range sizes. *The American Naturalist*, *161*(4), pp.553-566.

Wickham, H., Chang, W. and Wickham, M.H., 2016. Package ‘ggplot2’. *Create elegant data visualisations using the grammar of graphics. Version*, *2*(1), pp.1-189.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D.A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J. and Kuhn, M., 2019. Welcome to the Tidyverse. *Journal of open-source software*, *4*(43), p.1686.

Wickham, H., Çetinkaya-Rundel, M. and Grolemund, G., 2023. *R for data science*. " O'Reilly Media, Inc.".

Willig, M.R., Kaufman, D.M. and Stevens, R.D., 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual review of ecology, evolution, and systematics*, *34*(1), pp.273-309.

Willis, C.K., Cowling, R.M. and Lombard, A.T., 1996. Patterns of endemism in the limestone flora of South African lowland fynbos. *Biodiversity & Conservation*, *5*, pp.55-73.

Willis, K.J. and Niklas, K.J., 2004. The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1442), pp.159-172.

Wilson, A.M. and Jetz, W., 2016. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLoS biology*, *14*(3), p.e1002415.

Woodward, F.I., 1987. *Climate and plant distribution*. Cambridge University Press.

Wright, D.H., 1983. Species-energy theory: an extension of species-area theory. *Oikos*, pp.496-506.

Wüest, R.O., Boucher, F.C., Bouchenak-Khelladi, Y., Karger, D.N. and Linder, H.P., 2019. Dissecting biodiversity in a global hotspot: Uneven dynamics of immigration and diversification within the Cape Floristic Region of South Africa. *Journal of Biogeography*, *46*(9), pp.1936-1947.

Young, G.M., Minter, W.E.L. and Theron, J.N., 2004. Geochemistry and palaeogeography of upper Ordovician glaciogenic sedimentary rocks in the Table Mountain Group, South Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *214*(4), pp.323-345.

Zobel, M., 1992. Plant species coexistence: the role of historical, evolutionary and ecological factors. *Oikos*, pp.314-320.

Appendix

Multicollinearity test

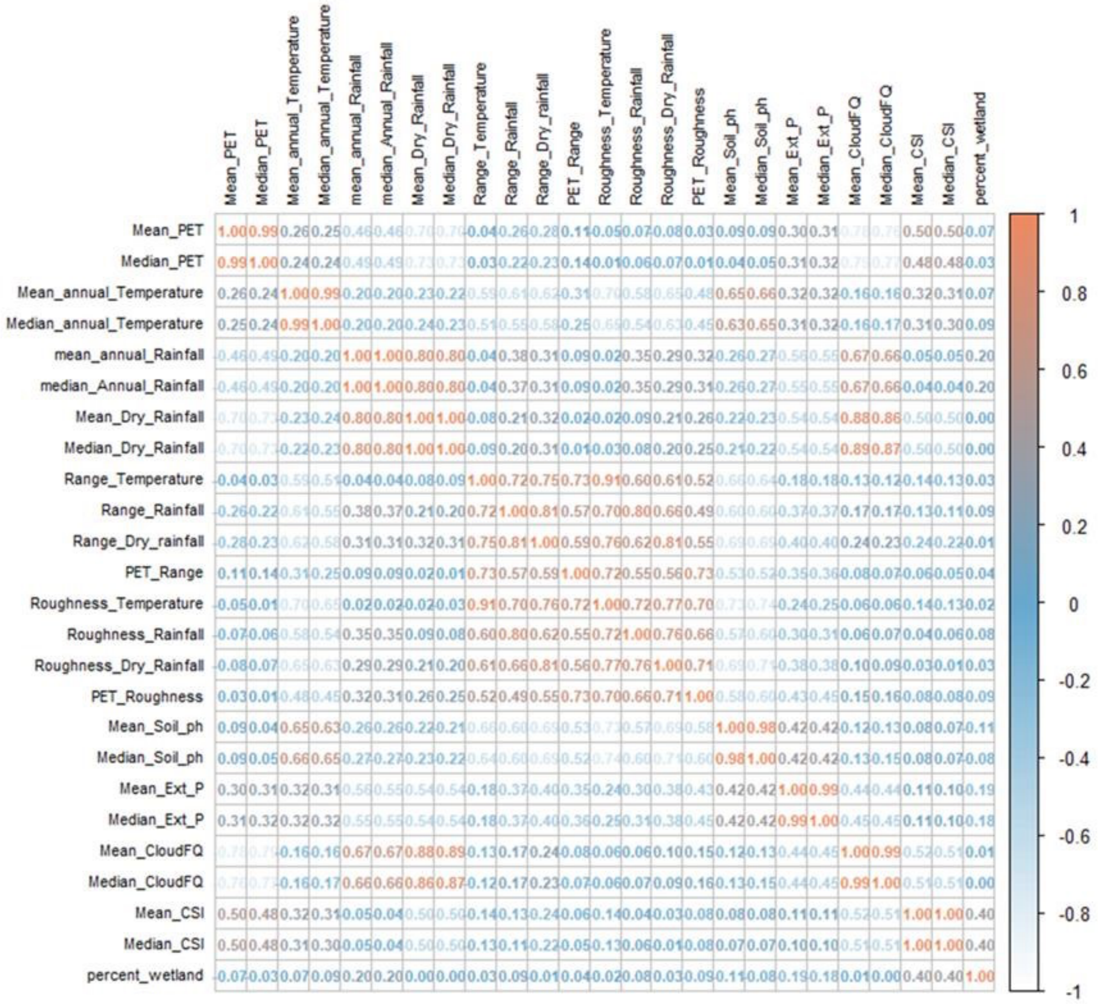

Figure A1: Correlation matrix of all covariates used in the study

Tests for spatial autocorrelation

Moran's I

Table A4: Moran’s I for all response variable for the different response variable to test for spatial autocorrelation

Response	Estimates	P-values
Species Richness	0.4	3.25×10^{-25}
Narrow range endemics	0.2	8.13×10^{-8}
Wetland-dependent narrow-range endemics	0.3	1.62×10^{-15}

Lagrange Multiplier’s test

Table A2: Lagrange Multiplier’s test to test for spatial autocorrelation in the spatial lag (LMlag), spatial error (LMerr), robust spatial lag (LMlag), robust spatial error (RLMerr) and the spatial moving average model (SARMA)

Test	P-value	Response
LMerr	0,03671	Species Richness
RLMerr	0,03612	Species Richness

LMlag	0,53165	Species Richness
RLMlag	0,51744	Species Richness
SARMA	0,0915	Species Richness
LMerr	0,09366	Narrow range endemics
RLMerr	0,07618	Narrow range endemics
LMlag	0,03514	Narrow range endemics
RLMlag	0,02892	Narrow range endemics
SARMA	0,02256	Narrow range endemics
LMerr	0,00112	Wetland-dependent Narrow range endemics
RLMerr	0,00046	Wetland-dependent Narrow range endemics

LMlag	0,24927	Wetland-dependent Narrow range endemics
RLMlag	0,08403	Wetland-dependent Narrow range endemics
SARMA	0,00111	Wetland-dependent Narrow range endemics

Bayesian Model Diagnostics

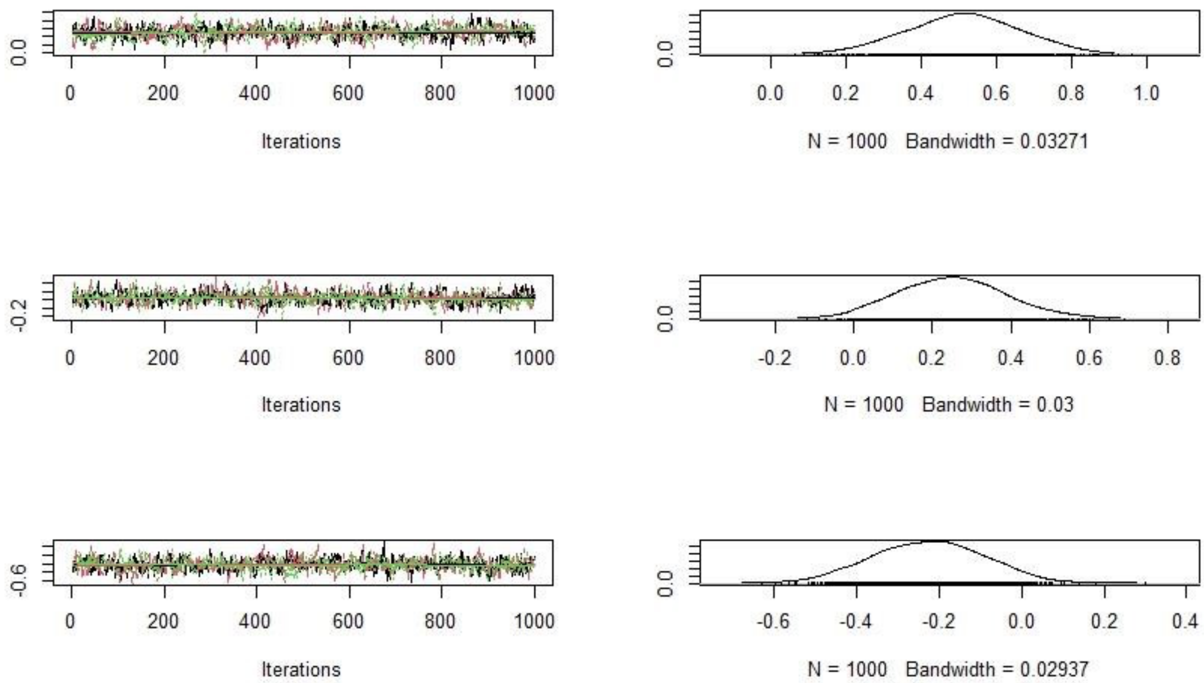

Figure A2: Bayesian Model diagnostic test for the total number of species in the CFR

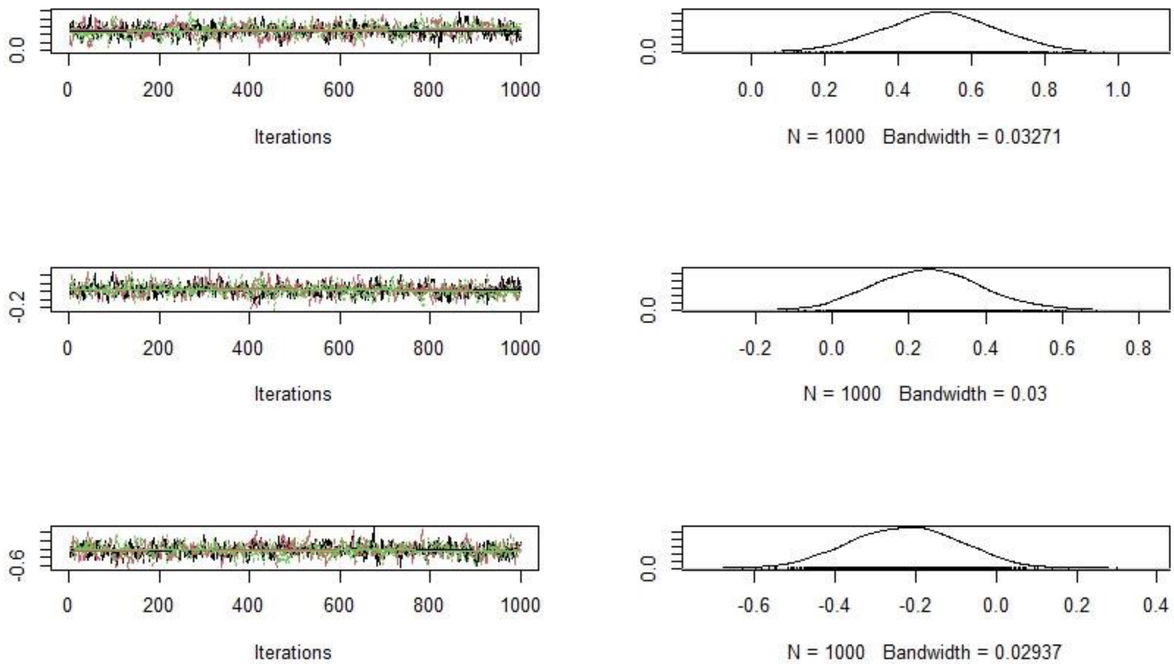

Figure A3: Bayesian Model diagnostic test for the narrow range endemics of species in the CFR

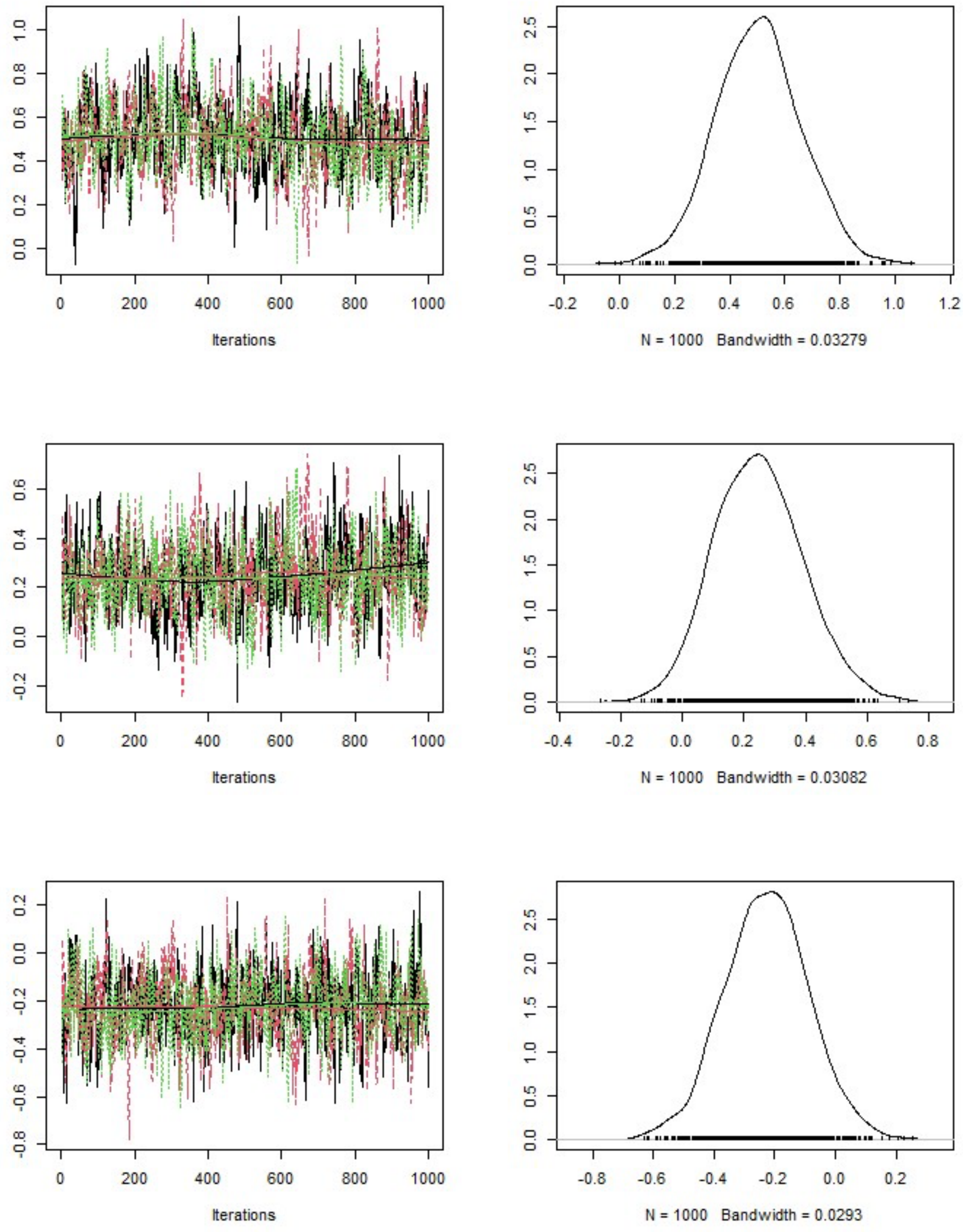

Figure A4: Bayesian Model diagnostic test for the wetland-dependent narrow range endemics in the CFR

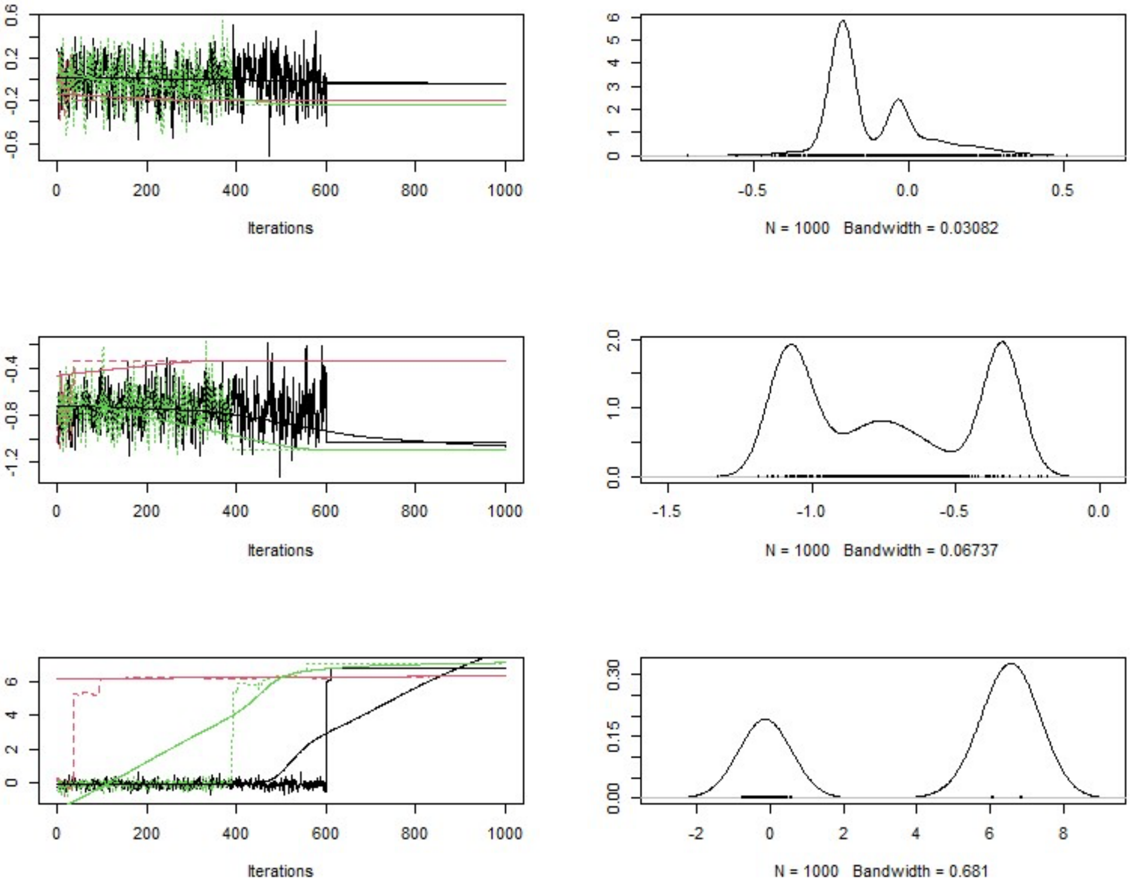

Figure A5: Bayesian Model diagnostic test for the wetland dependent narrow range endemics productivity model that did not converge.

Data Management plan

1. Data and software outputs

1.1 The data and software outputs your research will generate?

The data and software will produce spatial heatmaps of :

1) The total species richness,

2) Narrow range endemic plant species,

3) Widespread species

4) wetland-dependent narrow range endemic species.

1.2 When you intend to share your data and software?

These data will be accessible as soon as the dissertation is submitted and will be shared on ZivHub as a raster images and vector files.

1.3 Where your data and software will be made available?

The R scripts and the species data will be saved on Zendo (https://doi.org/10.5281/zenodo.14608156) repository and institutional repository (ZivaHub) respectively. The data, in the form of shapefiles as well as supplementary data from other such studies will be accessible on ZivaHub.

1.4 How will your data and software will be accessible to others?

The reproducible code will be made accessible to through Zendo, which is an online, data repository that serves as a publishing and access platform to research data and scholarly outputs. In this repository there will be a folder for clearly annotated R scripts and data folders with vector and raster image datasets clearly labelled.

1.5 Are any limits to data and software sharing required?

No, this study does not have any limits to data and software sharing.

1.6 How datasets and software will be preserved?

The raw data with geolocations of threatened and vulnerable species will be saved in a private repository while only QDS counts will be available to the public for usage.

2. Research materials

2.1 What materials your research will produce and how these will be made available?

This research will produce an endemism (NREs and WD-NREs) as well as a total species richness map of the Cape floristic region (CFR). The endemism maps are documented to relate to long term stability sites in the region. Therefore, this study will also produce a hydrological refugia map of the CFR. These species distribution data will be saved as a raster image (tif) file and will be accessible on: https://doi.org/10.5281/zenodo.14608156.

3. Intellectual property

3.1 What IP your research will generate?

My research will generate a map of the character and distribution of refugia in the CFR's fynbos Biome.

3.2 How IP will be protected?

The raw data will be saved on a private repository with clean data in the form of raster and vector files saved in the publicly accessible repository.

3.3 How IP will be used to achieve health benefits?

This IP will be used to inform conservation efforts in hydrologically stable sites that buffer rare and endemic species from the effects of climate change.

4. Resources required

4.1 People and skills

These data and code do not require a data manager or data scientist; however, they do require that some coding knowledge on R and some experience with spatial data for people that might need to use the data.

4.2 Storage and computation

The data will be stored a repository: https://doi.org/10.5281/zenodo.14608156

4.3 Access

No access is needed.

4.4 Deposition and preservation of data, software and materials

The study will store spatial data in different formats, with the rough R scripts saved in GitHub, where most of the version control occurs.