Title
The importance of defining measures of stability in macroecology and biogeography

Permalink
https://escholarship.org/uc/item/9wj0j9ct

Journal
Frontiers of Biogeography, 11(3)

Authors
McDonald-Spicer, Christiana J.
Moritz, Craig C.
Ferrier, Simon
et al.

Publication Date
2019

DOI
10.21425/F5FBG43355

License
CC BY 4.0

REVIEW

The importance of defining measures of stability in macroecology and biogeography

Christiana J. McDonald-Spicer[1] iD , Craig C. Moritz[1] iD , Simon Ferrier[2] and Dan F. Rosauer[1] iD

[1] Ecology and Evolution, Research School of Biology, The Australian National University, Acton, ACT, Australia

[2] CSIRO Land and Water, Canberra, ACT, Australia

Corresponding author: McDonald-Spicer, C.J.; christiana.mcdonald-spicer@anu.edu.au, https://christianamcdonaldspicer.wordpress.com

Abstract. Stability, the continuity of environments or habitats through space and time, is a widely used concept in macroecology and biogeography and is often invoked in studies attempting to explain the uneven spatial distribution of biodiversity. Stability can be measured in various ways and at various spatiotemporal scales; however, few studies explicitly define their use of the term. This makes interpreting and comparing studies difficult. We suggest an integrated approach to defining measures of stability in macroecology and biogeography. This approach addresses five key challenges concerning the biological, environmental and spatiotemporal scales at which stability is assessed, and how the complexity of change across time and space is summarised into a metric of stability. Using this approach allows for clarity around the choice, conceptualisation, communication and comparison of measures of stability.

Keywords: biodiversity; climate stability; habitat stability; refugia; scale; stability

Stability in macroecology and biogeography

The term "stability" appears widely, and in many different contexts, across ecology and evolutionary biology (Ives and Carpenter 2007, Grimm and Wissel 1997). In macroecology and biogeography, a wide range of studies have linked stability to the accumulation of biodiversity in specific areas and to processes such as evolution of the abiotic niche. For example, the relative climate stability of an area through time has been linked to high richness and endemism of species and genetic lineages, compared to less stable areas (Carnaval et al. 2009, Gavin et al. 2014, Cowling et al. 2015, Rosauer et al. 2015, Sandel et al. 2016).

In this context, stability is defined broadly as the continuity of environments, habitats, or populations through space and time. For instance, a site where a single habitat has occurred across millennia (e.g., rainforest) would be considered more stable than a site that has experienced multiple habitat switches (such as repeated shifts between rainforest and grassland) (Costa et al. 2018). Within this broad definition, measures of stability can vary in terms of the entity being measured (e.g., climate, species, or habitat), the spatiotemporal scale used (e.g., global studies over millions of years to local interannual studies), and the method of calculating it (e.g., the variance, mean or extremes). However, studies do not always

clearly define their measure, leading to ambiguity in the interpretation of results. This is the issue we seek to address. We focus on stability as measured over millennia across regional to landscape scales, a topic of many studies, although the framework we describe can be applied to many different spatial and temporal scales.

Most studies on stability as a potential cause of diversity seek to identify landscapes exhibiting higher stability as these areas tend to accumulate more biological diversity than areas with higher stochasticity or variability. This can occur through processes of speciation (as stability may promote speciation) and persistence (as stability may protect taxa from extinction). Stability can promote speciation over long time scales by allowing taxa more time to adapt to their local environment (Klopfer 1959, Fischer 1960) or by isolating populations in separate stable areas, allowing them to diverge (Haffer 1997). High species and functional diversity can then help to stabilise communities and biomes by buffering them against climatic changes over time, for example by increasing resistance as shown by Isbell et al. (2015).

The relatively short time scale of the analyses of stability we consider, typically from the Last Glacial Maximum (LGM, 21kya) or Last Interglacial (LIG, 120kya) to the present, means they are usually focused on persistence rather than speciation. Continuity of environments and habitats can allow diversity to

     1

persist by protecting older communities and lineages from extinction whilst they are lost in less stable areas. For example, stable regions in Africa are believed to have protected Gondwanan lineages such as ricinuleid spiders from extinction during the climatic changes following the breakup of Gondwana (Murienne et al. 2013). Persistence can be assessed explicitly, for instance using population genetic tests for sustained high population size or range expansion, often within the context of refugia – climatically stable regions that allow taxa to persist while the climate in surrounding areas is unsuitable (e.g., Carnaval et al. 2009). Such insights may be relevant to deriving recommendations and policies on how to manage ecosystems for resilience to future climate change (Reside et al. 2013), imbuing the assessment of stability with important practical implications. However, it is important to note that attributes such as increased diversity over time may constitute a biological response to stability but are not measures of stability in themselves.

Measures of stability

The broad perspective on stability adopted across macroecology and biogeography still leaves open a vast array of possible measures of stability that can be employed in any given study. Few authors justify their particular measure, and the measures from different studies vary markedly in terms of the attribute of the system for which stability is being assessed (e.g., climate, habitat), the spatial and temporal scales of analysis, and the strategy used to synthesise complex spatiotemporal results into a summary metric. This leads to ambiguity about what is being measured and how to interpret results. For example, terms such as "climate stability" and "habitat stability" are often used without definition and sometimes interchangeably (e.g., Faye et al. 2016). This issue has been identified before in relation to community ecology (e.g., Grimm and Wissel 1997), but no clear solution has emerged. With a rising number of macroecological studies invoking stability in its various forms (Figure 1), it is important to strive for greater clarity around the definition of measures of stability employed in such studies.

Here, we propose a framework for conceptualising stability within the context of particular biogeographic or macroecological hypotheses and for better understanding the choices which need to be made in defining an appropriate measure of stability in any given context. These choices correspond to five questions that must be answered when conceptualising stability:

1. What are we measuring stability of?

2. What is the spatial scale?

3. What is the temporal scale?

4. How is the interaction between space and time addressed?

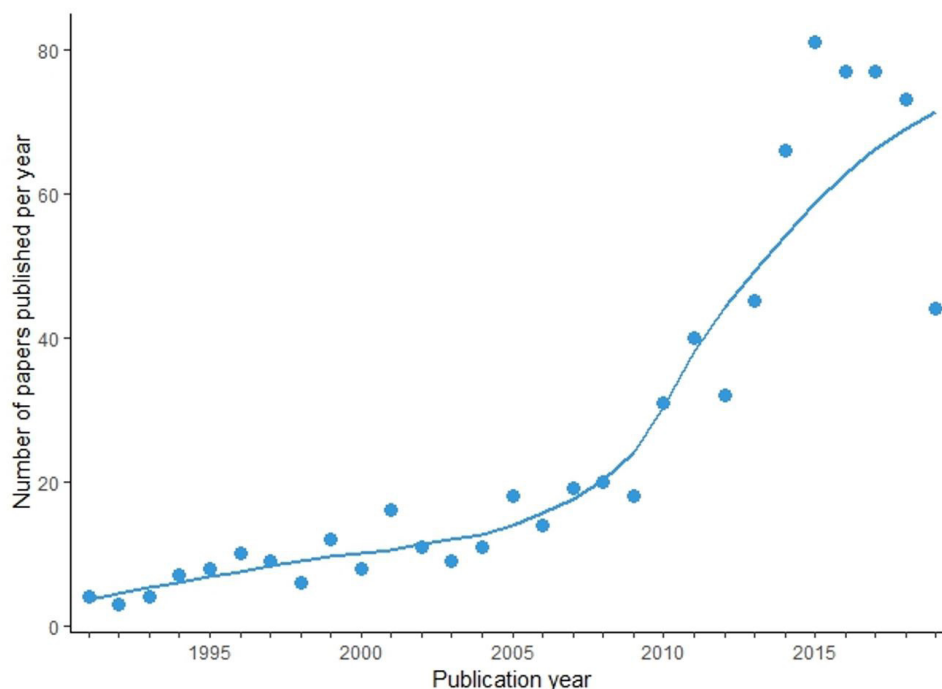5. How do we summarise temporal variation into a single measure of stability for a site?

Figure 1. The number of macroecological studies invoking stability is rising. This shows the publication date for papers in Web of Science using the term "stability" and either "macroecolog*" or "biogeograph*" in their title, abstract, or keywords on 13 August 2019. The line depicts the loess regression as a visual aid.

We hope this framework will assist authors in making informed decisions in selecting and defining measures of stability employed in their studies and will allow for more effective comparison of results across the body of research on this topic.

A framework for defining measures of stability

Confusion around the concept of stability has been acknowledged by various authors (e.g., Pimm 1984, Grimm and Wissel 1997, Donohue et al. 2013), but despite this there is little consistency in how stability is defined and measured.

Our framework focuses on one stability property: spatial continuity through time within geographic or climatic space, as it may be used for predicting alpha and beta diversity at biogeographic scales. It builds on the checklist described by Grimm and Wissel (1997), which sought to help ecologists clarify how the term stability was used. That paper defined three fundamental categories of stability concepts: persistence (persistence through time), resilience (returning to reference state after disturbance), and constancy (staying essentially unchanged). However, we view these as the biological manifestations of a single stability concept, that is, spatial continuity through time. Spatiotemporal continuity is then the driver of other forms of stability.

Additionally, in the two decades since the publication of Grimm and Wissel's (1997) checklist, there have been many studies focusing on the methodology of operationalising the concept of stability. However, this literature has yet to be unified. For example, recent studies have revealed the importance of using fine temporal resolution to capture climate fluctuations (Fordham et al. 2018) and the importance of considering the temporal extent when defining areas of stability (Ashcroft et al. 2012). Other studies have focused on describing metrics for quantifying stability, including climate velocity (Loarie et al. 2009, Brito-Morales et al. 2018) and how the relationship between space and time changes with different metrics (Garcia et al. 2014). Here, we combine each of these components to provide a comprehensive conceptual framework for selecting and defining measures of stability in future studies.

The general concept of stability employed in our framework is illustrated in Figure 2, which shows a variable changing through time, for example species' ranges expanding and contracting through time but maintaining continuity in space. To derive a basic measure of stability for any given point in this region we can create a line graph showing change over time at that point (Figure 2b). From Figure 2, we can see there are several questions and challenges that arise in relation to key features of any given measure of stability, summarised in Table 1.
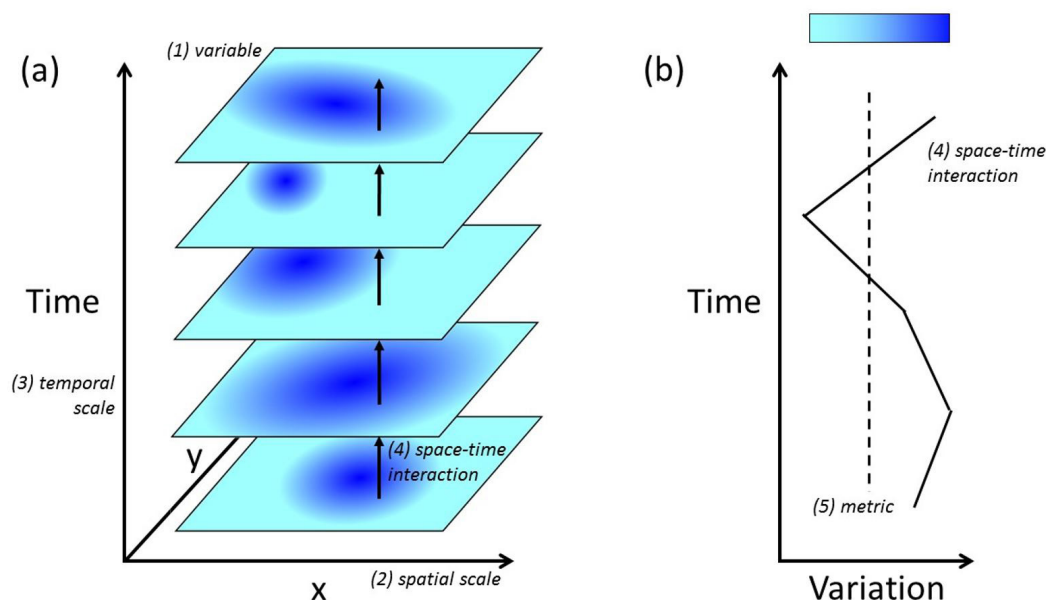

Figure 2. Measuring stability through space and time involves answering several key questions, as discussed in the text. (a) Shows a variable (such as temperature or habitat suitability) changing across space (x and y axes) and through time (vertical axis). The planes show a region at different time points, and the shading represents the variable being measured, for example habitat suitability. The arrows track a single site through time. The choices in metric design are shown in italics. In order to measure stability we need to choose (1) the variable being measured, (2) the spatial scale it is being measured across, (3) the temporal scale being measured across, and (4) the way of measuring the interaction between space and time. (b) Shows stability for the site tracked in (a), summarised into a line graph. To do so, we need to choose (4) the interaction between space and time and (5) the metric used to summarise patterns to a single measure of stability for that site (vertical dashed line), for example the arithmetic mean.

Table 1. Challenges in defining measures of stability

Feature of stability	Checklist question for this feature	Possible options
(1) Ecological attribute	What are we measuring stability of?	Environment, species, assemblage, biome
(2) Spatial scale	What spatial scale is being considered?	Size of the research area (extent), resolution of spatial data (grain)
(3) Temporal scale	What temporal scale is being considered?	Time-period being studied (extent), resolution of time slices (grain)
(4) Interaction between space and time	How is the interaction between space and time addressed?	No interaction (local stability), movement within adjacent regions (neighbourhood stability), dispersal-limited movement (dynamic landscape stability)
(5) Metric	How do we summarise temporal variation into a single measure of stability?	Average, extremes, variance, difference, presence, rate of change

What are we measuring stability of?

Stability can be measured in relation to many different attributes of a system or of a particular entity within that system. At the scales we are concerned with here, we can measure the variability of an environmental parameter such as annual mean temperature or the range of a species, assemblage or biome, measured for example using the bioclimatic envelope.

The simplest way to measure stability, and one of the most commonly used, is to examine it across one or more environmental parameters, such as a measurement of temperature or rainfall. This allows for the measurement of how much, or how fast, the environment has changed over time using measures such as climate velocity (the displacement rate of climate through time divided by the rate through space (Loarie et al. 2009)). Environmental or climate stability therefore refers to continuity in environmental variables at a specific location. For example, southern Africa has high climate stability as the rate of change of mean annual temperature over the last 21ky has been low, while in contrast central Europe has low climate stability as it has experienced large changes in its temperature since the LGM (Sandel et al. 2016). Measures of environmental stability can be used to encapsulate variation in the past or projected future (Garcia et al. 2014), or novel climates that have arisen, or are expected to arise (Williams et al. 2007). Environmental stability can also be used as a proxy for changes in the potential distribution of species or biomes if there are no distributional data available for the biological group of interest (Garcia et al. 2014).

Models of species-level stability rest on estimates of the changing spatial distribution of a species' abiotic niche over time or the continuity in the spatial location of a species range through time. For example, desert pupfish have stable ranges because they inhabit a limited number of desert ponds that have moved little (Brown and Feldmeth 1971). This contrasts with species whose ranges are shifting rapidly either due to direct human intervention (introductions) or to track climate change. Species-level stability can be used to look at questions of extinction or migration under past or future climates (Nogués-Bravo 2009). It is usually measured using correlative ecological niche models (ENMs), which are fitted to the current realised niche then projected into the time periods of interest. There are several issues with this method, including the assumptions that species are in equilibrium with their environment and that a taxon's realised and fundamental niches are equivalent. Correlative ENMs also do not generally account for biotic interactions, non-analog climates, or niche shifts (e.g., Pearson and Dawson 2003, Nogués-Bravo 2009, Fitzpatrick and Hargrove 2009). Hence, there is strong interest in applying more mechanistic models of species stability (e.g., Fordham et al. 2012, Mathewson et al. 2016), but this approach remains difficult to scale up to large numbers of taxa. For now, it seems that practicality dictates use of correlative models in most cases despite their well-known limitations (Wiens et al. 2009).

Compositional or assemblage stability relates to changes in community composition (beta diversity) over time. For example, the Serengeti Plains in eastern Africa have high compositional stability as the community has changed little over time, possibly due to the low rainfall and small species pool (Anderson 2008). This contrasts with areas where the community composition has changed rapidly, for example through species introductions or species range shifts associated

with climate change. Compositional stability is usually measured using macroecologically constrained "stacked" species distribution models (Guisan and Rahbek 2011), but distance matrix-based modelling techniques such as generalised dissimilarity models (GDM) (Catullo et al. 2015) are also used. These models assess the degree to which community composition has been stable over time.

Biome or ecosystem stability is analogous to species stability, but here the goal is to estimate the stability of the range of a biome rather than a species. It is measured with a particular regional assemblage in mind, usually by fitting ENMs to the realised niche of the biome, or sometimes using mechanistic dynamic vegetation models (Thuiller et al. 2008). These methods use models fitted in the present and projected into other time periods to assess the continuity and, hence, stability of the biome or vegetation type (e.g., Costa et al. 2018). Biome stability has been studied in a variety of systems, with clearest results for those with well-defined climatic boundaries such as rainforests (Graham et al. 2010, Rosauer et al. 2015) or regional forest to savanna transitions (Hirota et al. 2011).

These different types of biological stability are interlinked. Compositional stability is impacted by biome stability, as when a biome retreats or expands it affects the community composition at a site. Habitat stability is in turn affected by environmental stability, depending on how broad a climatic tolerance the ecosystem has (West and Salm 2003). This close interaction may explain why many studies looking at climate or habitat stability are unclear about which they are studying despite the concepts being quite distinct (Ashcroft 2010).

Environmental variables

Most studies of stability include a measure of environment, whether explicitly or in models such as ENMs. The term "environment" is very broad. For the current purpose, it comprises the abiotic variables describing a region, including its climate, geology, and topography. These variables can be looked at in two ways: as raw or as transformed variables. Raw variables are those directly measured in the environment, for example annual precipitation as measured by a weather stations, or inferred through a model or proxy, such as annual mean temperature derived from a paleoclimate model. Estimating stability using these variables would directly measure changes in the abiotic environment. Alternatively, measures of stability can be derived using environmental variables which have first been statistically transformed to better reflect observed present-day patterns in the turnover of the species composition of communities across these gradients. For example, methods such as GDM and Gradient Forest use available biological data to statistically transform each of a set of raw environmental variables such that distances within the multivariate space defined by these transformed variables correlate as closely as possible with observed

dissimilarities in present-day species composition between sampled sites (Ferrier et al. 2007, Ellis et al. 2012). This approach scales the relative effect that changes in different environmental variables are expected to have on compositional turnover (e.g., the relative importance of temperature versus precipitation), along with scaling variation in this effect at different points along any given gradient (e.g., a higher rate of turnover per unit change in precipitation at the low versus high end of a precipitation gradient). This scaling of environmental space also allows changes over time to be expressed in terms of the compositional dissimilarity expected between two time points as a function of changes in multiple environmental variables (Blois et al. 2013).

Using either raw or transformed variables, the variables that are most important will depend on the physiology, niche, and ecological interactions of the biological entity of interest (Williams et al. 2012). Regions that are stable for one species or entity may not be stable for another. The best way to identify informative variables, at least for studying single entities, is to include physiological and ecological data, such as those obtained from performance trials and experimental or extensive field studies. However, for many systems these data are not available and are impractical to obtain, such that realised distributions are used as a surrogate. When direct physiological data are not available, data on the ecology of the taxa can be combined with environmental layers and presence/absence data (Williams et al. 2012).

What is the spatial scale?

The issue of scale has been discussed widely in ecological literature since at least the 1970s, with several comprehensive reviews published (e.g., Wiens 1989, Levin 1992, Chave 2013). The importance of conducting studies at an appropriate spatial scale is well-known (e.g., Chase and Leibold 2002, Williams et al. 2002, Cavender-Bares et al. 2006) as processes and correlates that are important at one scale may not be important at others. For example, biotic interactions tend to be important in describing species distributions at local scales, with decreasing importance as the scale increases. In contrast, climate is classically viewed as being an important driver of diversity at regional scale and above and less so at a local scale. However, recent work has shown the importance of microclimates for environmental filtering at local scales, with the mechanisms by which drivers influence biogeographic patterns also changing with scale (Chase and Leibold 2002, Hortal et al. 2010, D'Amen et al. 2017).

Our framework recognises two major components of spatial scale, extent and resolution, both of which need to be chosen carefully based on the patterns and processes being studied. Spatial resolution, also known as grain or focus, relates to the size of the individual spatial units being analysed (Turner et al. 1989, Whittaker et al. 2001). These may be plots of

a few square metres or grid cells of 100 kilometres. As the size of the spatial units increases, variation between cells decreases because more variation is captured in each individual cell (Levin 1992). This means that some patterns, such as micro-refugia, will be more apparent at a fine resolution that captures more variation between cells (e.g., Ashcroft et al. 2012).

Extent refers to the overall size of the analysis region, such as a protected area, biogeographic region, country or global scale (Wiens 1989). A greater extent generally captures more variance between the cells. It is important to note that very few systems are completely closed, so processes and patterns outside the chosen extent may still impact the results (Wiens 1989). Taxa perceive and interact with their environment at different scales, so using a priori behavioural and ecological data will assist in choosing an appropriate scale (Wiens 1989, Rahbek 2004, Anderson et al. 2010).

What is the temporal scale?

Like spatial scale, the temporal scale of a study needs to be defined in terms of both extent and resolution. The temporal extent considered will depend largely on the question being considered. For instance, looking at the stability of an area over a month would give a very different response to looking over a millennial timescale, with the location of areas of stability varying based on the time frame considered (Ashcroft et al. 2012). Without attention to the temporal scale, studies addressing the same question may be mistakenly compared despite measuring very different things. Most studies invoking stability focus on millennial time scales, often since the last interglacial or LGM, although some consider smaller temporal extents, including down to intra-annual time scales (e.g., Martin and Ferrer 2015, Gainsbury and Meiri 2017).

Temporal resolution refers to the number and spacing of time periods considered, represented in Figure 2 by the number of time slices included. A study comparing only the LGM to the present would have a different result to one considering the same temporal extent, but with modelled data for every 100 years, with higher temporal resolution leading to greater accuracy (Fordham et al. 2018). If, for example, the modelled range of a population became regionally extinct at one time but was later re-established, it would not have maintained continuity through time, so it would not be considered stable. However, if one considered only two time points, before and after this discontinuity, this break in continuity would not be identified.

Studies at different temporal scales may not be comparable (Wiens 1989). Different processes operate at different scales, with a gradual shift from ecological to evolutionary processes as the temporal extent lengthens (Chave 2013). Yet, there is a link between variation at different scales, such as between annual temperature range and longer term temperature fluctuations (Janzen 1967). Studies at a large spatial scale often (though not always) use a large temporal scale as well (Wiens 1989). This means that the appropriate temporal scale for a study will depend on the processes being studied and the spatial scale chosen, as well as any time lags between the process and response (Anderson et al. 2010) and the generation times of the organisms being studied, if any (Levin 1992).

How is the interaction between space and time addressed?

Another challenge in describing stability in a region is considering how changes over both space and time interact. How can changes through time across the surrounding landscape be addressed in assessing stability for a single site? Three possible ways of doing this are local stability, neighbourhood stability, or dynamic landscape stability.

The simplest case is local or static stability, where a single site in a region is compared to itself through time (Graham et al. 2010). A stable area would be one that has remained continuously suitable or similar through time. This approach does not take the conditions in adjacent cells into account, although the spatial scale is still important. Local stability is the most commonly measured type of stability.

Neighbourhood stability considers the spatially dynamic nature of environments, whereby a species or biome may persist by moving locally to track changes in the environment. In neighbourhood stability, a single cell is compared to the surrounding cells in the region, to look for analogous environments. Climate change velocity uses this method, comparing change in climate over time to that over space (Sandel et al. 2011).

In more complex dynamic landscape stability models, entities such as species or biomes can shift to track changes across the landscape through time. The size of the surrounding region considered can be scaled depending on the question and organism of interest. The maximum distance allowed from the original cell of interest to a surrounding analogous cell depends on the capacity of the organism or biome to disperse, being larger for a high-dispersal organism such as a bird compared to a low-dispersal organism such as a lizard (Sandel et al. 2011).

The method chosen to combine space and time will have a significant impact on the final measure of stability, as shown in Figure 3. In this example, a site becomes completely unsuitable at one time, suggesting local extinction using a static stability model. However, when using a dynamic stability model (Graham et al. 2010), which allows species or biomes to track contiguous suitable environments through the landscape, changes are much less pronounced.

How do we summarise temporal variation into a single measure of stability for a site?

Having resolved the first four challenges, a final decision is choosing a metric to summarise temporal variation for a site into a single measure. There are
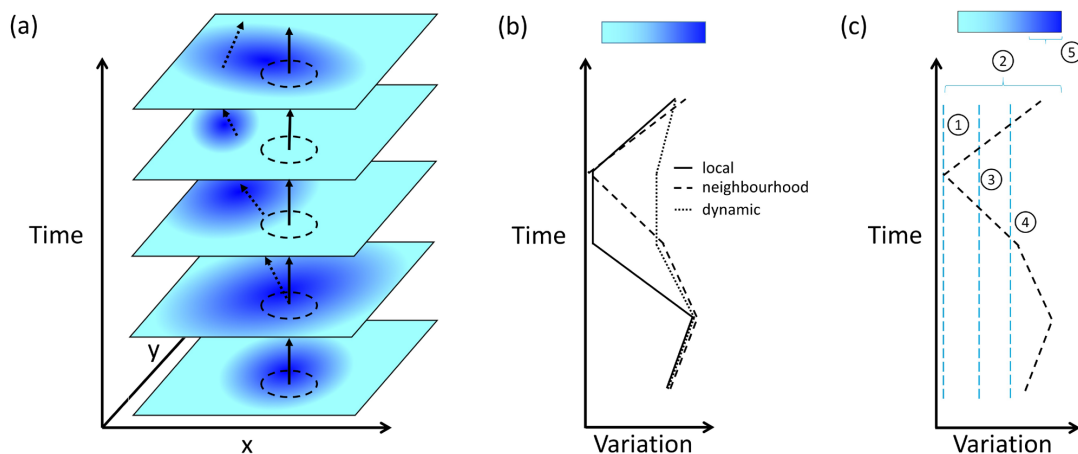
Figure 3. Stability through time for a site: (a) is a variant of Figure 2, showing a variable (for example, habitat suitability for a species) changing over space (x and y axes) and time (vertical axis). Here we show different methods of combining space and time when measuring stability for a site: local stability (red arrow), neighbourhood stability within a radius of the original site (red arrow combined with the circle around the site), and fully dynamic landscape stability allowing for tracking across the landscape (yellow arrow). (b) Shows how stability for that site would be measured across time using all three methods for combining space and time. The shaded bar represents the value of the variable being measured (e.g., habitat suitability for the site), and each line in the plot represents a method of combining space and time. (c) The final step of measuring stability is to obtain a single value for the stability at each site. This illustrates some metrics for doing this. Possible metrics include (1) extremes (shown as the minimum), (2) difference or anomalies (shown as the difference between one end of the time series and the extremes), (3) geometric mean, (4) arithmetic mean, or (5) percentage of time in a given range of values (with the bracket indicating a hypothetical range of values).

Table 2. Commonly used metrics for summarising stability. Biological meanings are defined assuming that stability is being measured for climate, but similar interpretations apply for other levels of stability.

Metric	Definition	Examples of specific metrics	Biological question
Difference between time periods	The amount of change that has occurred between time periods.	Climate anomalies (e.g., Sonne et al. 2016)	How similar is the current environment/available niche to environments in other time periods?
Mean	The climate or suitability of a location averaged across time.	Arithmetic mean; geometric mean (e.g., Graham et al. 2010)	What climatic conditions have taxa had to adapt to?
Rate of change of environment	The speed at which the environment has changed over time.	Climate velocity (Ma et al. 2016)	How well can taxa track the changes in climate?
Extremes	The most extreme conditions or suitability experienced over time.	Maximum temperature; minimum suitability	Could taxa have consistently occurred at this location across time?
Presence in all time periods	The predicted presence of the attribute of interest (e.g., climate, biome or taxa) across all time periods.	Percentage of time in which conditions have been similar to the present; presence of a biome/taxa in all time periods (e.g. Terribile et al. 2012)	How well does the current climate represent the conditions taxa have experienced?
Variance	The variance (for example, in suitability or temperature) between time points.	Standard deviation (e.g. Brown et al. 2014)	How much climatic variability have taxa experienced?

six commonly used classes of metric (see Table 2): difference between time periods, mean, rate of change, extremes, presence in all time periods, and variance between time periods.

Different metrics emphasise different biological processes, so their choice should be driven by the system and question being studied. For example, extremes such as very low suitability may indicate bottlenecks in a population, while the geometric mean is useful in showing whether a region was continuously suitable through time. Some metrics rely on decisions made in other steps. Climate velocity, for example, is a measure of the rate of change of the environment but assumes some form of dynamic stability (where entities can track changes across the landscape) (Ma et al. 2016).

Implementing the framework

Together, these five challenges make up a framework for designing and communicating measures of stability at the biogeographic scale. By working through each of these challenges sequentially, a more robust measure of stability that is relevant to the hypothesis being tested will be designed and communicated. Explicitly considering the variable being measured will ensure that the results can be interpreted in a biologically meaningful way. The choice of spatial and temporal scales will affect the drivers and mechanisms that can be tested for. How stability is summarised into a single number for each site – through both the choice of how space and time interact and the choice of metric – will change the biological meaning of the result and which hypotheses can be tested. Figure 4 summarises the challenges and the options available for each.

Unfortunately, while there have been a few studies measuring the impact of one specific aspect of stability, for example temporal resolution (Fordham et al. 2018) or dynamic and static stability (Graham et al. 2010), there have been no studies systematically altering how stability is measured across the five dimensions

of stability. This gap in the literature means that, while explicitly considering how stability is measured is important from a conceptual and communication perspective, it is difficult to know what impact the current lack of clarity has on the results of studies. Future studies systematically investigating this will allow the impact of consciously choosing a stability measure to be measured.

Despite this lack, some insight can be gained in comparing the results of studies investigating the same region but using different measures of stability. For example, there has been a lot of research on stability of the Australian Wet Tropic rainforests, starting with some of the earliest spatial models of paleoclimate (Nix and Switzer 1991). While this is an intensively studied region, with broad patterns of stability well-established from both paleomodeling and paleoecological data (Vanderwal et al. 2009), variation in the details of results occurs. Much of this is due to differences in the stability metrics used. For example, using dynamic stability consistently shows greater connectivity between refugial areas compared to using static stability (Graham et al. 2010, Rosauer et al. 2015). Changing the spatial extent can make a large difference to predictions of refugia (e.g., Vanderwal et al. 2009). Similarly, the differences in the refugia identified by Bell et al. (2010) and Moussalli et al. (2009) are likely due to a combination of the taxa chosen (widespread versus montane skinks) and the metrics used to summarise across time, specifically the geometric mean versus the product of suitability.

As can be seen in this example and in Box 1, the framework offers a clear foundation for choosing the most appropriate way of measuring stability based on a given hypothesis. Doing so, and clearly communicating the choices made and reasons behind them, will help to enhance interpretation and comparison across multiple studies in this field, while future research will help clarify the quantitative importance of these decisions.
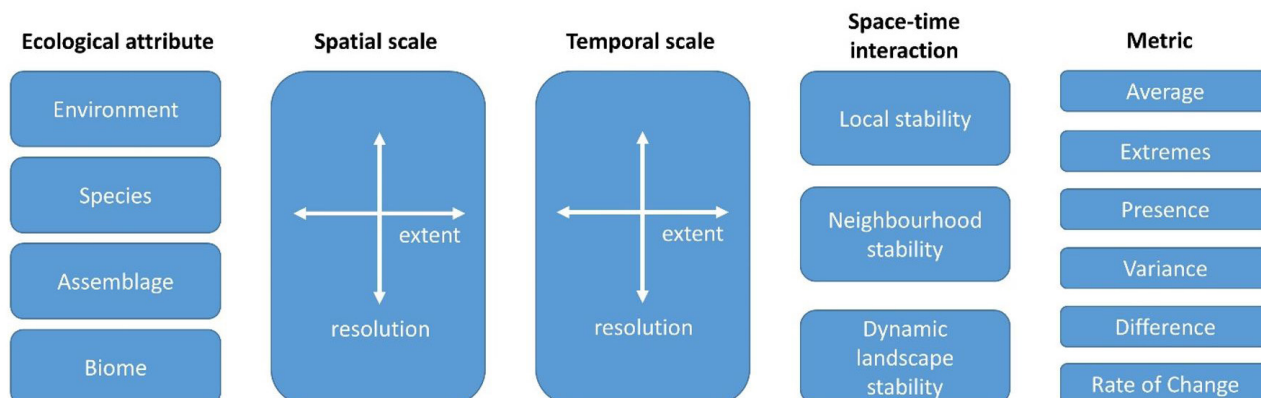

Figure 4. A methodological framework for analysing stability in macroecology and biogeography, showing the five challenges to resolve when defining measures of stability as it relates to diversity at biogeographic scales and the possible options for each.

Box 1 – An example of the framework

Here, we give two examples of how this framework can be used to determine possible approaches to measuring stability (see Figure 5).

Our first example uses stability to test the hypothesis "Areas that acted as refugia through the last glacial-interglacial period have shaped patterns of phylogenetic diversity (PD) in the Australian Wet Tropics (AWT) Rainforest". For testing this hypothesis, the variable being measured would be the biome as a whole, particularly as tropical rainforest as a biome is well-defined climatically and so can be readily modelled using a small number of environmental variables (Hilbert et al. 2007). The spatial scale would ideally be a local to regional resolution, to allow for the identification of fine patterns of PD at the same resolution, with an extent slightly larger than the AWT, including a buffer to allow for past climatic changes and reduce edge effects. The temporal scale would be an extent of the present to the LGM, with as fine a resolution as possible given the available data, and the generation time of the taxa. A common practice is to use only a few time periods – the present, the LGM, and one or two intermediary points in the Holocene, representing the variability observed in pollen records (Kershaw and Nix 1988). While this reduces computation time, having such a low resolution means that key features, such as periods of high velocity, could be missed. Thus, temporal resolution would ideally be of centuries or even decades (e.g., Fordham et al. 2018). Allowing space and time to interact through dynamic landscape stability allows the biome to shift and track suitable climatic conditions (Graham et al. 2010). There are several appropriate metrics that could be used to identify refugia, for example the rate of change (e.g., climate velocity) or the minimum suitability over time. In contrast, the average suitability over time would not be appropriate as areas that have been moderately unsuitable but stable could get the same score as areas that have fluctuated between being highly suitable and highly unsuitable.

Our second example uses stability to identify current microrefugia for a low-dispersal endangered species with a shrinking range induced by climate change. Here, the variable being measured is species stability. The spatial scale would be a local extent with fine resolution in order to incorporate microclimate observations (e.g., Ashcroft et al. 2012). Temporal scale would likely be an extent of fifty to one hundred years, possibly including future projections, with a resolution of years (e.g., Cheddadi et al. 2017). Static stability may be appropriate here as the aim is to identify areas to focus conservation efforts on. Finally, the most appropriate metric would likely be the presence of the species at a site in all time periods.
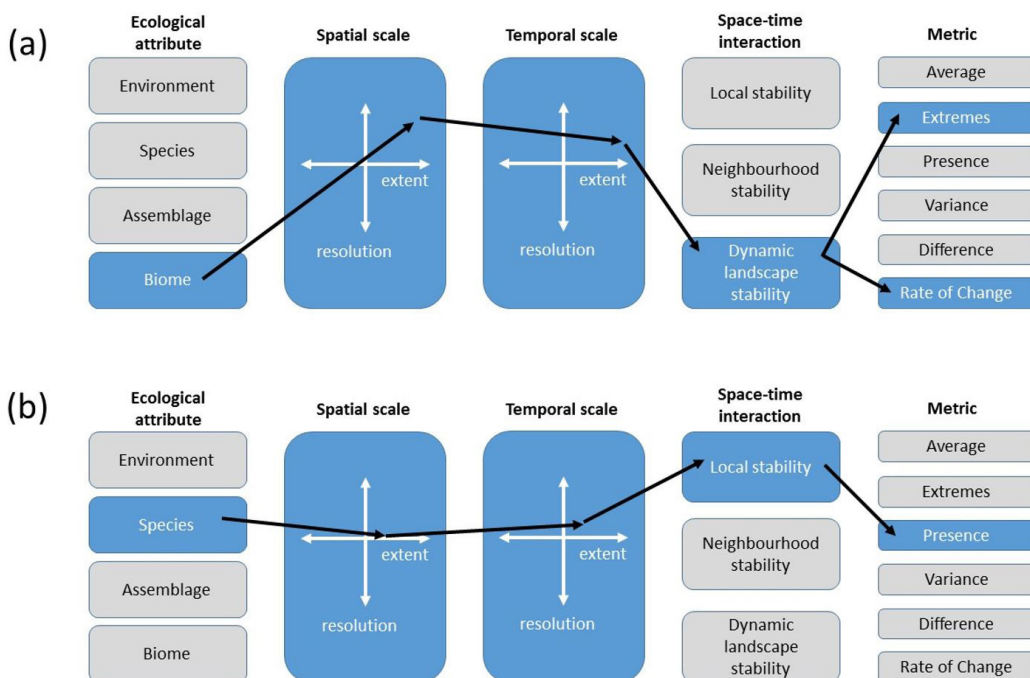

Figure 5. Two examples of how the methodological framework for stability in macroecology and biogeography can be used. (a) Shows appropriate choices for measuring stability when testing the hypothesis "Areas that acted as refugia through the last glacial-interglacial period have shaped patterns of phylogenetic diversity in the Australian Wet Tropics". (b) Shows appropriate choices for measuring stability over much smaller spatiotemporal scales when aiming to identify current microrefugia for an endangered species with a shrinking range.

Implications for future projections and conservation

While the concept of stability has traditionally been used to study the past, an increasing number of studies use the concept of stability to identify areas that may act as refugia under future climate change. These can then be used to evaluate current reserve systems and incorporated into conservation planning (Reside et al. 2013), with refugia now being considered in the creation of government policy as well. For example, the Australian Government's Biodiversity Conservation Strategy explicitly references the need to "identify and protect climate change refugia" (Natural Resource Management Ministerial Council 2010).

With such direct, practical implications, it is even more vital that stability is clearly defined and that an appropriate measure be used. Multiple studies have shown the identification of future refugia, and, hence, appropriate reserve choices are heavily dependent on the methodological choices made (Ashcroft et al. 2012, Keppel et al. 2012, Reside et al. 2013). Employing our framework in studies of future climate change will ensure that sound conservation recommendations can be made.

Acknowledgements

Author contributions: CJM led conceptualisation and writing. All other authors contributed equally to both conceptualisation and manuscript preparation.

References

Anderson, C.D., Epperson, B.K., Fortin, M.-J., Holderegger, R., James, P.M.A., Rosenberg, M.S., Scribner, K.T. & Spear, S. (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. Molecular Ecology, 19, 3565–3575.

Anderson, T.M. (2008) Plant compositional change over time increases with rainfall in Serengeti grasslands. Oikos, 117, 675–682.

Ashcroft, M.B. (2010) Identifying refugia from climate change. Journal of Biogeography, 37, 1407–1413.

Ashcroft, M.B., Gollan, J.R., Warton, D.I. & 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, 1866–1879.

Bell, R.C., Parra, J.L., Tonione, M., Hoskin, C.J., MacKenzie, J.B., Williams, S.E. & Moritz, C. (2010) Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. Molecular Ecology, 19, 2531–2544.

Blois, J.L., Williams, J.W., Fitzpatrick, M.C., Ferrier, S., Veloz, S.D., He, F., Liu, Z., Manion, G. & Otto-Bliesner, B. (2013) Modeling the climatic drivers of spatial patterns in vegetation composition since the Last Glacial Maximum. Ecography, 36, 460–473.

Brito-Morales, I., Garcia Molinos, J., Schoeman, D.S., et al. (2018) Climate velocity can inform conservation in a warming world. Trends in Ecology & Evolution, 33, 441–457.

Brown, J.H. & Feldmeth, C.R. (1971) Evolution in constant and fluctuating environments: thermal tolerances of desert pupfish (Cyprinodon). Evolution, 25, 390–398.

Brown, J.L., Cameron, A., Yoder, A.D. & Vences, M. (2014) A necessarily complex model to explain the biogeography of the amphibians and reptiles of Madagascar. Nature Communications, 5, 5046.

Carnaval, A.C., Hickerson, M.J., Haddad, C.F.B., Rodrigues, M.T. & Moritz, C. (2009) Stability predicts genetic diversity in the Brazilian Atlantic Forest hotspot. Science, 323, 785–789.

Catullo, R.A., Ferrier, S. & Hoffmann, A.A. (2015) Extending spatial modelling of climate change responses beyond the realized niche: estimating, and accommodating, physiological limits and adaptive evolution. Global Ecology and Biogeography, 24, 1192–1202.

Cavender-Bares, J., Keen, A. & Miles, B. (2006) Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. Ecology, 87, 109–122.

Chase, J.M. & Leibold, M.A. (2002) Spatial scale dictates the productivity–biodiversity relationship. Nature, 416, 427–430.

Chave, J. (2013) The problem of pattern and scale in ecology: what have we learned in 20 years? Ecology Letters, 16, 4–16.

Cheddadi, R., Henrot, A.-J., François, L., Boyer, F., Bush, M., Carré, M., Coissac, E., De Oliveira, P.E., Ficetola, F., Hambuckers, A., Huang, K., Lézine, A.-M., Nourelbait, M., Rhoujjati, A., Taberlet, P., Sarmiento, F., Abel-Schaad, D., Alba-Sánchez, F. & Zheng, Z. (2017) Microrefugia, Climate

Change, and Conservation of *Cedrus atlantica* in the Rif Mountains, Morocco. Frontiers in Ecology and Evolution, 5, 114. doi:10.3389/fevo.2017.00114.

Costa, G.C., Hampe, A., Ledru, M.-P., Martinez, P.A., Mazzochini, G.G., Shepard, D.B., Werneck, F.P., Moritz, C. & Carnaval, A.C. (2018) Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling. Global Ecology and Biogeography, 27, 285–297.

Cowling, R.M., Potts, A.J., Bradshaw, P.L., et al. (2015) Variation in plant diversity in Mediterranean-climate ecosystems: the role of climatic and topographical stability. Journal of Biogeography, 42, 552–564.

D'Amen, M., Rahbek, C., Zimmermann, N.E. & Guisan, A. (2017) Spatial predictions at the community level: from current approaches to future frameworks. Biological Reviews, 92, 169–187.

Donohue, I., Petchey, O.L., Montoya, J.M., Jackson, A.L., Mcnally, L., Viana, M., Healy, K., Lurgi, M., O'Connor, N.E. & Emmerson, M.C. (2013) On the dimensionality of ecological stability. Ecology Letters, 16, 421–429.

Ellis, N., Smith, S.J. & Roland Pitcher, C. (2012) Gradient forests: Calculating importance gradients on physical predictors. Ecology, 93, 156–168.

Faye, A., Deblauwe, V., Mariac, C., Richard, D., Sonk?, B., Vigouroux, Y. & Couvreur, T.L.P. (2016) Phylogeography of the genus *Podococcus* (Palmae/Arecaceae) in Central African rain forests: Climate stability predicts unique genetic diversity. Molecular Phylogenetics and Evolution, 105, 126–138.

Ferrier, S., Manion, G., Elith, J. & Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Diversity and Distributions, 13, 252–264.

Fischer, A.G. (1960) Latitudinal variations in organic diversity. Source: Evolution, 14, 64–81.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. Biodiversity and Conservation, 18, 2255–2261.

Fordham, D.A., Resit Akçakaya, H., Araújo, M.B., et al. (2012) Plant extinction risk under climate change: Are forecast range shifts alone a good indicator of species vulnerability to

global warming? Global Change Biology, 18, 1357–1371.

Fordham, D.A., Saltré, F., Brown, S.C., Mellin, C. & Wigley, T.M.L. (2018) Why decadal to century timescale palaeoclimate data are needed to explain present-day patterns of biological diversity and change. Global Change Biology, 24, 1371–1381.

Gainsbury, A. & Meiri, S. (2017) The latitudinal diversity gradient and interspecific competition: no global relationship between lizard dietary niche breadth and species richness. Global Ecology and Biogeography, 26, 563–572.

Garcia, R.A., Cabeza, M., Rahbek, C. & Araújo, M.B. (2014) Multiple dimensions of climate change and their implications for biodiversity. Science, 344, 1247579–1247579.

Gavin, D.G., Fitzpatrick, M.C., Gugger, P.F., et al. (2014) Climate refugia: Joint inference from fossil records, species distribution models and phylogeography. New Phytologist, 204, 37–54.

Graham, C.H., VanDerWal, J., Phillips, S.J., Moritz, C. & Williams, S.E. (2010) Dynamic refugia and species persistence: Tracking spatial shifts in habitat through time. Ecography, 33, 1062–1069.

Grimm, V. & Wissel, C. (1997) Babel, or the ecological stability discussions: An inventory and analysis of terminology and a guide for avoiding confusion. Oecologia, 109, 323–334.

Guisan, A. & Rahbek, C. (2011) SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. Journal of Biogeography, 38, 1433–1444.

Haffer, J. (1997) Alternative models of vertebrate speciation in Amazonia: An overview. Biodiversity and Conservation, 6, 451–476.

Hilbert, D.W., Graham, A. & Hopkins, M.S. (2007) Glacial and interglacial refugia within a long-term rainforest refugium: The Wet Tropics Bioregion of NE Queensland, Australia. Palaeogeography, Palaeoclimatology, Palaeoecology, 251, 104–118.

Hirota, M., Holmgren, M., Van Nes, E.H. & Scheffer, M. (2011) Global resilience of tropical forest and savanna to critical transitions. Science, 334, 232–235.

Hortal, J., Roura-Pascual, N., Sanders, N.J. & Rahbek, C. (2010) Understanding (insect) species distributions across spatial scales. Ecography, 33, 51–53.

Isbell, F., Craven, D., Connolly, J., et al. (2015) Biodiversity increases the resistance of ecosystem productivity to climate extremes. Nature, 526, 574–577.

Ives, A.R. & Carpenter, S.R. (2007) Stability and diversity of ecosystems. Science, 317, 58–62.

Janzen, D.H. (1967) Why mountain passes are higher in the tropics. The American Naturalist, 101, 233–246.

Keppel, G., Van Niel, K.P., Wardell-Johnson, G.W., Yates, C.J., Byrne, M., Mucina, L., Schut, A.G.T., Hopper, S.D. & Franklin, S.E. (2012) Refugia: Identifying and understanding safe havens for biodiversity under climate change. Global Ecology and Biogeography, 21, 393–404.

Kershaw, A.P. & Nix, H.A. (1988) Quantitative palaeoclimatic estimates from pollen data using bioclimatic profiles of extant taxa. Journal of Biogeography, 15, 589–602.

Klopfer, P.H. (1959) Environmental determinants of faunal diversity. The American Naturalist, 93, 337–342.

Levin, S.A. (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. Ecology, 73, 1943–1967.

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

Ma, Z., Sandel, B. & Svenning, J.-C. (2016) Phylogenetic assemblage structure of North American trees is more strongly shaped by glacial-interglacial climate variability in gymnosperms than in angiosperms. Ecology and Evolution, 6, 3092–3106.

Martin, B. & Ferrer, M. (2015) Temporally variable environments maintain more beta-diversity in Mediterranean landscapes. Acta Oecologica, 68, 1–10.

Mathewson, P.D., Moyer-Horner, L., Beever, E.A., Briscoe, N.J., Kearney, M.R., Yahn, J.M. & Porter, W.P. (2016) Mechanistic variables can enhance predictive models of endotherm distributions: the American pika under current, past, and future climates. Global change biology, 23, 1048-1064.

Moussalli, A., Moritz, C., Williams, S.E. & Carnaval, A.C. (2009) Variable responses of skinks to a common history of rainforest fluctuation: Concordance between phylogeography and palaeo-distribution models. Molecular Ecology, 18, 483–499.

Murienne, J., Benavides, L.R., Prendini, L., Hormiga, G. & Giribet, G. (2013) Forest refugia in Western and Central Africa as "museums" of Mesozoic biodiversity. Biology letters, 9, 20120932.

Natural Resource Management Ministerial Council (2010) Australia's Biodiversity Conservation Strategy 2010-2030, Canberra: Commonwealth Government of Australia.

Nix, H.A. & Switzer, M.A. (1991) Rainforest animals: Atlas of vertebrates endemic to Australia's wet tropics, Canberra: Australian National Parks and Wildlife Service.

Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches. Global Ecology and Biogeography, 18, 521–531.

Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? Global Ecology and Biogeography, 12, 361–371.

Pimm, S.L. (1984) The complexity and stability of ecosystems. Nature, 307, 321–326.

Rahbek, C. (2004) The role of spatial scale and the perception of large-scale species-richness patterns. Ecology Letters, 8, 224–239.

Reside, A.E., VanDerWal, J., Phillips, B.L., et al. (2013) Climate change refugia for terrestrial biodiversity: defining areas that promote species persistence and ecosystem resilience in the face of global climate change, Gold Coast: National Climate Change Adaptation Research Facility.

Rosauer, D.F., Catullo, R.A., VanDerWal, J., Moussalli, A., Hoskin, C.J. & Moritz, C. (2015) Lineage range estimation method reveals fine-scale endemism linked to Pleistocene stability in Australian rainforest herpetofauna. PLoS One, 10, e0126274.

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

Sandel, B., Monnet, A.-C., Govaerts, R. & Vorontsova, M. (2016) Late Quaternary climate stability and the origins and future of global grass endemism. Annals of Botany, 119, mcw178.

Sonne, J., Martín González, A.M., Maruyama, P.K., et al. (2016) High proportion of smaller ranged hummingbird species coincides with ecological

specialization across the Americas. Proceedings of the Royal Society, 283, 20152512.

Terribile, L.C., Lima-Ribeiro, M.S., Araújo, M.B., et al. (2012) Areas of climate stability of species ranges in the Brazilian Cerrado: Disentangling uncertainties through time. Natureza a Conservacao, 10, 152–159.

Thuiller, W., Albert, C., Araújo, M.B., et al. (2008) Predicting global change impacts on plant species' distributions: Future challenges. Perspectives in Plant Ecology, Evolution and Systematics, 9, 137–152.

Turner, M.G., O'Neill, R. V, Gardner, R.H. & Milne, B.T. (1989) Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology, 3, 153–162.

Vanderwal, J., Shoo, L.P. & Williams, S.E. (2009) New approaches to understanding late Quaternary climate fluctuations and refugial dynamics in Australian wet tropical rain forests. Journal of Biogeography, 36, 291–301.

West, J.M. & Salm, R. V. (2003) Resistance and resilience to coral bleaching: Implications for coral reef conservation and management. Conservation Biology, 17, 956–967.

Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: Towards a general, hierarchical theory of species diversity. Journal of Biogeography, 28, 453–470.

Wiens, J.A. (1989) Spatial scaling in ecology. Functional Ecology, 3, 385–397.

Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A. & Snyder, M.A. (2009) Niches, models, and climate change: Assessing the assumptions and uncertainties. Proceedings of the National Academy of Sciences USA, 106, 19729–19736.

Williams, J.W., Jackson, S.T. & Kutzbach, J.E. (2007) Projected distributions of novel and disappearing climates by 2100 AD. Proceedings of the National Academy of Sciences USA, 104, 5738–5742.

Williams, K.J., Belbin, L., Austin, M.P., Stein, J.L. & Ferrier, S. (2012) Which environmental variables should I use in my biodiversity model? International Journal of Geographical Information Science, 26, 1–39.

Williams, S.E., Marsh, H. & Winter, J. (2002) Spatial scale, species diversity, and habitat structure: small mammals in Australian tropical rain forest. Ecology, 83, 1317–1329.