Tsai, Cheng-Han (2020) ***Variation in relative species abundance of reef fishes: drivers and ecosystem impacts.*** **PhD Thesis, James Cook University.**

https://doi.org/10.25903/hmc4%2Dca16

Variation in relative species abundance of reef fishes: drivers and ecosystem impacts

Thesis submitted by

Cheng-Han Tsai

For the degree of Doctor in Philosophy
AIMS@JCU, College of Science and Engineering,
and ARC Centre of Excellence for Coral Reef Studies,
James Cook University

Apr 2020

Acknowledgements

Firstly, I would like to thank Sean Connolly, for his support throughout my PhD. The honour is mine, and I am extremely lucky, as an international student, to have him as my mentor helping me to smooth the life transitions in Australia. G'day Sean, you are the best and thank you.

I would like to thank my co-supervisor, Hugh Sweatman, for his kindly support and patience, especially his participation in each of my milestone presentations. Thanks heaps, Hugh.

I would also like to thank members of the Ecological Modelling Research Group at James Cook University (Loïc Thibaut, Martino Malerba, Mariana Álvarez-Noriega, Kevin Bairos-Novak, Katie Peterson, Alfonso Ruiz, Jessica Zamborain-Mason, Joanne Moneghetti, Kevin Go, Jess Hopf, Shane Blowes, Saskia Juriaans, Mia Hoogenboom and Danielle Christine) for their feedback throughout my PhD years. I would also like to thank my previous supervisors, Chih-hao Hsieh and Takeshi Miki, for their comments on the early version of my PhD thesis, and my secondary supervisor, David Bellwood, for his insights of classification of reef fish functional groups. I am very thankful to AIMS@JCU, James Cook University and the ARC Centre of Excellence for Coral Reef Studies for the financial support and for the many opportunities provided.

Last but not least, I would like to thank Lik-Theng Ho, my dear family and friends (e.g., Ciemon Caballes, C-Y Kuo, Andreas Dietsel, Maria Palacios, Renato Morais, Collin Wen, Man Lok Yu, Daniel Tong, Gladys Chua, Andy Chu, Alexander Sun, Juno Siu, Jungling Wang, Xiaonan Li, Hanaka Mera, Tracy MacKeracher, Arun Oakley, Kristina Pahang, and Jeremy Horowitz) for their many forms of support throughout my PhD life.

Statement of contribution of others

This thesis was done under the supervision of Prof. Sean Connolly and Dr. Hugh Sweatman. The data used was collected by Long-term Monitoring Program of Australian Institute of Marine Science (LTMP). Contribution to individual chapters can be found in Table I.

Financial support was provided by AIMS@JCU, James Cook University, and the ARC Centre of Excellence for Coral Reef Studies.

Table I. Statement of contribution to thesis chapters. CHT is Cheng-Han Tsai, SRC is Sean R. Connolly, HS is Hugh P. Sweatman.

Chapter	**Statement of contribution**
Chapter 1: General introduction	CHT wrote the chapter with feedback from SRC, HS.
Chapter 2: Increasing environmental volatility erodes niche structure, but not diversity, in coral reef fish assemblages	CHT, SRC, and HS conceived the research. CHT wrote the chapter with feedback from SRC, HS.
Chapter 3: Robustness test for variance partitioning of relative species abundance from simulated community time-series data	CHT and SRC conceived the research. CHT wrote the chapter with feedback from SRC.
Chapter 4: Portfolio effect and diversity-stability relationship depend on how communities are structured across coral reefs	CHT, SRC, and HS conceived the research. CHT wrote the chapter with feedback from SRC, HS.
Chapter 5: General discussion	CHT wrote the chapter with feedback from SRC, HS.

Thesis abstract

On coral reefs, drivers of the variation in relative species abundance of reef fishes are not well understood. Proximately, these drivers can be conceptualized as partitioned between the variation dependent on species' characteristics (such as niche traits) that maintain the relative abundances of species deterministically over time, versus the variation due to stochastic effects of demographic and environmental fluctuations. Ultimately, the relative importance of these deterministic and stochastic drivers can vary regionally and among habitats, depending on biotic and abiotic environmental conditions. In this thesis, I used an analytical approach, variance partitioning of relative species abundance (VPRSA), and long-term reef fish community data across the Great Barrier Reef (i) to estimate the proximate importance of species' niche differences, relative to demographic and environmental fluctuations, as drivers of patterns of commonness and rarity (*Chapter 2*), (ii) to understand whether and how the relative importance of these deterministic versus stochastic components ultimately depend on environmental conditions (*Chapter 2*), (iii) to determine how the above proximate and ultimate drivers influence community-level stability and the diversity-stability relationship on coral reefs (*Chapter 4*), and (iv) to evaluate the robustness of inferences about niche versus stochastic process drawn from VPRSA to violation of those simplifying assumptions of the theoretical model from which the partitioning approach is derived (*Chapter 3*). In *Chapter 2*, I found that the variation in relative abundance of reef fishes is strongly niche structured (above i), and that this deterministic niche structure is eroded on reefs with high temporal variability in coral cover (above ii). in contrast, geographical variation in coral cover variability has no detectable relationship to diversity patterns such as richness and evenness. In *Chapter 3*, I confirmed the robustness of inferences about the importance of deterministic niche versus stochastic process drawn from this VPRSA approach (above iv). In *Chapter 4*, I found that total fish community abundance is more stable on more species-rich reefs of the Great Barrier Reef. However, regional variation in the relative importance of deterministic niche versus stochastic components of the species abundance distribution has an equally large effect on fish community stability (above iii). Overall, my thesis establishes the first process-based framework to understand and tease apart proximate and ultimate drivers of variation in relative abundance of reef fishes, as well as the corresponding ecosystem impacts, across natural coral reef gradients at a regional scale.

Table of contents

List of tables

Chapter tables

Appendix tables

List of figures

Chapter figures

Appendix figures

Chapter 1: General introduction

1.1 Thesis overview

Species-rich communities typically consist of a small number of highly-abundant species and many rare species, i.e., the frequency distribution of species abundances (species-abundance distribution, hereafter SAD) is approximately lognormal (Preston 1948; Sugihara 1980; Hubbell 1997; Magurran & Henderson 2003; McGill *et al.* 2007; Ulrich *et al.* 2010; Sæther *et al.* 2013; Connolly *et al.* 2014; Matthews & Whittaker 2015). However, SADs are not static. Depending on the community dynamics underlying SADs, the extent to which species are common or rare can vary substantially over space and time (Fung *et al.* 2016; Engen *et al.* 2017). The extent to which variation in SADs are driven by persistent differences in species' intrinsic traits, responses to environmental fluctuations, and demographic stochasticity (chance variation in the fates of individuals) has important implications for understanding biodiversity maintenance, ecosystem functioning, and to anticipate how species-rich community structure, including alpha- and beta-diversity patterns, may respond to changing environmental regimes or management interventions (McGill *et al.* 2007; Sæther *et al.* 2013; Matthews & Whittaker 2015).

In the last few decades, there has been a proliferation of process-based models proposed to explain the drivers of variation in SADs (Hubbell 1997; McGill *et al.* 2007; Chisholm & Pacala 2010; Kalyuzhny *et al.* 2015; Engen *et al.* 2017). These process-based models differ from phenomenological statistical distributions (e.g., Fisher et al. 1948; Pielou 1975; Šizling et al 2009; White et al. 2012; Mathews et al. 2014; Connolly et al. 2017) whose parameters do not necessarily correspond to directly measurable ecological quantities, such as demographic rates. However, process-based models vary in the extent to which they explicitly represent the details of demographic and ecological differences among species. Some models explicitly characterize such differences on a species-by-species basis, such as accounting for species differences in competition for resources or differences in how particular species' demographic rates respond to environmental conditions (Ellner *et al.* 2016). Other models omit species-specific attributions: for example, in neutral community dynamics theory, all species are assumed to have the same birth and death rates, and equivalent per-capita dispersal propensities and competitive abilities (Hubbell 1997; Volkov *et al.* 2007) Between these two extremes, community dynamics theory developed by Engen and coworkers (Engen & Lande 1996; Engen *et al.* 2002; Lande *et al.* 2003; Sæther *et al.* 2013) allows for species differences by specifying distributions for demographic rates and responses to environmental fluctuations, without attributing particular values to particular species.

Species-by-species modelling approaches allow more mechanistic detail in explanations of community structure and dynamics, including application of, e.g., "modern coexistence theory" (hereafter MCT) (Ellner *et al.* 2016), but they are intractable in species-rich systems, due to what is called the "many body problem" in statistical mechanics (Volkov *et al.* 2007). The problem in the language of ecology might be expressed as follows: the number of parameters needed for species-by-species community dynamics models increases faster than the amount of information available, as the number of species increases. Specifically, the number of between-species interaction terms and covariance parameters (which describe how correlated different species' responses to environmental fluctuations are) in species-by-species models typically increases with the square of the number of species in the community, precluding estimation of those parameters from either experimental manipulations, or from fits of models to time series data (i.e., the "curse of dimensionality") (Ives *et al.* 2003; Ovaskainen *et al.* 2017). Moreover, in the ecological context, much of the species pool of high-diversity assemblages consists of rare species, for which even baseline demographic rates like birth and death rates may be difficult to obtain (Hubbell 1997, 2001; Volkov *et al.* 2007; Chisholm & Pacala 2010). By contrast, the more high-level modelling approaches can only be used to explain aggregate community patterns, such as SADs, but they cannot explain why particular species are more abundant or widespread than others in detail (Kalyuzhny *et al.* 2015; Connolly *et al.* 2017)

To date, "static" SADs have been used commonly for testing predictions from both phenomenological and process-based models. However, many models, both phenomenological and process-based, describe SAD patterns almost equally well (McGill *et al.* 2007; Ulrich *et al.* 2010; Connolly *et al.* 2014). This lack of discrimination has raised concerns about testing between alternative models for community structure by identifying the best-fitting model for SADs (McGill *et al.* 2007). Recent studies suggest that analysing temporal dynamic patterns of SADs can provide critical information about the underlying processes that cannot be otherwise obtained from analysing static SADs (Sæther *et al.* 2013; Fung *et al.* 2016; Engen *et al.* 2017). For instance, one might be able to identify the dynamic drivers of SAD by examining the temporal correlations in species-abundance patterns (Fung *et al.* 2016; Engen *et al.* 2017).

One dynamical approach is to decompose the overall variance in SADs into separate contributions of community processes: specifically, the role of deterministic differences between species (e.g., interspecific variation in long-term mean abundances due to differences in demographic traits), the role of environmental stochasticity that induces fluctuations in relative species abundance, and the roles of demographic stochasticity (i.e., chance variation in the fates of individuals, regardless of species) and sampling effects due to e.g., local aggregation. This variance partitioning approach (hereafter, "variance partitioning of relative species abundance", VPRSA) was developed by Engen and co-workers as an extension of their stochastic-dynamic theory of community structure (Engen & Lande 1996; Engen *et al.* 2002; Lande *et al.* 2003). Although VPRSA can be readily applied to commonly-collected survey data (Sæther *et al.* 2013), it has received far less attention in the literature compared to alternative community dynamics theories, such as neutral theory for tropical tree and coral reef communities (Hubbell 1997, 2001; Volkov *et al.* 2007; Bode *et al.* 2012; Fung *et al.* 2016).

The VPRSA approach is especially useful for analysing temporal dynamics of SADs in species-rich assemblages. For instance, there has been an extensive debate about the importance of niche versus "neutral" processes (i.e., pure demographic stochasticity) in structuring high-diversity assemblages like tropical coral reefs (Hubbell 1997; Dornelas *et al.* 2006; Volkov *et al.* 2007; Chisholm & Pacala 2010; Engen *et al.* 2017). In marine systems and coral reefs in particular, this debate has largely been informed by the analysis of static rather than dynamic SADs (Connolly *et al.* 2005; Dornelas *et al.* 2006; Volkov *et al.* 2007; Bode *et al.* 2012). However, even when the overall shape of SADs are static, or when species richness and evenness (i.e., alpha-diversity) are constant over time, the temporal correlations in species abundances (i.e., temporal beta-diversity) may vary substantially depending on how much the variation in a SAD is driven by long-term species differences, versus for instance differential responses to environmental fluctuations (Fung *et al.* 2016; Engen *et al.* 2017). By analysing temporal autocorrelation in relative abundances, VPRSA is therefore capable of estimating whether, and to what extent, the patterns of commonness and rarity in species-rich communities are structured by deterministic versus stochastic processes.

On coral reefs, the relative importance of niche structure, neutral processes, and environmental fluctuations in driving SAD and diversity patterns are not well understood (Roughgarden 1974; Connell 1978; Anderson *et al.* 1981; Chesson & Warner 1981; Hubbell 1997; Connolly *et al.* 2005; Dornelas *et al.* 2006; Volkov *et al.* 2007; Bode *et al.* 2012). Proximately, the overall variance in SADs depends on deterministic, species-differentiated trait values that maintain relative abundances over time, as well as on the varying effects of environmental and neutral fluctuations that generate additional variance in species abundances (see Box 1. for definitions and terms). Ultimately, across reefs, the relative importance of these drivers can vary with macro-ecological environmental variables, such as mean and variability in coral cover, latitude, and cross-shelf position. However, the relationships among environmental variables and the variance components of reef community structure as described above are unclear.

The broader ecosystem consequences of the processes that drive dynamics of species' relative abundances are not well understood either. For instance, most previous studies suggest that community-level stability (e.g., temporal stability in community total abundance) is strongly determined by species richness and evenness via the "portfolio effect" (see Box 1. for definitions and terms) (Doak *et al.* 1998; Tilman 1999; Hooper *et al.* 2005; Hector *et al.* 2010; Thibaut & Connolly 2013; Schindler *et al.* 2015). The evidence for such diversity-dependent stabilizing effects comes overwhelmingly from manipulative experiments, in which the dynamics of species' relative abundances are controlled (Campbell *et al.* 2011; Gross *et al.* 2014). By contrast, in nature, diversity typically covaries with other properties of ecological communities, such as relative abundance, along environmental gradients (Sankaran & McNaughton 1999; Srivastava & Vellend 2005; Wardle 2016). When these naturally-occurring assemblages are compared, evidence for a consistently stabilizing effect of diversity is much more ambiguous (Wardle 2016; Blüthgen *et al.* 2016). The extent to which the proximate and ultimate determinants of species' relative abundances discussed above may influence or obscure the relationship between diversity and stability is not known.

1.2 Thesis outline

The overall aim of this thesis is to develop and implement novel applications of the VPRSA approach to understand the drivers of regional-scale variation in species abundances in reef fish assemblages of the Great Barrier Reef, and their consequences for community stability. Specifically, I will:

1. Estimate the relative contributions of neutral process, species' responses to environmental variability, and persistent niche-related differences among species to patterns of commonness and rarity.
2. Determine how much those contributions of community structuring process vary at the regional scale, and how much of that variation can be explained by coral cover dynamics and environmental gradients of latitude and cross-shelf position.
3. Test whether and how the relative contribution of community structuring processes influences the diversity-stability relationship.
4. Because the VPRSA approach was explicitly derived from a process-based theory that makes important simplifying assumptions about community dynamics, I also evaluate the robustness of inferences about community structure drawn from VPRSA to violation of those simplifying assumptions.

Chapter 2 focuses on the first two aims. Therein, I test whether coral cover fluctuations, as a proxy for disturbance dynamics, affect reef fish diversity patterns and the importance of "niche structure" (see Box 1). I find that relative abundances in reef fish assemblages are, in general, highly deterministically structured, with on average 80% of the variance in relative species abundances attributable to persistent species differences in average abundances. However, this deterministic niche structuring is eroded on reefs where fluctuations in coral cover are higher. In contrast, aggregate properties of community structure, such as species richness and unevenness, are comparatively insensitive to coral cover fluctuations; instead, they respond to more conventional macroecological variables such as latitude and cross-shelf position. Separate analyses of three species-rich functional groups indicates that the disproportionate importance of deterministic species differences, and the macroecological variation in these species-abundance variance components, are due to broadly consistent responses across functional groups, rather than by macroecological changes in functional composition, although the strengths of these relationships vary somewhat among functional groups. My findings suggest that the long-term impacts of disturbance regime on reef community dynamics might have been underestimated in previous studies, because most of those studies have focused on short-term changes in community structure [e.g., comparing before-and-after change in composition and relative abundance in a short time period; (Wilson *et al.* 2006; Hughes *et al.* 2018b), or a time snapshot of community structure across spatial gradient; (Sandin *et al.* 2008; Ruppert *et al.* 2018)] and have focused on differences between static measures (such as richness, evenness, and static species-abundance distribution) rather than temporal variations in relative abundance patterns.

Chapter 3 tests, for the first time as far as I am aware, the robustness of estimates from VPRSA to relaxation of important simplifying assumptions of the original underlying theory from which VPRSA was derived (aim 4 above). Specifically, I apply VPRSA to simulated community data that violate several of these assumptions. I also evaluate a random-effects approach of VPRSA to estimating variance components for time series that are spatially replicated but relatively short in duration, in order to determine whether parameter estimates obtained in this way are better than those obtained by estimating variance components independently for each time series. I show that estimated variance components are robust to violations of the original VPRSA assumptions, and that adopting a random-effects approach to estimating these variance components is an effective way to leverage spatial replication when time series are short. Moreover, I show that estimates of the variance components of community structure are robust, even when the interpretation of these diversity components in terms of particular parameters of the stochastic-dynamic theory may not hold. This chapter concludes that VPRSA indeed provides a much more robust estimate of the relative importance of deterministic and stochastic processes in high-diversity assemblages than one might assume based on the assumptions of the original theory from which the approach was derived.

In *Chapter 4*, I test whether and how the relative contribution of stochastic and deterministic processes influences the diversity-stability relationship on coral reefs (aim 4 above). Specifically, based on the analytical framework of *Chapter 2* and *Chapter 3*, I develop a “niche structure index” to measure the relative determinism/stochasticity of community structure by estimating the relative contribution of intrinsic species differences versus environmental and demographic stochasticity (i.e., the VPRSA approach). If the community is highly structured by species differences in intrinsic traits, then the value of the niche structure index is higher. Conversely, a lower value of the niche structure index represents increased stochasticity in community structuring and increased temporal community turnover. Equipped with this index, I test the simultaneous effect of changes in alpha-diversity (measured by species richness and evenness) and temporal-beta diversity (measured by niche structure index, NSI) on community stability. Specifically, the simultaneous effects of richness, evenness, and the niche structure index are incorporated within an expanded diversity-stability framework (see Box 1), providing an integrated assessment of the broader ecosystem consequences of alpha- and temporal beta-diversity change. In this chapter, I show a strong stabilizing effect of species richness and evenness, which occurs by reducing the synchrony in abundance fluctuations rather than by reducing average population variability. Also, the niche structure index affects community stability, but this effect occurs by mediating average population variability rather than synchrony. This chapter is the first regional-scale documentation of the diversity-stability relationship on coral reefs and, importantly, it suggests that regional variation in community structuring processes has a stabilizing effect that is comparable in magnitude to those of richness and evenness.

Table 1.1 Definitions and Terms

Variance components of community structure

Throughout the thesis, "variance components of community structure" refer to the following factors that can change the relative species abundance of a community through time. Specifically, components are classified as follows:

1. ***Heterogeneity in species' intrinsic traits and carrying capacities:*** this encompasses any factors that tend to maintain species' relative abundances through time (i.e., that generate persistent differences in species' long-term mean abundances, for instance due to niche packing or other persistent demographic differences between species). This is termed "persistent species differences" or "niche structure" in the chapters.
2. ***Environmental stochasticity or species' responses to environmental perturbations:*** any external stochastic perturbations to population growth rates that cause species' relative abundances to deviate from their expected deterministic trajectories towards their respective carrying capacities.
3. ***Demographic stochasticity:*** random variation in births and deaths of individuals. Unlike environmental stochasticity, the variance induced by demographic stochasticity is inversely proportional to population size (i.e., it has proportionately larger effects on small populations, and decreases rapidly in importance as population size increases).
4. ***Sampling stochasticity:*** variation in observed species abundances due to random sampling effects. Like demographic stochasticity, this process is also proportionately larger for smaller populations

The latter two sources of stochasticity (i.e., demographic and sampling stochasticity) are not distinguished from each other in this study because their effects on relative species abundance dynamics are not statistically distinguishable for data such as that analysed in this thesis.

Diversity-stability relationship

1. ***Community stability:*** temporal stability is measured as the inverse of coefficient of variation (CV) in aggregate multispecies abundance (Tilman 1999, Tilman et al. 2014).
2. ***Portfolio effect:*** an analogy of portfolio selection in finance (Markowitz 1952; Doak et al. 1998), generally indicating that the variability of an aggregate property of fluctuating components decreases in volatility with increasing number of different components. Also termed "statistical averaging" in previous studies (Doak et al 1998; Cottingham et al 2001).
3. ***Temporal mean-variance scaling of population variability:*** the tendency for temporal variance in a species' abundance to increase with its mean abundance in time series. Specifically, this tendency follows a power scaling function (i.e., Taylor's power law), with a scaling exponent between 1 and 2, so that both temporal variance and coefficient of variation of abundance tend to increase with mean abundance (Tilman 1999; Kilpatrick and Ives 2003). Tilman 1999 demonstrates that species diversity tends to increase temporal variance in species abundance fluctuations, whereas it decreases temporal variance in community total abundance (i.e., diversity decreases population stability but increases community stability).
4. ***Synchrony/asynchrony of population fluctuations:*** a measure of how different species' abundances are correlated with each other over time (e.g., Loreau and de Mazancourt 2008, Gross et al. 2014). It varies from zero for perfectly compensating fluctuations (total community abundance is constant through time) to unity when all species' fluctuations are perfectly positively correlated.

Chapter 2: Increasing environmental volatility erodes niche structure, but not diversity, in coral reef fish assemblages

2.1 Introduction

The problem of how deterministic and stochastic factors shape patterns of commonness and rarity of species has engaged ecologists for over a century, and this problem remains enigmatic, particularly for species-rich systems (Preston 1947; MacArthur 1957; Pielou 1975; Engen and Lande 1996; Hubbell 2001; Magurran and Henderson 2003; Sugihara et al. 2003; Tilman 2004; McGill et al. 2007; Chisholm and Pacala 2010; Saether et al. 2013; Connolly et al. 2014; Kalyuzhny et al. 2015). Research into this problem was revolutionized by the development of neutral theory of biodiversity, which attributes variation in species abundances to only demographic stochasticity (chance variation in births and deaths of individuals) and dispersal limitation, and thus offers a tractable set of models that can be confronted directly with abundance data in highly speciose systems (Hubbell 2001; Volkov et al. 2007; Rosindell et al. 2011; Chust et al. 2013). However, neutrality alone has proved inadequate to explain numerous important features of patterns of commonness and rarity in nature, including temporal drift in species abundances (McGill et al. 2005; Dornelas et al. 2014; Engen et al. 2017), spatial patterns of community similarity (Dornelas et al. 2006; Connolly et al. 2017), and the overall extent of heterogeneity in species' abundances (Connolly et al. 2014).

More recently, several extensions to neutral theory have been developed which might be collectively termed "quasi-neutral theory". Quasi-neutral models assume species to have the same demographic rates on average, but species differences are implicitly allowed by modelling their dynamics independently of one another (Volkov et al. 2007), or by allowing their demographic rates to fluctuate in different ways (Kalyuzhny et al. 2015). Such implicit species differences can give species a demographic advantage as they become rare, facilitating population recovery, reducing extinction rates and supporting higher levels of biodiversity than neutral models (Chesson 2000; Kalyuzhny et al. 2015). Consequently, species differences in response to stochastic fluctuations maintain the variance in the distribution of commonness and rarity through time, although species' dominance within that distribution can interchange.

Empirically, coral reefs have been a focal system for tests of neutral theory as well as assessments of the importance of species/niche differences, due to their high levels of biodiversity. An early test of neutral theory applied to coral assemblages showed that neutral model failed to account for the high variability in patterns of community similarity (Dornelas et al. 2006). That test was subsequently challenged by another neutral model (Volkov et al. 2007), although that alternative neutral model has also now been shown to perform poorly (Connolly et al. 2017). Both reef coral and reef fish assemblages appear to have more heterogeneity in species abundances (i.e., many rare species coupled with common species that are highly numerically dominant) than neutral models can produce (Connolly et al. 2009, Connolly et al. 2014). However, these analyses do not reveal the extent to which this heterogeneity is due to persistent, deterministic differences in species' abundances (i.e., a tendency for some species to be substantially more common than others), versus interspecific differences in responses to environmental fluctuations that amplify the variability in species' relative abundances.

In particular, coral reefs are profoundly influenced by stochastic environmental disturbances (Connell 1978; Karlson & Hurd 1993; Hughes et al. 2018b), which induce substantial changes in species composition and relative abundance over space and time. Differences in species' susceptibility to those disturbances and differences in their rates of recovery from them suggest that environmental stochasticity could play an important role in determining patterns of commonness and rarity on reefs (Karlson & Hurd 1993). On the other hand, there is also evidence for persistent differences in species' mean abundances over quite long time periods: for example, *Acropora cervicornis* and *Acropora palmata* remained substantially more abundant than other coral species throughout the Pleistocene in Barbados (Bode et al. 2012). Indeed, both species differences and environmental stochasticity have been invoked as major drivers of biodiversity maintenance on coral reefs (e.g., Connell 1978; Anderson *et al.* 1981; Chesson & Warner 1981; Dornelas *et al.* 2006). Consequently, the existing body of knowledge calls for a framework for the analysis of commonness and rarity that can determine the relative contributions of the stochastic and deterministic factors, rather than a comparison of models that strictly omit one or the other.

To undertake evaluations of the relative importance of factors such as deterministic species differences versus environmental fluctuations to species abundances and biodiversity, some recent studies have turned to the analytical framework of modern coexistence theory [MCT: (Chesson 2000; Ellner *et al.* 2016; Usinowicz *et al.* 2017)]. These are detailed models of species interactions in which particular features that promote coexistence can be turned on or off to evaluate their effects on species abundances or on overall species richness. However, MCT requires detailed, species-level information about the demographic rates of species, and the precise nature of density-dependence, rendering the approach unfeasible for many high-diversity systems like coral reefs, where time series are commonly limited to counts of individuals through time (Dornelas et al. 2006; Sweatman et al. 2005; Blowes et al. 2019), and estimates of individual demographic rates are limited to only a small subset of the assemblage on reefs (Madin et al. 2014; Alvarez-Noriega et al. 2016; Dornelas et al. 2017).

An alternative to MCT is to employ a stochastic-dynamic theory of biodiversity which allows use of the temporal evolution of relative abundances (i.e., temporal beta diversity) in order to partition the variation in species abundances into the contributions of long-term differences in species' mean abundances (i.e., species traits that tend to cause species to have high or low local abundances), the contribution of environmentally-induced fluctuations in species' demographic rates, and the contribution due to other factors, such as demographic and sampling stochasticity (Engen and Lande 1996; Engen *et al.* 2002; Lande *et al.* 2003; Saether *et al.* 2013). This approach cannot be used to draw some inferences that are possible with MCT: for instance, partitioning a species' low-density growth rate into contributions of different coexistence mechanisms, such as the storage effect or relative nonlinearity of competition (Chesson 2000; Adler *et al.* 2007). However, it has the advantage of requiring information about abundances only, not individual demographic rates, and thus can be employed in many contexts where MCT cannot.

Despite the potential utility, the variance-partitioning approach mentioned above (hereafter VPRSA, for "variance partitioning of relative species abundances") has not been employed in species-rich coral reef systems. Here, I apply the VPRSA approach to quantify the extent to which patterns in commonness and rarity are due to stochastic fluctuations in species' relative abundance, versus long-term differences in mean abundances of species (hereafter "deterministic species differences" or "niche structure") among coral reef fishes on the Great Barrier Reef. I also test how those contributions vary at the regional scale in response to environmental gradients, and to the dynamics of coral cover. I compare the broad-scale, macroecological variation in the relative importance of these dynamic diversity components with the variation in two classical, static diversity measures: species richness and evenness. My goals are (i) to determine, broadly, how deterministically-structured versus stochastically-structured reef fish assemblages are, and (ii) to estimate the magnitude of regional-scale variation in deterministic vs. stochastic components of community structure, in relation to variation in static biodiversity measures such as species richness and evenness over comparable spatial scales. This allows me to move beyond tests of neutral theory, by quantifying the relative importance of deterministic and stochastic factors that influence species' relative abundances.

In addition to the above objectives, I also aim (iii) to understand whether and how these community characteristics respond to differences in the temporal dynamics of local coral cover, as well as how they vary with latitude and with distance from the coastline along the continental shelf. By comparing the community structure of reefs across gradients in coral cover dynamics, latitude, and cross-shelf position, I aim to determine how the underlying forces that regulate coral reef biodiversity may be reshaped by the rapidly changing disturbance regime that these ecosystems are experiencing. Finally, I (iv) test whether the results from (i, ii and iii) are consistent when species-rich functional groups of fishes are analysed separately, in order to determine whether any ecologically significant macroecological variation in community structuring processes is driven by consistent relationships within functional groups, or whether they instead indicate shifts among functional groups. My expectation is that persistent, deterministic species differences will explain a smaller, but still substantial, proportion of the variation in species abundances within functional groups, compared to the fish assemblage as a whole.

2.2 Methods

2.2.1 Long-term fish community data and environmental covariates on the Great Barrier Reef

I use data from the Australian Institute of Marine Science Long-Term Monitoring Program (LTMP), which consists of extensive visual surveys of reef benthos and reef fish communities encompassing more than 20 years and 10 degrees of latitude on the Great

Barrier Reef (Sweatman *et al.* 2005). The surveys themselves are hierarchically structured, in which 3 sites are selected usually on the NE faces of 40 reefs, and 5 permanent transects are laid within each site (see Halford and Thompson 1996 for a detailed description of the survey protocol). Here I extract and compile the community dynamics data from 1994 to 2004 (11 years) for analysis, because this time window contains annual surveys of each reef (the frequency of reef surveys and study sites changed after the 2004 rezoning of the Great Barrier Reef Marine Park).

The statistical analyses focus on counts of fish identified to species and percentage cover of live coral at each survey reef. Fish species are counted visually for a prescribed list of species representing 10 families, including the parrotfishes, damselfishes, surgeon fishes, butterflyfishes, and some others. All species examined here are largely non-cryptic and easily identified underwater, and thus cryptic species groups, such as gobies, were excluded. A full list of species observed each year are included in the appendices of each LTMP status report (Sweatman et al. 2005). Corals were identified to relatively broad taxonomic and morphological categories, but I consider only total hard coral cover in the analyses in this paper. I pool fish community and coral cover data at the scale of the entire reef, summing abundances over all 15 transects laid on each reef. Percentage cover is similarly averaged across transects and sites within reefs. I adopt this approach to reduce stochastic sampling error, thereby obtaining more precise estimates of the community structure statistics that are of interest in this study.

Because the small-sized fish taxa (mainly *Pomacentridae*) were surveyed in narrower transects than other, larger fish taxa (Halford and Thompson 1996), I use subsampling to rescale the abundances of large-sized species to standardize sampling effort. Each fish counted on the wider transects is given a 20% probability of appearing in the sub-sample (because the small-fish transects cover only 20% of the area of large-fish transects). It is the sub-sampled data that are used for the analyses.

For each reef, I extract the temporal average (11-yr mean), standard deviation (SD) and coefficient of variation (CV) in coral cover as proxies for disturbance-induced coral cover variability. I also extract each reef's latitude and cross-shelf position, where latitude is measured by degrees from the equator and cross-shelf position is the standardized distance to the nearest continental shelf boundary (i.e., 0 represents the shelf boundary and 1 represents the coast, respectively). I use latitude and cross-shelf position to represent major environmental gradients because community structure on the Great Barrier Reef is known to vary strongly along both gradients, and because they serve as good proxies for other variables, such as mean and variability in temperature, and variation in terrestrial input (Harriott & Banks 2002; Mellin *et al.* 2010).

2.2.2 Partitioning variance in relative species abundance: theoretical framework

I use the partitioning approach of Engen et al. (2002) to quantify the contribution of deterministic species differences, relative to environmental and demographic stochasticity, in driving the "total variance" in relative species abundances. These variance components can be estimated by how the correlation in a community's log-abundances decays over time, i.e., the temporal autocorrelation in relative species abundance (Engen *et al.* 2002):

$\rho_t = \left(\rho_0 - \rho_\infty\right) e^{-\delta t} + \rho_\infty$ (eq. 2.1)

where ρ_t represents the correlation coefficient of log species-abundances of a community at time lag *t* (i.e., it is a measure of community similarity between species' log-abundances in two different years). This quantity is modelled as an exponential function of time lag *t*. That is, for all pairs of years on a reef, I estimate the correlation coefficient between these two samples, and then analyse how the strength of this correlation decreases as a function of the time elapsed between the two samples (here, time lag or interval between two samples ranges from 1 to 10 years). Parameter ρ_∞ represents asymptotic similarity. For a community with no persistent niche structure in abundance (e.g., where all species have the same mean abundance, and variation in species abundances is due entirely to their different responses to environmental fluctuations, alongside demographic stochasticity), ρ_∞ would be zero. For a community in which environmental fluctuations play no role, ρ_∞ would be large (and in the absence of demographic or sampling stochasticity as well, it would tend to unity). Parameter ρ_0 is the intercept (i.e., the expected correlation in species' log-abundances for a time lag of zero), and thus $1-\rho_0$ represents variation that is inversely proportional to population size (including demographic stochasticity of neutral process, and sampling effects due to variation at scales smaller than the site scale, such as aggregation). δ measures the strength of density-regulation in the system: it would be larger in assemblages that revert quickly toward their long-term mean relative abundances after a disturbance (see section 2.2.3 of this chapter, as well as Engen et al. 2002, 2011 for further explanation and derivation of eq. 2.1).

The correlation coefficients ρ_t are estimated by fitting the bivariate Poisson-lognormal distribution to all possible pairs of surveys at each site, and these correlation coefficients are then modelled as a function of the amount of time separating the samples following eq. 2.1. This distribution assumes that the two surveys represent random samples of individuals from two communities whose species abundances follow lognormal distributions with correlation coefficient ρ_t. That is, the correlation coefficient parameter estimates the true underlying correlation in species' log-abundances between the communities at the two sampling times, taking account of the fact that each abundance distribution in the data represents an incomplete sample from the community (Engen et al. 2002). This model is justified because the static species-abundance distributions of these reef fish assemblages have been shown to be well-described by the Poisson-lognormal distributions (Connolly et al. 2014). As a further check, I conducted parametric bootstrap tests (N=100 for each fitted bivariate Poisson-lognormal distribution) to verify that the bivariate Poisson-lognormal is an adequate distribution for these data (Appendix A; Figure A1).

After fitting eq. 1 to the pairwise correlation coefficients as a function of time lag according to eq. (1), variance components of relative species abundance can be estimated as:

$\pi_r = \rho_\infty$ (eq. 2.2)

$\pi_e = \rho_0 - \rho_\infty$ (eq. 2.3)

$\pi_d = 1 - \rho_0$ (eq. 2.4)

where π_r represents the proportional variation in relative abundance that is due to persistent niche differences among species, π_e represents the proportional variation in relative abundance due to species' responses to environmental fluctuations, and π_d represents the proportional (residual) variation in relative abundance due to other processes, such as neutral demographic stochasticity and sampling error.

2.2.3 Stochastic community dynamics model of reef fish communities

Originally, the partitioning approach described above was explicitly derived from a stochastic theory of community dynamics. This theory characterizes the temporal dynamics of abundance of *S* species according to the stochastic Ornstein-Uhlenbeck process (Lande *et al.* 2003):

$$dX_i = \left(r_i - \delta X_i\right)dt + \sigma_e dW_i \quad \text{(eq. 2.5)}$$

$$r_i \sim N\left(\mu_r, \sigma_r^2\right) \quad \text{(eq. 2.6)}$$

where X_i represents the abundance of species *i* on a logarithmic scale, r_i is the intrinsic population growth rate of species *i*, δ measures the strength of density-dependence, σ_e scales the magnitude of environmental fluctuations in the growth rate (i.e., larger σ_e implies larger fluctuations), and dW_i models the fluctuations themselves as a Brownian motion (Wiener) process. Eq. (2.6) specifies that intrinsic growth rates, r_i, vary among species according to a normal distribution with mean μ_r and variance σ_r^2. Eq. 2.5 can be interpreted as a continuous-time analogue of discrete-time Gompertz-type community dynamics (Ives *et al.* 2003; Dennis & Ponciano 2014).

Analysis of the model in eqns (2.5)-(2.6) shows that each species' abundance fluctuates around a species-specific equilibrium "carrying capacity", $e^{\frac{r_i}{\delta}}$, and both the carrying capacities, and the abundances themselves, follow lognormal distributions among species. Notably, the stationary distribution of species' abundances in the community remains lognormal, even in the presence of some violations of the model's simplifying assumptions, such as the incorporation of correlations in species' responses to environmental fluctuations, and of inter-specific interactions (Sæther *et al.* 2013). The model is therefore consistent with the reef fish data in this study whose species-abundance distributions are well-described as discrete, random samples of individuals from lognormal abundance distributions (i.e., Poisson-lognormal distributions: Connolly et al. 2014). In addition, previous work suggests that the Gompertz form of density-dependence is appropriate for these data (Thibaut et al. 2012). For this model, the total variance in relative species abundance (hereafter σ_{total}^2) can be analytically partitioned into additive components as follows:

$$\sigma_{total}^2 = \frac{\sigma_r^2}{\delta^2} + \frac{\sigma_e^2}{2\delta} + \theta^2 = V_r + V_e + V_d \quad \text{(eq. 2.7)}$$

and thus the proportional variance components π_r, π_e, and π_d would be equal to V_r / σ^2_{total}, V_e / σ^2_{total} and V_d / σ^2_{total}, respectively.

I acknowledge that some of the assumptions of the stochastic dynamics model in eq. (2.5) are somewhat restrictive. Most notably, it assumes that the strength of density regulation and the magnitude of environmentally-induced fluctuations (on a logarithmic scale, i.e., proportional variance in population growth rate due to environmental fluctuations) are the same for all species, and it also assumes that there are no species interactions, nor covariation in species' responses to environmental fluctuations. All of these assumptions are likely to be violated to some degree in my study system. For a community undergoing Gompertz-type dynamics, the lognormal stationary distribution of abundances is quite robust to these simplifying assumptions (Engen and Lande 1996; Ives et al. 2003). However, the robustness of variance components estimates has not yet been investigated, so I address this issue in considerable detail in Chapter 3.

2.2.4 Estimating species richness and unevenness of reef fish community

I also use the Poisson-lognormal distribution to estimate (time-averaged) richness and unevenness of reef fishes as follows. I fit this distribution to each of the 440 species-abundance distributions (40 reefs × 11 years), using the method of maximum likelihood (Pielou 1975; Engen 1978; Connolly et al. 2005; Saether et al. 2013). Fitting this distribution yields maximum likelihood estimates of the standard deviation of log-abundance $\hat{\sigma}$ for the underlying community from which the data are a sample. The skewness of a lognormal distribution is a function of only this parameter, implying that the skewness (i.e., the unevenness) of abundances in the fish community from which the data are a sample is a monotonically increasing function of $\hat{\sigma}$ (i.e., larger values indicate more uneven communities) (Saether et al. 2013). This can be seen as a generalization of the evenness metric based on the variance in log abundances among species (Smith and Wilson 1996), which makes the evenness metric more robust to incomplete sampling (i.e., the $\hat{\sigma}$ parameter is independent of sampling effort: Connolly et al. 2009; Saether et al. 2013). Additionally, Poisson-lognormal fits can be used to estimate the total number of species in the community, by producing an estimate of the probability that a species is present in the community but not observed in the sample, $\hat{p}_0$. An estimate of total community richness is simply the number of observed species divided by $1-\hat{p}_0$ (Connolly et al. 2005; Saether et al. 2013).

I estimate the variance of the species abundance distribution for each year, and then I calculate the mean of these values (σ^2_{total}) across years for each reef to use as my reef-scale measure of unevenness. Similarly, I use the mean (across years) of the estimated total community richness for each reef as my reef-scale estimate of species richness.

2.2.5 Parameter Estimation of variance partitioning of relative species abundance

To estimate the three variance components that completely partition the variation in relative abundances of fishes over time for each reef, I fit a mixed-effects version of eq. (2.1) to the data set. This approach outperforms fitting each reef individually, for simulated data with time series lengths, sample sizes, species richness levels and σ^2_{total} values comparable to those of these data (see Chapter 3). Specifically, I fit a family of nonlinear mixed-effects models, in which each of the fitted parameters (ρ_∞, ρ_0, δ) may be fixed constants for all reefs or they may randomly vary among reefs according to latent normal distributions (Appendix B; Table B1). These models are parameterized by Template Model Builder (TMB) in R (Kristensen et al. 2016). I fit models with different combinations of the three temporal autocorrelation function parameters as fixed versus random, and I ranked model fits by AIC. I then checked for numerical stability of the model fits (i.e., I confirmed that the model's random effects parameter estimates were valid), and I chose the best-fitting model (by AIC) that yielded a numerically stable fit as our basis for inference. Specifically, I used fixed and reef-level random effects estimates to calculate variance components for each reef according to eqs 2.2-2.4.

2.2.6 Explaining regional variation in community structure

I use ordinary least-squares (OLS) regression to investigate the extent to which variation among reefs in coral dynamics (mean, standard deviation [SD] and coefficient of variation [CV] of coral cover at each reef), latitude, and cross-shelf position explain variation in fish community structure (variance components of community structure, species richness, and evenness). I choose latitude and cross-shelf position as my abiotic explanatory variables because many factors likely to affect community structure on coral reefs, such as temperature, turbidity, and nutrient loading vary strongly with one or both of these spatial variables. SD and CV of coral cover were both right-skewed, so they were log-transformed to reduce the heterogeneity of leverage values associated with the different reefs in our analysis (model selection yielded the same best-fitting models when they were untransformed, and r^2 values are higher, making the results below conservative with respect to the conclusions drawn). log(SD) and log(CV) of coral cover were very strongly collinear ($r \sim 0.9$), and model selection always preferred models using SD rather than CV when analogous models were compared (e.g., AIC favours a multiple regression model with interactive effects of log(SD) and mean coral cover over a model with interactive effects of log(CV) and mean coral cover), so I have not presented results for the models using log(CV). None of the remaining explanatory variables were highly correlated with one another (Appendix A; Table A1). I used AIC and adjusted R^2 for model comparisons.

Lastly, because the reef-level proportional variance components are estimated from fits of another statistical model, rather than being directly observed, I also performed a sensitivity analysis to test the robustness of my results to parameter uncertainty. Specifically, I estimated the uncertainty in my estimates of proportional variances by parametric bootstrap, resampling random effects values for my 40 reefs 10000 times from the variance-covariance matrix of the fitted nonlinear mixed-effects model based on TMB. Then, each bootstrap set of reef-level proportional variance component values were re-analysed using the OLS regressions repeated on each of the bootstrapped data sets. I evaluated the robustness of my model selection procedure by quantifying the percentage of bootstrapped data sets for which each model was selected as the best model by AIC.

2.2.7 Sensitivity analysis of community structure across trophic/functional groups

Reef fish community data were separated into major trophic/functional groups: herbivores, planktivores and benthic invertivores. These trophic groups were classified according to previously published functional classifications of reef fishes, supplemented by communication with experts in the field (see Appendix A; Table A3 for details). Of these functional groups, only three (i.e., herbivores, planktivores, and benthic invertivores) were sufficiently species-rich to analyse separately. I repeated all of the analyses conducted on the overall dataset, as described above, for each of the groups. I then compared the distribution of reef-scale estimates of variance components for the three trophic groups with each other, and with those for the mixed (i.e., the whole data, regardless of functional groups) assemblages, to estimate the magnitude of the difference between variance components. These comparisons were performed by paired t-test with Bonferroni correction for multiple comparisons.

To examine the potential influence of coral cover dynamics on variance components of relative species abundance within functional groups, I fitted relationships between variance components and coral cover dynamics using OLS regression. In the full model, temporal mean and standard deviation (SD) in coral cover, and trophic group were considered as explanatory variables. AIC was used for model selection, beginning with a model including all main effects and interactions.

2.3 Results

The model of temporal autocorrelation of community similarity in which all three parameters (ρ_∞, ρ_0, δ) had both fixed and random components was numerically unstable: that is, the estimates from methods of Laplace approximation and importance sampling did not converge (Appendix B; Figure B1). Of the remaining models, the best-fitting model by AIC included random effects on ρ_∞ and ρ_0, but not δ (Appendix B; Table B1). This model exhibited good statistical behaviour (Appendix B; Figure B2), so I used it to calculate variance components for all analyses below.

Overall, variation in the commonness and rarity of reef fishes is highly deterministically structured on the Great Barrier Reef (Figure 2.1). Deterministic (intrinsic) species differences explain a larger proportion of variation in reef fish community structure ($\pi_r = 0.76$, on average [parametric bootstrap 95% CI: 0.73-0.8]), compared to environmental stochasticity ($\pi_e = 0.17$ on average [parametric bootstrap 95% CI: 0.14-0.2]). Only ~0.07 [parametric bootstrap 95% CI: 0.06-0.07] of the variance was attributable to other sources of variance, such as demographic stochasticity and sampling error (Figure 2.1E).

Nevertheless, there is substantial variation in these variance components of community structure among coral reefs: the relative importance of deterministic species differences in driving reef fish abundances varies strongly with coral cover dynamics (Figure 2.2). Specifically, the importance of the environmental stochasticity variance component was higher on reefs with more variable coral cover, especially when mean coral cover was high, while the species difference variance component changed in a compensatory direction (Figure 2.2). The interactive effects of the mean and standard deviation of coral cover explain about 40% of the variation in these two relative variance components of fish assemblage structure (R^2=0.39, P<0.001 for the variance component of species differences, and R^2=0.4, P<0.001 for the variance component of environmental stochasticity; Table 2.1). Model selection favours this interactive model over all alternatives considered (Table 2.1; also P=0.001 for both models of species difference and environmental stochasticity being selected in 1000 times bootstrap tests).

In marked contrast to the dynamic quantities represented by the proportional variance components, coral cover variables explain very little variation in time-averaged richness and evenness of the fish assemblages (R^2~0 in all cases; Table 2.1). Instead, these quantities vary strongly and interactively with latitude and cross-shelf position, which together explain 55% and 71% of the variation in richness and evenness, respectively. Specifically, species richness increases and unevenness decreases (i.e., evenness increases) towards the equator, but the increases in richness and decreases in unevenness are much steeper on the inner shelf than the outer shelf of the Great Barrier Reef (Figure 2.3). Model selection strongly favours this interactive model over the alternatives considered, for both richness and evenness (Table 2.1).

The among-reef patterns of fish community structure are similar across trophic/functional groups that contain more ecologically-similar species, suggesting these patterns are not produced by shifts in the relative dominance of different trophic/functional groups, but rather represent parallel responses across functional groups. In line with the patterns of mixed (the whole data, regardless of functional groups) assemblages described above, deterministic species differences remain the most important factor driving patterns of relative species abundance in trophic/functional groups: surprisingly, it explains ~65-80% (i.e., approximately the same magnitude) of the variance in relative species abundance of trophic groups, comparable to that of the mixed assemblage (Appendix C; Figure C1). By contrast, environmental stochasticity explains only ~5-25% of variance in relative species abundance of trophic groups (Appendix C; Figure C1), again similar to the original analysis. Demographic and sampling stochasticity together explain much less species-abundance variations (~2-12%) of trophic groups (Appendix C; Figure C1). The above variance components are statistically significantly different from each other for some combinations of between-group difference, but the magnitude of this between-group variation is relatively small (Appendix C; Table C1-Table C3). As in mixed assemblages, coral cover dynamics, but not latitude or cross-shelf position, explain the relative importance of deterministic versus stochastic variance components across trophic/functional groups (Appendix C; Figure C2, Table C4). Latitude and cross-shelf position explain the richness pattern of trophic groups among reefs, as in mixed assemblages, but only latitude explain the evenness pattern of trophic groups among reefs (Appendix C; Figure C3-C5, Table C4).

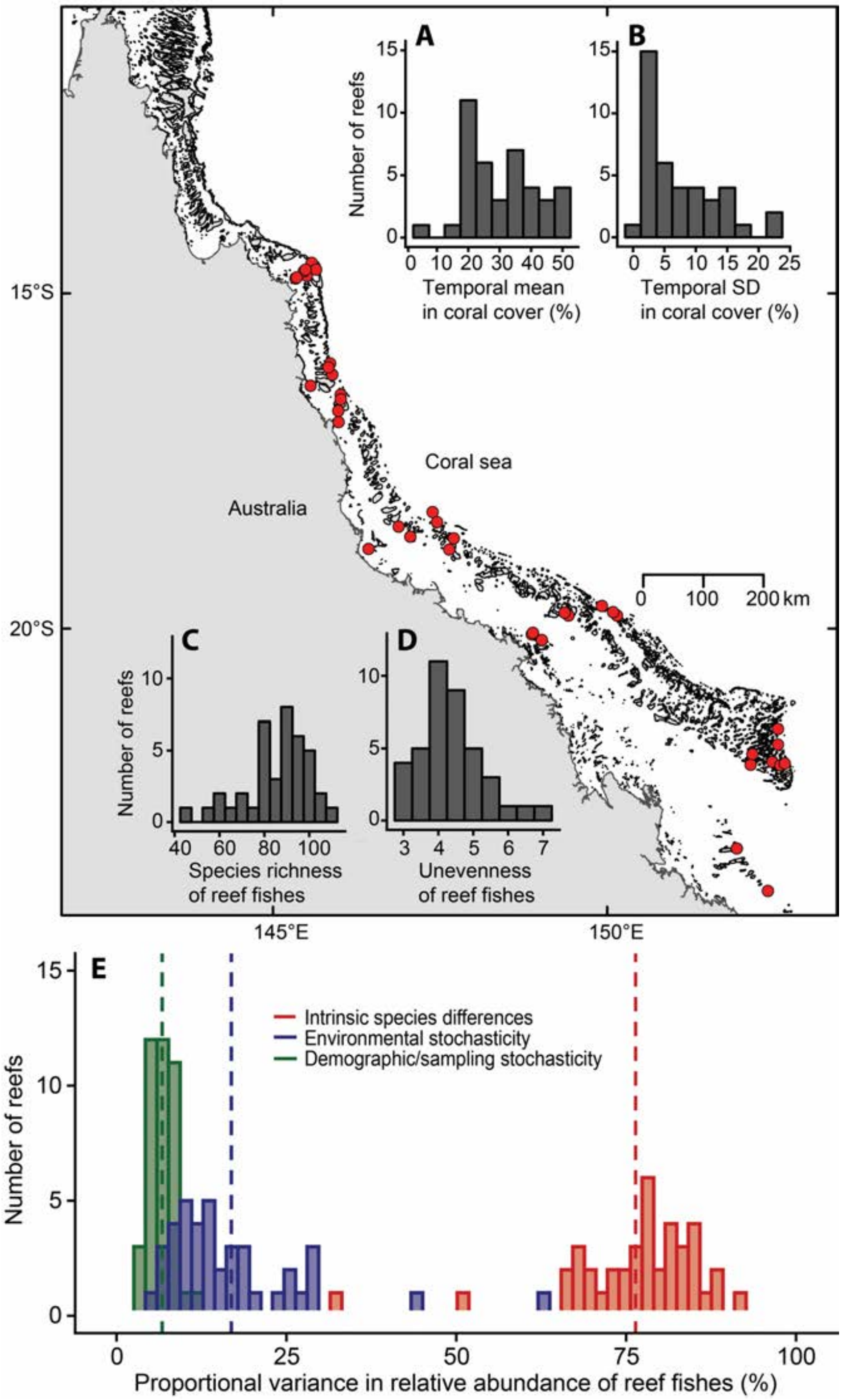

Figure 2.1 Map showing reefs included in all analyses, along with frequency distributions of explanatory and response variables. On the map, red circles show locations of the 40 reefs used in this study. (**A-B**) Frequency distribution of the temporal mean and standard deviation in coral cover that are proxies for environmental volatility across study sites. (**C-D**) Frequency distribution of (time-averaged) species richness and unevenness of reef fish communities across study sites. (**E**) Frequency distribution of the proportional variance in relative species abundances attributable to deterministic intrinsic

species differences that produce differences in long-term mean abundances (red bars; n=40), to environmentally-induced stochastic fluctuations in species' growth rates (blue bars; n=40), and to residual effects such as demographic stochasticity and sampling effects such as local aggregation (green bars; n=40). Color-coded, dashed lines indicate the mean value of the corresponding variance component.

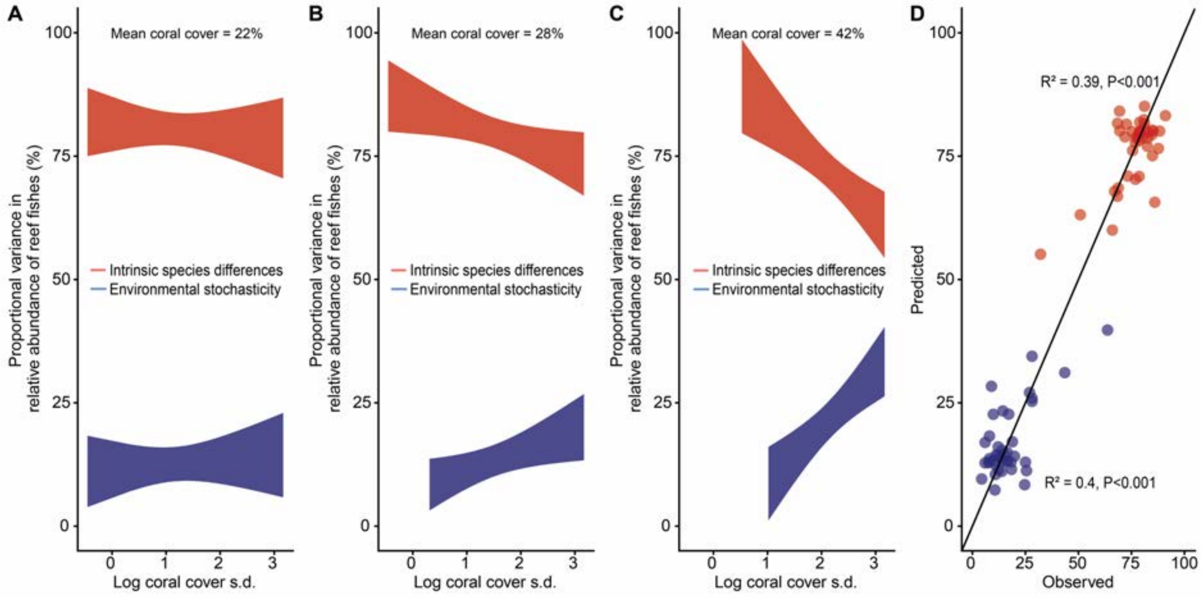

Figure 2.2 Relationship between coral cover dynamics and variance components of reef fish community structure. (**A-C**) Relationships between the reef-scale coral cover variables (temporal SD and mean of coral cover for each reef) and the relative importance of variance components in structuring fish species-abundances across reefs (n=40 reefs). The relationships are plotted using parameter estimates for the lowest-AIC models, with interactive effects of the mean and SD of coral cover as explanatory variables, and variance components of fish community structure as response variables (Table 1). The red bands represent the 95% C.I. of the proportional variance attributable to persistent species or niche differences, while the blue band represents the 95% C.I. of the proportional variance attributable to environmental stochasticity. To illustrate the interactive relationships, the 1st, median and 3rd quartiles of mean coral cover are fixed in panels (A), (B), and (C), respectively, and the fitted relationship between SD of coral cover and variance component values are plotted for the corresponding value of mean coral cover. (**D**) The relationship between observed and predicted values from the corresponding OLS regression models whose fits are plotted in panels (A-C). The solid line is the unity line (observed=predicted).

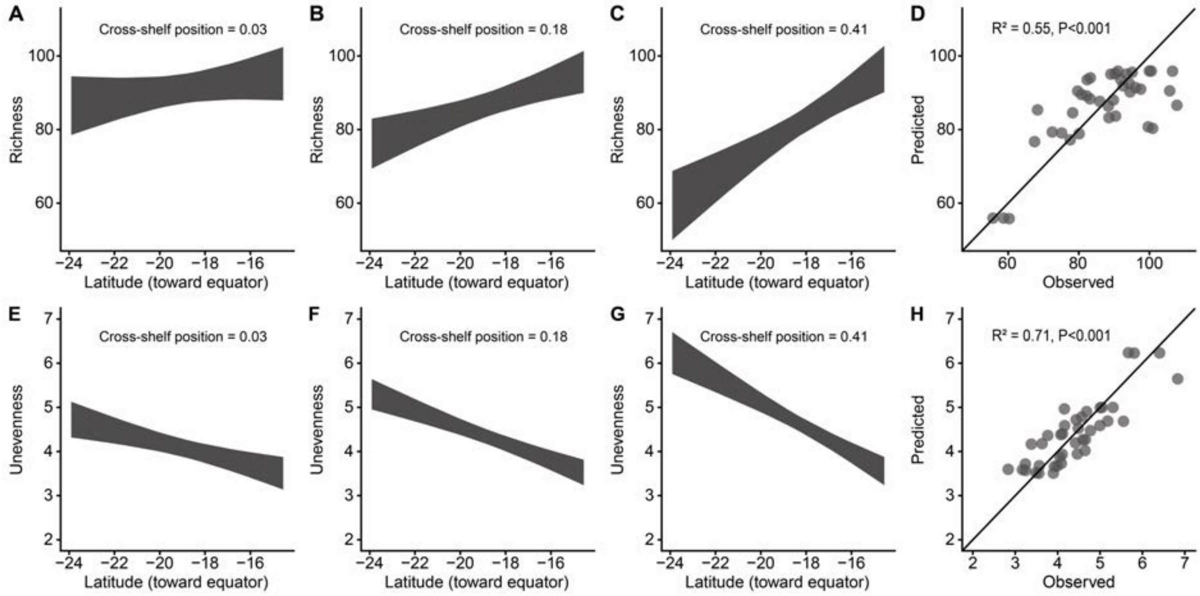

Figure 2.3 Richness and unevenness depend on latitude and cross-shelf position. (**A-C**) Relationship between time-averaged fish species richness and the interaction of latitude with cross-shelf position. (**E-G**) Relationship between time-averaged unevenness and the interaction of latitude with cross-shelf position. Grey bands are the 95% confidence intervals predicted from the lowest-AIC models for richness and unevenness (n=40 reefs; Table 1). To better illustrate the interactive relationships, the 1st, median, and 3rd quartiles of cross-shelf positions are fixed in panels (A, E), (B, F), and (C, G), respectively, and the relationship between richness or unevenness as a function of latitude are shown for the corresponding value of cross-shelf position. Note that variable cross-shelf position increases towards the coast. **(D, H)** The relationship between observed and predicted values from the OLS regression models corresponding to the fitted relationships in (A-C) and E-G), respectively. The solid line is the unity line (observed=predicted).

Table 2.1 OLS regression models for proportional variance components, richness, and unevenness. For explanatory variables, ‘mC’ and ‘log(sdC)’ represent the long-term (11-yr) mean and log-transformed standard deviation in annual coral cover fluctuations, respectively. ‘Lat’ and ‘Shelf’ represents latitude and cross-shelf position, respectively. ‘Intercept’ represents the regression model that contains only an intercept (i.e., no explanatory variables). Cross symbols indicate models that include main effects and interactions, whereas plus symbols denote models including only main effects (i.e., additive effects of the explanatory variables).

Response variable	Explanatory variable	adjR2	AIC	ΔAIC
Species differences	log(sdC)×mC	0.39	-80.93	0
Species differences	log(sdC)	0.19	-71.64	9.29
Species differences	mC	0.16	-70.28	10.65
Species differences	log(sdC)×mC+Lat×Shelf	0.34	-75.66	5.27
Species differences	log(sdC)+Lat×Shelf	0.19	-68.83	12.1
Species differences	mC+Lat×Shelf	0.12	-65.59	15.34
Species differences	Lat×Shelf	0	-59.87	21.06
Species differences	Lat	0	-62.24	18.69
Species differences	Shelf	0	-63.78	17.15
Species differences	Intercept	0	-64.2	16.73
Environmental stochasticity	log(sdC)×mC	0.4	-77.47	0
Environmental stochasticity	log(sdC)	0.19	-67.43	10.04
Environmental stochasticity	mC	0.16	-65.91	11.56
Environmental stochasticity	log(sdC)×mC+Lat×Shelf	0.35	-71.96	5.51
Environmental stochasticity	log(sdC)+Lat×Shelf	0.18	-64.39	13.08
Environmental stochasticity	mC+Lat×Shelf	0.12	-61.05	16.42
Environmental stochasticity	Lat×Shelf	0	-55.55	21.92
Environmental stochasticity	Lat	0	-57.97	19.5
Environmental stochasticity	Shelf	0	-59.4	18.07
Environmental stochasticity	Intercept	0	-59.85	17.62
Richness	log(sdC)×mC	0	332.25	32.91
Richness	log(sdC)	0	330.47	31.13
Richness	mC	0.01	329.39	30.05
Richness	log(sdC)×mC+Lat×Shelf	0.54	303.61	4.27
Richness	log(sdC)+Lat×Shelf	0.54	301.18	1.84
Richness	mC+Lat×Shelf	0.55	299.83	0.49
Richness	Lat×Shelf	0.55	299.34	0
Richness	Lat	0.1	325.51	26.17
Richness	Shelf	0.24	318.69	19.35
Richness	Intercept	0	328.96	29.62
Unevenness	log(sdC)×mC	0.02	107.83	47.38
Unevenness	log(sdC)	0.03	105.65	45.2
Unevenness	mC	0.02	106.16	45.71
Unevenness	log(sdC)×mC+Lat×Shelf	0.69	63.80	3.35
Unevenness	log(sdC)+Lat×Shelf	0.7	60.58	0.13
Unevenness	mC+Lat×Shelf	0.7	60.45	0
Unevenness	Lat×Shelf	0.71	60.78	0.33

Unevenness	Lat	0.27	94.14	33.69
Unevenness	Shelf	0.22	96.93	36.48
Unevenness	Intercept	0	106.03	45.58

2.4 Discussion

Across the Great Barrier Reef, patterns of commonness and rarity in reef fish assemblages are, in general, highly niche-structured, with approximately 76% of the variance in relative species abundances attributable to persistent, deterministic species differences (i.e., to differences among species in long-term mean abundance). However, this proportion is strongly dependent on coral cover dynamics, decreasing as volatility in coral cover increases. That relationship is most pronounced when average coral cover is high, whereas, when coral cover is low, reef fish communities may shift to an alternative pattern of relative abundances that is more resistant to environmental fluctuations (e.g., due to changed functional composition). Such resistant fish communities may have fewer species that depend strongly on living corals (e.g., coral feeders; Coker *et al.* 2014), or fewer species that are influenced by lagged effect of coral cover declines on recruitments (a previous study showed about 60% of fish species were affected in this way; Jones *et al.* 2004). By contrast, the reef fish diversity measures of species richness and unevenness are comparatively insensitive to coral cover dynamics. Instead, they respond to the more classical macroecological variables of latitude and distance from the coast along the continental shelf. Broadly, these relationships remain when trophic/functional groups are considered separately: all indicate patterns of commonness and rarity that are driven mainly by persistent species differences, and variance components of relative species abundance respond to coral cover, latitude, and cross-shelf position in similar ways.

Despite the well-documented ecological importance of episodic disturbances on coral reefs, my findings suggest that deterministic niche structure plays a major role in allowing species to be consistently common or rare through time. This has important implications for the debate about the relative importance of niche vs. neutral processes in structuring high-diversity assemblages (Hubbell 1998; Dornelas et al. 2006; Volkov et al. 2007; McGill 2010; Chisholm and Pacala 2010; Connolly et al. 2014; Mathew and Whitaker 2014; Kalyuzhny et al. 2015; Fung et al. 2016; Engen et al. 2017). For marine systems in general and coral reefs in particular, the niche-neutrality debate has been largely informed by the analysis of static rather than dynamic species-abundance distributions (Dornelas et al. 2006; Volkov et al. 2007; Connolly et al. 2014). Here, by analysing temporal dynamics of species-abundance distributions, I am able to evaluate the importance of deterministic species differences, relative to stochastic fluctuations in population growth rates, in driving patterns of commonness and rarity on coral reefs. Indeed, in the context of the relative importance of niche structure, the proportion of variance in species abundance that is attributable to what I term “deterministic (intrinsic) species differences” is an estimate of the extent to which the variance in species abundances is driven by differences in the long-term equilibrium mean abundances of different species. Differences in species’ niche or trait characteristics may also lead them to respond differently to environmental fluctuations (e.g., quasi-neutral dynamics and MCT; Chesson 2000; Lande et al. 2003; Miller et al. 2011; Thibaut et al. 2012), and such effects might be captured by the variance component of environmental stochasticity in our analysis (Engen et al. 2002). Thus, the term for proportional variance of species differences is likely to be a conservative estimate of the contribution that differences in species characteristics make to their relative abundance.

This chapter is also the first regional-scale study, from any system, that evaluates whether and how the relative importance of variance components in relative species abundance changes regionally in response to biotic and abiotic environmental variables. The evaluation suggests that much of this geographical variation is potentially explainable: for these data, nearly half the variation in deterministic niche structure of fish communities over 10 degrees of latitude can be explained by just two explanatory variables linked to coral cover dynamics (Table 2.1). Although stochastic fluctuations in population growth make a comparatively small contribution to overall variation in relative abundances of reef fishes, this contribution varies systematically across the Great Barrier Reef in a manner that is also closely linked to coral cover dynamics (Figure 2.2). Importantly, increased volatility in coral cover is associated with increased importance of stochastic fluctuations in population growth rates in structuring reef fish abundance patterns. In other words, reefs with high coral cover volatility also have increased fish community turnover and thus temporal beta-diversity. The deterministic component of species relative abundances is reduced, particularly when coral cover is high (Figure 2.2C). Conversely, coral cover has a much weaker, statistically non-significant relationship with alpha-diversity measures such as species richness and unevenness (Table 2.1). This suggests that the effects of volatility in coral cover on fish assemblages may be masked by apparent stability in the kinds of macroecological variables typically used in regional-scale analyses, which are overwhelmingly based on static snapshots rather than time series. This finding underscores the importance of utilizing dynamic aspects of community structure for anticipating how those communities may respond to changed disturbance regime and volatility in environmental fluctuations (Mathews and Whittaker 2015; Helmut et al. 2017; Engen et al. 2017).

Surprisingly, I found that latitude and cross-shelf position, which can serve as proxies for a range of regional-scale environmental gradients (such as mean and variability in annual temperature and solar radiation, as well as turbidity, nutrient enrichment, and concentrations of environmental toxins) (Wolanski & Bennett 1983; Bellwood & Hughes 2001; Fabricius *et al.* 2008; Malcolm *et al.* 2010b; Schiller *et al.* 2015) explained almost none of the regional-scale variation in the relative importance of deterministic species differences. This suggests that longer-term macroevolutionary, microevolutionary, and community assembly processes affect patterns of commonness and rarity in a consistent fashion. This is true for high-diversity tropical reefs as well as lower-diversity subtropical reefs, and on high-turbidity, nutrient-enriched inshore reefs as well as offshore reefs that are much less influenced by coastal processes. In other words, species-rich reef fish communities along the Great Barrier Reef are structured in similarly predictable ways, despite the well-established, substantial differences in species composition that occur along these north-south and cross-shelf gradients (Hoey & Bellwood 2008; Emslie *et al.* 2010, 2012; Cheal *et al.* 2012; Hoey *et al.* 2013; Goatley *et al.* 2016).

In marked contrast, the two alpha-diversity measures explored here, i.e., richness and unevenness, were relatively insensitive to coral cover dynamics (Table 2.1), but they varied strongly with latitude and cross-shelf position (Figure 2.3). The relatively poor relationships between alpha-diversity and coral cover dynamics found here is consistent with some previous studies of reef fish communities, which have also found that diversity does not always respond to external disturbances (Wilson et al. 2009; Bellwood et al. 2012). Patterns of species richness have long been known to covary with both latitude and cross-shelf position (Hoey & Bellwood 2008; Malcolm *et al.* 2010a; Bennett 2011; Bode *et al.* 2011; Cheal *et al.* 2012; Mellin *et al.* 2019), variables that are strongly associated with important environmental gradients on the Great Barrier Reef. The stronger latitudinal gradient for inshore locations may partially reflect the latitudinal distribution of coastal impacts (Devlin

& Brodie 2005; Fabricius *et al.* 2016), which tend to be greater in the southern half of the Great Barrier Reef than the northern half, as well as the fact that the continental shelf is wider further from the equator, potentially reducing dispersal between nearshore and offshore reefs (Bode *et al.* 2011). However, to my knowledge, this chapter is the first study to document latitudinal and onshore-offshore gradients in evenness, which are actually stronger than those for richness (R^2=0.71 versus R^2=0.55) on coral reefs. This is particularly striking since the measure of unevenness used here is the (averaged) total variance in species' relative abundances, which is what is partitioned by the variance components (eq. 7). That is, total variability in species abundances responds to macroecological-scale environmental gradients, while the proportion of that variability attributable to persistent species differences versus stochasticity responds instead to coral cover dynamics.

This chapter leverages a unique, regional-scale, highly spatially-replicated time series of relative abundances of fishes at species level on the Great Barrier Reef to reveal that patterns of commonness and rarity are structured disproportionately by deterministic factors. This implies that it is characteristics of species that determine whether they tend to be common or rare in a particular environment. Moreover, this dominance of deterministic structure holds from the high-diversity north to the lower-diversity south, and from turbid, nutrient-enriched inshore reefs to reefs close to the outer shelf, of the Great Barrier Reef, even though which species have high versus low abundances differs profoundly among those environments (Williams 1982; Russ 1984b, a; Hoey & Bellwood 2008; Emslie *et al.* 2010, 2012). However, the deterministic structure is eroded on reefs with large fluctuations in coral cover (Figure 2.2). In this system, regional variation in the volatility of coral cover over this period was driven substantially by episodic disturbances, particularly cyclones and crown-of-thorns starfish outbreaks (De'ath *et al.* 2012), whose frequency or severity may be linked to anthropogenic environmental change (Wooldridge & Brodie 2015; Cheal *et al.* 2017).

Moreover, as climate change accelerates, coral bleaching may overtake other disturbances as a key driver of increased coral cover volatility (Hughes *et al.* 2018a, 2019; Wolff *et al.* 2018) by causing mass mortality and potentially driving a shift to coral species with more "weedy" (i.e., fast growth and high fecundity) demographic characteristics (Pandolfi *et al.* 2011; Hughes *et al.* 2018b). Consequently, accelerating anthropogenic impacts could erode niche structure of reef assemblages in coming decades, most dramatically on the reefs with the highest levels of average coral cover. Moreover, the marked differences in how the deterministic and stochastic factors of community structure respond to environmental gradients and coral cover dynamics imply that setting conservation priorities based principally on classical measures of alpha-diversity or species composition, without explicitly considering the drivers of species' relative abundance dynamics, might fail to detect important changes in the processes structuring ecological assemblages (see examples in Mathews and Whittaker 2015; Helmut et al. 2017). Given that biodiversity and community structure are maintained by a combination of deterministic and stochastic factors, dynamic partitioning of relative species abundance as demonstrated here may help to provide an early warning of longer-term threats to biodiversity (e.g., decreased deterministic niche structure) that are not apparent from richness or evenness alone.

Chapter 3: Robustness test for variance partitioning of relative species abundance from simulated community time-series data

3.1 Introduction

In Chapter 2, I found that reef fish assemblages on the Great Barrier Reef are, in general, highly niche structured, with over 75% of the variance in relative species abundances, on average, attributable to deterministic species difference (i.e., among-species differences in long-term mean abundance), and less than 25% attributable to environmental, demographic, or sampling stochasticity. Moreover, I found that the relative importance of these variance components differs among reefs, depending on the local dynamics of coral cover. These inferences were all based on applying the variance partitioning of relative species abundance (VPRSA) method, which is derived from Engen and Lande's stochastic-dynamic theory of biodiversity (Engen & Lade 1996; Engen *et al.* 2002; Lande *et al.* 2003; also see General Introduction). However, the VPRSA approach makes some important simplifying assumptions that may commonly be violated in real communities, including the LTMP data (Chapter 2), so it is important to evaluate whether inferences about proportional variance components are sensitive to those assumptions.

The assumptions of VPRSA flow from the fact that the approach is derived from a model in which the temporal abundance dynamics of *S* species follows the Ornstein-Uhlenbeck process (Engen *et al.* 2002; Dennis & Ponciano 2014):

$dX_i = \left(r_i - \delta X_i\right)dt + \sigma_e dW_i$ (eq. 3.1)

$r_i \sim N\left(\mu_r, \sigma_r^2\right)$ (eq. 3.2)

where the subscript *i* indexes the species, and X_i represents the log abundance of species *i*, such that log-abundances of species fluctuate around their deterministic equilibrium values, which are determined by species-specific intrinsic growth rate (r_i) and a common density dependent parameter (δ). Intrinsic growth rates are perturbed by environmental stochasticity, which is modeled as Wiener (Brownian) process noise (W_i) with variance σ_e^2. Intrinsic growth rate r_i varies among species according to a normal distribution with mean μ_r and variance σ_r^2.

Simplifying assumptions of this model include that: (i) species' intrinsic growth rates vary according to a normal distribution, (ii) the strength of intra-specific density dependence is the same for all species (note that assumptions i and ii together, given the log-scaling of abundance in eq. 3.1, imply a lognormal distribution of equilibrium population sizes), and (iii) the responses of species' intrinsic growth rates to environmental stochasticity (environmental variance) are assumed to be independent and equal in magnitude (i.e., they fluctuate from year to year with the same variance; because of the log-scaling of abundance in this model, this implies that fluctuations in population growth have a variance that is proportional to the mean). Samples from communities whose abundances follow these assumptions generate a static species-abundance pattern that follows a Poisson-lognormal distribution, consistent with what is commonly observed in data (McGill *et al.* 2007; Saether *et al.* 2013). Importantly, the assumptions also allow the variance in relative species log-abundances to be partitioned into components explained by species differences in intrinsic rate of increase, species' responses to environmental stochasticity, and variance due to other factors, such as demographic and sampling stochasticity, when time series of species-abundance distributions are available.

The partitioning, which I term VPRSA, can be expressed as follows:

$$\sigma_{total}^2 = \frac{\sigma_r^2}{\delta^2} + \frac{\sigma_e^2}{2\delta} + \theta^2 = V_r + V_e + V_d . \quad \text{(eq. 3.3)}$$

where σ_{total}^2 represents the overall variance in log-abundance values among species and the three terms in the middle of the equation follow the notation of eqs. (3.1-3.2) and Engen et al. (2002). The first of these terms (denoted V_r) on the right-hand side of eq. (3.3) is the variance in species' equilibrium abundances (i.e., the variance in their equilibrium carrying capacities, measured on a log scale). I term this the "deterministic species differences" or "intrinsic species differences" variance component. It can be conceived of as the variance due to niche or trait differences among species (Engen *et al.* 2002; Engen *et al.* 2011). The second term (V_e) is the variance in species' relative abundances due to their responses to environmental fluctuations, and I term this the "environmental stochasticity" variance component. The last term (V_d) represents the remaining variance in log-abundance that is not explained by the other two terms, and incorporates the effects of sampling stochasticity, as well as small-scale spatial variation and demographic stochasticity. I term this "demographic/sampling stochasticity". These variance terms can be used to derive proportional variance estimates by dividing the variance terms of V_r, V_e and V_d by the overall variance σ^2_{total} in eq. 3.3. By doing so, as per Chapter 2, proportional variances from VPRSA provide an estimate of the relative importance of deterministic versus stochastic factors in shaping patterns of commonness and rarity (Engen *et al.* 2002; Lande *et al.* 2003; Engen *et al.* 2011).

The community dynamics model of eqs. (3.1)-(3.2) implies a particular analytical form of temporal autocorrelation in log-abundance of species. That is, the correlation coefficient of the bivariate Poisson-lognormal distribution of abundances decays exponentially with the amount of time separating the two species-abundance samples (Engen *et al.* 2002):

$$\rho_\tau = (\rho_0 - \rho_\infty)e^{-\delta\tau} + \rho_\infty \quad \text{(eq. 3.4)}$$

where ρ_τ is the correlation between species' log-abundance values separated by time lag τ , and the parameters ρ_0 and ρ_∞ are linked to the variance components in eq. (3) as follows:

$$V_r = \rho_\infty \cdot \sigma^2_{total} \qquad \text{(eq. 3.5)}$$

$$V_e = (\rho_0 - \rho_\infty)\cdot \sigma^2_{total} \quad \text{(eq. 3.6)}$$

$$V_d = (1-\rho_0)\cdot \sigma^2_{total} \quad \text{(eq. 3.7)}$$

In addition to the statistical challenge of obtaining well-behaved estimates of the autocorrelation parameters, it is not clear whether eqs. (3.4)-(3.7) constitute reasonable estimates of variance components when the assumptions of the underlying stochastic community dynamics model (eqs. 3.1-3.3) are violated. When some assumptions are violated, the interpretation of these variance components in terms of particular parameters of the original stochastic-dynamic theory may not hold. For instance, the deterministic "intrinsic species difference" component captures the proportional variance in log-abundance due to species differences in equilibrium abundance, but this will no longer be directly proportional to variance in intrinsic growth rates (as in the V_r term of eq. 3.3) when species interactions or among-species heterogeneity in density-dependence is present.

Some of the theory's underlying assumptions are more reasonable than others. For instance, the assumption of Gompertz-type density dependence is consistent with many previous studies, which have found that this model characterizes the functional form of density dependence well and performs better than, or as well as, other forms such as the

Ricker or Logistic forms (MacNally et al. 2010; Knape and de Valpine 2011; Thibaut et al. 2019). More specifically, in a closely-related previous study focusing on reef fish functional group dynamics on the Great Barrier Reef (Thibaut et al. 2012), the Gompertz-type density dependence was found to fit data better than other forms of density dependence. Another assumption is that of a normal distribution of intrinsic growth rates. Because of the log-scaling of species abundances in eq. (3.1), this implies a lognormal distribution of geometric growth factors. A strongly right-skewed distribution of this quantity, such as a lognormal, is consistent with the few studies of variation in population growth at the assemblage-level, which show that most species are relatively slow growing, with a long tail of few fast-growing species. (e.g., the freshwater fish: Wang et al. 2019; fishery species: Myers et al. 1999).

In contrast, the assumptions of equal strength of density dependence, and equal proportional magnitude of environmentally-induced fluctuations in abundance, seem unlikely to hold in nature (Thibaut et al. 2012; Mori et al. 2013; Bonin et al. 2015). Between these extremes, the assumptions that interspecific interactions are negligible, and that species respond independently to environmental fluctuations, are common in biodiversity models, but controversial. For instance, there is some evidence that between-species interactions tend to be weak, particularly for high-diversity systems (e.g., Freckleton & Lewis 2006; Thibaut et al. 2012; Gellner and McCann 2016), and species' responses to fluctuations tend to be relatively independent on average (e.g., Lande et al. 2003; Loreau 2010). The lognormal shape of the static species abundance distribution has been shown to be robust to violation of these assumptions (Engen & Lande 1996; Saether et al. 2013). However, whether variance components estimated from the temporal evolution of such species-abundance distributions are equally robust is unknown.

VPRSA is a potentially powerful way to draw inferences about the importance of deterministic versus stochastic drivers of species abundance for high-diversity assemblages, where most species are too rare for species-by-species approaches to be effective. However, to date it has been applied by fitting the bivariate Poisson-lognormal to all pairs of times at an individual site, and then fitting eq. (3.4) by least squares (hereafter the 'Individual Maximum Likelihood', or IML, approach) (Engen et al. 2002, 2018). An important limitation this approach is that time-series data from assemblage surveys are usually too noisy and short to obtain reliable estimates in this fashion. A potential solution is to use multiple time series (i.e., spatial replicates of short time series) from the same community type as random effects in a single analysis, instead of estimating parameters independently for each time series. The rationale is that, if one can assume that replicates of time series across space represent similar community dynamics, then VPRSA can benefit from leveraging information across multiple communities.

Here, I simulate different scenarios of community dynamics to test the robustness of variance components estimated from VPRSA, and I also investigate the potential to leverage multiple time series via random effects modelling to obtain more reliable estimates of variance components for short time series (hereafter "Mixed-Effects Model" or MM: the approach adopted in Chapter 2). Specifically, I conduct VPRSA analysis using both MM and IML on simulated community dynamics data that systematically violates the assumptions of VPRSA, and I compare estimated variance components with approximate "true" variance components based on the analytical solutions and (known) underlying parameters of the simulated communities. R codes, including IML, MM approaches and Gompertz community dynamics simulations, are open access at https://github.com/TsaiCH/simsEngenVPRSA.

3.2 Methods

3.2.1 Community dynamics model

I use state-space models that incorporate different assumptions about community dynamics to produce simulated data, which I then analyse using the VPRSA approach applied in the previous chapter (Chapter 2). This allows me to evaluate both the performance of my mixed model (MM) versus the independent maximum likelihood (IML) approach, and to evaluate the robustness of VPRSA estimates to violation of the assumptions of the community dynamics model from which it was derived. The R code for simulations of community dynamics, and VPRSA estimation, is available and open access at https://github.com/TsaiCH/simsEngenVPRSA.

First, I simulate the abundance dynamics of species i=1..S in a community according to the discrete-time multivariate Gompertz model (Ives et al. 2003):

$$N_{i,t+1} = N_{i,t} \exp\left[a_i + \left(b_{ii} - 1 \right) \log N_{i,t} - \sum\nolimits_{j \notin i} b_{ij} \log N_{i,t} \right] + \varepsilon_{d,t} + \varepsilon_{e,t} \quad \text{(eq. 3.8)}$$

where $N_{i,t}$ is the abundance of species *i* at time *t*, a_i is the species-specific intrinsic growth rate, and b_{ii} and b_{ij} are coefficients related to intra- and inter-species density dependence, respectively. b_{ii}-1 is the strength of intra-specific density-dependence, while larger values of b_{ij} indicate stronger inter-specific interactions. Here, I constrain these values to be between 0 and 1 (b_{ii}<0 represents strongly over-compensatory interactions, where increases in N_t produce decreases in N_{t+1}, whereas b_{ii}=1 indicates density-independent dynamics. Additionally, the species-specific intrinsic growth rates (a_i) are assumed to vary among species according to a normal distribution (*sensu* r_i in eqs.1-2). $\epsilon_{d,t}$ and $\epsilon_{e,t}$ are species-specific process noises of demographic and environmental stochasticity, assumed to follow normal distributions (see definitions as follows).

Let $\log N_t = \mathbf{X_t}$ and take the natural logarithm of both sides of eq. (3.8) in order to facilitate expressing the Gompertz-type community dynamics in matrix form. By doing so, the community dynamics model becomes an order-one multivariate autoregressive model with two components of process noise as follows:

$\mathbf{X_{t+1}} = \mathbf{A} + \mathbf{B}\mathbf{X_t} + \mathbf{D_t} + \mathbf{E_t}$ (eq. 3.9)

$\mathbf{E_t} \sim MVN(\mathbf{O}, \mathbf{\Sigma})$ (eq. 3.10)

$\mathbf{D_t} \sim MVN(\mathbf{O}, \mathbf{\Lambda})$ (eq. 3.11)

where $\mathbf{X_t}$ is a vector containing log abundance for each species *i* at time *t*, $\mathbf{A}$ is a vector containing species-specific intrinsic growth rates (a_i), and $\mathbf{B}$ is the interaction matrix where the diagonal (b_{ii}) and off-diagonals (b_{ij}) are coefficients in related to intra- and inter-specific density dependence (keeping in mind that the intra-specific density-dependence is b_{ii}-1, whereas the intra-specific density dependence is b_{ij}). $\mathbf{E_t}$ is a vector of random variables containing species' responses to environmental fluctuations (i.e., the perturbations to the intrinsic growth rate due to environmental stochasticity) and follows a multivariate normal distribution (*MVN*) with zero means and variance-covariance matrix $\mathbf{\Sigma}$ (eq. 3.10). $\mathbf{D_t}$ is a vector of random variables containing perturbations due to demographic stochasticity. These quantities also follow a multivariate normal distribution (*MVN*) with zero means and variance matrix $\mathbf{\Lambda}$ (by definition, the covariances of this matrix are all zero). Because less abundant species are more prone to demographic stochasticity than abundant species, I follow previous work and model the demographic variances in log-abundance (the diagonal in $\mathbf{\Lambda}$) as inversely proportional to the square root of species abundance (eq. 3.11) (Lande *et al.* 2003; Boettiger 2018). That is, the diagonal elements of $\mathbf{\Lambda}$ follow:

$$\Lambda_{ii} = \frac{\sigma_d^2}{\sqrt{\exp\left(X_{i,t}\right)}} \quad \text{(eq. 3.12)}$$

Finally, let $\boldsymbol{\lambda}$ represent a vector containing the expected relative species abundance in a random sample of the species-abundance distribution, such that:

$$\boldsymbol{\lambda}_{\mathbf{t}} = c\exp\left(\mathbf{X}_{\mathbf{t}}\right) \text{ (eq. 3.13)}$$

where c is a measure of sampling intensity. Hence the (simulated) sampled species abundances Y_t will be a Poisson sample of $\boldsymbol{\lambda}_{\mathbf{t}}$:

$$\mathbf{Y}_{\mathbf{t}} \sim Poisson(\boldsymbol{\lambda}_{\mathbf{t}}) \text{ (eq. 3.14)}$$

The final model of community dynamics represents a discrete-time multivariate state-space model with normally-distributed equilibrium log-abundances, normally distributed process noise, and Poisson-distributed observation error. I therefore model the sampled abundance values y_t as following a Poisson-lognormal distribution (Engen & Lande 1996).

3.2.2 Simulating empirically constrained community dynamics data

Simulated data were constrained to have similar numbers of locations ("reefs"), time series lengths, species richness, and numbers of individuals as the LTMP data analysed in Chapter 2. Specifically, I generated 100 simulated data sets, each of which consisted of 40 simulated time series ("reefs") sampled annually for 11 years, to correspond to the time series for the 40 annually-sampled reefs in the LTMP. For each reef, true total species richness was fixed at S=100 in all simulations. The level of "true" total richness used in simulations is close to the upper bound of estimated total richness at the reef scale in the LTMP (see Figure 2.1 in Chapter 2). Then, for each year at each reef, I simulated a Poisson random sample with a mean of 1500 individuals (i.e., the sampling intensity c was set so that the sum of the Poisson mean abundance across all species was equal to 1500), since this was close to the median sample size in the LTMP. If any simulated samples had fewer than 40 observed species (i.e., species with sampled abundance greater than zero), that sample was discarded, and a new random sample was taken from the community for that reef and year. This threshold of 40 observed species was used to prevent unrealistically low representation of the community (in the LTMP data, no reefs had fewer than 40 observed species in any year). Simulations where this occurred were extremely rare (approximately 2% of simulations), so it is unlikely that this culling process has biased the conclusions of my analyses.

For each simulated data set, the communities on all reefs were assumed to have the same community dynamics parameters for eqs. (3.9-3.11), except for the strength of environmental stochasticity (i.e., σ_e in eq. 1 and the diagonal $[\sigma_e^2]$ of $\mathbf{\Sigma}$ in eq. 10). This last quantity was varied systematically, in order to produce a data set in which the relative importance of deterministic species differences versus environmental stochasticity varied widely among reefs, which could then be used to evaluate how well the IML and MM methods resolved these differences. Specifically, the environmental variance term (i.e., σ_e^2) was varied from 0.025 to 0.5 in equal sized-increments across the simulated reefs (e.g., one reef had $\sigma_e^2 = 0.025$, another had $\sigma_e^2 = 0.0372$, and so on up to $\sigma_e^2 = 0.5$). This created a true distribution of variance components among reefs that was uniform and extended almost all the way to zero, and thus deviated markedly from the Gaussian distribution assumed by the MM model (Figure C1). Consequently, it represented a particularly stringent test of random-effects estimates from that model. For the effect of deterministic, intrinsic species differences (cf. eq. 3.1-3.2 and eq. 3.9), the species-specific intrinsic growth rates were modelled as varying among species according to a normal distribution with mean μ_r=1.5 and standard deviation σ_r= 0.25, based on a meta-analysis of global fishery stock assessments (Myers *et al.* 1999). Demographic stochasticity was simulated as process noise ($\mathbf{D}_t$), where the demographic variances (the diagonal of $\mathbf{\Lambda}$ in eq. 3.11) were scaled by the value $\sigma_d^2 = 0.5$ (this value is used as comparable with the magnitude of environmental stochasticity, though realised demographic stochasticity scales with species abundances). Other parameter values varied among simulation scenarios, as specified below.

3.2.3 Scenarios of simulated community dynamics data

I simulated five scenarios of community dynamics, which are empirically constrained as described above, to test the robustness of variance components (i.e., the relative importance of deterministic versus stochastic factors in eq. 3.3) estimated by VPRSA. The values of σ_e^2 and σ_d^2 specified above for all simulations, and the additional parameters specified in scenarios (i)-(v) below, were chosen so that the simulations produced frequency distributions of $\widehat{\sigma_{tot}^2}$ (variance of log-abundance in the communities), and sample completeness (measured as the fraction of the species pool observed at each site in each year) that were similar to those produced when the Poisson-lognormal was fitted to the LTMP data (Figure C1, also see Chapter 2). In addition, to ensure a stationary distribution of population sizes (i.e., all species coexisting, and fluctuating around their species-specific equilibrium abundances), the complex norm of eigenvalues of interaction matrix **B** were constrained to be less than one. This constrained the overall strength of interspecific competition in scenario (iii) (i.e., if interactions were too strong, species would be unable to coexist).

The details of parameters and scenarios of simulated community data are as follows:

(i) *Baseline*: These simulations were run to conform with the assumptions of the stochastic community-dynamic theory from which VPRSA was derived. Specifically, intra-specific density dependence was the same for all species (b_{ii}=b), the interaction matrix **B** contained no inter-specific density dependence (i.e., the off-diagonals of $\mathbf{B}$ were zero in eq. 3.9) and responses to environmental fluctuations were independent and equal in variance (i.e., the off-diagonals of $\mathbf{\Sigma}$ were zero in eq. 3.10, and the variances were all equal to the reef-specific values of σ_e^2 specified above: $\sigma_{e,ii}^2 = \sigma_e^2$). In these simulations, I set b=0.84 (i.e., the strength of density dependence was 0.16).

(ii) *Varied intra-specific interactions (varIntra)*: This scenario introduces between-species variation in intra-specific density dependence. Specifically, values on the diagonal of $\mathbf{B}$ in (eq. 3.9) were drawn from a normal distribution with mean 1.8 and standard deviation 0.4, and then inverse-logit transformed to yield values between 0 and 1. This produces random coefficients of the diagonal of **B** with mean values close to 0.84 (i.e., $E[b_{ii}] \approx 0.84$, implying average strength of density-dependence $1 - E[b_{ii}]$ ≈0.16), and standard deviations close to 0.06, implying a coefficient of variation of density-dependent strength of about 0.37. Because equilibrium abundance is $\exp\left(\frac{r_i}{b_{ii}}\right)$, this approach produced unrealistically large variance in the total variance of log-abundance, σ_{tot}^2. Therefore, the diagonal elements of $\mathbf{B}$ were reordered to increase with species' intrinsic growth rates (i.e., elements of vector **A** in eq. 3.9), so that species with strong density dependence also had high intrinsic growth rates. This yielded more realistic variances of log-abundance (see below).

(iii) *Inter-specific interactions (varInter)*: This scenario introduces diffuse inter-specific density dependence by drawing the off-diagonals of $\mathbf{B}$ from a uniform distribution between 0 and 0.002 (mean=0.001). This yielded an average summed effect of interspecific interactions (across the other 99 species in the community) that was approximately 60% the strength of intra-specific density-dependence (i.e., $0.001 \times 99 \approx 0.6 \times 0.16$). All other parameter values were the same as in the Baseline simulation. Scenarios with stronger interspecific interactions are not considered here. Mean interaction strength values slightly above those employed here (0.001-0.003) tended to produce distributions of observed richness values that differed notably from the data (lower observed richness levels, and more strongly right-skewed abundance distributions). Moreover, mean interaction strengths above about 0.003 tended to produce assemblages lacking a stable coexistence equilibrium (Figure C2).

(iv) *Unequal environmental variances (varEnv)*: This scenario introduces heterogeneity among species in sensitivity to environmental fluctuations. Specifically, species-specific environmental variances (the diagonal of Σ in eq. 3.10) were drawn from a uniform distribution between 0 and 2 (mean=1), and then multiplied by the reef-specific σ_e^2 term, as specified above. This ensured that the average value of environmental variance ranged from 0.025 to 0.5, as in the *Baseline* scenario, and thus continued to yield realistic σ_{tot}^2 values. All other parameter values were as in the *Baseline* scenario.

(v) *Unequal environmental covariances (varcovEnv)*: This scenario introduces covariance in species' responses to environmental fluctuations. In order to draw random matrices that were valid variance-covariance matrices (i.e., that were positive-definite symmetric matrices), I used the Cholesky decomposition method implemented in the "Matrix" package in R for matrix algebraic computations (see https://github.com/TsaiCH/simsEngenVPRSA for R codes of simulations of this scenario). Specifically, I generated a lower-triangular matrix $\mathbf{L}$ whose elements were drawn from a normal distribution with mean 0 and standard deviation 0.25. I then produced a covariance matrix $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$. The elements of $\mathbf{\Sigma}$ were subsequently standardized by the mean of the diagonal elements, and then the entire matrix multiplied by the reef-specific environmental variance term σ_e^2. This yielded a matrix of unequal variances and covariances among species, whose correlation coefficients ranged from -0.5 to 0.5, with mean 0, and with the mean of the diagonal elements equal to the reef-specific value σ_e^2 (and thus comparable in environmental variability to the other scenarios).

3.2.4 Estimating VPRSA through simulated community dynamics data

I tested the robustness of statistical inferences drawn from VPRSA by analysing the simulated data described above. Following the original estimation procedure of VPRSA, I use the R package "poilog" to fit the bivariate Poisson-lognormal distributions to paired assemblage at different time lags (Grøtan & Engen 2008). This involves fitting the bivariate Poisson lognormal to all possible pairs of times by maximum likelihood, for each simulated reef's time series, and then analysing the correlation coefficients for all of these pairs as functions of the time elapsed between them, as in Chapter 2. (Engen et al. 2002; Engen et al. 2011).

After obtaining the correlation coefficients from Poisson-lognormal fits, I estimated the three parameters (ρ_0, ρ_∞, and δ) of the autocorrelation function of (eq. 3.4). Here, I compare two different approaches: Individual Maximum Likelihood (IML) and Mixed-effects Model (MM). As per the original IML approach of Engen et al. (2002), I fit the autocorrelation function (eq. 3.4) to the estimated correlation coefficients from the bivariate Poisson-lognormal fits, independently for each reef. That is, I fit eq. (3.4) with a Gaussian error term and fixed parameters to the correlation coefficients as a function of time lag. I then compare this approach with a random-effects version of eq. (3.4) fitted to the simulated community time series from all 40 reefs simultaneously (as employed in Chapter 2). This random-effects approach can be formulated as follows:

$$\rho_{k,\tau} = \left(\rho_0^{R_k} - \rho_\infty^{R_k}\right)\exp\left(-\delta\tau\right) + \rho_\infty^{R_k} + \varepsilon \text{ (eq. 3.15)}$$

$$\rho_0^{R_k} \sim \mathrm{LN}(\mu_{\rho_0}, \sigma_{\rho_0}^2) \text{ (eq. 3.16)}$$

$$\rho_\infty^{R_k} \sim \mathrm{LN}(\mu_{\rho_\infty}, \sigma_{\rho_\infty}^2) \text{ (eq. 3.17)}$$

$$\varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2) \text{ (eq. 3.18)}$$

where $\rho_{k,\tau}$ represents the correlation coefficient of relative abundances for two samples from reef *k* as a function of the time lag τ between them. $\rho_0^{R_k}$, $\rho_\infty^{R_k}$, and δ represent the three parameters of the autocorrelation function (eq. 3.4), in which superscript R_k denotes a reef-specific value for reef *k* due to the random effect of reef. Density dependence is assumed to be a fixed effect (as in Chapter 2) in eq. (3.15), because analyses of LTMP data found numerical instability when the density dependence parameter is instead modelled as a random effect (Appendix B, Chapter 2). The other random effects of parameters (eqs. 3.16-3.17) are assumed to follow lognormal distributions, while the residual error is assumed to follow a normal distribution (eq. 3.18). Visual inspection of model fits tended to indicate acceptable agreement with model assumptions (also see Appendix B, Chapter 2), and the numerical stability and convergence is guaranteed for all estimations.

For both the IML and MM approaches described above, I use the estimated parameters of the autocorrelation function to calculate the proportion of the total variance of log-abundance attributable to deterministic species differences, responses to environmental fluctuations, and demographic stochasticity and sampling effects, as in Chapter 2 (eqs. 3.3-3.7 in Introduction). I use R package "bbmle" (Bolker and R Development Core Team 2017) to compute the IML estimates, and the R package "TMB" (Kristensen *et al.* 2016) to compute the MM estimates.

3.2.5 Testing the robustness of statistical inferences drawn from VPRSA

I compare the performance of the MM and IML estimates with the *Baseline* scenario, since the "true" proportional variance terms (i.e., the relative importance of deterministic versus stochastic factors in eq. 3.3) can be calculated directly from the parameters used in the process-based model (eq. 3.8-3.14). Specifically, for each reef in each simulation, there is an estimate of the proportional variance components from IML and from MM. I compare these estimates with the analytical expected variance components, which I calculated using the process model parameters of the *Baseline* simulation: variance in intrinsic growth rates (σ_r^2), environmental variance (σ_e^2), and intra-specific density dependence (δ) according to eq. 3.3. Specifically, for the density-dependent parameter, I used $\delta \equiv b$-1 (i.e., the strength of intra-specific density-dependent parameters used in the simulations), to ensure that the equilibrium abundances matched their predicted values (see below). Because the environmental variance parameter varied from one reef to the next in each simulated data set, estimator performance was assessed visually by plotting both estimated and analytical variance components as functions of the reef-specific environmental variance.

Once a method of estimation was identified as superior (either IML or MM), I then applied that method to assess the robustness of estimated proportional variance components to violations of model assumptions (scenarios ii through iv). For these simulations, I used average values (across species) for computing "expected" proportional variances in scenarios of varied intra- and inter-specific interactions (scenarios ii and iii) and environmental variances and covariances (scenarios iv and v), because the density dependence and environmental variance parameters are not fixed constants applied to all species in these scenarios. For analytical simplicity, the "expected" proportional variance explained by demographic/sampling stochasticity is ignored in computations of eq. 3.3 (i.e., assuming V_d is negligible in eq. 3.3). This assumption is reasonable given that the proportional variance of demographic/sampling stochasticity is quite small in both simulations in this chapter, and in the LTMP data analysed in Chapter 2.

In addition, I calculated, analytically, an alternative measure of expected proportional variances to take more explicit account of between-species heterogeneity, species interactions, and covariances in response to environmental fluctuations. I did this by exploiting general analytical solutions (e.g., Ives et al. 2003) for the environmental variance and variance in equilibrium abundance for the discrete time, stochastic, multivariate Gompertz model as follows:

$$v_r = \mathrm{var}\left(\log \hat{\mathbf{N}}\right) = \mathrm{var}\left[\left(\mathbf{I}-\mathbf{B}\right)^{-1}\mathbf{A}\right] \text{ (eq. 3.21)}$$

$$v_e = \bar{\mathbf{V}}_e = avg\left[diag\left[\left(\mathbf{I}-\mathbf{B}\otimes\mathbf{B}\right)^{-1} vec\left(\mathbf{\Sigma}\right)\right]\right] \text{ (eq. 3.22)}$$

$$P_{niche} = \frac{v_r}{v_r + v_e} \text{ (eq. 3.23)}$$

$$P_{env} = \frac{v_e}{v_r + v_e} \text{ (eq. 3.24)}$$

where v_r is the among-species variance of equilibrium population sizes (on a logarithmic scale), and v_e represents the average species-level variance of log-abundance due to environmental stochasticity. In (eq. 3.21), $\log \hat{\mathbf{N}}$ represents the species' abundances at stationary or equilibrium states, "var" represents the variance operator, **B** is the interaction matrix (as per eq. 3.9), **A** is a vector of intrinsic growth rates (as per eq. 3.9). In (eq. 3.22), $\bar{\mathbf{V}}_{\mathrm{e}}$ is the average of the diagonal of the environmental variance-covariance matrix at stationary states, "avg" represents the arithmetic mean function, and "diag" and "vec" are the diagonal and vectorization operators (Searle 1982). The symbol $\otimes$ represents the Kronecker or tensor product. From equations (3.23)-(3.24), the two variance components are generalized to account for heterogeneity in intra-specific and inter-specific density dependences through the interaction matrix **B**, as well as heterogeneity in environmental variances and covariances through the environmental variance-covariance matrix $\mathbf{\Sigma}$.

Importantly, under the assumptions of the *Baseline* scenario, the approximate measures of ν_r and ν_e above (eqs. 3.21-3.24) collapse to discrete-time analogous of Engen and colleagues' functional forms of V_r and V_e (eq. 3.3) (where the density-dependent parameter $\delta \equiv 1 - b_{ii}$). However, once species interactions or heterogeneity in density-dependence are incorporated (e.g., scenarios ii and iii), σ_r^2 deceases to be directly proportional to the variance in species' equilibrium log-abundances, so this modified version of ν_r (eq. 3.21) might be a better measure of the relative importance of niche structure than V_r from the original theory (eq. 3.3). Similarly, in the presence of environmental covariances or heterogeneity in environmental variances among species (e.g., scenarios iv and v), the functional relationship between V_e in eq. (3.3) and the overall contribution of environmentally-mediated population fluctuations to variance in species-abundances may also break down, rendering eq. (3.22) a more robust measure. For example, environmental covariances and compensatory interspecific interactions would likely change the linear functional relationship between r_i and equilibrium log-abundance. Consequently, for scenarios (ii)-(v), I test estimated variance components from the IML and MM fits (eqs. 3.5-3.7) against theoretical variance components calculated according to both the original theory (eq. 3.3), and the generalized forms above. However, the model from which these generalized forms are derived (eqs. 3.21-3.22) omits demographic stochasticity; thus, if the contribution of V_d (eq. 3.3) is non-negligible, then analytically-calculated proportional variance components from (eqs. 3.23-3.24) will be biased. Consequently, to maximize the comparability of these quantities ν_r and ν_e with the original VPRSA forms, I normalized the variance components of the former V_r and V_e (eqs. 3.5-3.7) as follows:

$$\hat{V_r} = \frac{V_r}{V_r + V_e} \equiv \frac{V_r}{\sigma_{\mathrm{tot}}^2 - V_d} \quad \text{(eq. 3.25)}$$

$$\hat{V_e} = \frac{V_e}{V_r + V_e} \equiv \frac{V_e}{\sigma^2_{\mathrm{tot}} - V_d} \quad \text{(eq. 3.26)}$$

3.3 Results

Incorporating spatial replicates via MM estimation substantially improved estimates of proportional variance (i.e., V_r , V_e , and V_d divided by the overall variance; eq. 3.3), compared to IML estimation. Specifically, for both the mixed model (MM) and independent model (IML) fits, as the magnitude of environmental variance increases along the spatial replicates of community dynamics, the proportional variance explained by deterministic "intrinsic species differences" decreases (red points in Figure 3.1), and the proportional variance explained by "environmental stochasticity" correspondingly increases (blue points in Figure 3.1).

However, the MM method has much lower sampling variance (spread of colored points around colored lines in Figure 3.1A), and the analytical expectation is less biased (compare black lines with colored lines in Figure 3.1A), compared to the IML-based estimates (Figure 3.1B) (see Figure C3 for various scenarios with IML). In particular, the IML estimates appear bimodal, with some time series erroneously converging on "quasi-neutrality" (i.e., blue points along the top of Figure 3.2B, indicating estimated proportional variance explained by deterministic species difference ~ 0 and proportional variance explained by environmental stochasticity ~1), even when the true proportions are nearly the opposite (Figure 3.1B). Interestingly, this behaviour is also observed when the IML approach is applied to the LTMP data (results not shown), so my simulations strongly suggest that this bimodality is an artefact of poor estimator performance.

Focusing on the better-performing MM-based estimates, I find that the central tendency of VPRSA estimates deviate detectably from analytical prediction (Figure 3.1A and Figure 3.2). The discrepancy is slightly more pronounced when intraspecific density-dependence varies among species (Figure 3.2A) compared to the other scenarios (Figure 3.2B-3.2D). Robustness of IML estimates are qualitatively similar to those of the MM results, but, as with the baseline scenario, the estimates exhibit considerably greater sampling variability than the corresponding MM estimates (c.f. Figure 3.2 and Figure C3).

To understand this discrepancy, it is important to note that the variance attributable to "deterministic species differences" is the variance in deterministic equilibrium population sizes (on a logarithmic scale). In Engen et al.'s (2002) original model, as in the baseline scenario, this is equal to σ_r^2/δ^2 or ρ_∞ with equations (3.3)-(3.7). However, when there is heterogeneity among species in intra-specific density dependence (i.e., the "varIntra" scenario; Figure 3.2A) or species interactions (i.e., the "varInter" scenario; Figure 3.2B), or when there are different species responses to environmental fluctuations (i.e., the "varEnv" and "varcovEnv" scenario), this is no longer true, and thus the original analytical expectation (eq. 3.3) itself may be biased. However, VPRSA estimates provide even less biased estimates of the proportional variance in log-abundance attributable to "deterministic species differences", when those proportions are calculated by taking the additional heterogeneity among species into account with the appropriate multivariate Gompertz expressions (eqs. 3.21-3.24) (Figure 3.3 and Figure 3.4). Interestingly, the modified expressions (eqs. 3.21-3.24) also perform better in the baseline scenario (i.e., when species differ only in intrinsic growth rates), compared to the original analytical expressions (eq. 3.3) (Figure C4).

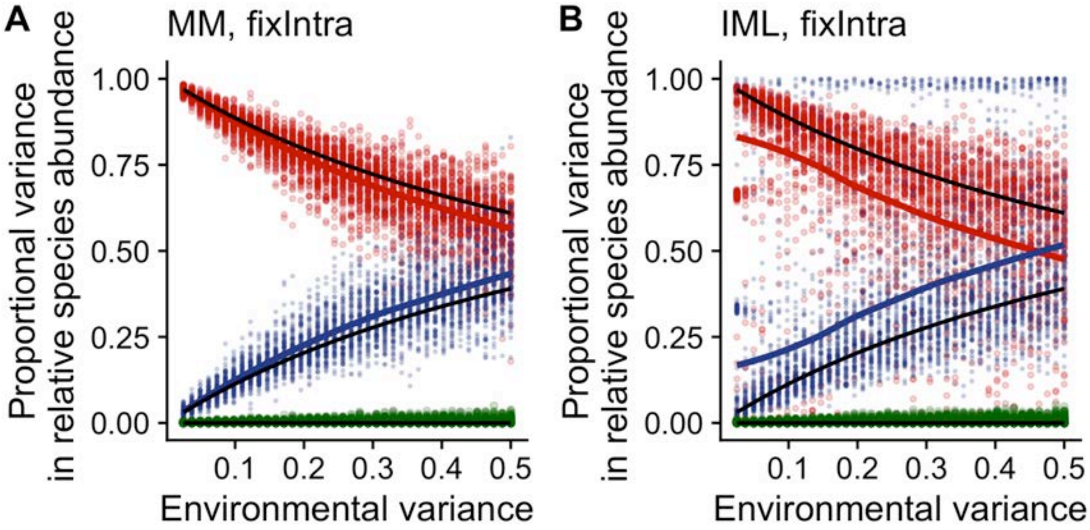

Figure 3.1 Relationships between true environmental variance and VPRSA-estimated variance components of deterministic species differences, environmental stochasticity, and a combination of demographic and sampling stochasticity for the "Baseline" scenario. (**A**) Mixed-effects model (MM) estimates (**B**) Individual maximum likelihood (IML) model estimates. Red, blue, and green colors represent the VPRSA-estimated

proportional variance driven by deterministic species differences (niche structure), environmental stochasticity, and demographic/sampling stochasticity, respectively. Each point, irrespective of color, represents one simulated time series for one spatial replicate (reef) (i.e., for each color, n = 40 spatial replicates with varied environmental variance × 100 simulations = 400 points). All community dynamics data (points) are simulated from the baseline scenario. Red, blue and green colored lines represent the kernel smoothing of proportional variance estimates, respectively, obtained using local polynomial regression fitting. Black lines represent the analytical predictions of Engen et al. 2002 (eq. 3.3) using the "true" parameters from the simulations (and thus are identical in the two panels).

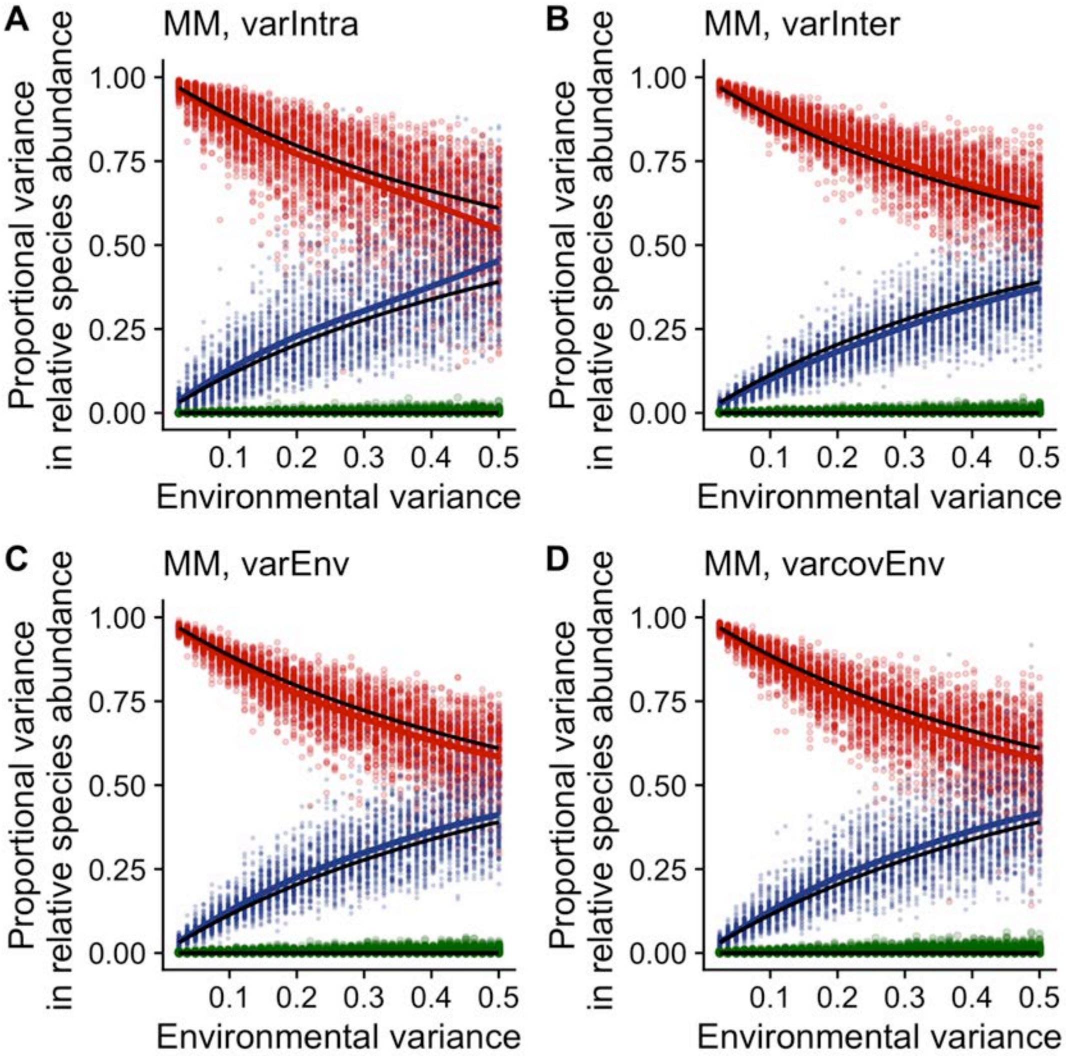

Figure 3.2 Relationships between true environmental variance and MM-estimated variance components under different community dynamics scenarios. Red, blue, and green colors represent the MM-estimated proportional variance components due to deterministic species differences, environmental stochasticity, and to other sources of

variance (e.g., demographic and sampling stochasticity). Colored (red, blue, and green) lines represents the kernel smoothing of proportional variance estimates, respectively, obtained using local polynomial regression fitting. Black lines represent the analytical prediction of Engen et al. 2002 (eq. 3.3) using the true parameters from the simulations. (**A**) The "*varIntra*" scenario, which includes species differences in intra-specific density dependence. (**B**) The "*varInter*" scenario, which includes species differences in inter-specific density dependence. (**C**) The "*varEnv*" scenario, which includes species differences in the magnitude of environmental variance. (**D**) The "*varcovEnv*" scenario, in which species' responses to environmental fluctuations covary.

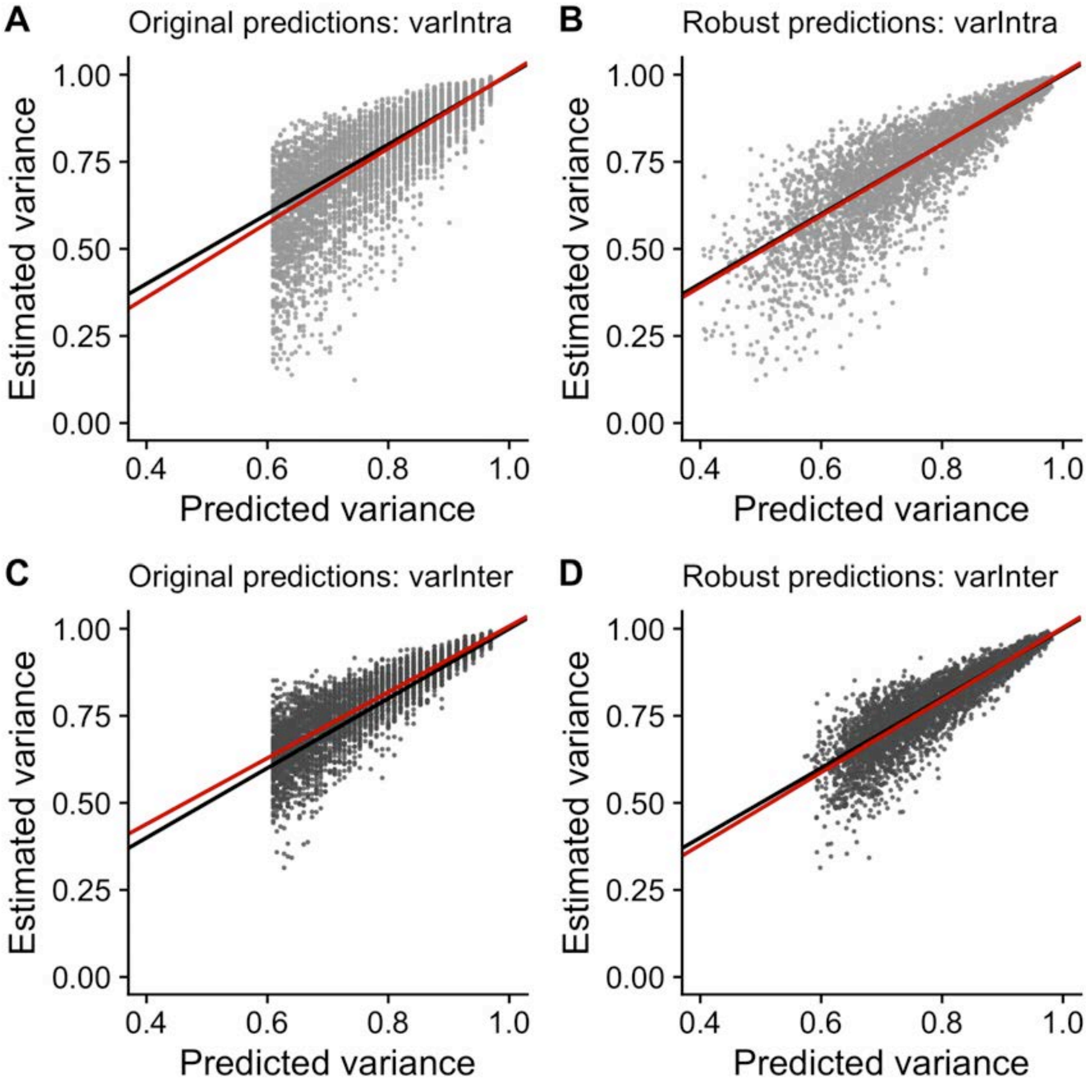

Figure 3.3 Relationships between variance estimate of deterministic species differences and its analytical prediction. The black line is the unity line indicating perfect agreement between VPRSA estimates and analytical predictions. The red line is a quantile regression

through the median of the MM-based VPRSA estimates of variance components of deterministic species differences. (**A, C**) The relationship between VPRSA estimates and the original analytical prediction of Engen et al. 2002 (eq. 3.3), and (**B, D**) the relationship between VPRSA estimates and the generalized analytical prediction from the discrete-time multivariate Gompertz model (eqs. 3.21-3.24) under the (**A, B**) “varIntra” and (**C, D**) “varInter” community dynamics scenarios.

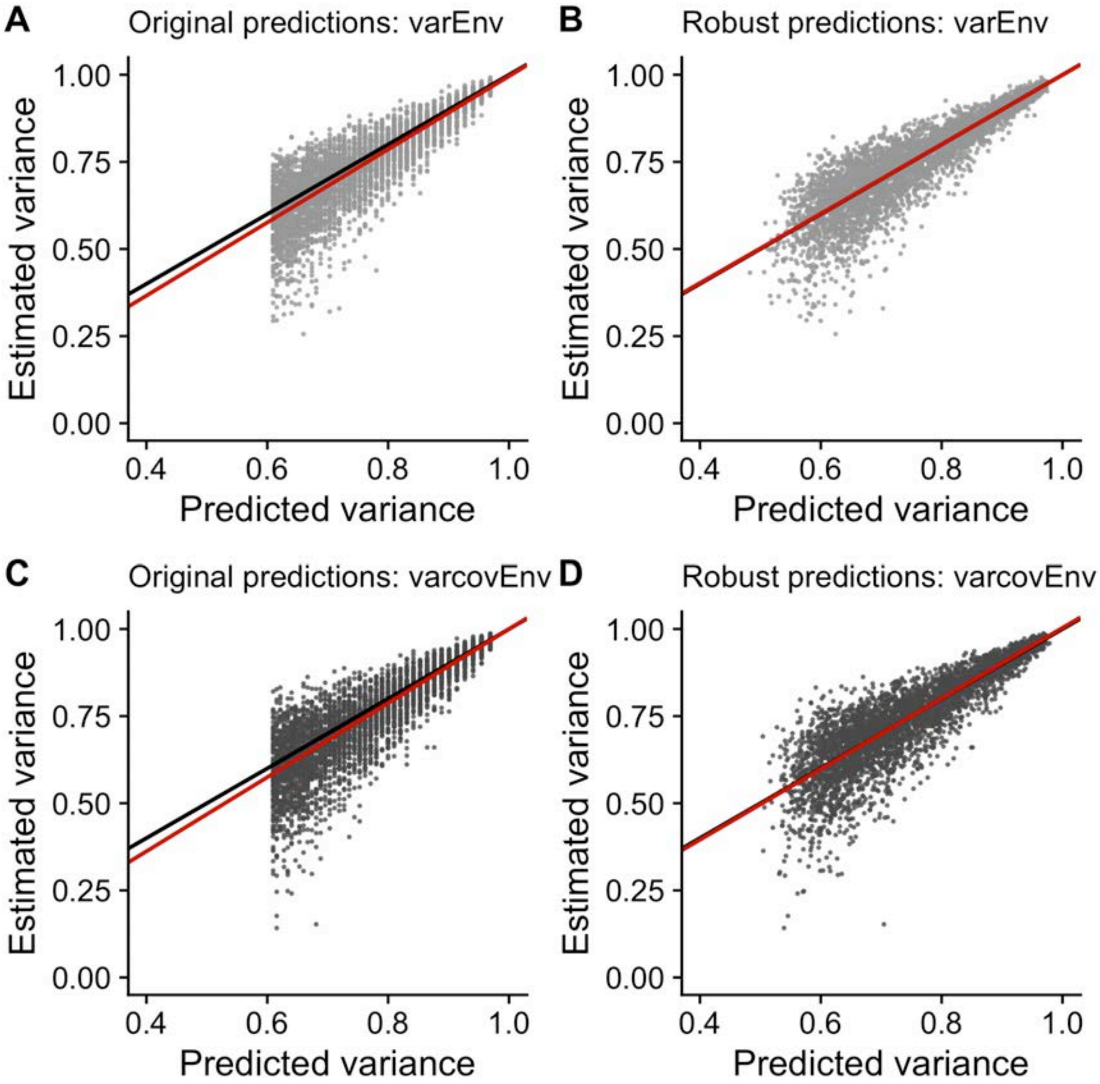

Figure 3.4 Relationships between variance estimate of deterministic species differences and its analytical prediction. The black line is the unity line indicating perfect agreement between VPRSA estimates and analytical predictions. The red line is a quantile regression through the median of the MM-based VPRSA estimates of variance components of deterministic species differences. (**A, C**) The relationship between VPRSA estimates and the original analytical prediction of Engen et al. 2002 (eq. 3.3), and (**B, D**) the relationship

between VPRSA estimates and the generalized analytical prediction from the discrete-time multivariate Gompertz model (eqs. 3.21-3.24) under the (**A, B**) “varEnv” and (**C, D**) “varcovEnv” community dynamics scenarios.

3.4 Discussion

My results suggest that proportional variance estimates from VPRSA provide robust information about the relative importance of deterministic species differences versus stochastic environmental fluctuations in shaping patterns of commonness and rarity among species, even when key simplifying assumptions about community dynamics made by the original theory are violated. Additionally, the MM method generally outperforms the original IML method for estimating variance components in temporal dynamics of species’ relative abundances, at least for the level of spatial replication and time series length considered here, which correspond to the characteristics of the LTMP data analysed in Chapter 2. These results suggest that the VPRSA approach may be more broadly applicable than previously realized.

First, my simulations clearly show that MM-based VPRSA performs better than IML-based VPRSA. MM outperforms the original IML by shrinking the estimation error of VPRSA for levels of spatial and temporal replication comparable to that of the LTMP data (Figure 3.1, also see Figure 3.2 and Figure C3). MM-based VPRSA shrinks variance estimates toward the ensemble mean through inclusion of spatial latent variables (i.e., random effects due to varied magnitudes of environmental stochasticity in community dynamics). This is because MM, but not IML, has the mathematical advantage of “shrinkage”: a tendency to shrink the magnitude of estimates of individual effects (e.g., reef-level deviations from the overall mean variance component values) by penalizing estimates that are too far from the mean. However, the advantage comes at the price of assuming a particular distributional form for the random effects (eqs. 3.15-3.18). My results suggest that better estimates are produced by leveraging information across time series on the order of 10 years, even where the true distribution of random effects lacks an internal mode, as in the simulations employed here (for which the distribution is uniform). Of course, as time series length increases, and among-site heterogeneity in variance components increases, this tradeoff could change. Regardless, simulation studies like this one provide a way to evaluate the relative plausibility of estimates obtained using MM versus IML approaches.

Second, my results support the robustness of inferences drawn from VPRSA to violations of simplifying assumptions of the underlying theory (Figure 3.2). That is, VPRSA provides robust estimates of the proportion of variance in species log-abundances explained by deterministic species differences and stochastic factors. However, this robustness does not extend without limitation to all process parameters. For example, VPRSA estimates are robust estimates of the contribution of deterministic species differences (or variance in carrying capacities at deterministic equilibria) to the overall variance in relative species abundance, which can be expressed as V_r / σ^2_{total} or $\mathrm{var}(\log \hat{\mathrm{N}}) / \sigma^2_{total}$. The numerator $\mathrm{var}(\log \hat{\mathrm{N}})$ can be interpreted as the deterministic variance component due to persistent species differences, which is exactly equivalent to the V_r term in eq. 3 when the underlying assumptions of VPRSA are met. I initially expected that, when key assumptions are violated (particularly the inclusion of interspecific interactions), equilibrium abundances would become decoupled from their VPRSA theory-predicted relationships with σ_r^2, σ_e^2 and δ, rendering the expressions in eq. (3.3) inaccurate (Figure 3.2). Under such conditions, I expected that the expressions in eqs. (3.21)-(3.24), which explicitly account for species interactions, covariation in responses to environmental fluctuations, and among-species heterogeneity in density-dependence and sensitivity to environmental fluctuations, would perform better. And, indeed, these alternative forms were less biased than the VPRSA estimates (Figure 3.3 and Figure 3.4). However, because this reduced bias was also apparent in the baseline simulations, where VPRSA assumptions were met, this better performance may have more to do with the fact that VPRSA is based on a continuous-time stochastic differential equation model, whereas the stochastic simulations, and our generalized variance component expressions, assume discrete-time dynamics. In any case, my findings suggest that VPRSA provides much more robust estimates of the relative importance of deterministic and stochastic processes in high-diversity assemblages than one might have assumed based on the assumptions of the original theory from which the expressions for these variance components was derived.

Third, my simulations were necessarily limited to a subset of possible departures from model assumptions (scenarios i-v). Consequently, I cannot rule out the possibility that

VPRSA is more sensitive to ecologically plausible violations of model assumptions not considered here. For example, I did not consider "modularized" interaction networks, where small subsets of species interact very strongly with each other but weakly or not at all with the other species in the assemblage (McCann *et al.* 1998; Berlow 1999; Wootton & Emmerson 2005). Obviously, an exhaustive exploration of all possible combinations of assumptions is logistically infeasible. Nevertheless, the results presented here provide evidence that the approach is much more robust than one might have assumed, based on the restrictive assumptions of the original theory. Moreover, further theoretical development may make possible a formal derivation of the variance due to environmental fluctuations from the general multivariate Gompertz model, analogous to the variance in equilibrium population sizes used here (eqs. 3.21-3.24), enhancing the robustness and strengthening the theoretical foundations of the VPRSA approach.

Fourth, my results have expanded the range of empirical cases for which VPRSA is feasible, but some additional characteristics of data sets are also important. In particular, VPRSA's statistical power is subject to sampling completeness and the natural variation in community dynamics across spatial replicates. In this chapter, spatial replicates and time-series length in community simulations were specifically designed to be comparable to the LTMP data. However, when applied to other data sets (e.g., bioTime global datasets; Dornelas et al. 2014, Bowles et al. 2019), the number of spatial replicates needed to produce reliable MM-based VPRSA estimates may depend on the magnitudes of both process and observation noise in the data. For example, if sample sizes are low, such that only a small proportion of the fauna is represented in each sample, spatial replicates may be indistinguishable from a collection of very noisy random samples from the same underlying community. This could render MM estimates uninformative. Alternatively, strong violation of VPRSA assumptions about community dynamics could also increase unknown process noise in VPRSA estimation from individual time series (e.g., model mis-specification error in species interactions). If so, VPRSA's first-order (Gompertz-type) approximation of community dynamics may deviate from true ones across spatial replicates. Consequently, MM-based VPRSA might fail to converge due to undistinguishable noises that behave like observation or process error in samples, and the sample size of spatial replicates might need to increase. The issues above warrant further simulation studies, with other community datasets as empirical baselines, in order to evaluate VPRSA's statistical power.

Lastly, my results suggest that VPRSA is well-suited to studies that focus on assessing the relative importance of deterministic versus stochastic processes in structuring ecological communities, where richness is too high for species-by-species approaches to be feasible. VPRSA has two major advantages over other approaches for such species-rich communities: it is derived from process-based theory, and it requires only species-abundance data in space and time. For instance, the recent empirical operationalization of modern coexistence theory (MCT) (e.g., Chesson 2000, Angert et al. 2009, Chu and Adler 2015, Ellner et al. 2016) explicitly incorporates niche structure and environmental stochasticity into process-based models to explain variation in species abundance. However, unlike VPRSA, MCT relies on species-by-species measures of inter-specific interactions and demographic parameters (e.g., Chu and Adler 2015). For such approaches, much more detailed species-level data, rather than just numerical abundances, are necessarily required, limiting the applicability of MCT to high-diversity communities, except in extraordinary cases where such data exist (Usinowicz *et al.* 2017). The fact that VPRSA requires only abundance data does restrict the inferences that can be drawn from its application, compared to MCT. However, in high-diversity systems, such as tropical coral reefs (Chapter 2), where detailed species-level data limit the applicability of MCT, VPRSA, particularly if it can be combined with functional and phylogenetic approaches, has considerable potential to advance our understanding of how deterministic and stochastic aspects of community dynamics interact to shape patterns of commonness, rarity, and biodiversity.

Chapter 4: Portfolio effect and diversity-stability relationship depend on how communities are structured across coral reefs

4.1 Introduction

Species diversity can play an important role in stabilizing aggregate community properties such as total community abundance (hereafter, “community stability” refers to temporal stability in aggregate community properties) (Hooper *et al.* 2005; Hector *et al.* 2010; Loreau 2010; Tilman *et al.* 2014). Ever since the diversity-stability relationship was first articulated (McNaughton 1977; Tilman 1999), theoretical models and manipulative experiments have generally supported the hypothesis that species diversity can increase community stability (Hughes & Roughgarden 2000; Jiang & Pu 2009; Campbell *et al.* 2011; Gross *et al.* 2014). Such stabilizing effects are broadly termed “portfolio effects” in analogy with the tendency for more diverse financial portfolios to show more stable asset returns over time (Doak *et al.* 1998; Thibaut & Connolly 2013; Schindler *et al.* 2015).

One general approach to partition diversity-stability relationships (or portfolio effects) is according to whether diversity reduces the overall synchrony in fluctuations of a community’s constituent populations (Loreau & de Mazancourt 2008a; Hector *et al.* 2010), or whether diversity reduces the average fluctuations of the individual populations themselves (Tilman 1999; Thibaut & Connolly 2013). In particular, reduced fluctuations of a community’s constituent species can arise from species interactions or responses to environmental fluctuations (Hughes & Roughgarden 2000; Thibaut *et al.* 2012; Loreau & de Mazancourt 2013), depending on how population-level variability changes systematically with the mean (Doak *et al.* 1998; Tilman 1999; Gross *et al.* 2014) and whether species unevenness responds to changes in richness (Wayne *et al.* 2007; Sasaki & Lauenroth 2011; Ma *et al.* 2017) (see Thibaut & Connolly 2013 for a unified definition). The established consensus that diversity generally stabilizes ecological communities has led to calls for protecting “diversity” itself as a means of securing stability in the provision of ecosystem goods and services (Hooper *et al.* 2005; Cardinale *et al.* 2012; Schindler *et al.* 2015).

There exists, nevertheless, a longstanding controversy over whether the above mentioned diversity-stability relationships and portfolio effects occur along natural richness gradients at the regional scale (Sankaran & McNaughton 1999; Srivastava & Vellend 2005; Wardle 2016), which obviously has implications for whether such relationships can inform priority-setting in biodiversity conservation (Srivastava & Vellend 2005; Gonzalez *et al.* 2016; Wardle 2016; Vellend 2017). In particular, the evidence for diversity-stability effects comes overwhelmingly from manipulative experiments, where the richness gradient is manipulated and other confounding sources of variation in community structure and stability (such as variability in species composition, relative abundance, and environmental fluctuations) are controlled (Sankaran & McNaughton 1999; Srivastava & Vellend 2005; Wardle 2016). Conversely, in natural systems, such additional sources of variation in community structure could mediate or counteract the stabilizing effect of diversity (Sankaran & McNaughton 1999; Blüthgen *et al.* 2016; Ma *et al.* 2017), potentially explaining cases in which the sign and strength of diversity on community stability is negative (Valdivia & Molis 2009; Veen *et al.* 2018), or undetectable (Petchey *et al.* 2002; Romanuk *et al.* 2009; Xu *et al.* 2015).

Regional-scale (spatial) variation in community structuring processes, rather than diversity, could be strongly associated with community stability. Within assemblages, interspecific variation in species abundance is driven both by persistent, deterministic species differences that create heterogeneity in long-term mean abundance, and by asynchronous fluctuations in species abundances, due to factors such as differential responses to environmental stochasticity (Lande et al. 2003; also see variance components defined in Chapters 2 and 3). Moreover, the relative importance between deterministic species differences and environmental stochasticity has the potential to influence overall community stability. Deterministic niche structure, for instance, underpins species' differences in intrinsic growth rates and intra- and inter-specific interaction strengths, which in turn affects species' long-term mean of abundances (Hughes & Roughgarden 2000; Engen *et al.* 2002; Loreau & de Mazancourt 2008a). These deterministic factors have been known to influence community stability by means of demographic trait differentiation, resource partitioning, enhanced intra-specific versus interspecific density dependence, and asynchronous species dynamics (Tilman 1999; Hughes & Roughgarden 2000; Loreau & de Mazancourt 2008b; Loreau 2010). Alternatively, stochastic fluctuations, arising from environmental and demographic stochasticity, is expected to reduce community stability through amplifying population-level variability (Hughes & Roughgarden 2000).

It is tempting to expect that a relationship between the variance components of community structure (deterministic species differences versus stochasticity; Chapter 2-3) and community stability (this chapter) would be trivial: that is, where fluctuations are more important, abundances will be more variable, and this will destabilize overall community abundance. However, there are several reasons why this might not be the case. Firstly, regional variation in the ratio of V_r (variance due to deterministic species differences; Chapter 2-3) to V_e (variance due to stochastic environmental fluctuations; Chapter 2-3) may not be driven by variation in V_e; that is, more "niche-structured" species abundance distributions may not exhibit reduced fluctuations of constituent species. Also, V_r depends on the variance in log-abundance; on this scale, a doubling of abundance from 10 to 20 is a fluctuation comparable in magnitude to a doubling of abundance from 1000 to 2000, and because rare species constitute the overwhelming majority of species in high-diversity assemblages, it is fluctuations in the abundance of rare species that will dominate estimates of V_e in the VPRSA approach employed in preceding chapters of this thesis. In contrast, community stability is measured on an arithmetic scale, such that it is overwhelmingly influenced by the dynamics of the most abundant few species.

In this chapter, I estimate the magnitude of regional variation in community stability of fish assemblages on the Great Barrier Reef, and I test whether regional variation in diversity explains this variation in stability. I also test whether regional-scale variation in the relative importance of stochastic fluctuations in determining interspecific variation in species abundances is associated with regional-scale variation in community stability using the VPRSA approach used in the previous chapters (Chapters 1-3). Specifically, I define a niche structure index (hereafter, NSI) that quantifies the relative importance of persistent niche differences versus stochastic fluctuations in community dynamics. I focus on the relative, rather than the absolute, importance of these two factors in order to decouple NSI from species unevenness (i.e., the overall variance in log-abundance among species: Chapter 2). I quantify how much the variation in community-level stability across coral reefs can be explained by the simultaneous effects of species diversity, unevenness and NSI, and if so, how much these effects occur due to effects on synchrony versus average population variability. In addition, I ask whether environmental variables associated with the community structure variables of richness, evenness, and NSI explain regional variation in community structure better, or worse, than the community structure variables themselves, to gain insight into the plausibility of the hypothesis that diversity or NSI are causally or coincidentally related to community stability.

4.2 Methods

4.2.1 Long-term reef fish community data and environmental covariates

Using data from the Long-term Monitoring Program of the Australian Institute of Marine Science (Sweatman *et al.* 2005), I analyse annual time-series (11 years) of reef fish abundances and associated environmental variables at 40 reefs across much of the length of the Great Barrier Reef (see Figure 2.1 in Chapter 2). I quantify each response and explanatory variable on each reef and treat reefs as replicates in regression analyses at the regional scale. Specifically, to consider the potential influence of environmental factors, I use latitude and cross-shelf position as proxies because these two geographical variables covary strongly with important macroecological variables, such as sea surface temperature and coastal inputs of nutrients, and because fish community composition also varies strongly with both of these variables. The latitudinal gradient is measured in degrees from the equator. The cross-shelf position is measured as standardized distance from the location of the reef to the nearest edge of the continental shelf (i.e., 0 represents the shelf boundary and 1 represents the coast). Because of their importance in Chapter 2, I also use the long-term average (11-yr mean) and standard deviation (s.d.) in coral cover over the same time period, as proxies for the availability and variability of coral habitat.

4.2.2 Quantifying portfolio effect via synchrony and average population variability

Throughout this chapter, I use the coefficient of variation of total abundance in reef fish assemblages over time (hereafter, "community variability" or CV_C) as an inverse measure of "community stability")(Tilman 1999; Loreau 2010). Following a unified approach of Thibaut and Connolly (2013), I partition community variability (CV_C) into two components: (i) *synchrony* representing how correlated the fluctuations in abundance of different species are, and (ii) *average population variability,* which is the average coefficient of variation of species-level abundance, weighted by mean abundance. The mathematical relationship between the effects of synchrony and average population variability can be expressed as follows (Thibaut & Connolly 2013):

$$CV_C = \sqrt{\varphi} \cdot CV_S \qquad \text{(eq. 4.1)}$$

where φ and CV_S represents the synchrony index and the (weighted-) average population variability, respectively.

The synchrony index (φ) is defined as follows (Loreau & de Mazancourt 2008a):

$$\varphi = \frac{\sum_{ij} \sigma_{ij}^2}{\left(\sum_i \sigma_{ii}\right)^2} \quad \text{(eq. 4.2)}$$

where σ_{ii} and σ_{ij} represents the square-roots of the elements of the variance-covariance matrix for the temporal fluctuations in species-abundance observations [i.e., σ_{ii}^2 is the variance (over time) of the abundance of species i, and σ_{ij}^2 is the covariance of the fluctuations of species i and species j]. The numerator is the variance of total community abundance, and the denominator represents the maximum possible total variance of a community with the same constituent species variances (i.e., where species' fluctuations are perfectly positively correlated). As such, the synchrony index (φ) is standardized to range between zero (perfectly constant total community abundance) and one (perfectly positively correlated fluctuations of all species) (Hector *et al.* 2010).

Average population variability (CV_S) is defined as follows (Thibaut and Connolly 2013):

$$CV_S = \sum_i \left(\frac{\mu_i}{\sum_j \mu_j}\right)\left(\frac{\sigma_i}{\mu_i}\right) \quad \text{(eq. 4.3)}$$

where σ_i and μ_i represent the temporal standard deviation and mean abundance of species i, respectively. Eq. 4.3 represents abundance-weighted average coefficient of variation: the term in the first set of brackets inside the sum is the relative abundance of species i, and the term in the second set of brackets is species i's coefficient of variation.

4.2.3 Estimating species diversity (richness) and unevenness

I use time-averaged observed species richness to represent species diversity of reef fishes for each reef, to maintain consistency with previous studies of diversity-stability relationships (Campbell *et al.* 2011; Morin *et al.* 2014). Observed richness is highly correlated with maximum-likelihood estimates of true species richness obtained by fitting the Poisson-lognormal distribution to observed species abundances (R^2=0.95, P<0.001; Figure D1), a distribution that characterizes these data very well (Chapter 2). Consequently, variation in observed richness is highly likely to be representative of variation in true species richness for these data.

To quantify unevenness in relative species abundance, I use maximum-likelihood estimates of the Poisson-lognormal species-abundance distribution. Poisson-lognormal fits can be used to estimate unevenness through the variance parameter σ^2. Specifically, this parameter is equivalent to σ^2_{total} , i.e., the overall variance in log abundance among species defined in Chapter 2-3. This quantity has two important advantages over other evenness metrics: it is robust to variation in sampling (i.e., it estimates the variance of log-abundance in the underlying community from which the data are a sample), and it completely specifies unevenness independent of species richness (Sæther *et al.* 2013). That is, the skew of a lognormal distribution is entirely specified by the parameter σ^2, where larger σ^2 implies larger unevenness.

4.2.4 Estimating niche structure index (NSI) from stochastic community dynamics

Here, I propose the “niche structure index” (NSI) to quantify the relative importance of niche differences versus stochastic fluctuations in driving variation of fish community structure across coral reefs. This is simply the ratio of the variance components of relative species abundance presented in the previous chapters (Chapter 2-3). To recap, under the assumptions of stochastic Gompertz-type community dynamics, the (stationary) total variance in relative species abundance can be partitioned into the variance components explained by persistent species differences (what Engen et al. 2002 term “species heterogeneity”), environmental stochasticity, and an additional term representing residual sources of variance, such as demographic and sampling stochasticity (Engen *et al.* 2002):

$$\sigma^2_{total} = \frac{\sigma^2_r}{\delta^2} + \frac{\sigma^2_e}{2\delta} + \theta^2 = V_r + V_e + V_d \qquad \text{(eq. 4.4)}$$

where σ^2_{total} represents the overall variance in log abundance among species (i.e., the unevenness measure used in this chapter), and V_r, V_e and V_d represent the variance in species’ equilibrium abundances, the variance due to environmentally-induced fluctuations in species year-to-year intrinsic growth rates, and the variance due to other factors such as sampling error and demographic stochasticity, respectively. The other parameters in the middle part of eq. 4.4 are from mathematical analysis of stochastic Gompertz-type community dynamics theory (Engen *et al.* 2002), which are described in Chapters 2 and 3.

Using the three additive terms in (eq. 4.4), I define the niche structure index, NSI, as follows:

$$\mathrm{NSI} = \frac{V_r}{V_e + V_d} \qquad \text{(eq. 4.5)}$$

NSI is dimensionless and varies from zero (for completely stochasticity-structured assemblages, where all species have the same equilibrium abundance) to infinity (for completely niche-structured assemblages), where NSI = 1 indicates that the stochastic and deterministic components of community structure are equal in magnitude. That is, NSI measures the relative strength of deterministic species differences versus stochasticity in community turnover. I use NSI, rather than the absolute magnitude of the variance components (i.e., V_r, V_e and V_d), to ensure that this quantity is not statistically confounded with unevenness σ^2_{total} (and indeed it is not: the correlation coefficient between these two quantities, after log-transforming to meet assumptions of parametric correlation, is r=0.07, P=0.65). To estimate NSI, I use the variance component estimates obtained in the previous chapter from analysis of the temporal autocorrelation of relative species abundances (Chapter 2).

4.2.5 Testing pathways of portfolio effect and community stability

I use regression analyses to test for the stabilizing effects of species diversity (richness), unevenness, and NSI, via the pathways of synchrony and average population variability. I evaluate the sign and strength of the effect of species diversity, unevenness and NSI on the two pathways. Specifically, in my regression analyses, I first model community variability (CV_C) as a function of its two components: synchrony and (weighted-) average population variability. I do this on a log scale to linearize the expected relationships among the terms: that is, on a log-scale, eq. 1 is:

$$\log(\mathrm{CV_c}) = \frac{1}{2}\log(\phi) + \log(\mathrm{CV_s})$$

(Note that, the above relationship is a mathematical identity for a given community, but a regression of $\mathrm{CV_c}$ on ϕ and $\mathrm{CV_s}$ across communities is not a mathematical identity). Then I model ϕ and $\mathrm{CV_s}$ as functions of fish community structure variables (species richness, unevenness, and NSI) on log scale. The total effect of community structure variables on community variability, via pathways of synchrony and average population variability, is computed by multiplying the effect magnitudes along the pathway (i.e., path analysis). Additionally, I consider important environmental variables in Chapter 2 (i.e., latitude, cross-shelf position, mean coral cover, and coral cover variability), instead of the fish community structure variables, to determine which set of variables best explains synchrony and average population variability, and therefore their total effects on community stability.

Ordinary-least-squares (OLS) regressions and R^2 are used for selecting the best single-variable model for explaining synchrony and average population variability. Firstly, community structure variables (i.e., richness, evenness and NSI) and environmental variables (i.e., latitude, shelf position, mean coral cover, and volatility in coral cover) are separated initially into two model selection procedures. Secondly, for my four model selection procedures (i.e., community structure and environmental explanatory variable groups, each paired with the response variables synchrony and average population variability), I find the single-variable model that best explains the synchrony and average population variability in terms of R^2. All variables are log-transformed, and they are standardized to unit mean and standard deviation prior to the path analysis. Nonparametric bootstrapping is used to compute the standard errors of effect magnitudes.

Note that, the abovementioned procedure selects a best community structure model and a best environmental model (single variable) for both synchrony and weighted-average population variability. Then, to test among the community structure variables and environmental variables, I considered the best environmental and community structure explanatory variables together in a final model selection procedure using a “full” model including the best environmental and community structure variables, as well as their interactions, and nested models that omit particular terms. I used AICc (i.e., AIC with bias-correction for small sample sizes) as my model selection statistic. The objective of this analysis is to determine whether covarying environmental and community structure variables may be explaining the same variation, and, if so, whether one or more such variables are redundant and should not be included in a final model. Firstly, for synchrony as a response variable, I find the best community structure model out of the following: NSI, richness, unevenness, including possible pairwise interactions between NSI and richness or between NSI and unevenness. However, explanatory variables are not included in the same model, if the correlation between them is too high. I thus separated a pair of explanatory variables when Pearson correlation between two explanatory variables is larger than 0.7 (it turns out only the correlation between richness and evenness is larger than this threshold). Secondly, again for synchrony as a response variable, I find the best environmental model out of the following: latitude, cross-shelf position, mean coral cover and volatility in coral cover, and their respective interactions (unless any of the paired variables have correlation that is larger than 0.7 to include in the same model). Thirdly, for a combined model selection, I take the best variables from the community structure model selection and the environmental model selection. Then, I pooled them together (with pairwise interactions) for final model selection (ranked by AICc) to see whether community structure variables are retained in the model selection, or whether they are redundant with covarying environmental variables. Third-order and higher interactions are not considered due to the lack of degrees of freedom for such models. Finally, I repeat the same procedure as above for selecting best community structure, environmental, and combined models for explaining average population variability.

4.3 Results

On the Great Barrier Reef, community variability of reef fishes is driven more by synchrony than average population variability (Figure 4.1). The greater variation in synchrony index means that this quantity explains much more of the regional variation in community variability/stability (R^2=0.73, P<0.001; Figure 4.1A) than average population variability does (R^2=0.17, P=0.008; Figure 4.1B). There is no statistically significant correlation between the two quantities (Figure 4.1C).

Both species richness and evenness, considered separately, have weakly stabilizing effects by reducing synchrony (R^2=0.1, P=0.04; Figure 4.2A and R^2=0.13, P=0.02; Figure 4.2B), but they are not significantly related to average population variability (R^2=0, P=0.72;

Figure 4.2D and R^2=0.01, P=0.51; Figure 4.2E). Due to the high correlation between richness and evenness for these data (r = -0.77, P<0.001; Figure S1, also see Chapter 2), I do not consider the effects of these variables jointly; that is, I cannot causally distinguish effects of evenness from those of richness for these data. In contrast, the niche structure index (NSI) has a marginally non-significant relationship with synchrony (R^2=0.09, P=0.07; Figure 4.2C), but a much stronger relationship with average population variability: high NSI is associated with the stabilizing effect of low average population variability (R^2=0.4, P<0.001; Figure 4.2F).

Single-variable model selection produces distinct community structure variables and environmental variables for explaining synchrony index and average population variability among reefs. For community structure variables, the single-variable models show synchrony index is best explained (with highest R^2) by evenness (with the strongly covarying effect of richness providing only slightly poorer fit), while average population variability is best explained by NSI (Table 4.1). For environmental variables, the single-variable models show synchrony index is best explained by cross-shelf position, while average population variability is best explained by coral cover variability (Table 4.1).

Based on the results of single-variable models, path analysis reveals that the magnitude of the net stabilizing effect of niche structure index (NSI) is comparable to that of species diversity (richness) and evenness (Figure 4.3A): a one standard deviation change in each is associated with an approximately 0.3-0.35 standard deviation change in community variability (Figure 4.3C). However, a complementary path analysis reveals that the strength of the relationship between cross-shelf position and population synchrony is substantially stronger than that of either richness or evenness (Figure 4.3B), such that a one standard deviation change in environmental effect of cross-shelf position is associated with an approximately 0.5 standard deviation change in community variability (Figure 4.3C). In contrast, the effect of coral cover variability is associated with an approximately 0.25 standard deviation change in community variability (Figure 4.3C), slightly lower than that of NSI.

Model selection of final models ranked by AICc (reporting only those models with ΔAICc < 2), that include community structure variables (CMS; Table E1), environmental variables (ENV; Table E1) and pooled variables (CMS+ENV; Table E1) in both additive and interactive ways, supports potentially joint effects of community structure and environmental variables on average population variability, but not on synchrony index (Table E1). For explaining variation in synchrony index among reefs, the best model includes only cross-shelf position as an explanatory variable (Table E1). Specifically, synchrony index increases on reefs that are geographically closer to the coastline of the Great Barrier Reef (or farer from edges of the continental shelf) (Table E2). By contrast, for weighted-average population variability among reefs, the best model includes NSI, as well as both the mean and variability of coral cover, and an interaction of NSI and coral cover variability (Table E1). Specifically, weighted-average population variability decreases on reefs that have higher NSI in fish assemblages and higher mean coral cover, whereas it increases on reefs that have higher variability of coral cover. Consistently, the negative impact of NSI on average population variability is detected from both the best single-variable model and the best final model (Table E2). Interestingly, NSI and variability of coral cover interact to reduce the weighted-average population variability (Table E2), in which increasing NSI reduces weighted-average population variability more on reefs that have higher variability of coral cover (Figure E2).

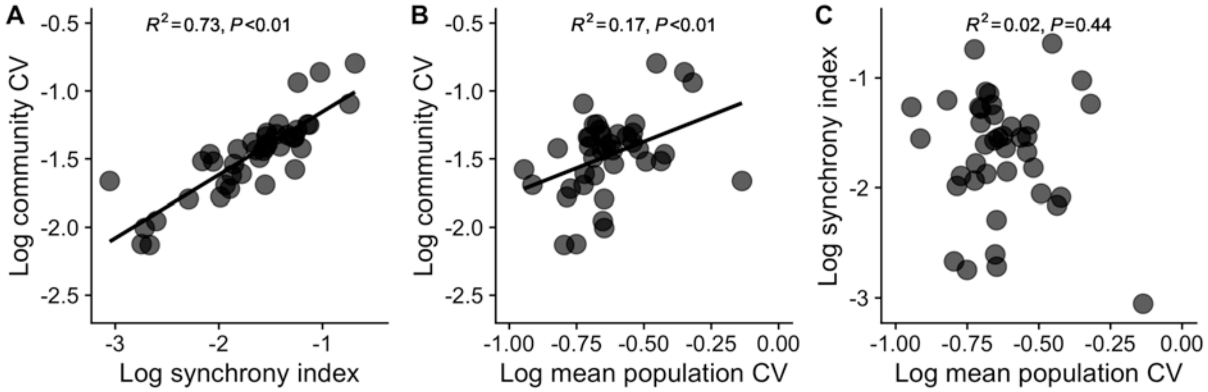

Figure 4.1 Community variability determined by synchrony and weighted-average population variability in reef fish assemblages across the Great Barrier Reef. Each point represents a reef (n=40). Lines indicate estimated relationships from ordinary-least-squares regression. (**A**) The relationship between community variability and synchrony index. (**B**) The relationship between community variability and weighted-average population variability.

(**C**) The relationship between two portfolio effect components: synchrony index and weighted-average population variability.

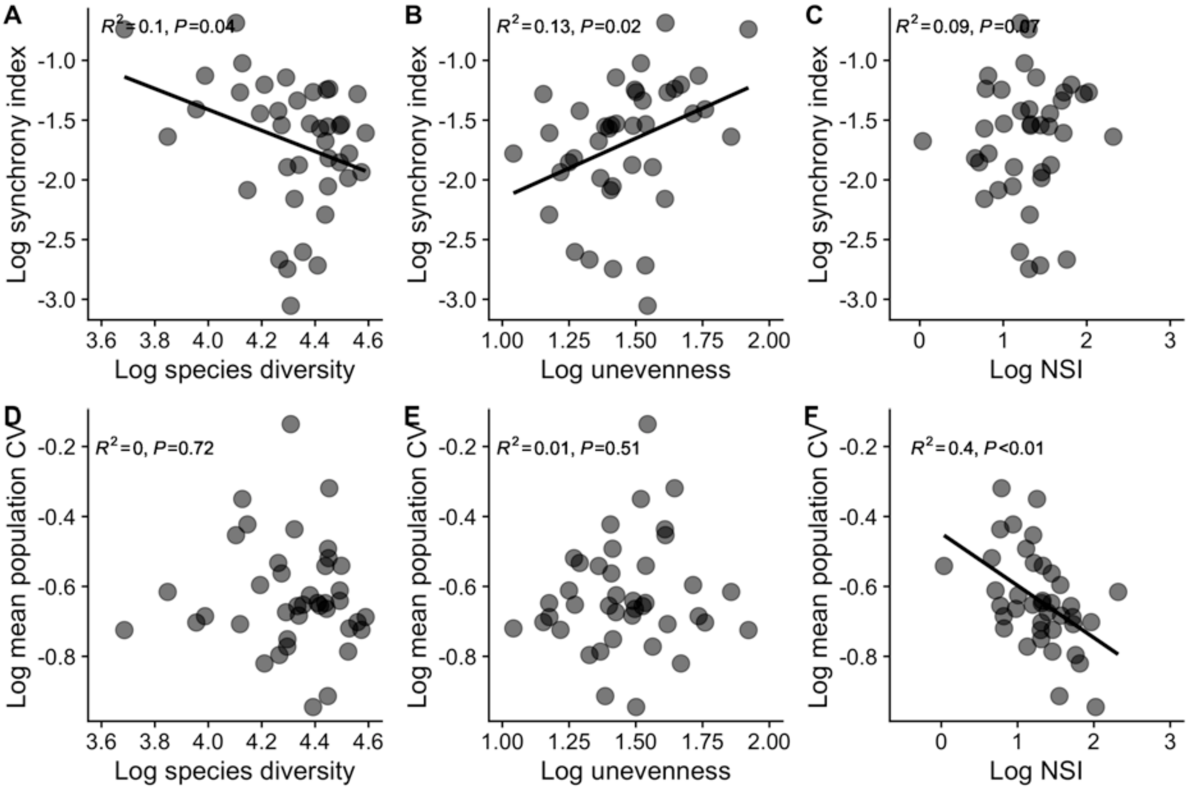

Figure 4.2 Stabilizing effects of species diversity (richness), evenness and niche structure index (NSI) on synchrony and weighted-average population variability. Each point represents a reef (n=40). Lines indicate estimated relationships from ordinary-least-squares regression, for only those relationships that met the conventional statistical threshold for significance (P<0.05). (**A, B, C**) Synchrony index of reef fish communities in relation to species diversity (richness), unevenness and niche structure index (NSI). (**D, E, F**) Weighted-average population variability of reef fish communities in relation to species diversity, unevenness and niche structure index (NSI).

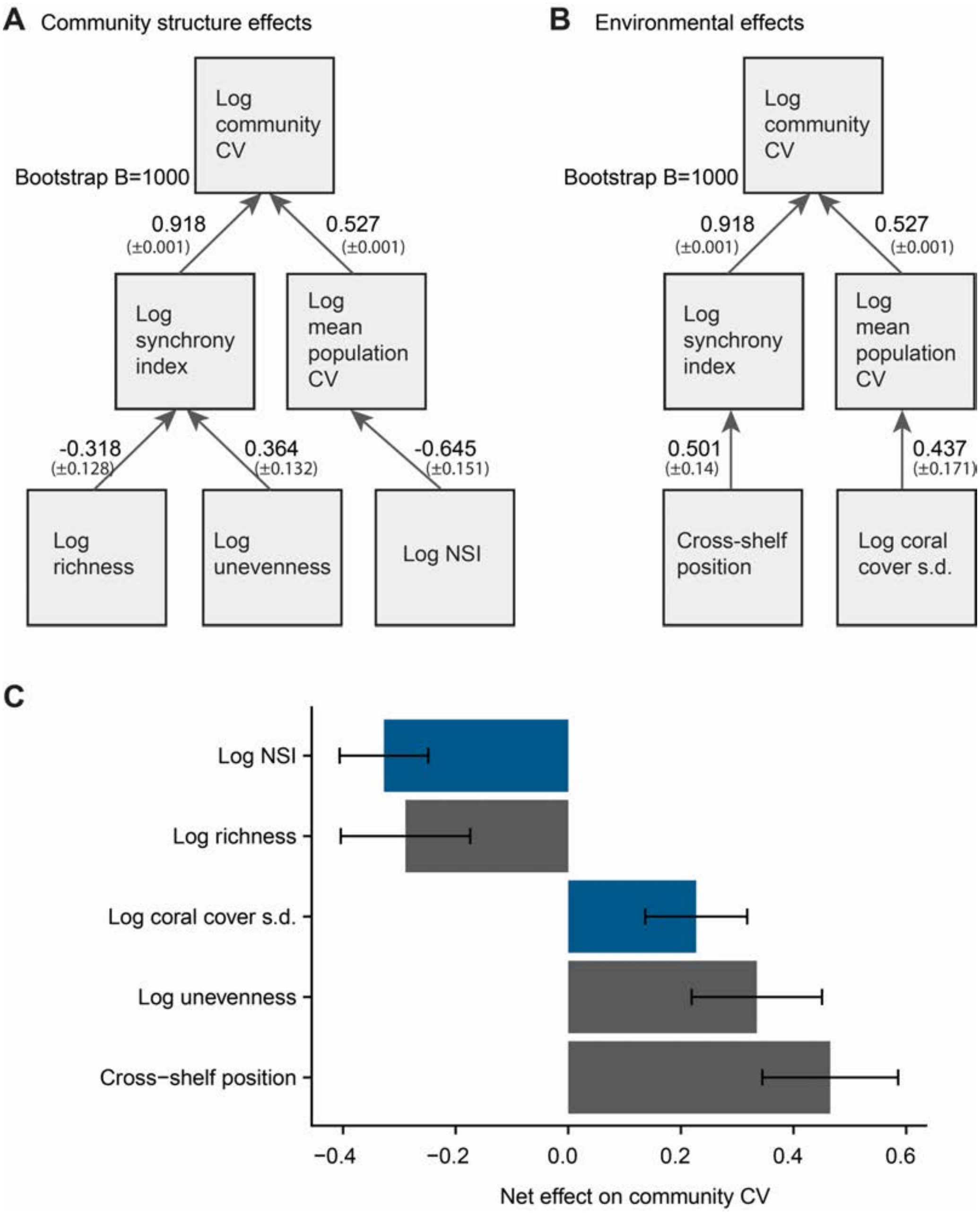

Figure 4.3 Community variability depends on community structure and environmental effects, via pathways of synchrony and weighted-average population variability. Standard errors (numbers in parentheses) and error bars of effect magnitudes are computed from bootstrapping (B=1000). (**A**) Path analysis of effects of fish community structure variables on synchrony and average population variability. This pathway is termed “community structure effect”. (**B**) Path analysis of biotic and abiotic environmental variables on synchrony and average population variability. This pathway is termed “environmental effect”. (**C**) Overall stabilizing effects, arising from species diversity, unevenness, niche structure index (NSI), coral cover variability and cross-shelf position, via pathways of synchrony and average population variability. The effect magnitude is computed by

multiplying the effect pathway of middle and bottom layers of panel (A) and (B), respectively. Blue represents the stabilizing effect via average population variability, and gray represents the stabilizing effect via synchrony.

Table 4.1 Results of single-variable model selection. For OLS regressions, terms to the right-hand side of the tilde (~) represent the explanatory variable, while the term on the left-hand side of the tilde represents the response variable. “sync” and “cvs” represent the synchrony index and weighted-average population variability, respectively. “richness”, “unevenness”, and “NSI” are community structure variables representing species richness, unevenness, and the niche structure index, respectively. “lat”, “shelf”, “mC”, and “sdC” are environmental variables representing latitude, cross-shelf position, mean coral cover, and coral cover variability, respectively.

Candidate variables	Single-variable OLS model	R^2	Selected
log(richness), log(uneveness), logNSI	log(sync) ~ log(richness)	0.1	no
	log(sync) ~ log(unevenness)	0.13	yes
	log(sync) ~ logNSI	0.09	no
	log(cvs) ~ log(richness)	0	no
	log(cvs) ~ log(unevenness)	0.01	no
	log(cvs) ~ logNSI	0.4	yes
lat, shelf, mC, log(sdC)	log(sync) ~ lat	0.01	no
	log(sync) ~ shelf	0.24	yes
	log(sync) ~ mC	0	no
	log(sync) ~ log(sdC)	0	no
	log(cvs) ~ lat	0.01	no
	log(cvs) ~ shelf	0	no
	log(cvs) ~ mC	0.01	no
	log(cvs) ~log(sdC)	0.19	yes

4.4 Discussion

To my knowledge, this study is the first documentation of multiple, distinct stabilizing effects, arising from both diversity and community structuring processes (i.e., deterministic species differences versus stochasticity), in a highly species-rich assemblage. My findings suggest that positive diversity-stability relationships exist at the regional scale in a tropical coral reef system. I find that the two key components of community stability, synchrony and weighted-average population variability, are statistically independent of one another geographically, and that geographical variation in synchrony explains 4-5 times more variation in community stability than weighted-average population variability (Figure 1). Despite that, the existence of a strong relationship between average population variability and the relative importance of deterministic species differences versus stochasticity (measured by NSI), combined with a much weaker relationship between synchrony and diversity, suggest that these two assemblage variables have comparably strong overall relationships with community stability (Figure 2 and Figure 3). Moreover, I find strong evidence suggesting that NSI, coral cover dynamics and their interaction, together, are important stabilising components of weighted-average population variability among reefs (Table E1, Figure E2); however, I find no evidence supporting an unique effect of richness or unevenness that can be teased apart from the environmental covarying effect of cross-shelf position (such as variation in species composition and nutrient loadings) on synchrony index among reefs (Table E1).

The diversity effect on community stability on the Great Barrier Reef is comparable in sign and magnitude to those of previous studies that examined such relationships along natural diversity gradients (c.f., Campbell et al. 2011; Houlahan *et al.* 2018). Interestingly, the principal stabilizing effect of diversity (i.e., via the synchrony pathway) is confounded with the effect of cross-shelf position on stability, whose statistical relationship with stability is approximately equal or stronger to that of diversity. Consequently, I cannot unambiguously disentangle these effects or partition their causal contributions to the regional variation in synchrony. My results, therefore, imply that environmental variables related to cross-shelf gradients, such as nutrient and pollution discharge from coastal human activities or inshore-offshore hydrodynamics, might be an important underlying driver of geographical variation in both diversity and stability on coral reefs, rather than being a direct causal relationship between diversity and stability (cf. Table 4.1 and Table E1). Indeed, covarying effects of environmental variables on both diversity and community stability are often observed in studies of natural systems (Shurin *et al.* 2007; Bulleri *et al.* 2012), and may overwhelm the diversity effect itself (Blüthgen *et al.* 2016).

The relative statistical weakness of the relationship between diversity and stability (Figure 4.3C and Table E1) discovered here is also typical of previous observational studies (Campbell et al. 2011; Houlahan *et al.* 2018), and contrasts with the stronger relationships often found in manipulated experiments (e.g., Tilman 1999; Campbell et al. 2011; Gross et al. 2014). Interestingly, the diversity-stability relationship occurs overwhelmingly via synchrony index, with no statistically detectable relationship between diversity and average population variability (Figure 4.2). Additionally, the diversity-synchrony relationship is largely independent of population variability in determining community stability across coral reefs (Figure 4.1C). These results agree well with the consensus view: a diversity-dependent synchrony effect can easily arise from species' compensatory dynamics through competitive interactions and differential responses to environmental fluctuations (Loreau & de Mazancourt 2008; Hector et al. 2010).

The lack of relationship between diversity and average population variability suggests an absence of diversity-stability mechanisms that would operate through this pathway, such as overyielding or dominance/unevenness (Thibaut & Connolly 2013; Gross *et al.* 2014; Turnbull *et al.* 2016). One possible explanation for the lack of a diversity-dependent effect on average population variability is that this effect is obscured by the larger effect of the niche structure index (NSI) on average population variability (Figure 4.2F). Such an effect would not appear in experimental diversity-stability relationship studies, because manipulated experiments deliberately initialize communities from identical species pools and with consistent relative abundance patterns (usually perfect evenness) (Tilman 1999; Gross *et al.* 2014; Winfree *et al.* 2015). However, it is possible that overyielding is particularly weak or negligible in coral reef fish assemblages. Overyielding might arise more commonly in systems with a small number of shared limiting resources, e.g., temperate grasslands (Marquard *et al.* 2009; Hector *et al.* 2010), whereas it might not occur in systems like coral reefs that are trophically and functionally diverse and the effect of the presence/absence of one species on other species is less consistent (Sale 1977; Dunne & Williams 2004; Casey *et al.* 2019). For example, a previous study in intertidal marine systems has shown dominant patterns of under-yielding and limited evidence for species complementarity (Emmerson *et al.* 2001).

In addition to the diversity-stability relationship discussed above, the relative importance of deterministic species differences (i.e., NSI derived from VPRSA as a proxy for niche structure) is associated with greater community stability. This effect is mediated by the relationship between NSI and reduced weighted-average population variability, rather than reduced synchrony. This relationship is stronger than that between weighted-average population variability and other environmental variables, such as coral cover variability, providing modest evidence of a more direct link between average population variability and deterministic niche structure than the environmental variables that are correlated with them both (Table 4.1). Unlike the relationship between diversity and synchrony, evidence for a relationship between NSI and population variability is retained, even when accounting for the effects of environmental variables, where NSI, mean coral cover, and variability in coral cover all contribute to explaining variations in average population variability among reefs (Table E1). In particular, the directions of these effects make ecological sense. For instance, weighted-average population variability is lower on reefs that have higher mean coral cover (Table E2), suggesting increased habitat availability stabilizes intraspecific fluctuations in fish abundance. By contrast, increased coral cover variability has a negative impact on weighted-average population variability as expected, because it indicates greater environmental stochasticity that I would expect to induce larger fluctuations in fish species' abundances. However, the interaction between NSI and coral cover variability implies that this negative effect is most pronounced when NSI is low and vanishes when NSI is high. In other words, increasing NSI reduces weighted-average population variability on reefs with high volatility in coral cover more strongly than it does on low-volatility reefs (Figure E2). This might imply that deterministic differences in species' demographic or life-history traits (or niche structure in a sense of traditional niche concept) that influence long-term mean abundance may disproportionally absorbs volatile fluctuations in species' intrinsic growth rates in extremely good or bad years.

Finally, my results suggest protecting diversity *per se* might be a good proxy for stabilising coral reef assemblages on the Great Barrier Reef, because such protection decreases synchrony in abundance fluctuations among species against environmental perturbations (Figure 4.2 and Figure 4.3). However, my results could also imply that conservation practitioners may find that preserving the highest level of species diversity (e.g., richness hotspots over seascape), a common target of conservation planning, may not protect the most stable assemblages of reef fishes depending on the covarying effect of community determinism (measured as NSI) and environmental variables, such as cross-shelf position and coral cover (Wang & Loreau 2016; Helmut *et al.* 2017). For instance, community determinism has a stabilizing effect that is stronger, or at least comparable in magnitude to those of richness or evenness, but realizations of these stabilizing effects might ultimately depend on the environmental context (Figure 4.3, Table E2, Figure E2). Specifically, on the world's largest coral reef system analysed here, the stabilizing effect of fish diversity occurs almost entirely via its effects on synchrony (Figure 4.3), and may in fact not be causal at all, but rather a coincidence of covarying cross-shelf patterns, e.g., in functional composition and human impacts (Table E1 and Table E2). Therefore, protecting distinct pathways of community stability (e.g., synchrony mechanisms vs. average population variability mechanisms), identifying the covarying environmental effects (e.g., nutrient discharge from coast, cross-shelf hydrodynamics, habitat availability and variability), and understanding different effects of community structure (e.g., richness, evenness, species differences in demographic traits) might provide a foundation for a more comprehensive approach to assist spatial prioritization than focusing exclusively on areas of highest richness. Given recently increased variability in assemblages and environmental conditions on the world's coral reefs (Hughes *et al.* 2018b, a), conservation strategies that simultaneously target the drivers of alpha-diversity (e.g., richness and evenness) and drivers of temporal beta-diversity (e.g., community turnover in relative species abundance and composition) could better mitigate the volatility in diversity-dependent ecosystem goods and services.

Chapter 5: General discussion

Overall, my thesis uses VPRSA to quantify proximate and ultimate drivers of relative abundance of reef fishes, as well as their impacts on regional variation in assemblage-level stability of total community abundance on the Great Barrier Reef. Specifically, I found:

1. Deterministic niche-related differences among species drive much more of the patterns of commonness and rarity in reef fish assemblages, compared to stochastic environmental fluctuations and other sources of variation, such as neutral processes.
2. The contributions of the above community structure components (i.e., persistent differences among species, environmental fluctuations, neutral process) vary at the regional scale, and such variation can be explained by in-situ coral cover dynamics, but not by classical environmental gradients such as latitude and onshore-offshore position.
3. The conclusions mentioned above (1 and 2) hold within broadly-defined trophic/functional groups, as well as for the reef fish assemblage as a whole. That is, the patterns for the fish fauna as a whole are not driven by changes in the functional composition of the fish assemblage across environmental gradients (at least not exclusively), but rather by shifts occurring within multiple constituent functional groups.
4. VPRSA is generally robust to violation of several key underlying assumptions of the theory on which it is based, namely the assumptions of homogeneous intraspecific interactions, negligible inter-specific interactions, and homogeneous responses to environmental fluctuations.
5. There is a positive relationship between diversity and stability of reef fish assemblages on the Great Barrier Reef. The stabilizing effect of diversity operates through reducing synchrony of fluctuations in abundance among species.
6. The relative importance of species differences versus environmental stochasticity influences community stability in fish assemblages. Specifically, the relative importance of persistent differences among species, relative to that of stochastic environmental and neutral fluctuations, (i.e., niche structure

index, NSI) stabilizes communities by reducing weighted-average population variability.

7. The stabilizing effects of NSI and species richness have comparable magnitudes, although they occur via different pathways. However, the extent to which these relationships represent causal links, versus correlated responses to common environmental drivers, is not clear. My analysis suggests that NSI and coral cover dynamics act in concert to influence average population variability, whereas the effect of richness (or evenness) is too confounded with cross-shelf variation in other community properties or environmental conditions to confidently attribute a causal relationship between diversity and stability.

Conclusion 1 indicates that patterns of commonness and rarity in coral reef fish assemblages are driven by persistent differences between species in long-term average abundances (i.e., niche-related traits that are associated with long-term abundance), to a substantially greater extent than they are driven by stochastic fluctuations in the relative abundances of different species (Chapter 2). This study builds on previous work, which showed that neutral theory was not a sufficient theory for coral reef systems (Dornelas *et al.* 2006; Connolly *et al.* 2009, 2014), and suggests that the mechanisms included in neutral theory in fact play a very minor role in determining relative abundance patterns on coral reefs, relative to the species differences that neutral theory specifically excludes. However, the contribution of differential responses to environmental fluctuations is substantial (around 20-25% of the total variance in species log-abundances), suggesting considerable scope for fluctuation-mediated mechanisms (such as the storage effect) in species coexistence (Chesson 2000). Because VPRSA is based solely on abundance data, whereas tests of modern coexistence theory (MCT) require information about which demographic rates are density-dependent, and what the functional form of that density-dependence is (Ellner *et al.* 2016), VPRSA does not allow a direct assessment of the relative importance of fluctuation-independent versus fluctuation-dependent mechanisms of species coexistence *per se*.

In previous studies on tropical coral reefs, debates over niche versus neutral/quasi-neutral community dynamics have focused on analysing static frequency distributions of relative species abundance (i.e., snapshots of relative abundances in time). Invariably, such approaches focus on whether neutral models are sufficient to explain commonness and rarity on coral reefs. VPRSA, in contrast, allowed us to understand how persistent relative abundance relationships are in time, and thereby take the additional step of estimating the relative importance of deterministic species differences versus stochastic fluctuations (both neutral and non-neutral) as drivers of relative abundance, and thus makes possible an assessment of whether and how those components vary in relative importance along ecological gradients. By exploiting this potential, I showed that the relative importance of niche-related species differences versus stochastic environmental and demographic fluctuations varies with coral cover volatility (Conclusion 2). To my knowledge, this is the first such application of VPRSA in the literature, for any system, and its success suggests that this approach may yield insights in other systems where community level time series exist along environmental gradients, e.g., the bioTime database (Dornelas *et al.* 2014, 2018; Blowes *et al.* 2019).

Additionally, the functional composition of reef fishes is known to change along geographical gradients, such as latitude and cross-shelf position (Hoey & Bellwood 2008; Emslie *et al.* 2010, 2012; Bennett 2011; Cheal *et al.* 2012; Hoey *et al.* 2013; Goatley *et al.* 2016). Specifically, the geographical variation in composition and abundance of a given functional group (e.g., herbivores; Cheal *et al.* 2012) or clade (e.g., Chaetodontidae; Emslie *et al.* 2010) has been the principal focus of previous studies. Conclusion 3 suggests that important geographical variation in community structure is not confined to compositional changes in the relative abundances of different functional groups. Rather, the extent to which relative abundance patterns persist over time is explained by within-functional-group dynamics, which exhibit similar relationships with the explanatory variables considered in my analyses. Therefore, it might be interesting to further investigate why effect of environmental variables on the relative importance of deterministic versus stochastic community processes is so consistent across such diverse functional groups.

Conclusion 3 also highlights potentially fruitful areas for further work. Firstly, functional groups identified in this thesis (Table A3) are more or less subject to experts' classification reflecting their experiences. It will be important to determine the extent to which the consistency across functional groups that I found in this thesis may be sensitive to how species are categorized functionally or trophically. Secondly, functional diversity metrics are not analysed in this thesis, but it would be interesting to investigate whether functional diversity metrics (e.g., Bellwood *et al.* 2003; Villéger *et al.* 2008, 2010; Stuart-Smith *et al.* 2013; D'agata *et al.* 2014; Mouillot *et al.* 2014; Brandl & Bellwood 2016) can be used along with Engen *et al.*'s VPRSA approach to assess the relative importance of deterministic vs. stochastic community dynamics. For example, it might be expected that sites with greater diversity in resource acquisition traits might also possess more deterministic structure in relative species abundance, whereas diversity in environmental response traits might be associated with more stochastic-driven communities.

Conclusion 4 suggests that Conclusions 1-3 are robust to at least some important ways in which real assemblage dynamics might differ from those assumed in the theory from which the VPRSA approach is derived (Chapter 3). Additionally, it suggests that VPRSA may be more broadly applicable than one might have assumed, based on the restrictive nature of some of those initial assumptions. On the other hand, my investigation of robustness is not exhaustive (and indeed, an exhaustive investigation of assumption violations would be impossible). For example, I have not investigated the extent to which modularized interaction networks, where species have a small number of strong interactions and many interactions that are zero, or nearly so, as I noted in my Chapter 3 Discussion. Similarly, I have not explicitly investigated departures from the Gompertz form of density-dependent interactions. Thus, there is scope for further work exploring the robustness of the VPRSA approach.

More generally, the VPRSA approach does not reveal which species traits are responsible for the abundance variation observed in the community. Consequently, there is considerable potential for VPRSA to complement and enrich the insights obtained from other approaches. For example, phylogenetic and functional approaches seek to explain variation in abundances as a function of phylogenetic or functional similarity. "Null model" versions of this approach (Webb et al. 2002; Kraft et al. 2007; Cavender-Bares et al. 2009; Freilich & Connolly 2015; Gerhold et al. 2015) cannot be used to apportion variation in the relative contribution of niche vs. stochastic factors, because they only test for significant departure from abundances randomized with respect to phylogeny or functional traits. This approach is in contrast to classical "ordination" or "eigen-decomposition" approaches (Legendre et al. 2009; Dumbrell et al. 2010; Smith & Lundholm 2010; Caruso et al. 2011; Tuomisto et al. 2012; Yeh et al. 2015; Ford & Roberts 2018) that explicitly model covariance in abundance as a function of similarity in environmental variables. However, in both cases, the effects of unmeasured or phylogenetically unstructured species traits that influence abundance would be missed, and the niche or stochastic contribution to abundance variation would be correspondingly underestimated. By complementing such analyses with a VPRSA estimate of the overall variation due to persistent species differences, the extent to which phylogeny or measured functional traits explain those persistent aspects of community structure could be potentially quantified. Indeed, where time series are available, this opens up the possibility of estimating the contribution of phylogeny or functional traits to both long-term variation in abundance, and to the patterns of interannual fluctuations in abundance.

Conclusions 4-6 provide a framework for simultaneously exploring the effects of diversity and the relative importance of deterministic vs. stochastic community assembly process on community-level stability. The framework also represents a novel example, derived from species-rich coral reef systems, showing how the degree of determinism in relative species abundance patterns might be linked to diversity-stability relationships (see Turnbull et al. 2016 for a review). Moreover, it empirically extends the previous theoretical model of Thibaut and Connolly 2013 by showing that the variance (dynamical) components of relative species abundance are associated with weighted-average population variability component of community stability. Importantly, conclusions 4-6 seem to suggest that average population variability is more strongly related to community determinism (measured as NSI) than to richness (Tilman 1999; Hector *et al.* 2010; Gross *et al.* 2014) or evenness (Sasaki & Lauenroth 2011; Thibaut & Connolly 2013). This is especially true when comparing the covarying effect of NSI and coral cover variables, showing a portion of variation in community stability is explained uniquely by NSI and its interaction with environmental conditions (Table E1, Table E2, and Figure E2; Chapter 4). Intriguingly, such a strong effect of community determinism is independent of effect of richness or evenness, suggesting separate stabilising pathways that have not been previously identified (Thibaut & Connolly 2013; Gross *et al.* 2014; Tilman *et al.* 2014). Amid more frequent bleaching episodes (Hughes et al. 2018a, b), changes in coral reef assemblages could have profound ecosystem consequences: Conclusions 4-6 imply a potential destabilisation of coral reef fish communities whenever diversity declines (which would be expected to increase community-level synchrony in abundance fluctuations) and/or niche structure erodes (via increased weighted-average population variability).

Examination of diversity-stability relationships along natural diversity gradients have the advantage of addressing these relationships in real systems, at the spatial scales relevant to management and conservation (Srivastava & Vellend 2005). However, they inherit the limitations that characterize such "natural experiments" (Wardle 2016). Most importantly, all the detected effects involving diversity measures and deterministic vs. stochastic components of community structure are correlational. Consequently, environmental factors that covary along these gradients might be the true drivers of the covariation between diversity, community assembly process and community stability. Such causal relationships cannot be teased apart unambiguously using the statistical analyses employed in this thesis, because the differences in the explanatory power (R^2) of environmental variables vs. assemblage properties is comparable (Table 4.1; Chapter 4). Particularly, I find that the detected stabilising effect of richness (or evenness) might be an artefact of its covariation with cross-shelf position (Table E1; Chapter 4). By contrast, effects of NSI and environmental variables that are related to coral cover dynamics are both important, and the two effects interact to determine population and community stability (Table 4.1, Table E1, Figure E2; Chapter 4). These findings suggest that regional variation in community stability in this system may depend more on variation in NSI than on variation in richness or evenness. Nevertheless, a positive diversity-stability relationship is observable for fish communities on the Great Barrier Reef, even though the causal nature of such a relationship is complicated by other community properties and environmental factors that vary systematically with position on the continental shelf. Meanwhile, a clear NSI-stability relationship among reefs is strongly supported by my analyses, even after accounting for environmental influences.

Finally, throughout my thesis, several core topics in community ecology are connected by a common theme: variance partitioning of relative species abundance (or species abundance distribution, SAD). My work extends VPRSA theory in several important ways. Firstly, by employing it on an extensive network of time series at a regional scale, I show that there is statistically detectable variation in the importance of deterministic and stochastic components of community structure, and moreover that this variation is potentially explainable in terms of regional variation in environmental conditions (Chapter 2). Secondly, I show that VPRSA is surprisingly robust to violation of several important simplifying assumptions of the theory, particularly when the variance component due to deterministic species differences is interpreted more broadly as the component due to variance in species' long-term average abundances, rather than specifically to variance in intrinsic growth rates (Chapter 3). Lastly, the relative stability in species' abundances implied by a high niche structure index (NSI) correlates strongly with the "weighted average population variability" of Thibaut and Connolly (2013), and thus is conceptually linked to diversity-stability theory (Chapter 4). On the Great Barrier Reef, these extensions show that reef fish assemblages are highly niche structured, that this niche structure is eroded on reefs with high volatility in coral cover, and that the relative magnitudes of deterministic versus stochastic components of community structure, via their close relationship with weighted average population variability, explain regional variation in community stability to approximately the same extent as species richness or evenness. These findings highlight the potential of VPRSA to greatly expand the toolkit of community ecology and macroecology, particularly in high-diversity assemblages, where species-level estimation of community dynamics parameters is not feasible.

References

Anderson, G.R. V, Ehrlich, A.H., Ehrlich, P.R., Roughgarden, J.D., Russell, B.C. & Talbot, F.H. (1981). The community structure of coral reef fishes. *Am. Nat.*, 117, 476–495.

Bellwood, D.R., Hoey, A.S. & Choat, J.H. (2003). Limited functional redundancy in high diversity systems: resilience and ecosystem function on coral reefs. *Ecol. Lett.*, 6, 281–285.

Bellwood, D.R. & Hughes, T.P. (2001). Regional-Scale Assembly Rules and Biodiversity of Coral Reefs. *Science (80-.).*, 292, 1532 LP – 1535.

Bennett, S. (2011). Latitudinal variation in macroalgal consumption by fishes on the Great Barrier Reef . *Mar. Ecol. Prog. Ser.*, 426, 241–252.

Berlow, E.L. (1999). Strong effects of weak interactions in ecological communities. *Nature*, 398, 330–334.

Blowes, S.A., Supp, S.R., Antão, L.H., Bates, A., Bruelheide, H., Chase, J.M., *et al.* (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science (80-.).*, 366, 339 LP – 345.

Blüthgen, N., Simons, N.K., Jung, K., Prati, D., Renner, S.C., Boch, S., *et al.* (2016). Land use imperils plant and animal community stability through changes in asynchrony rather than diversity. *Nat. Commun.*, 7, 10697.

Bode, M., Bode, L. & Armsworth, P.R. (2011). Different dispersal abilities allow reef fish to coexist. *Proc. Natl. Acad. Sci.*, 108, 16317 LP – 16321.

Bode, M., Connolly, S.R. & Pandolfi, J.M. (2012). Species differences drive nonneutral structure in pleistocene coral communities. *Am. Nat.*, 180, 577–88.

Boettiger, C. (2018). From noise to knowledge: how randomness generates novel phenomena and reveals information. *Ecol. Lett.*, 21, 1255–1267.

Brandl, S.J. & Bellwood, D.R. (2016). Microtopographic refuges shape consumer-producer dynamics by mediating consumer functional diversity. *Oecologia*, 182, 203–217.

Bulleri, F., Cusson, M., Maggi, E., Arenas, F., Aspden, R., Bertocci, I., *et al.* (2012). Temporal stability of European rocky shore assemblages : variation across a latitudinal gradient and the role of habitat-formers, 1801–1809.

Campbell, V., Murphy, G. & Romanuk, T.N. (2011). Experimental design and the outcome and interpretation of diversity–stability relations. *Oikos*, 120, 399–408.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., *et al.*

(2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59.

Caruso, T., Chan, Y., Lacap, D.C., Lau, M.C.Y., McKay, C.P. & Pointing, S.B. (2011). Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *ISME J.*, 5, 1406–1413.

Casey, J.M., Meyer, C.P., Morat, F., Brandl, S.J., Planes, S. & Parravicini, V. (2019). Reconstructing hyperdiverse food webs: Gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods Ecol. Evol.*, 10, 1157–1170.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecol. Lett.*, 12, 693–715.

Cheal, A., Emslie, M., Miller, I. & Sweatman, H. (2012). The distribution of herbivorous fishes on the Great Barrier Reef. *Mar. Biol.*, 159, 1143–1154.

Cheal, A.J., MacNeil, M.A., Emslie, M.J. & Sweatman, H. (2017). The threat to coral reefs from more intense cyclones under climate change. *Glob. Chang. Biol.*, 23, 1511–1524.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.*, 31, 343–366.

Chesson, P.L. & Warner, R.R. (1981). Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.*, 117, 923–943.

Chisholm, R.A. & Pacala, S.W. (2010). Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proc. Natl. Acad. Sci.*, 107, 15821–15825.

Coker, D.J., Wilson, S.K. & Pratchett, M.S. (2014). Importance of live coral habitat for reef fishes. *Rev. Fish Biol. Fish.*, 24, 89–126.

Connell, J.H. (1978). Diversity in tropical rain forests and coral reefs. *Science (80-.).*, 199, 1302–1310.

Connolly, S.R., Hughes, T.P. & Bellwood, D.R. (2017). A unified model explains commonness and rarity on coral reefs. *Ecol. Lett.*, 20, 477–486.

Connolly, S.R., Hughes, T.P., Bellwood, D.R. & Karlson, R.H. (2005). Community Structure of Corals and Reef Fishes at Multiple Scales. *Science (80-.).*, 309, 1363–1365.

Connolly, S.R., MacNeil, M.A., Caley, M.J., Knowlton, N., Cripps, E., Hisano, M., *et al.* (2014). Commonness and rarity in the marine biosphere. *Proc. Natl. Acad. Sci. U. S. A.*, 111, 8524–8529.

D'agata, S., Mouillot, D., Kulbicki, M., Andréfouët, S., Bellwood, D.R., Cinner, J.E., *et al.* (2014). Human-mediated loss of phylogenetic and functional diversity in coral reef fishes. *Curr. Biol.*, 24, 555–560.

De’ath, G., Fabricius, K.E., Sweatman, H. & Puotinen, M. (2012). The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 17995–17999.

Dennis, B. & Ponciano, J.M. (2014). Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95, 2069–2076.

Devlin, M.J. & Brodie, J. (2005). Terrestrial discharge into the Great Barrier Reef Lagoon: nutrient behavior in coastal waters. *Mar. Pollut. Bull.*, 51, 9–22.

Doak, D.F., Bigger, D., Harding, E.K., Marvier, M.A., O’Malley, R.E. & Thomson, D. (1998). The statistical inevitability of stability-diversity relationships in community ecology. *Am. Nat.*, 151, 264–276.

Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D., *et al.* (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Glob. Ecol. Biogeogr.*, 27, 760–786.

Dornelas, M., Connolly, S.R. & Hughes, T.P. (2006). Coral reef diversity refutes the neutral theory of biodiversity. *Nature*, 440, 80–82.

Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., *et al.* (2014). Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344, 296–299.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A.H. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J.*, 4, 337–345.

Dunne, J.A. & Williams, R.J. (2004). Network structure and robustness of marine food webs . *Mar. Ecol. Prog. Ser.*, 273, 291–302.

Ellner, S.P., Snyder, R.E. & Adler, P.B. (2016). How to quantify the temporal storage effect using simulations instead of math. *Ecol. Lett.*, 19, 1333–1342.

Emmerson, M.C., Solan, M., Emes, C., Paterson, D.M. & Raffaelli, D. (2001). Consistent patterns and the idiosyncratic effects of biodiversity in marine ecosystems. *Nature*, 411, 73–77.

Emslie, M.J., Logan, M., Ceccarelli, D.M., Cheal, A.J., Hoey, A.S., Miller, I., *et al.* (2012). Regional-scale variation in the distribution and abundance of farming damselfishes on Australia’s Great Barrier Reef. *Mar. Biol.*, 159, 1293–1304.

Emslie, M.J., Pratchett, M.S., Cheal, A.J. & Osborne, K. (2010). Great Barrier Reef butterflyfish community structure: the role of shelf position and benthic community type. *Coral Reefs*, 29, 705–715.

Engen, S. & Lande, R. (1996). Population dynamic models generating the lognormal species abundance distribution. *Math. Biosci.*, 132, 169–183.

Engen, S., Lande, R., Walla, T. & DeVries, P.J. (2002). Analyzing spatial structure of communities using the two-dimensional Poisson lognormal species abundance model. *Am. Nat.*, 160, 60–73.

Engen, S., Solbu, E.B. & Sæther, B. (2017). Neutral or non-neutral communities: temporal dynamics provide the answer. *Oikos*, 126, 318–331.

Fabricius, K.E., De'ath, G., Puotinen, M.L., Done, T., Cooper, T.F. & Burgess, S.C. (2008). Disturbance gradients on inshore and offshore coral reefs caused by a severe tropical cyclone. *Limnol. Oceanogr.*, 53, 690–704.

Fabricius, K.E., Logan, M., Weeks, S.J., Lewis, S.E. & Brodie, J. (2016). Changes in water clarity in response to river discharges on the Great Barrier Reef continental shelf: 2002–2013. *Estuar. Coast. Shelf Sci.*, 173, A1–A15.

Ford, B.M. & Roberts, J.D. (2018). Latitudinal gradients of dispersal and niche processes mediating neutral assembly of marine fish communities. *Mar. Biol.*, 165, 94.

Freilich, M.A. & Connolly, S.R. (2015). Phylogenetic community structure when competition and environmental filtering determine abundances. *Glob. Ecol. Biogeogr.*, 24, 1390–1400.

Fung, T., O'Dwyer, J.P., Rahman, K.A., Fletcher, C.D. & Chisholm, R.A. (2016). Reproducing static and dynamic biodiversity patterns in tropical forests: the critical role of environmental variance. *Ecology*, 97, 1207–1217.

Gerhold, P., Cahill Jr, J.F., Winter, M., Bartish, I. V & Prinzing, A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better). *Funct. Ecol.*, 29, 600–614.

Goatley, C., González-Cabello, A. & Bellwood, D.R. (2016). Reef-scale partitioning of cryptobenthic fish assemblages across the Great Barrier Reef, Australia. *Mar. Ecol. Prog. Ser.*, 544, 271–280.

Gonzalez, A., Cardinale, B.J., H., A.G.R., Jarrett, B., K., A.E., G., B.D., *et al.* (2016). Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. *Ecology*, 97, 1949–1960.

Gross, K., Cardinale, B.J., Fox, J.W., Gonzalez, A., Loreau, M., Wayne Polley, H., *et al.* (2014). Species richness and the temporal stability of biomass production: a new analysis of recent biodiversity experiments. *Am. Nat.*, 183, 1–12.

Harriott, V. & Banks, S. (2002). Latitudinal variation in coral communities in eastern

Australia: a qualitative biophysical model of factors regulating coral reefs. *Coral Reefs*, 21, 83–94.

Hector, A., Hautier, Y., Saner, P., Wacker, L., Bagchi, R., Joshi, J., *et al.* (2010). General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology*, 91, 2213–2220.

Helmut, H., Bernd, B., T., B.E., M., C.J., A., D.J., Klemens, E.B., *et al.* (2017). Biodiversity change is uncoupled from species richness trends: consequences for conservation and monitoring. *J. Appl. Ecol.*, 55, 169–184.

Hoey, A.S. & Bellwood, D.R. (2008). Cross-shelf variation in the role of parrotfishes on the Great Barrier Reef. *Coral Reefs*, 27, 37–47.

Hoey, A.S., Brandl, S.J. & Bellwood, D.R. (2013). Diet and cross-shelf distribution of rabbitfishes (f. Siganidae) on the northern Great Barrier Reef: implications for ecosystem function. *Coral Reefs*, 32, 973–984.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., *et al.* (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol. Monogr.*, 75, 3–35.

Houlahan, J.E., Currie, D.J., Cottenie, K., Cumming, G.S., Findlay, C.S., Fuhlendorf, S.D., *et al.* (2018). Negative relationships between species richness and temporal variability are common but weak in natural systems. *Ecology*, 99, 2592–2604.

Hubbell, S.P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral reefs*, 16, S9–S21.

Hubbell, S.P. (2001). *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Princeton University Press.

Hughes, J.B. & Roughgarden, J. (2000). Species diversity and biomass stability. *Am. Nat.*, 155, 618–627.

Hughes, T.P., Anderson, K.D., Connolly, S.R., Heron, S.F., Kerry, J.T., Lough, J.M., *et al.* (2018a). Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. *Science*, 359, 80–83.

Hughes, T.P., Kerry, J.T., Baird, A.H., Connolly, S.R., Chase, T.J., Dietzel, A., *et al.* (2019). Global warming impairs stock–recruitment dynamics of corals. *Nature*.

Hughes, T.P., Kerry, J.T., Baird, A.H., Connolly, S.R., Dietzel, A., Eakin, C.M., *et al.* (2018b). Global warming transforms coral reef assemblages. *Nature*, 556, 492–496.

Ives, A.R., Dennis, B., Cottingham, K.L. & Carpenter, S.R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecol. Monogr.*, 73, 301–330.

Jiang, L. & Pu, Z. (2009). Different effects of species diversity on temporal stability in single-trophic and multitrophic communities. *Am. Nat.*, 174, 651–659.

Jones, G.P., McCormick, M.I., Srinivasan, M. & Eagle, J. V. (2004). Coral decline threatens fish biodiversity in marine reserves. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 8251–8253.

Kalyuzhny, M., Kadmon, R. & Shnerb, N.M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecol. Lett.*, 18, 572–580.

Karlson, R.H. & Hurd, L.E. (1993). Disturbance, coral reef communities, and changing ecological paradigms. *Coral reefs*, 12, 117–125.

Kraft, N.J.B., Cornwell, W.K., Webb, C.O. & Ackerly, D.D. (2007). Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities. *Am. Nat.*, 170, 271–283.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H. & Bell, B.M. (2016). TMB: Automatic Differentiation and Laplace Approximation. *J. Stat. Software; Vol 1, Issue 5* .

Lande, R., Engen, S. & Saether, B.-E. (2003). *Stochastic population dynamics in ecology and conservation*. Oxford University Press on Demand.

Legendre, P., Mi, X., Ren, H., Ma, K., Yu, M., Sun, I.-F., *et al.* (2009). Partitioning beta diversity in a subtropical broad-leaved forest of China. *Ecology*, 90, 663–674.

Loreau, M. (2010). *From populations to ecosystems: Theoretical foundations for a new ecological synthesis (MPB-46)*. Princeton University Press.

Loreau, M. & de Mazancourt, C. (2008a). Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. *Am. Nat.*, 172, E48–E66.

Loreau, M. & de Mazancourt, C. (2008b). Species Synchrony and Its Drivers: Neutral and Nonneutral Community Dynamics in Fluctuating Environments. *Am. Nat.*, 172, E48–E66.

Loreau, M. & de Mazancourt, C. (2013). Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.*, 16, 106–115.

Ma, Z., Liu, H., Mi, Z., Zhang, Z., Wang, Y., Xu, W., *et al.* (2017). Climate warming reduces the temporal stability of plant community biomass production. *Nat. Commun.*, 8, 15378.

Magurran, A.E. & Henderson, P.A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422, 714.

Malcolm, H.A., Jordan, A. & Smith, S.D.A. (2010a). Biogeographical and cross-shelf patterns of reef fish assemblages in a transition zone. *Mar. Biodivers.*, 40, 181–193.

Malcolm, H.A., Smith, S.D.A. & Jordan, A. (2010b). Using patterns of reef fish assemblages to refine a Habitat Classification System for marine parks in NSW, Australia. *Aquat. Conserv. Mar. Freshw. Ecosyst.*, 20, 83–92.

Marquard, E., Weigelt, A., Temperton, V.M., Roscher, C., Schumacher, J., Buchmann, N., *et al.* (2009). Plant species richness and functional composition drive overyielding in a six-year grassland experiment. *Ecology*, 90, 3290–3302.

Matthews, T.J. & Whittaker, R.J. (2015). On the species abundance distribution in applied ecology and biodiversity management. *J. Appl. Ecol.*, 52, 443–454.

McCann, K., Hastings, A. & Huxel, G.R. (1998). Weak trophic interactions and the balance of nature. *Nature*, 395, 794–798.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., *et al.* (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.*, 10, 995–1015.

McNaughton, S.J. (1977). Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. *Am. Nat.*, 111, 515–525.

Mellin, C., Bradshaw, C.J.A., Meekan, M.G. & Caley, M.J. (2010). Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Glob. Ecol. Biogeogr.*, 19, 212–222.

Mellin, C., Thompson, A., Jonker, M.J. & Emslie, M.J. (2019). Cross-shelf variation in coral community response to disturbance on the Great Barrier Reef. *Diversity*, 11, 38.

Morin, X., Fahse, L., Mazancourt, C., Scherer-Lorenzen, M. & Bugmann, H. (2014). Temporal stability in forest productivity increases with tree diversity due to asynchrony in species dynamics. *Ecol. Lett.*, 17, 1526–1535.

Mouillot, D., Villéger, S., Parravicini, V., Kulbicki, M., Arias-González, J.E., Bender, M., *et al.* (2014). Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl. Acad. Sci.*, 111, 13757–13762.

Myers, R.A., Bowen, K.G. & Barrowman, N.J. (1999). Maximum reproductive rate of fish at low population sizes. *Can. J. Fish. Aquat. Sci.*, 56, 2404–2419.

Ovaskainen, O., Tikhonov, G., Dunson, D., Grøtan, V., Engen, S., Sæther, B.-E., *et al.* (2017). How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proc. R. Soc. B*, 284, 20170768.

Pandolfi, J.M., Connolly, S.R., Marshall, D.J. & Cohen, A.L. (2011). Projecting coral reef futures under global warming and ocean acidification. *Science (80-.).*, 333, 418–422.

Petchey, O.L., Casey, T., Jiang, L., McPhearson, T.P. & Price, J. (2002). Species richness, environmental fluctuations, and temporal change in total community biomass. *Oikos*, 99, 231–240.

Preston, F.W. (1948). The commonness, and rarity, of species. *Ecology*, 29, 254–283.

Romanuk, T.N., Vogt, R.J. & Kolasa, J. (2009). Ecological realism and mechanisms by which diversity begets stability. *Oikos*, 118, 819–828.

Roughgarden, J. (1974). Species packing and the competition function with illustrations from coral reef fish. *Theor. Popul. Biol.*, 5, 163–186.

Ruppert, J.L.W., Vigliola, L., Kulbicki, M., Labrosse, P., Fortin, M.-J. & Meekan, M.G. (2018). Human activities as a driver of spatial variation in the trophic structure of fish communities on Pacific coral reefs. *Glob. Chang. Biol.*, 24, e67–e79.

Russ, G. (1984a). Distribution and abundance of herbivorous grazing fishes in the central Great Barrier Reef. I. Levels of variability across the entire continental shelf. *Mar. Ecol. Prog. Ser. Oldend.*, 20, 23–34.

Russ, G. (1984b). Distribution and abundance of herbivorous grazing fishes in the central Great Barrier Reef. II. Patterns of zonation of mid-shelf and outershelf reefs. *Mar. Ecol. Prog. Ser. Oldend.*, 20, 35–44.

Sæther, B., Engen, S. & Grøtan, V. (2013). Species diversity and community similarity in fluctuating environments: parametric approaches using species abundance distributions. *J. Anim. Ecol.*, 82, 721–738.

Sale, P.F. (1977). Maintenance of High Diversity in Coral Reef Fish Communities. *Am. Nat.*, 111, 337–359.

Sandin, S.A., Smith, J.E., DeMartini, E.E., Dinsdale, E.A., Donner, S.D., Friedlander, A.M., *et al.* (2008). Baselines and degradation of coral reefs in the Northern Line Islands. *PLoS One*, 3, e1548.

Sankaran, M. & McNaughton, S.J. (1999). Determinants of biodiversity regulate compositional stability of communities. *Nature*, 401, 691–693.

Sasaki, T. & Lauenroth, W.K. (2011). Dominant species, rather than diversity, regulates temporal stability of plant communities. *Oecologia*, 166, 761–768.

Schiller, A., Herzfeld, M., Brinkman, R., Rizwi, F. & Andrewartha, J. (2015). Cross-shelf exchanges between the Coral Sea and the Great Barrier Reef lagoon determined from a regional-scale numerical model. *Cont. Shelf Res.*, 109, 150–163.

Schindler, D.E., Armstrong, J.B. & Reed, T.E. (2015). The portfolio concept in ecology and evolution. *Front. Ecol. Environ.*, 13, 257–263.

Shurin, J.B., Arnott, S.E., Hillebrand, H., Longmuir, A., Pinel-Alloul, B., Winder, M., *et al.* (2007). Diversity–stability relationship varies with latitude in zooplankton. *Ecol. Lett.*, 10, 127–134.

Smith, T.W. & Lundholm, J.T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography (Cop.)*., 33, 648–655.

Srivastava, D.S. & Vellend, M. (2005). Biodiversity-ecosystem function research: is it relevant to conservation? *Annu. Rev. Ecol. Evol. Syst.*, 36, 267–294.

Stuart-Smith, R.D., Bates, A.E., Lefcheck, J.S., Duffy, J.E., Baker, S.C., Thomson, R.J., *et al.* (2013). Integrating abundance and functional traits reveals new global hotspots of fish diversity. *Nature,* 501, 539–542.

Sugihara, G. (1980). Minimal Community Structure: An Explanation of Species Abundance Patterns. *Am. Nat.*

Sweatman, H., Burgess, S., Cheal, A., Coleman, G., Delean, S., Emslie, M., *et al.* (2005). Long-term monitoring of the Great Barrier Reef. Status Report Number 7.

Thibaut, L.M. & Connolly, S.R. (2013). Understanding diversity–stability relationships: towards a unified model of portfolio effects. *Ecol. Lett.*, 16, 140–150.

Thibaut, L.M., Connolly, S.R. & Sweatman, H. (2012). Diversity and stability of herbivorous fishes on coral reefs. *Ecology*, 93, 891–901.

Tilman, D. (1999). The ecological consequences of changes in biodiversity: a search for general principles. *Ecology*, 80, 1455–1474.

Tilman, D., Isbell, F. & Cowles, J.M. (2014). Biodiversity and ecosystem functioning. *Annu. Rev. Ecol. Evol. Syst.*, 45, 471–493.

Tuomisto, H., Ruokolainen, L. & Ruokolainen, K. (2012). Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning results. *Ecography (Cop.)*., 35, 961–971.

Turnbull, L.A., Isbell, F., Purves, D.W., Loreau, M. & Hector, A. (2016). Understanding the value of plant diversity for ecosystem functioning through niche theory. *Proc. R. Soc. B Biol. Sci.*, 283, 20160536.

Ulrich, W., Ollik, M. & Ugland, K.I. (2010). A meta-analysis of species–abundance distributions. *Oikos*, 119, 1149–1155.

Usinowicz, J., Chang-Yang, C.-H., Chen, Y.-Y., Clark, J.S., Fletcher, C., Garwood, N.C., *et al.* (2017). Temporal coexistence mechanisms contribute to the latitudinal gradient in forest diversity. *Nature*, 550, 105–108.

Valdivia, N. & Molis, M. (2009). Observational evidence of a negative biodiversity–stability relationship in intertidal epibenthic communities . *Aquat. Biol.*, 4, 263–271.

Veen, C.G.F., van der Putten, W.H. & Bezemer, T.M. (2018). Biodiversity-ecosystem functioning relationships in a long-term non-weeded field experiment. *Ecology*.

Vellend, M. (2017). The biodiversity conservation paradox. *Am. Sci.*, 105, 94.

Villéger, S., Mason, N.W.H. & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290–2301.

Villéger, S., Miranda, J.R., Hernández, D.F. & Mouillot, D. (2010). Contrasting changes in taxonomic vs. functional diversity of tropical fish communities after habitat degradation. *Ecol. Appl.*, 20, 1512–1522.

Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450, 45.

Wang, S. & Loreau, M. (2016). Biodiversity and ecosystem stability across scales in metacommunities. *Ecol. Lett.*, 19, 510–518.

Wardle, D.A. (2016). Do experiments exploring plant diversity–ecosystem functioning relationships inform how biodiversity loss impacts natural ecosystems? *J. Veg. Sci.*, 27, 646–653.

Wayne, P.H., Wilsey, B.J. & Derner, J.D. (2007). Dominant species constrain effects of species diversity on temporal variability in biomass production of tallgrass prairie. *Oikos*, 116, 2044–2052.

Webb, C.O., Ackerly, D.D., McPeek, M.A. & Donoghue, M.J. (2002). Phylogenies and Community Ecology. *Annu. Rev. Ecol. Syst.*, 33, 475–505.

Williams, D.M. (1982). Patterns in the distribution of fish communities across the Central Great Barrier Reef. *Coral Reefs*, 1, 35–43.

Wilson, S.K., Graham, N.A.J., Pratchett, M.S., Jones, G.P. & Polunin, N.V.C. (2006). Multiple disturbances and the global degradation of coral reefs: are reef fishes at risk or resilient? *Glob. Chang. Biol.*, 12, 2220–2234.

Winfree, R., W. Fox, J., Williams, N.M., Reilly, J.R. & Cariveau, D.P. (2015). Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecol. Lett.*, 18, 626–635.

Wolanski, E. & Bennett, A.F. (1983). Continental shelf waves and their influence on the circulation around the Great Barrier Reef. *Mar. Freshw. Res.*, 34, 23–47.

Wolff, N.H., Mumby, P.J., Devlin, M. & Anthony, K.R.N. (2018). Vulnerability of the Great Barrier Reef to climate change and local pressures. *Glob. Chang. Biol.*, 24, 1978–1991.

Wooldridge, S.A. & Brodie, J.E. (2015). Environmental triggers for primary outbreaks of crown-of-thorns starfish on the Great Barrier Reef, Australia. *Mar. Pollut. Bull.*, 101, 805–815.

Wootton, J.T. & Emmerson, M. (2005). Measurement of interaction strength in nature. *Annu. Rev. Ecol. Evol. Syst.*, 36, 419–444.

Xu, Z., Ren, H., Li, M., Ruijven, J., Han, X., Wan, S., *et al.* (2015). Environmental changes drive the temporal stability of semi-arid natural grasslands through altering species asynchrony. *J. Ecol.*, 103, 1308–1316.

Yachi, S. & Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proc. Natl. Acad. Sci.*, 96, 1463 LP – 1468.

Yeh, Y.-C., Peres-Neto, P.R., Huang, S.-W., Lai, Y.-C., Tu, C.-Y., Shiah, F.-K., *et al.* (2015). Determinism of bacterial metacommunity dynamics in the southern East China Sea varies depending on hydrography. *Ecography (Cop.).*, 38, 198–212.

Appendices

Appendix A

Table A1. Pearson correlations between environmental variables. 'mC', 'log(sdC)', and 'log(cvC)' represents the long-term (11-yr) average, log-transformed standard deviation, and log-transformed coefficient of variation in coral cover annual fluctuations, respectively. 'Lat' and 'Shelf' represents the latitude and cross-shelf position, respectively. Symbol * indicates $P<0.05$.

	log(sdC)	mC	log(cvC)	Lat	Shelf
log(sdC)		0.53*	0.89*	-0.24	-0.04
mC			0.12	-0.16	-0.12
log(cvC)				-0.18	0.07
Lat					0.15
Shelf					

Table A2. Pearson correlations between measures of reef fish community structure. 'Richness' and 'Unevenness' represents time-averaged richness and unevenness from Poisson-lognormal fits of reef fish relative species abundance across coral reefs, respectively. π_r, π_e and π_d represents the proportional variance in relative abundances of reef fishes explained by species differences, environmental stochasticity, and sampling/demographic stochasticity, respectively. Symbol * indicates $P<0.05$.

	Richness	Unevenness	π_r	π_e	π_d
Richness		-0.67*	-0.08	0.11	-0.17
Unevenness			0.09	-0.07	-0.08
π_r				-0.98*	0.23
π_e					-0.39*
π_d					

Table A3. Classifications of trophic/functional groups of coral reef fishes.

Species name	Trophic group
Acanthochromis polyacanthus	Planktivore

Acanthurus albipectoralis	Planktivore
Acanthurus auranticavus	Herbivore
Acanthurus bariene	Herbivore
Acanthurus blochii	Herbivore
Acanthurus dussumieri	Herbivore
Acanthurus grammoptilus	Herbivore
Acanthurus lineatus	Herbivore
Acanthurus maculiceps	Herbivore
Acanthurus mata	Planktivore
Acanthurus nigricans	Herbivore
Acanthurus nigricauda	Herbivore
Acanthurus nigrofuscus	Herbivore
Acanthurus nigroris	Herbivore
Acanthurus olivaceus	Herbivore
Acanthurus pyroferus	Herbivore
Acanthurus spp	Herbivore
Acanthurus thompsoni	Planktivore
Acanthurus triostegus	Herbivore
Acanthurus xanthopterus	Herbivore
Amblyglyphidodon aureus	Planktivore
Amblyglyphidodon curacao	Planktivore
Amblyglyphidodon leucogaster	Planktivore
Amphiprion akindynos	Planktivore
Amphiprion chrysopterus	Planktivore
Amphiprion clarkii	Planktivore
Amphiprion melanopus	Planktivore
Amphiprion percula	Planktivore
Amphiprion perideraion	Herbivore
Anyperodon leucogrammicus	Piscivore
Aprion virescens	Piscivore
Bolbometopon muricatum	Herbivore
Caesio caerulaurea	Planktivore
Caesio cuning	Planktivore
Calotomus carolinus	Herbivore
Cephalopholis argus	Piscivore
Cephalopholis boenak	Piscivore
Cephalopholis cyanostigma	Piscivore
Cephalopholis microprion	Piscivore
Cephalopholis miniata	Piscivore
Cephalopholis urodeta	Piscivore

Cetoscarus bicolor	Herbivore
Chaetodon aureofasciatus	Corallivore
Chaetodon auriga	Benthic invertebrate feeder
Chaetodon baronessa	Corallivore
Chaetodon bennetti	Corallivore
Chaetodon citrinellus	Benthic invertebrate feeder
Chaetodon ephippium	Benthic invertebrate feeder
Chaetodon flavirostris	Corallivore
Chaetodon kleinii	Benthic invertebrate feeder
Chaetodon lineolatus	Benthic invertebrate feeder
Chaetodon lunula	Benthic invertebrate feeder
Chaetodon melannotus	Benthic invertebrate feeder
Chaetodon mertensii	Benthic invertebrate feeder
Chaetodon meyeri	Corallivore
Chaetodon ornatissimus	Corallivore
Chaetodon oxycephalus	Corallivore
Chaetodon pelewensis	Corallivore
Chaetodon plebeius	Corallivore
Chaetodon punctatofasciatus	Corallivore
Chaetodon rafflesii	Benthic invertebrate feeder
Chaetodon rainfordi	Herbivore
Chaetodon reticulatus	Corallivore
Chaetodon speculum	Corallivore
Chaetodon trifascialis	Corallivore
Chaetodon trifasciatus	Corallivore
Chaetodon ulietensis	Corallivore
Chaetodon unimaculatus	Corallivore
Chaetodon vagabundus	Corallivore
Cheilinus fasciatus	Benthic invertebrate feeder
Cheilinus undulatus	Benthic invertebrate feeder
Cheiloprion labiatus	Corallivore
Chelmon rostratus	Benthic invertebrate feeder
Chlorurus bleekeri	Herbivore
Chlorurus japanensis	Herbivore
Chlorurus microrhinos	Herbivore
Chlorurus sordidus	Herbivore
Choerodon fasciatus	Benthic invertebrate feeder
Chromis acares	Planktivore
Chromis agilis	Planktivore
Chromis amboinensis	Planktivore

Chromis atripectoralis	Planktivore
Chromis atripes	Planktivore
Chromis chrysura	Planktivore
Chromis flavomaculata	Planktivore
Chromis iomelas	Planktivore
Chromis lepidolepis	Planktivore
Chromis lineata	Planktivore
Chromis margaritifer	Planktivore
Chromis nitida	Planktivore
Chromis retrofasciata	Planktivore
Chromis ternatensis	Planktivore
Chromis vanderbilti	Planktivore
Chromis viridis	Planktivore
Chromis weberi	Planktivore
Chromis xanthochira	Planktivore
Chromis xanthura	Planktivore
Chrysiptera biocellata	Herbivore
Chrysiptera flavipinnis	Planktivore
Chrysiptera rex	Planktivore
Chrysiptera rollandi	Planktivore
Chrysiptera talboti	Planktivore
Coris gaimard	Benthic invertebrate feeder
Cromileptes altivelis	Piscivore
Ctenochaetus binotatus	Herbivore
Ctenochaetus spp	Herbivore
Dascyllus aruanus	Planktivore
Dascyllus melanurus	Planktivore
Dascyllus reticulatus	Planktivore
Dascyllus trimaculatus	Planktivore
Dischistodus melanotus	Herbivore
Dischistodus perspicillatus	Herbivore
Dischistodus prosopotaenia	Herbivore
Dischistodus pseudochrysopoecilus	Herbivore
Epibulus insidiator	Benthic invertebrate feeder
Epinephelus cyanopodus	Piscivore
Epinephelus fasciatus	Benthic invertebrate feeder
Epinephelus fuscoguttatus	Piscivore
Epinephelus merra	Benthic invertebrate feeder
Epinephelus ongus	Benthic invertebrate feeder
Epinephelus quoyanus	Benthic invertebrate feeder

Forcipiger flavissimus	Benthic invertebrate feeder
Forcipiger longirostris	Benthic invertebrate feeder
Gnathodentex aureolineatus	Benthic invertebrate feeder
Gomphosus varius	Benthic invertebrate feeder
Gymnocranius spp	Benthic invertebrate feeder
Halichoeres hortulanus	Benthic invertebrate feeder
Hemiglyphidodon plagiometopon	Herbivore
Hemigymnus fasciatus	Benthic invertebrate feeder
Hemigymnus melapterus	Benthic invertebrate feeder
Hemitaurichthys polylepis	Planktivore
Hipposcarus longiceps	Herbivore
Lethrinus atkinsoni	Benthic invertebrate feeder
Lethrinus erythracanthus	Benthic invertebrate feeder
Lethrinus harak	Benthic invertebrate feeder
Lethrinus laticaudis	Benthic invertebrate feeder
Lethrinus lentjan	Benthic invertebrate feeder
Lethrinus miniatus	Benthic invertebrate feeder
Lethrinus nebulosus	Benthic invertebrate feeder
Lethrinus obsoletus	Benthic invertebrate feeder
Lethrinus olivaceus	Piscivore
Lethrinus ornatus	Benthic invertebrate feeder
Lethrinus rubrioperculatus	Benthic invertebrate feeder
Lethrinus semicinctus	Benthic invertebrate feeder
Lethrinus xanthochilus	Benthic invertebrate feeder
Lutjanus adetii	Piscivore
Lutjanus argentimaculatus	Piscivore
Lutjanus biguttatus	Piscivore
Lutjanus bohar	Piscivore
Lutjanus boutton	Piscivore
Lutjanus carponotatus	Piscivore
Lutjanus fulviflammus	Piscivore
Lutjanus fulvus	Piscivore
Lutjanus gibbus	Piscivore
Lutjanus kasmira	Piscivore
Lutjanus lemniscatus	Piscivore
Lutjanus lutjanus	Piscivore
Lutjanus monostigma	Piscivore
Lutjanus quinquelineatus	Piscivore
Lutjanus rivulatus	Piscivore
Lutjanus russellii	Benthic invertebrate feeder

Lutjanus sebae	Piscivore
Lutjanus semicinctus	Piscivore
Lutjanus vitta	Piscivore
Macolor spp	Planktivore
Monotaxis grandoculis	Benthic invertebrate feeder
Naso lituratus	Herbivore
Naso tuberosus	Herbivore
Naso unicornis	Herbivore
Neoglyphidodon melas	Benthic invertebrate feeder
Neoglyphidodon nigroris	Herbivore
Neoglyphidodon polyacanthus	Planktivore
Neopomacentrus azysron	Planktivore
Neopomacentrus bankieri	Planktivore
Neopomacentrus cyanomos	Planktivore
Paracanthurus hepatus	Planktivore
Plectroglyphidodon dickii	Benthic invertebrate feeder
Plectroglyphidodon johnstonianus	Herbivore
Plectroglyphidodon lacrymatus	Herbivore
Plectropomus areolatus	Piscivore
Plectropomus laevis	Piscivore
Plectropomus leopardus	Piscivore
Plectropomus maculatus	Piscivore
Pomacentrus adelus	Herbivore
Pomacentrus amboinensis	Herbivore
Pomacentrus australis	Herbivore
Pomacentrus bankanensis	Herbivore
Pomacentrus brachialis	Planktivore
Pomacentrus chrysurus	Herbivore
Pomacentrus coelestis	Planktivore
Pomacentrus grammorhynchus	Herbivore
Pomacentrus imitator	Planktivore
Pomacentrus lepidogenys	Planktivore
Pomacentrus moluccensis	Planktivore
Pomacentrus nagasakiensis	Planktivore
Pomacentrus nigromarginatus	Planktivore
Pomacentrus philippinus	Planktivore
Pomacentrus reidi	Planktivore
Pomacentrus tripunctatus	Herbivore
Pomacentrus vaiuli	Benthic invertebrate feeder
Pomacentrus wardi	Herbivore

Pomachromis richardsoni	Planktivore
Pomadasys taeniatus	Benthic invertebrate feeder
Premnas biaculeatus	Planktivore
Sargocentron spiniferum	Benthic invertebrate feeder
Scarus altipinnis	Herbivore
Scarus chameleon	Herbivore
Scarus dimidiatus	Herbivore
Scarus flavipectoralis	Herbivore
Scarus forsteni	Herbivore
Scarus frenatus	Herbivore
Scarus ghobban	Herbivore
Scarus globiceps	Herbivore
Scarus longipinnis	Herbivore
Scarus niger	Herbivore
Scarus oviceps	Herbivore
Scarus psittacus	Herbivore
Scarus rivulatus	Herbivore
Scarus rubroviolaceus	Herbivore
Scarus schlegeli	Herbivore
Scarus spinus	Herbivore
Scarus spp	Herbivore
Siganus argenteus	Herbivore
Siganus corallinus	Herbivore
Siganus doliatus	Herbivore
Siganus fuscescens	Herbivore
Siganus javus	Herbivore
Siganus lineatus	Herbivore
Siganus puellus	Herbivore
Siganus punctatissimus	Herbivore
Siganus punctatus	Herbivore
Siganus spinus	Herbivore
Siganus vulpinus	Herbivore
Stegastes apicalis	Herbivore
Stegastes fasciolatus	Herbivore
Stegastes gascoynei	Herbivore
Stegastes nigricans	Herbivore
Variola albimarginata	Piscivore
Variola louti	Piscivore
Zanclus cornutus	Benthic invertebrate feeder
Zebrasoma scopas	Herbivore

Zebrasoma veliferum	Herbivore

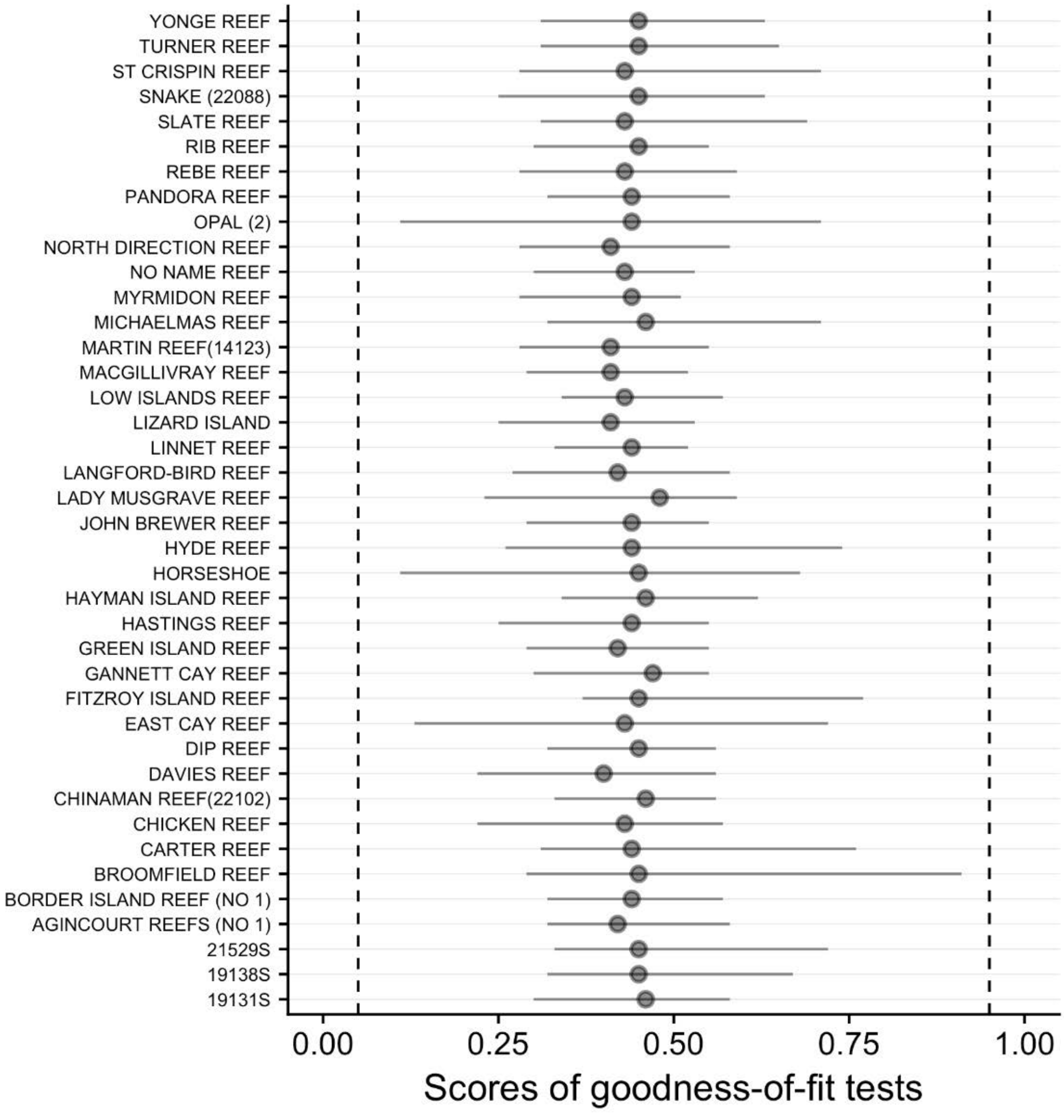

Figure A1. Scores (p-values) of parametric bootstrap tests under the null hypothesis of bivariate Poisson-lognormal species-abundance distributions over time lags for reef fish communities. For each time lag (n=55) on a reef, the scores (or p-values) are computed from 100-times bootstrapped likelihood values of bivariate Poisson-lognormal species-abundance distributions. The test statistic (scores) is the likelihood value of data compared to parametric bootstrapped likelihood values. Dashed line indicates the critical quantile values (5% and 95%) for rejecting the null hypothesis. Solid lines and circles represent the full range and the median of bootstrapped scores, respectively. The scores located within the interval between the two dashed lines indicate that the fitted bivariate Poisson-lognormal species-abundance distributions are not rejected by null hypothesis tests.

Appendix B

Numerical stability and AIC are used to select mixed-effects model of eq. 1 for reef fish communities (i.e., 40 reefs as random effects; Table B1). Although lowest-AIC model contains all three parameters as random effects, this model is numerically unstable - the corresponding parameters estimated from Laplace approximation and importance sampling do not converge to similar values (Table B1, Figure B1). By contrast, the model that contains only two random-effect parameters are numerically stable, where the fitted parameters from both algorithms are converged and the corresponding AIC value is close to lowest (Table B1, Figure B2).

Therefore, I used the later model of eq. 1 for variance partitioning of relative species abundance. The best mixed-effects model can be expressed as follows (also see Chapter 3):

$$\rho_{k,\tau} = \left(\rho_0^{R_k} - \rho_\infty^{R_k}\right)\exp\left(-\delta\tau\right) + \rho_\infty^{R_k} + \varepsilon$$

$$\rho_0^{R_k} \sim \mathrm{LN}(\mu_{\rho_0}, \sigma_{\rho_0}^2)$$

$$\rho_\infty^{R_k} \sim \mathrm{LN}(\mu_{\rho_\infty}, \sigma_{\rho_\infty}^2)$$

$$\varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2)$$

where notations are the same as eq. 1, except there are two random-effect parameters following lognormal distributions. Residual error is assumed to follow a normal distribution. Visual inspection of residual errors also agrees well with the assumptions (Figure B2).

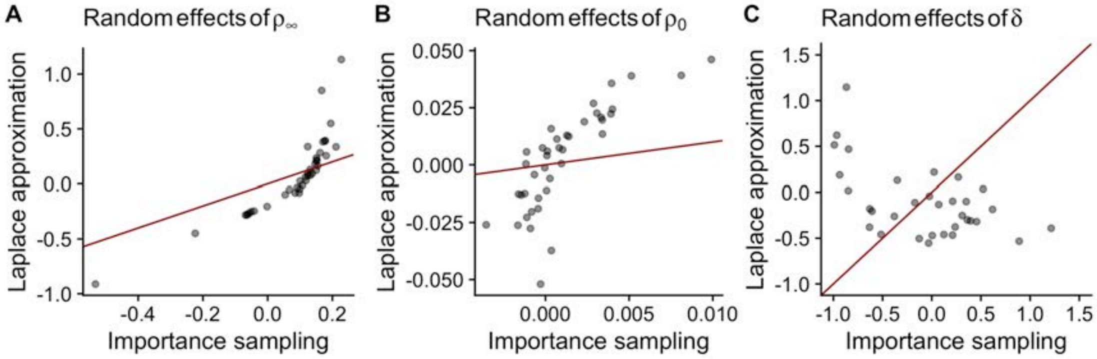

Figure B1. Relationship between random-effect parameters estimated from Laplace approximation and importance sampling. In this case, density dependence δ is assumed to be a random-effect parameter (i.e., δ^{R_k}) in addition to the other two parameters of eq. 1 (**A-C**). Red line represents a 1:1 relationship where estimates from both methods of Laplace approximation and importance sampling converge to the same values. Deviations from the red line represent the numerical instability.

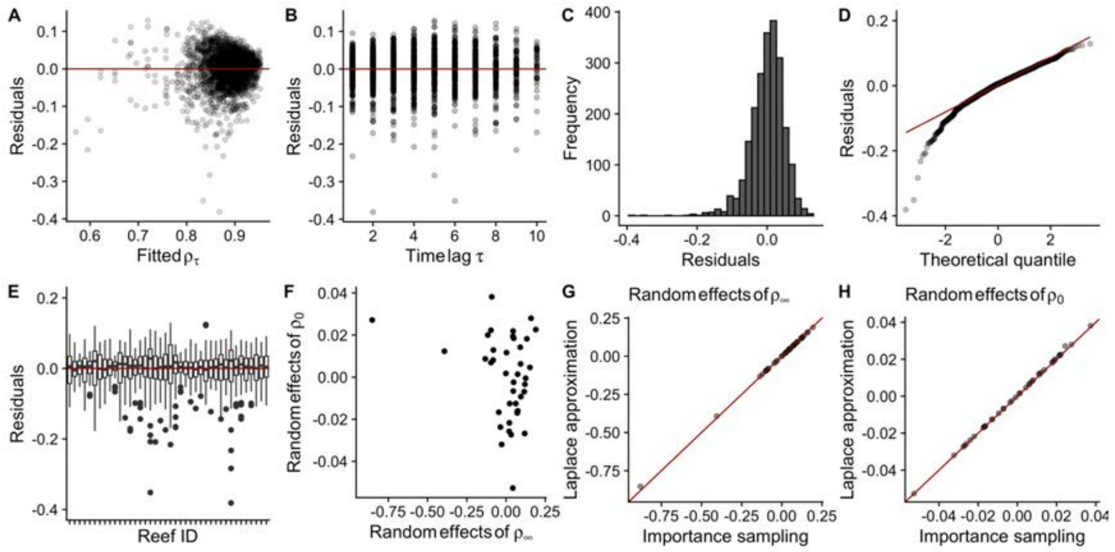

Figure B2. Visual inspections for best-fitted mixed-effects model. (**A-E**) Visual inspections of residual errors. (**F**) Estimated random-effect values are not correlated (R^2 = 0.05). (**G-H**) Both methods of Laplace approximation and importance sampling converge to almost the same random-effect estimates.

Appendix C

Table C1. Results of paired t-tests for differences between mean variance components of persistent species differences of the different trophic groups. The direction of the estimated mean difference is presented as row subtracted by column.

	Herbivore	Planktivore	Benthicinvertivore
Herbivore	NA	-1.8%	1.9%
Planktivore	NA	NA	3.7%
Benthicinvertivore	NA	NA	NA

* indicates *P* <0.05 and ** indicates *P* <0.01 after Bonferroni correction of *P* values

Table C2. Results of paired t-test for mean difference of variance components of environmental stochasticity between trophic groups. The direction of estimated mean difference is presented as row subtracted by column.

	Herbivore	Planktivore	Benthicinvertivore
Herbivore	NA	-4.2%	-11.2%**
Planktivore	NA	NA	-7%*
Benthicinvertivore	NA	NA	NA

* indicates *P* <0.05 and ** indicates *P* <0.01 after Bonferroni correction of *P* values

Table C3. Results of paired t-test for mean difference of variance components of demographic/sampling stochasticity between trophic groups. The direction of estimated mean difference is presented as row subtracted by column.

	Herbivore	Planktivore	Benthicinvertivore
Herbivore	NA	6.1%**	8.8%**
Planktivore	NA	NA	2.6%**
Benthicinvertivore	NA	NA	NA

* indicates *P* <0.05 and ** indicates *P* <0.01 after Bonferroni correction of *P* values

Table C4. OLS regression models for proportional variance components, richness, and unevenness in the context of trophic groups. For explanatory variables, 'mC' and 'log(sdC)' represents the long-term (11-yr) mean and log-transformed standard deviation in annual coral cover fluctuations, respectively. 'Lat' and 'Shelf' represents latitude and cross-shelf position, respectively. 'TG' represents the identity of trophic group as a categorical variable. Cross and plus symbols represent interactive and additive effects, respectively.

Response variable	Explanatory variable	adjR2	AIC	ΔAIC	Final model
Species differences	log(sdC)×mC×TG	0.15	-168.2	0	Yes
Species differences	log(sdC)×TG	0.07	-161.9	6.3	No
Species differences	mC×TG	0.05	-159.8	8.4	No
Species differences	Lat×Shelf×TG	0.05	-155.7	12.5	No
Species differences	Lat×TG	0	-154.5	13.7	No
Species differences	Shelf×TG	0.05	-160.5	7.7	No
Species differences	TG	0	-157.1	11.1	No
Environmental stochasticity	log(sdC)×mC×TG	0.26	-176.1	0	Yes

Environmental stochasticity	log(sdC)×TG	0.16	-165.1	11	No
Environmental stochasticity	mC×TG	0.15	-164.8	11.3	No
Environmental stochasticity	Lat×Shelf×TG	0.16	-160.2	15.9	No
Environmental stochasticity	Lat×TG	0.11	-159.3	16.8	No
Environmental stochasticity	Shelf×TG	0.15	-164.9	11.2	No
Environmental stochasticity	TG	0.11	-162.4	13.7	No
Richness	log(sdC)×mC×TG	0.74	730.9	57.5	No
Richness	log(sdC)×TG	0.72	731.4	58	No
Richness	mC×TG	0.74	725.7	52.3	No
Richness	Lat×Shelf×TG	0.84	673.4	0	Yes
Richness	Lat×TG	0.74	724.3	50.9	No
Richness	Shelf×TG	0.77	708.5	35.1	No
Richness	TG	0.73	727.1	53.7	No
Unevenness	log(sdC)×mC×TG	0.7	438.4	23.6	No
Unevenness	log(sdC)×TG	0.72	428.6	13.8	No
Unevenness	mC×TG	0.69	441.4	26.6	No
Unevenness	Lat×Shelf×TG	0.74	425.2	10.4	No
Unevenness	Lat×TG	0.75	414.8	0	Yes
Unevenness	Shelf×TG	0.65	453.7	38.9	No
Unevenness	TG	0.66	448	33.2	No

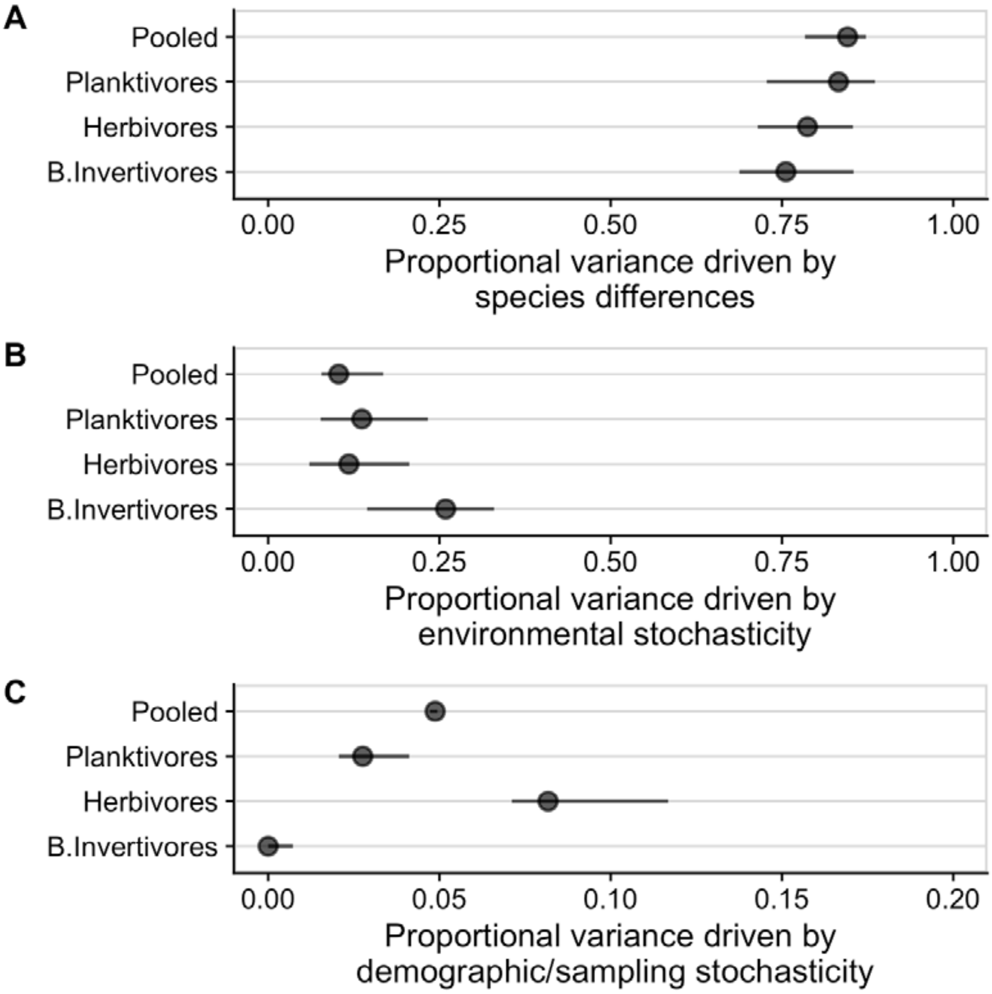

Figure C1. Estimated variance components of community structure of reef fish functional groups. Points and horizontal lines represent the median and the 1st and 3rd quantiles of estimates (n=40 reefs). (**A**) Estimated variance component due to persistent species differences. (**B**) Estimated variance component driven by stochastic fluctuations in population growth rates. (**C**) Estimated variance component driven by demographic or sampling stochasticity. Note that the scale of x-axis in (C) is different from that in (A-B), and that the inter-quartile range is smaller than the size of the points, so the horizontal lines are not visible.

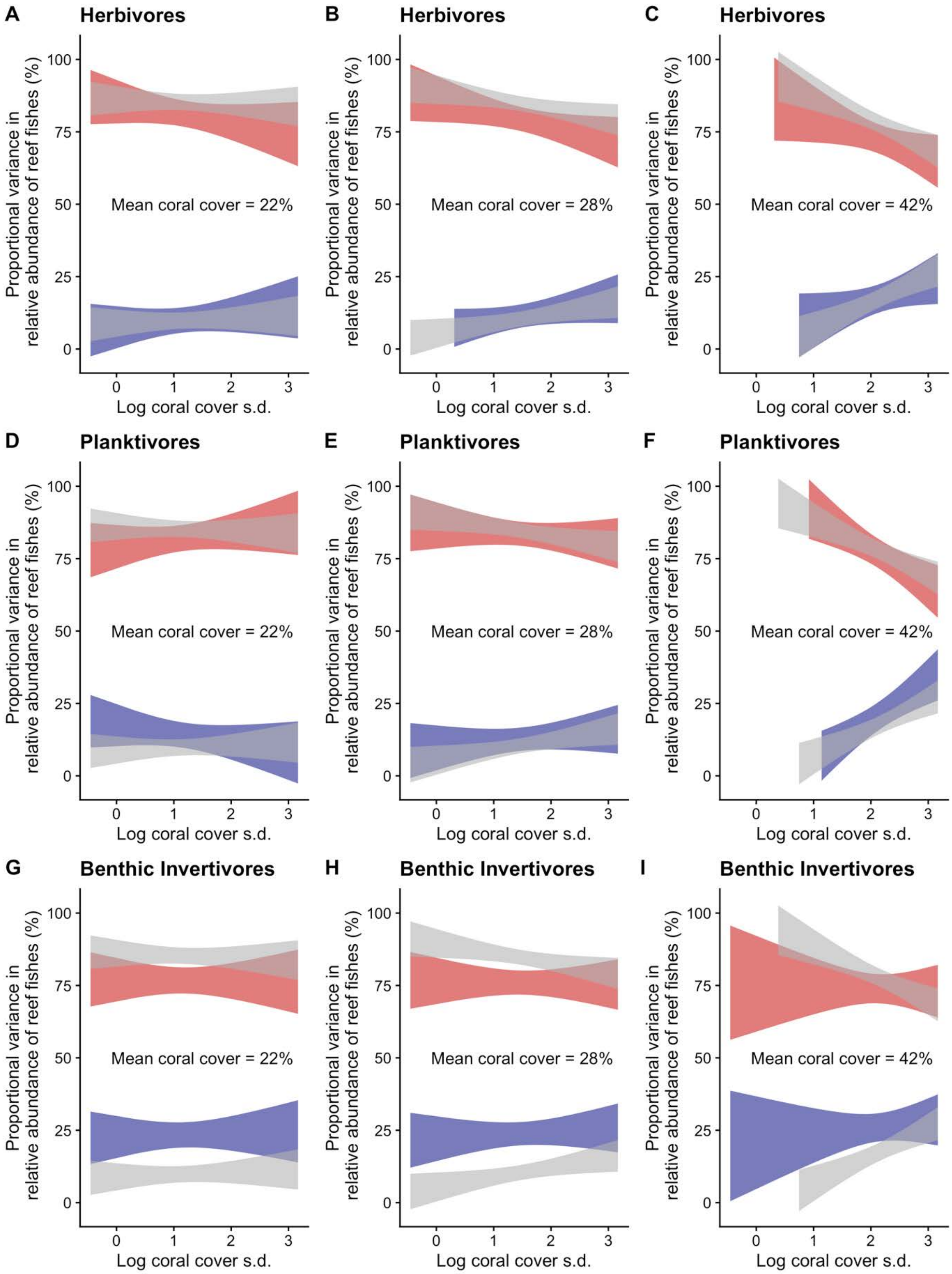

Figure C2. Relationship between coral cover dynamics and variance components of community structure of reef fish functional groups. Relationships between the reef-scale coral cover variables (temporal SD and mean of coral cover for each reef) and the relative importance of variance components in structuring species-abundances of fish trophic groups across reefs (n=40 reefs). (**A-C**) Herbivores. (**D-F**) Planktivores. (**G-I**) Benthic invertivores. The relationships are plotted using the lowest-AIC models, with interactive effects of the mean and SD of coral cover as explanatory variables, on variance components of fish community structure as response variables (Table A6). The red band represents the 95% C.I. of the proportional variance attributable to persistent species or niche differences, while the blue band represents the 95% C.I. of the proportional variance attributable to environmental stochasticity. To illustrate the interactive relationships, the 1st, median and 3rd quartiles of mean coral cover are fixed in panels (A-C), (D-F), and (G-I), respectively, and the relationship between SD of coral cover and variance component values plotted for the corresponding value of mean coral cover.

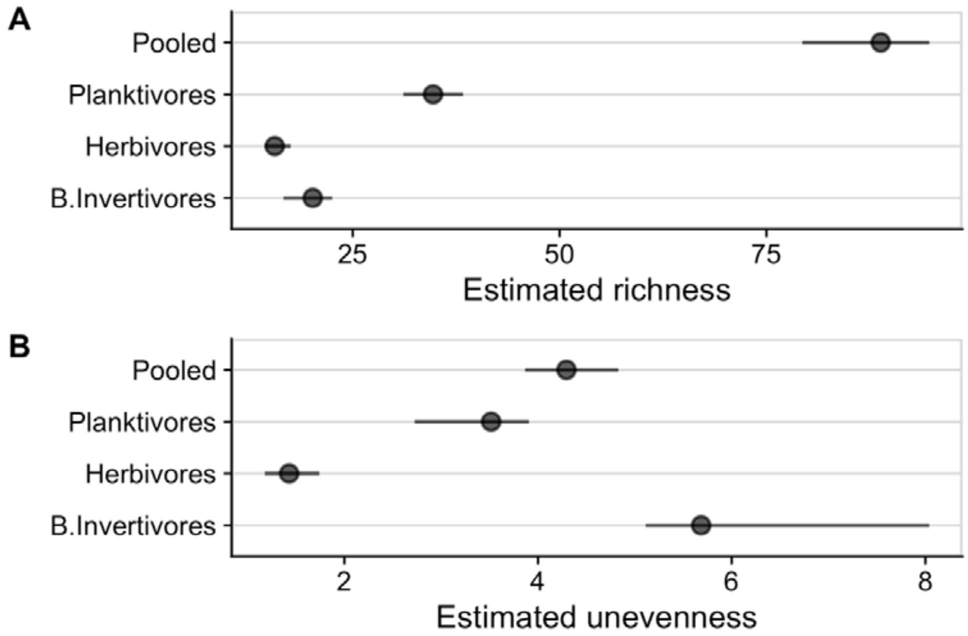

Figure C3. Estimated richness and unevenness of reef fish functional groups on the Great Barrier Reef. Points and horizontal lines represent the median and the 1st and 3rd quantiles of estimates (n=40 reefs). (**A**) Estimated species richness of trophic groups. (**B**) Estimated unevenness of trophic groups.

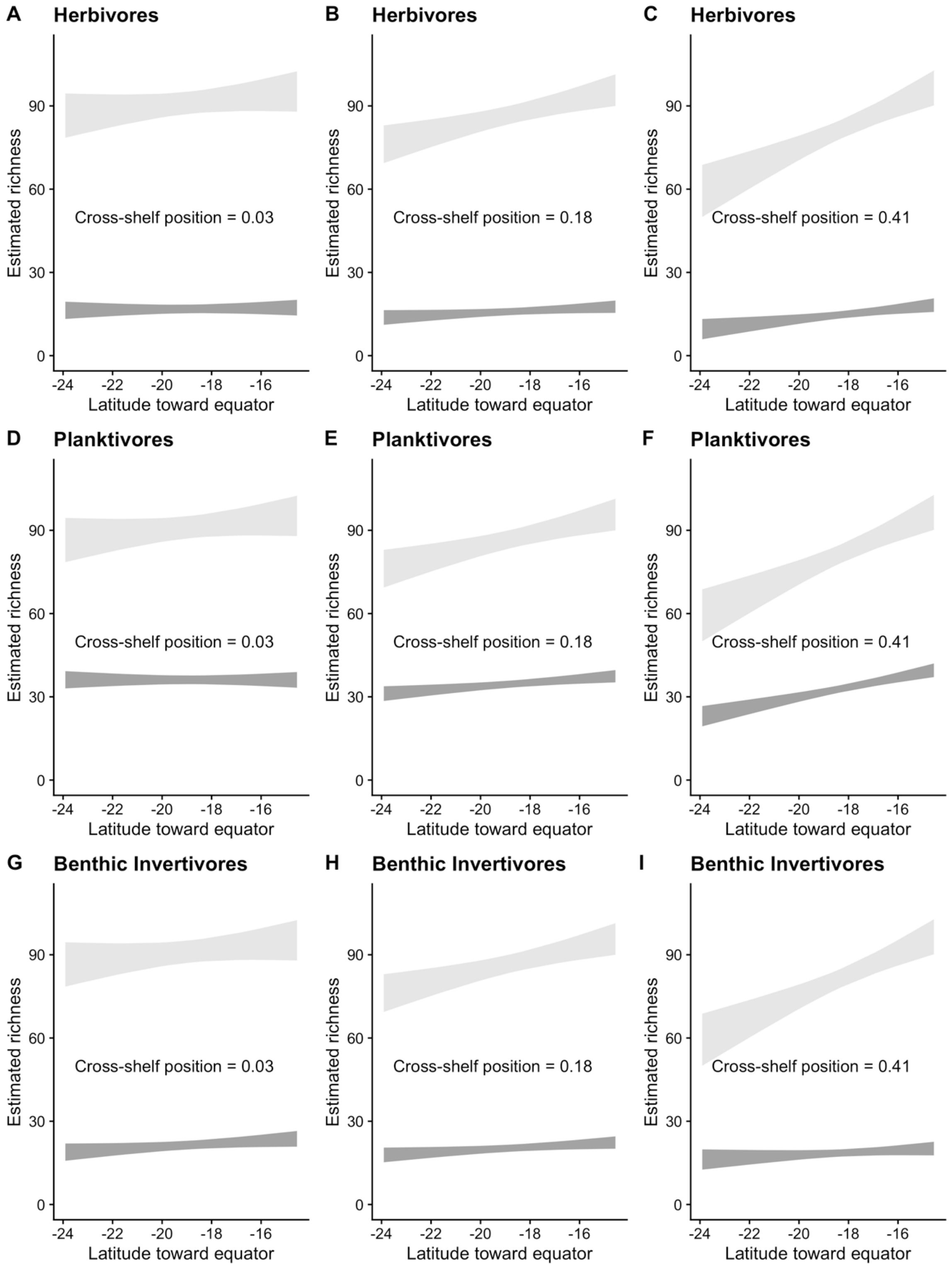

Figure C4. Richness of reef fish functional groups depend on latitude and cross-shelf position. Relationship between time-averaged species richness of fish trophic groups and the interaction of latitude with cross-shelf position. (**A-C**) Herbivores. (**D-F**) Planktivores. (**G-I**) Benthic invertivores. Grey bands are the 95% confidence intervals predicted from the lowest-AIC models for richness (n=40 reefs; Table A6). To better illustrate the interactive relationships, the 1st, median, and 3rd quartiles of cross-shelf positions are fixed in panels (A-C), (D-F), and (G-I), respectively, and the relationship between species richness as a function of latitude show for the corresponding value of cross-shelf position. Note that cross-shelf position increases towards the coast.

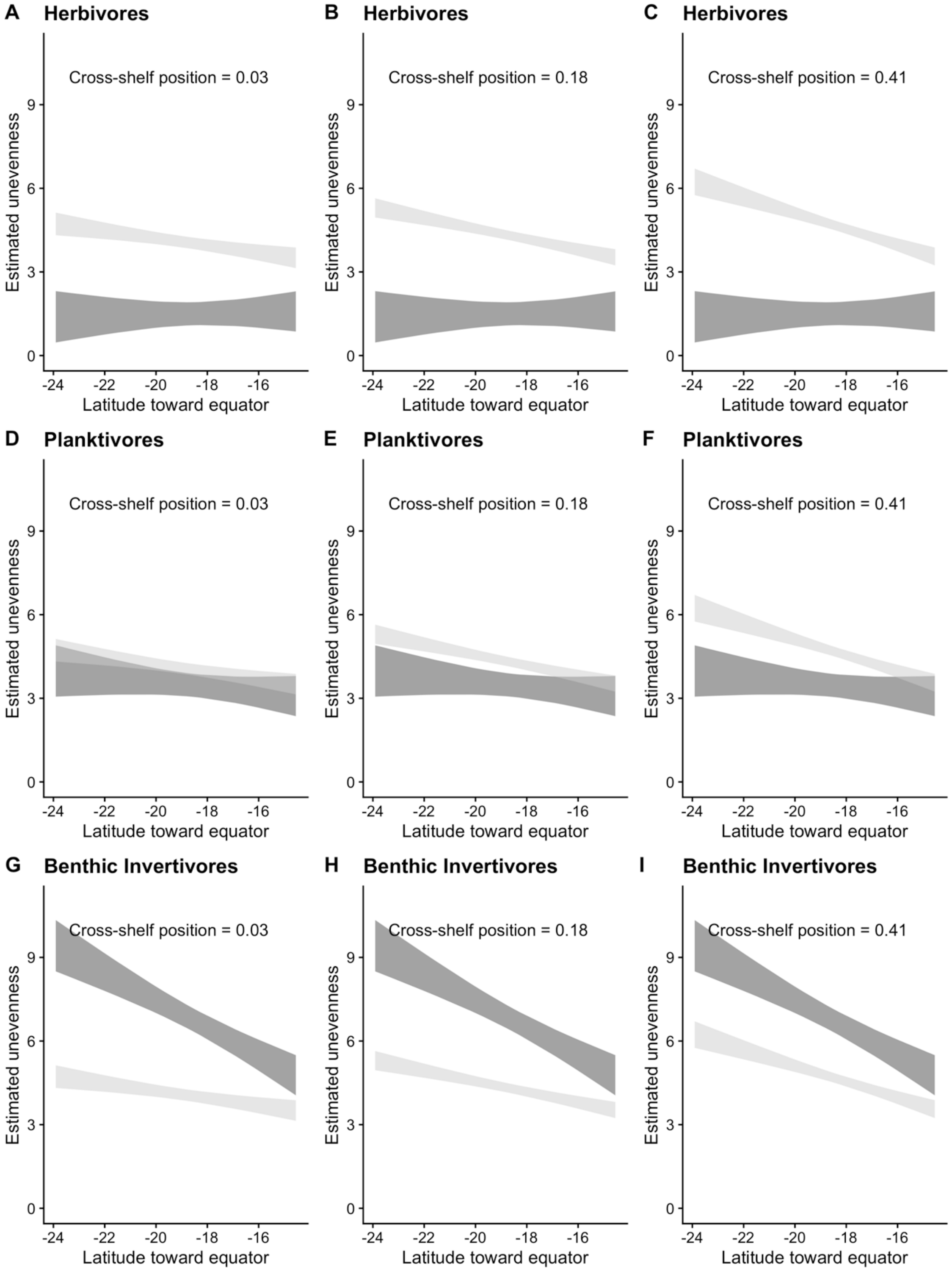

Figure C5. Unevenness of reef fish functional groups depend on latitude and cross-shelf position. Relationship between time-averaged unevenness of fish trophic groups and the interaction of latitude with cross-shelf position. (**A-C**) Herbivores. (**D-F**) Planktivores. (**G-I**) Benthic invertivores. Grey bands are the 95% confidence intervals predicted from the lowest-AIC models for unevenness (n=40 reefs; Table A6). To better illustrate the interactive relationships, the 1st, median, and 3rd quartiles of cross-shelf positions are fixed in panels (A-C), (D-F), and (G-I), respectively, and the relationship between unevenness as a function of latitude show for the corresponding value of cross-shelf position. Note that cross-shelf position increases towards the coast.

Appendix D

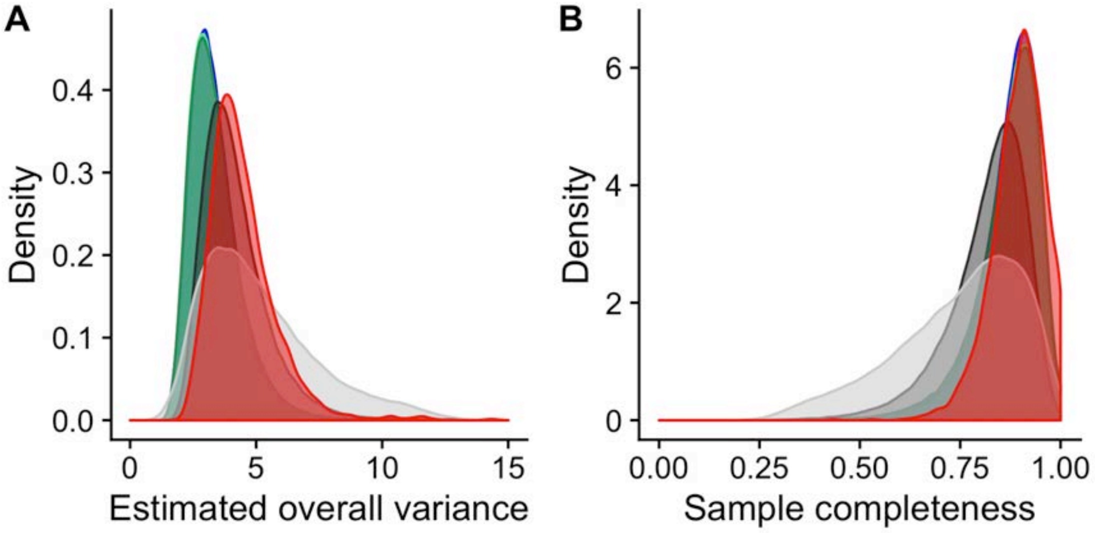

Figure D1. Density distributions of overall variance in relative species abundance and sample completeness estimated from empirical and simulated community time series data. (**A**) Density distribution of estimated overall variance of Poisson-lognormal species-abundance distributions. (**B**) Density distribution of sample completeness measured as the ratio of observed (sampled) species richness to estimated total species richness from Poisson-lognormal fits. Curves are probability density distributions. Red curves represent the empirical LTMP data. Green curves (light green, intermediate, and dark green) are '*baseline*','*varcovEnv*', and '*varEnv*' scenarios, respectively. Light- and dark-gray curves are '*varIntra*' and '*varInter*' scenarios, respectively.

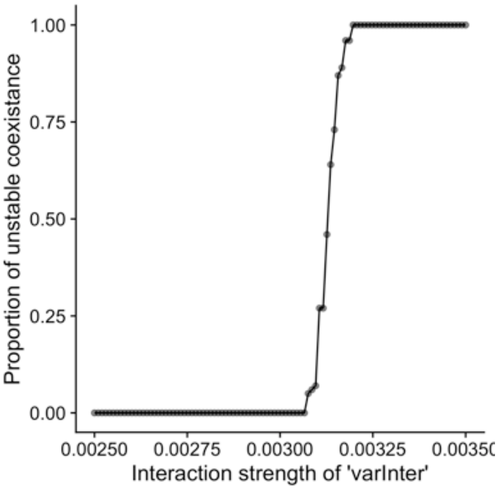

Figure D2. Relationship between interaction strength and probability of unstable equilibria under the scenario “varInter”. The proportion of unstable equilibria is defined as the number of interaction matrices B in eq. 3.9 whose largest eigenvalue is greater than 1 (i.e., the complex norm is within the unit circle) in every 100 simulations.

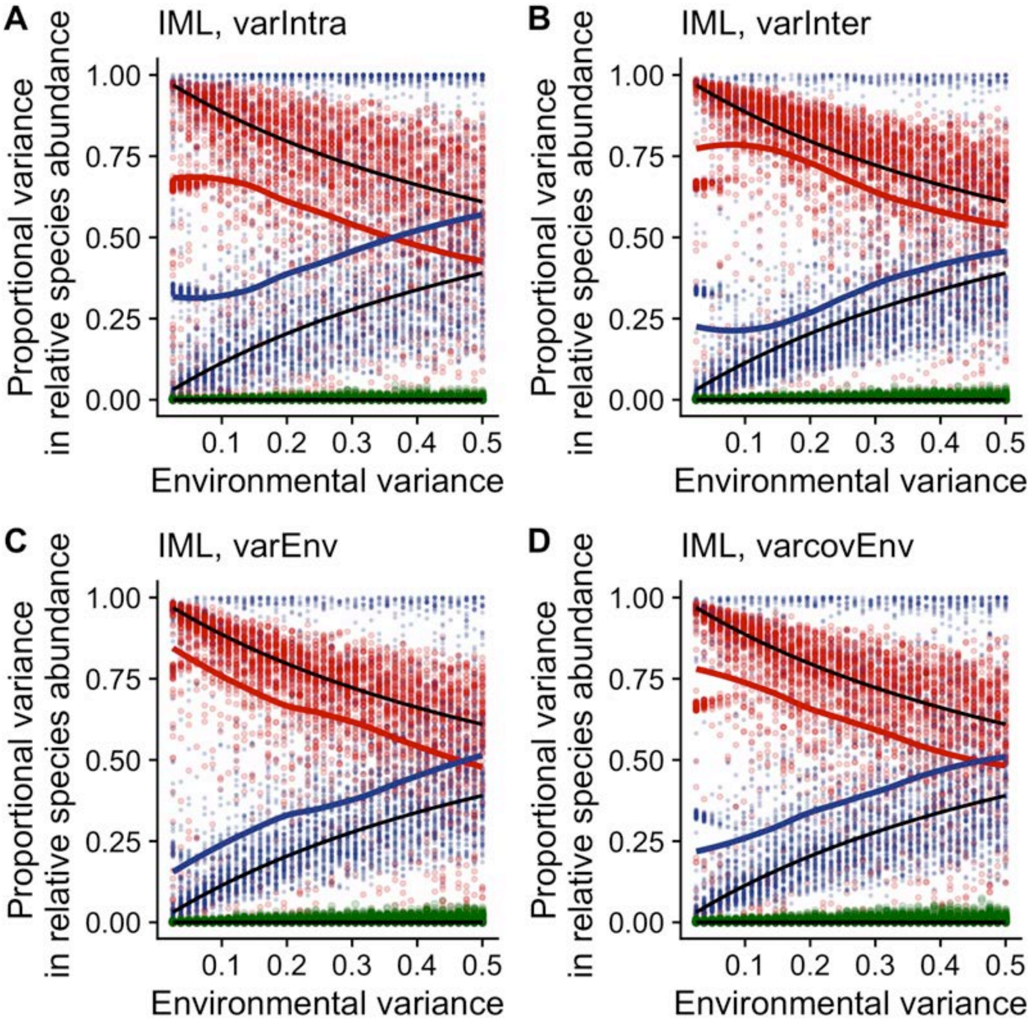

Figure D3. Relationships between simulation-controlled environmental variance and IML-estimated variance components under different scenarios of community dynamics (cf. Figure 3.2). The variance components estimated by IML approach (i.e., the estimation method from the original theory). Red, blue, and green points represent the IML-estimated proportional variance, for each reef in each simulation (4000 points per color in total) driven by deterministic species differences (or "niche structure"), environmental stochasticity, and the remaining variance attributable to demographic and sampling stochasticity, respectively. Colored (red, blue, and green) lines represent the kernel smoothing of proportional variance estimates, obtained using local polynomial regression fitting. Black lines represent the analytical prediction of Engen et al. 2002 (eq. 3.3) using the true parameters from the simulations. (**A**) "*varIntra*" scenario, in which species differences are introduced into intra-specific density dependence. (**B**) "*varInterS*" scenario, in which species differences are introduced into inter-specific density dependence. (**C**) "*varEnv*" scenario, in which species differences are introduced into environmental variance. (**D**) "*varcovEnv*" scenario, in which species differences are introduced into environmental covariance.

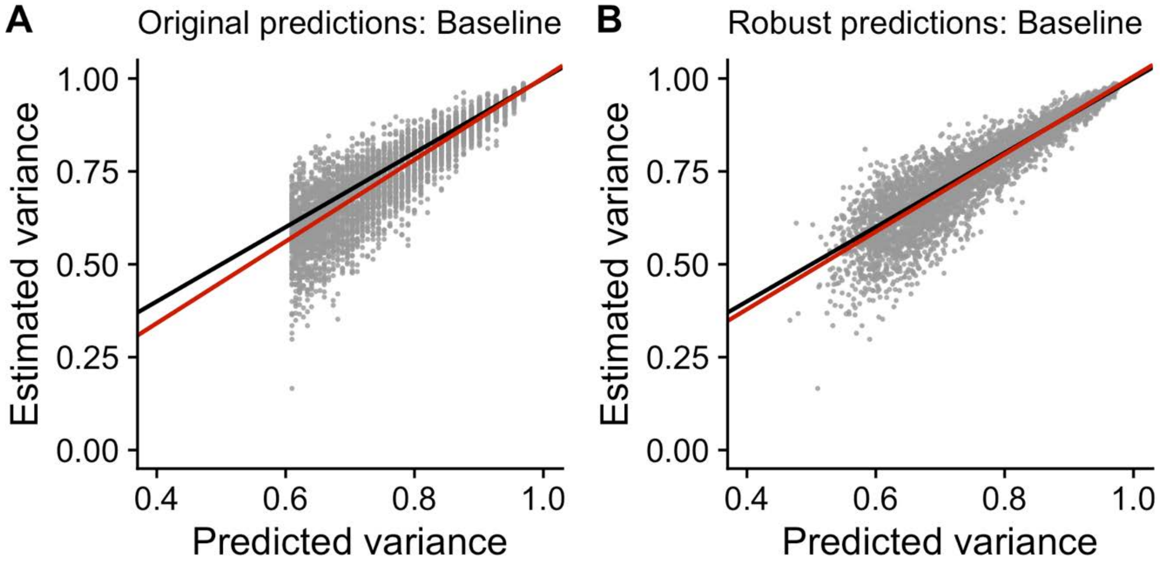

Figure D4. Relationships between variance estimate of deterministic species differences and its analytical prediction under baseline scenario. The black line is the unity line indicating perfect agreement between VPRSA estimates and analytical predictions. The red line is a quantile regression through the median of the MM-based VPRSA estimates of variance components of deterministic species differences. (**A**) The relationship between VPRSA estimates and the original analytical prediction of Engen et al. 2002 (eq. 3.3), and (**B**) the relationship between VPRSA estimates and the generalized analytical prediction from the discrete-time multivariate Gompertz model (eqs. 3.19-3.22).

Appendix E

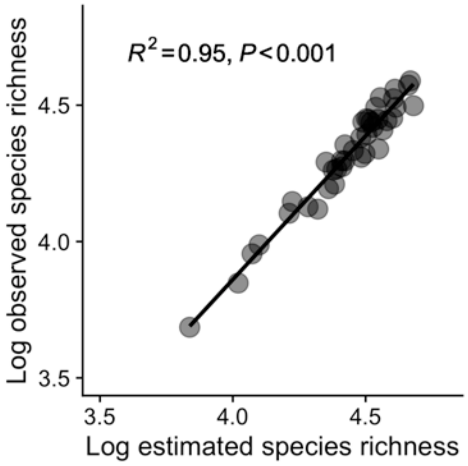

Figure E1. Relationship between observed and estimated species richness.

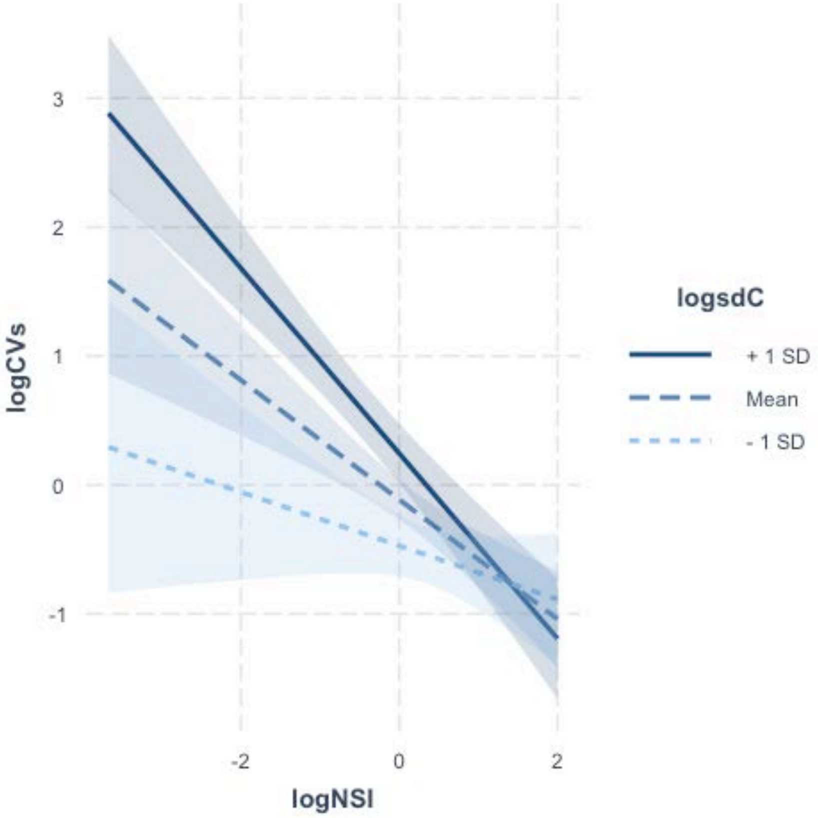

Figure E2. Relationship between NSI and weighted-average population variability, depending on volatility in coral cover. Predictions are from the best model for explaining weighted-population variability. Intervals represent 80% CI of NSI effects, given different levels of volatility in coral cover. Symbols and notations are the same as in Tables E1 and E2.

Table E1. Results of model selection ranked by AICc and ΔAICc. For OLS regressions, "CMS", "ENV" and "CMS+ENV" represents the regression that analyses community structure variables, environmental variables, and both of the variables sets, respectively. Global models are shown to compare with selected subset models. Terms to the right-hand side of the tilde (~) represent the explanatory variable, while the term on the left-hand side of the tilde represents the response variable. (+) and (*) represent additive and interaction terms. "log" represents log transformation. "Paiwise_Interaction" represents the pairwise interaction terms between individual variables. "sync" and "CVs" represent the synchrony index and weighted-average population variability, respectively. "richness", "unevenness", and "NSI" are community structure variables representing species richness, unevenness, and the niche structure index, respectively. "lat", "shelf", "mC", and "sdC" are environmental variables representing latitude, cross-shelf position, mean coral cover, and coral cover variability, respectively. "AICc" is the AIC corrected for small sample size. Only those subset models (and their AICc weights) that have ΔAICc < 2 are reported.

Candidate variables	Global model	Selected subset model	AICc	ΔAICc	weight
CMS	log(sync)~log(richness)+logNSI+log(richness)*logNSI	log(sync)~log(richness)+logNSI	114.4	0	0.42
		log(sync)~log(richness)	114.8	0.4	0.35
		log(sync)~logNSI	115.6	1.2	0.23
	log(sync)~log(unevenness)+logNSI+log(richness)*logNSI	log(sync)~log(unevenness)+logNSI	112.96	0	0.58
		log(sync)~log(unevenness)	113.64	0.68	0.42
ENV	log(sync)~lat+shelf+log(sdC)+mC+Pairwise_Interactions	log(sync)~shelf	107.86	0	1
CMS+ENV	log(sync)~log(richness)+logNSI+shelf+log(richness)*shlef+logNSI*shelf	log(sync)~shelf	107.86	0	0.55
		log(sync)~logNSI+shelf	108.26	0.4	0.45
	log(sync)~log(unevenness)+logNSI+shelf+log(richness)*shlef+logNSI*shelf	log(sync)~shelf	107.86	0	0.36
		log(sync)~logNSI+shelf	108.26	0.4	0.3
		log(sync)~log(unevenness)+shelf	109.09	1.24	0.19
		log(sync)~logNSI+log(unevenness)+shelf	109.61	1.75	0.15
CMS	log(CVs)~log(richness)+logNSI+log(richness)*logNSI	log(CVs)~logNSI	98.41	0	0.62

	log(CVs)~log(unevenness)+logNSI+log(unevnness)*logNSI	log(CVs)~logNSI+log(richness)	99.35	0.94	0.38
		log(CVs)~logNSI	98.41	0	0.55
		log(CVs)~logNSI+log(unevenness)	98.82	0.42	0.45
ENV	log(CVs)~lat+shelf+log(sdC)+mC+Pairwise_Interactions	log(CVs)~log(sdC)+mC+log(sdC)*mC	107.59	0	1
CMS+ENV	log(CVs)~logNSI+log(sdC)+mC+log(sdC)*mC+logNSI*log(sdC)+logNSI*mC+logNSI*log(sdC)*mC	log(CVs)~logNSI+log(sdC)+mC+logNSI*log(sdC)	91.99	0	1

Table E2. Estimated coefficients, CI, and P values of the best OLS models for synchrony index and weighted-average population variability. Symbols and notations are the same as in Table E1.

	log(sync)			**log(CVs)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.28 – 0.28	1.000	-0.11	-0.35 – 0.12	0.339
shelf	0.50	0.21 – 0.78	**0.001**			
logNSI				-0.46	-0.75 – -0.18	**0.002**
log(sdC)				0.36	0.08 – 0.63	**0.013**
mC				-0.37	-0.64 – -0.10	**0.009**
logNSI*log(sdC)				-0.25	-0.47 – -0.04	**0.020**

Observations (N)	40	40
R^2 / R^2-adjusted	0.246 / 0.227	0.584 / 0.536