Coral assemblages and neutral theory

Thesis submitted by

Maria Ana Azeredo de Dornelas

in June 2006

For the Degree of Doctor of Philosophy

in Marine Biology

within the School of Marine Biology

James Cook University

STATEMENT OF ACCESS

I, the undersigned, the author of this thesis, understand that James Cook University will make this thesis available for use within the University Library and, via the Australia Digital Theses network (unless granted an exemption), for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and;

I wish this work to be embargoed until November 2007

Maria Dornelas

27.06.2006

STATEMENT ON SOURCES

DECLARATION

I declare that this thesis is my own work and has not been submitted in any other form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given

Maria Dornelas

27.06.2006

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library is an accurate copy of the print thesis submitted, within the limits of the technology available.

_____ _____

Signature Date

Acknowledgments

A number of people provided priceless support during this project, and I am grateful to them all. Firstly, I want to say a huge thank you to Sean Connolly. Sean was everything I could have wished for in a supervisor, providing invaluable guidance, constant availability, and freedom to disagree. Secondly, I am profoundly grateful to Terry Hughes for his precious advice, for giving me the opportunity to work with his incredible dataset, and for listening to my (sometimes) mad ideas.

I want to thank the staff at the High Performance Computing Section, James Cook University, for all their assistance. Wayne Mallet, in particular, provided great support on using hydra and borg (the two super computers), and even translated some of my MATLAB code into C++, to speed the parametric bootstraps. I am very grateful to Brian McGill and Rampal Etienne for sharing their code and giving helpful advice on fitting neutral models. I also want to thank to Anne Magurran for friendly hospitality in her lab during the final stages of this thesis.

I want to thank everybody that, over the years, participated in the Theoretical Ecology Discussion Group, as well as Laura Castell for listening to and/or reading several versions of this work and providing great feedback. A special thanks to Mia, Ailsa and Matt who also provided feedback, support and company out of the Friday meetings.

I thank H. Cornell, R. Karlson and staff and students of the ARC Centre of Excellence for Coral Reef Studies for collecting, under the leadership of Terry Hughes, the dataset I used in Chapters 3 and 4.

A huge thank you also to Abbi McDonald, Ailsa Kerswell, Alex Kerr, Andrew Baird, Jackie Wolstenholme, Marie Kospartov, Maria Joao Rodrigues, Mia Hoogenboom, Matt Kosnik, Naomi Gardiner and Scott Burgess for helping me collect the dataset I present in Chapter 5. They were not only far more efficient than I could have hoped for, but also lots of fun to work with. A special thanks to Jackie for sharing her precious coral identification skills with me. I'm also very grateful to Anne Hoggett, Lyle Vail, Marianne and Lance Pearce, for providing making the Lizard Island Research Station a great place to do fieldwork.

I have no words to describe how grateful I am to Miguel Barbosa, for all the help he gave me along the way, and for being here for me always. A big, big thank you to Mum and Dad for everything, and to my Grandmother, who made me fall in love with biology.

Finally, I am most grateful to the Fundação para a Ciência e a Tecnologia, Portugal, for supporting me during this PhD.

Abstract

Neutral theory explains patterns of biodiversity based solely on speciation, demographic stochasticity, and dispersal limitation. The validation of this controversial theory depends on empirical support and it has been largely untested in marine communities. Coral assemblages have been repeatedly invoked as the animal communities most likely to conform to the assumptions of neutral theory. This thesis tested the hypothesis that neutral theory explains the macroecological structure of coral assemblages.

Firstly, I assessed whether neutral models can accurately characterise coral species abundance distributions across multiple scales. Simulation-based and analytical neutral models were fitted to a hierarchical dataset of coral species abundance distributions from across the Indo-Pacific gradient of biodiversity. The dataset has three replicate habitats (slope, crest and flat), and three spatial scales (site, island and region). Both models exhibit significant lack of fit to empirical data at the site and island scales, but not at the region scale. The neutral model consistently underestimates the number of rare species, and overestimates the number of common species. Additionally, the neutral model fits coral abundance distributions less accurately than the poisson-lognormal at all scales. Using two formulations of neutral theory, and two goodness-of-fit tests, along with comparisons with the lognormal distribution, ensures that the inferences about coral assemblages and neutral dynamics are robust. Neutral model predictions are consistently and significantly different from observed coral species abundance distributions.

Secondly, I developed a novel test of neutral theory that examines variability between communities of species relative abundances. In neutral communities, species

relative abundances are determined by demographic stochasticity or "ecological drift". Thus, communities diverge through time, and are expected to have low community similarity. In contrast, niche apportionment mechanisms have been invoked to argue that higher levels of community similarity should be observed under niche assembly than under neutral dynamics. These contrasting predictions provide an ideal opportunity to test neutral models against empirical data. Relative abundances of species across local communities differ markedly from neutral theory predictions: coral communities exhibit community similarity values that are far more variable, and lower on average, than neutral theory can predict. Surprisingly, empirical community similarities deviate from the neutral model in a direction opposite to that suggested in previous critiques of neutral theory. Instead, the results support spatio-temporal environmental stochasticity as a major driver of community structure at the macroecological scale.

Thirdly, I unveiled a coral local community species abundance distribution. Community structure patterns are notoriously sensitive to sampling issues, and a comprehensive characterization of such patterns requires extremely large sample sizes. Consequently, the fit of biodiversity models to species abundance distributions, and parameter estimates in particular may be sensitive to sample size. To address these questions, over 44,000 corals were counted and identified to species at an exposed crest in Lizard Island, Great Barrier Reef. A neutral model was fitted to the species abundance distribution of the total dataset, and to sub-samples of various sizes. Parameter estimates and fit of the neutral model at different sample sizes were compared. The unveiled species abundance distribution appears to be multimodal. Parameter estimates are not affected by sample size.

These results strongly indicate that the limited suite of ecological and evolutionary processes included in neutral theory do not suffice to explain diversity patterns in coral assemblages. In combination, the three approaches included in this thesis suggest that neutral theory is most useful as a null model for community structure. Furthermore, the thesis highlights differences in species' responses to environmental fluctuations as a potential major driver of species abundance patterns.

Table of Contents

List of Figures

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 1: General Introduction

Understanding the processes that govern biodiversity has long been one of the central objectives of community ecology (Hutchinson 1959; Whittaker 1972; MacArthur 1975). Since effective conservation efforts depend on understanding how ecological processes affect community structure, there are now also pressing practical implications associated with this endeavour. These questions are more relevant than ever as diversity is currently being lost at a rate unprecedented in the last 65 million years (Chapin et al. 2000). Coral reefs in particular are increasingly threatened, and urgent measures are needed in order to ensure the sustainability of these ecosystems (Hughes et al. 2003; Bellwood et al. 2004). The main objective of this thesis was to contribute to the understanding of the mechanisms that determine coral community structure.

Throughout this thesis the term community is defined as a group of species that are taxonomically similar and compete for resources (i.e., belong to the same trophic group) (Hubbell 2001). For example, on a coral reef, the corals and fishes belong to different communities according to this definition. Community structure, thus defined, refers to the patterns of diversity, and of species' relative abundances. Diversity indices are often used to characterize community structure. These indices aim to summarize community structure taking into account both the number of species in the community and their abundance (Magurran 2004). Diversity, however, is not an univariate linear property of a community, and different indices measure different properties of community structure (Hurlbert 1971). Alternatively, analysing the distribution of species abundance can provide more detailed information about community structure than diversity indices do. Therefore, the analyses in this thesis

focus on patterns of species abundances, particularly on how they vary among locations, as descriptors of community structure.

Two statistical distributions are particularly prominent candidates as good descriptive models for species abundances: the log-series and the lognormal. The log-series is derived from the negative binomial and accommodates a large number of rare species and a decreasing number of increasingly abundant species (Fisher et al. 1943). It arises as the result of random sampling from a community with heterogeneous species abundances (Fisher et al. 1943). The lognormal is a normal distribution in a log scale (Preston 1948). Statistically the lognormal arises as a consequence of the central limit theorem: many different random variables interacting multiplicatively are expected to generate lognormal distributions (May 1975). Preston (1948) also remarked that ecological data are usually a sample from a community, and we seldom have information about entire communities. Preston proposed that limitations of sampling would impose a "veil" on species abundance distributions, so that the rarest species would not be present in the sample. This veil would cause an apparent resemblance with a log-series distribution, which should disappear if sample size is increased. In fact, a sampling model for the lognormal (the Poisson lognormal) produces SADs that vary from log-series-like to lognormal-like depending on sample size (Pielou 1975; Lande et al. 2003a).

Species abundance distributions (SADs) have remarkable similarities across ecological communities. Empirical data usually have SADs that resemble either logseries or lognormal distributions. Communities dominated by a few highly competitive species, with low to moderate diversity, or relatively small samples from a large community often have log-series SADs (May 1975; Hubbell 2001; Magurran

2004). However increasing sample size often reveals an internal mode, as a lognormal-like distribution is unveiled. This "unveiling" has been demonstrated in communities as varied as, for example, birds (Nee et al. 1991), trees (Hubbell 1997b), estuarine fish (Magurran and Henderson 2003), and reef fishes and corals (Connolly et al. 2005). The two distributions are similar in most of their range (Hughes 1984), and therefore fitting a lognormal to samples that do not have an internal mode has been questioned (Hughes 1986). However, quantitative comparisons between the fit of the log-series and the sampling distribution from a lognormal to empirical data can still be made, even in the absence of an internal mode (Connolly et al. 2005). In general, species abundance distributions from a large variety of different communities seem to be well described by lognormal distributions.

The regularity of patterns in species abundance distributions suggests that in general the same processes regulate community structure. The classical theoretical explanation for these patterns is based on species differences in niche (Hutchinson 1959; MacArthur 1960; Sugihara 1980; Tilman 1982; Tokeshi 1990). Each species is adapted to certain environmental conditions, and is most efficient (and therefore has a competitive advantage) in places where the environment is similar to its optimal conditions. The species that live in any particular location must compete for the resources available. Niche theory suggests that the fraction of resources each species consumes is related to its competitive ability (Tilman 1982) and each species consumes a fraction of the resources left available from its superior competitors. Finally if a species' abundance is proportional to the fraction of resource it consumes, the distribution of species abundances can be predicted by this partition of resources (MacArthur 1960).

There are a number of ways that resources can be partitioned, which correspond to different predicted SADs. For example, if each species consumes a fixed fraction of the resources left available from its superior competitors (hierarchical model) the predicted SAD is a geometric series (Motomura 1932 in (Whittaker 1972). Alternatively, the broken stick distribution, which is similar to but more even than the lognormal, arises from the random partition of resources (MacArthur 1960). In this model the resource is partitioned at random, in S fractions of random size, where S is the number of species in the community. Finally, a hierarchical random partition of resources produces the lognormal distribution (Sugihara 1980). Resource partitioning in this model is often referred to as sequentially random, as species are sorted by competitive rank, and each species is sequentially allocated a random fraction of the resources left available by its superior competitors. This last model is particularly relevant, given the near-ubiquity of the lognormal distribution as a statistical description of empirical SADs.

Numerous studies have provided evidence for niche theory. Examples include the comparison of empirical SADs with distributions obtained from the random assortment of species as well as with different niche models (Tokeshi 1990). More sophisticated studies examine multiple patterns, such as the proportionality between species abundances and the fraction of resources they consume when isolated (Tilman 1990; Harpole and Tilman 2006), or between species abundances and niche similarities (Sugihara et al. 2003). The mechanisms that promote species coexistence under niche theory have been thoroughly studied (Tilman and Pacala 1993) and niche assembly is generally considered to be an essential driver of community structure (Chesson 2000). Niche theory has, thus, been the paradigm of community ecology for the past half century.

In spite of this popularity there are some limitations to niche theory. Niche models are static and based on the abstract partitioning of resources. Ideally, theoretical models used to explain patterns of community structure should rest on basic ecological processes and allow exploration of community dynamics through time and space. Yet, developing dynamic models based on species niches has proven extremely difficult: even for communities of modest diversity, the models require far too many parameters (e.g. Schwilk and Ackerly 2005).

An alternative explanation for community structure – neutral theory (Hubbell 1997b; Bell 2000; Bell 2001; Hubbell 2001) – has recently gained increasing attention. Neutral theory is based on demographic stochasticity and dispersal limitation. The neutrality assumption at the core of this theory means that all individuals, regardless of species, are demographically identical. This is in direct contradiction to niche theory, for which community structure is determined by species differences in resource use and local adaptation. Additionally, the neutrality assumption sits uncomfortably within community ecology, much of which is concerned with quantifying and explaining inter-specific differences in demographic rates, abundance, and distribution. Hence, it is not surprising that neutral theory is controversial (Abrams 2001; Brown 2001; Mazancourt 2001; Bengtsson 2002; Dial 2002; Silander 2002; Harte 2003; Ricklefs 2003; Chave 2004; Chisholm and Burgman 2004; Chase 2005; Gaston and Chown 2005). However, unlike niche models, neutral models are dynamic and explicitly incorporate fundamental demographic processes, such as births, deaths, and dispersal. Neutrality may be a reasonable simplifying assumption if inter-specific demographic differences are obscured by intra-specific variability, and thus species differences can be ignored when studying community structure. Most importantly, in spite of its simplicity,

neutral theory can produce community structure patterns that are surprisingly similar to empirical patterns (Hubbell 2001). Consequently, assessing whether neutral theory is capable of explaining observed community structure patterns is critical for the debate between niche and neutral theories. In this thesis, I test how well neutral theory can explain the community structure of scleractinian corals in the Indo-Pacific. As a highly diverse group with limited scope for resource partitioning within communities, scleractinian corals are ideally suited as a test case for neutral theory.

The thesis starts with a review and comparison of the most prominent neutral models that have been explicitly fitted to empirical data (Chapter 2). I show that in spite of differences in ancillary assumptions (i.e., assumptions other than the "core" assumptions that species have identical resource requirements and demographic rates), the neutral models compared predict similar patterns. Thus, conclusions drawn from empirical tests of neutral theory are likely to be robust to the choice of which particular neutral model is used.

Chapter 3 presents the first of three approaches used to test neutral theory: assessing the fit of neutral models to coral SADs across multiple scales, habitats and a biodiversity gradient. In this test I used an extensive data set of coral community composition that is housed at the ARC Centre of Excellence for Coral Reef Studies dataset (kindly provided by T.P. Hughes). The data include three habitats, and five regions distributed along a 10,000 km biodiversity gradient. In this Chapter, two neutral models are used to show that absolute goodness-of-fit tests indicate rejection of neutral theory as an explanation for coral community structure. Additionally, neutral models are rejected in a relative goodness-of-fit comparison with the lognormal.

Tests of macroecological theory based exclusively on curve-fitting to SADs are increasingly criticized as being relatively weak tests (McGill et al. 2006). Such tests do provide an important first test of the ability of a model to reproduce a pattern, but it is often impossible to use these results to link pattern and process, because several models can generate similar patterns (Magurran 2005). In particular, the superiority of the lognormal is difficult to interpret because, unlike neutral theory, it is not a mechanistic model. Therefore, in Chapter 4, I move beyond the analyses in Chapter 3 and present a novel test of neutral theory that is based on between community patterns. I analyse frequency distributions of community similarity and show that coral communities are more variable than neutral theory predicts. The nature of the differences between the data and the model predictions also highlight the role of environmental stochasticity in shaping reef corals community structure.

The shape of SADs is highly sensitive to sampling effort. Extremely large sample sizes are needed to know the true underlying SAD of a community. Thus, the fit of models to these kind of data is likely to be affected by sample size. In neutral models, parameter estimates, and immigration rates in particular, may be affected by sample size. To address these questions, I collected the largest coral species abundance dataset from a single location. In Chapter 5, I unveil a local coral community's SAD, and I determine whether neutral model parameter estimates are sensitive to sample size. The dataset reveals that increasing sample size does not unveil a lognormal distribution. Instead, the resulting distribution seems to be multi-modal. Neutral model analyses also indicate that parameter estimates are robust to variation in sample size.

The thesis finishes with the overall conclusions and a brief discussion of the implications of the results for community ecology and conservation (Chapter 6).

Chapter 2: A Review of neutral models

2.1. The origins of neutral models

Biodiversity neutral theory has its roots in population genetics. Neutral theory was initially developed to explain the dynamics of alleles with equal fitness consequences (neutral alleles) in population genetics (Kimura 1968; King and Jukes 1969). Most mutations are base substitutions that have little or no phenotypic effect, and hence a negligible influence on fitness. Such "neutral" mutations are thus not subject to natural selection, and genetic drift plays the principal role in evolution at the molecular level. Like the alleles in population genetics, in biodiversity neutral models, species are neutral. The neutrality assumption means that competition between individuals follows a lottery process that is independent of species identities. That is, all individuals in a community are competitively identical (use the same resources in the same amounts, and have the same demographic rates). In analogy with genetic drift, biodiversity neutral models propose that demographic stochasticity, in combination with dispersal limitation, drives community dynamics.

Population genetic neutral models (Ewens 1972; Karlin and McGregor 1972; Watterson 1974) were first analysed in the context of community ecology in the 70's (Caswell 1976; Hubbell 1979). Caswell (1976) advocated the use of neutral models as a scale of reference against which empirical patterns should be compared to estimate the effects of biological interactions on community structure. However, because immigrants were assumed to come from an infinite source of species, these models could not generate species abundance distributions (SADs) with the ubiquitous lognormal distribution. In contrast, Hubbell's neutral model (1979) emphasized dispersal limitation and combined neutral dynamics with island biogeography theory

(MacArthur and Wilson 1967). This model generated lognormal-like SADs. However, most ecologists continued to focus on niche partitioning as an explanation for community structure (MacArthur 1960; Whittaker 1972; Sugihara 1980), and thus neutral models were largely ignored. It was not until recently that neutral models were proposed as a general explanation for biodiversity patterns (Hubbell 1997b; Hubbell 1997a; Hubbell 2001) and gained prominence.

Hubbell's (2001) claim that neutral theory was adequate to explain many empirical generalities in biogeography and biodiversity has caused a great deal of interest and controversy (Abrams 2001; Brown 2001; Mazancourt 2001; Bengtsson 2002; Dial 2002; Silander 2002; Harte 2003; Ricklefs 2003; Chave 2004; Chisholm and Burgman 2004; Chase 2005; Gaston and Chown 2005). Since then, there have been numerous theoretical developments and empirical tests of neutral models. In this Chapter I review the literature regarding biodiversity neutral models in general, and I compare the three main theoretical approaches to modelling neutral dynamics in particular. I start by discussing the assumptions of a broad suite of neutral models, and I highlight differences in ancillary assumptions that could potentially give rise to discrepancies in model predictions. Then I focus on the three main modelling approaches that can be, and have been, explicitly fitted to empirical data, and I describe them in some detail. To finalise I compare model predictions, to show that differences are minimal, and thus neutral models can be used interchangeably according to convenience. To conclude I discuss some practical considerations regarding the use of the different models.

2.2. Neutral model assumptions

There are currently several neutral models including both simulation-based (Bell 2000; Bell 2001; Hubbell 2001; Chave and Leigh 2002; Chave et al. 2002), and analytical models (Volkov et al. 2003; Etienne and Olff 2004; McKane et al. 2004; Etienne 2005; He 2005). By definition, neutral models assume that all individuals have the same probability of dying, producing offspring, speciating and immigrating from the metacommunity. Although *per capita* demographic rates are the same for all individuals in a community, species vary in their total mortality, birth and immigration rates: the more abundant a species is, the more likely it is that one individual of this species will provide a local birth or an immigrant, but also the more likely it is to die. In general, neutral models include two spatio-temporal scales: the metacommunity and the local community. At the metacommunity scale, speciation rate and the total number of individuals determine species richness and species abundance. At the local community scale, random deaths, births and immigration determine community structure.

Both Hubbell's (2001) simulation, and Etienne's (2005) analytical neutral models have the additional assumption of community saturation, also known as the zero-sum assumption. This assumption states that every individual that dies is replaced either by a local birth or an immigrant, and therefore community size is constant. Because the number of individuals is unlikely to be constant in most communities, the applicability of neutral theory to unsaturated communities has been questioned (Silander 2002). However, this assumption is not present in some analytical neutral models (McKane et al. 2000; Vallade and Houchmandzadeh 2003; Volkov et al. 2003; McKane et al. 2004; He 2005). Comparing predictions made by

these different model versions should allow inferences about what, if any, is the effect of the saturation assumption.

Neutral models also differ in terms of dispersal dynamics. In most simulation models local communities have complete mixing within them, but dispersal from the metacommunity is filtered through a migration probability (but see Chave and Leigh (2002) for a model with spatially explicit local communities). In contrast, it has been argued that analytical models can be conceptualised as a continuous landscape (Etienne and Alonso 2005), with dispersal limitation occurring at each point in space. Given the importance dispersal limitation plays in neutral models, it is important to understand the effects of these differences in dispersal dynamics for neutral theory predictions.

Finally neutral models differ on the species abundance distribution (SAD) of the source pool of immigrants. Caswell's neutral model has an infinite source of immigrant species (Caswell 1976). Bell's neutral model (Bell 2000) as well as He's neutral model (He 2005) sample immigrants from a distribution with a constant number of species and a uniform distribution. These three models do not make assumptions about how species originate, but most other neutral models (Hubbell 2001; Volkov et al. 2003; McKane et al. 2004) assume that species arise by point mutation (i.e. instantaneously with the initial abundance of one individual). This speciation mechanism generates a logseries metacommunity SAD from which immigrants are sampled. This speciation mechanisms has been criticized as being unlikely to be the predominant type of speciation, and for generating too many species with short life-spans (Ricklefs 2003). Other speciation mechanisms can be

used for neutral dynamics which generate different metacommunity SADs (Hubbell 2001). However, such formulations are not yet available as tractable models.

The neutral models currently available can be classified into three main modelling approaches: simulation models, analytical models based on a mean field approximation, and analytical models based on a genealogical approach. These modelling approaches share the neutrality assumption, but, as discussed above, vary slightly in ancillary assumptions. Here I describe each approach separately, compare the patterns they predict, and discuss the advantages and disadvantages of using each modelling approach to test neutral model predictions.

2.3. Simulation neutral models

Biodiversity neutral models were firstly developed as simulation algorithms. There are several formulations that differ slightly in dynamics (Bell 2000; Bell 2001; Hubbell 2001; Chave and Leigh 2002; Chave et al. 2002). However, the most cited and commonly used simulation neutral model is Hubbell's (2001). This is the simulation model I used in the following Chapters, and thus it is described here in detail (henceforth SNM).

In the SNM, a local community is composed of J individuals. These individuals are initially randomly drawn from a metacommunity with J_m individuals. Metacommunity diversity is determined by both the per-time-step speciation rate (ν) and J_m. Under point mutation speciation, these two parameters always appear combined in the parameter θ (the "fundamental biodiversity number").

If more than one speciation event is allowed to occur at each time step, θ is defined as:

$$\theta = 2Jm\nu \tag{2.1}$$

(Hubbell 2001; Volkov et al. 2003).

Alternatively if only one species can arise per time step, θ is defined as:

$$\theta = \frac{\nu'}{1-\nu'}(Jm-1) \tag{2.2}$$

(Vallade and Houchmandzadeh 2003), where the per time step speciation rate (ν) relates to the per capita speciation rate (ν')

$$\nu = \frac{\nu'}{1-\nu'} \tag{2.3}$$

In the SNM, θ determines the number of species in a sample of the metacommunity, because it corresponds to the probability that an individual belongs to a species not previously sampled. The metacommunity is the source of immigrants for local communities. The distribution of abundances in the metacommunity is similar to Fisher's logseries, and, in practice, immigrants for the SNM are drawn from a large pool (relative to local community size) generated by a sampling algorithm (Ewens 1972; Hubbell 2001).

Local community dynamics occur in discrete time. At each time step, D individuals randomly die. Each empty site is either occupied by an immigrant from the metacommunity with probability of immigration m, or by the offspring of an adult randomly chosen from within the community, with probability 1-m. In simulations,

SADs usually reach a noisy equilibrium after about 50 turnovers of the community (i.e., a total number of births that exceeds J by a factor of about 50) (McGill 2003c). Allowing enough turnovers for communities to stabilize is extremely important, as it affects not only the shape of SADs (McGill 2003c; Chisholm and Burgman 2004), but also the degree of divergence between communities (Maurer and McGill 2004). In this thesis, I adopt a conservative approach and run simulation models for 500 turnovers of the community (50 000 time steps with D = 1% of J). The shape of the SAD is determined by J, m and θ, and is not affected by the size of the metacommunity (J_m) or the number of individuals replaced at each time step (D).

2.4. Mean field neutral model

The first analytical approach to biodiversity neutral models to be developed applies mean field approximations to neutral models (McKane et al. 2000; McKane et al. 2004). Instead of examining the interactions among the species that compose a community individually, this approach assumes that a species interacts with all the other species in the community combined. Grouping all but the focal species greatly simplifies the dynamics involved, as only the interactions between the focal species and the rest of the community need to be taken into account, instead of the interactions between each pair of species in the community. Thus, grouping the effects of all the other species (using the "mean field approximation") allows inferring an analytical expression for expected SADs. Several authors have followed this approach with equivalent results (McKane et al. 2000; Vallade and Houchmandzadeh 2003; Volkov et al. 2003; Alonso and McKane 2004; McKane et al. 2004). Here, and in subsequent Chapters, I use Volkov's (2003) terminology (henceforth MFNM).

For the MFNM the probability that a species has abundance n (p_n) is:

$$p_n = \frac{\theta}{S}\frac{J!}{n!(J-n)!}\frac{\Gamma(\gamma)}{\Gamma(J+\gamma)}\int_0^x \frac{\Gamma(n+x)}{\Gamma(1+x)}\frac{\Gamma(J-n+\gamma-x)}{\Gamma(\gamma-x)}e^{(\frac{-x\theta}{\gamma})}dx \qquad (2.4)$$

where θ, J and m are as defined for the SNM, S is the number of species, Γ is the gamma function, and $\gamma = m(J-1)/(1-m)$ (Volkov et al. 2003). The expected number of species with abundance n is equal to p_nS.

2.5. Genealogical neutral model

An alternative analytical approach returns to neutral theory's roots on population genetics, specifically to Ewens sampling formula (Ewens 1972). It is based on developing the genealogical tree of the local community by tracing back each individual to its ancestor that immigrated from the metacommunity (Etienne and Olff 2004; Etienne 2005; Etienne and Alonso 2005) (henceforth GNM). In the GNM, the probability of observing a SAD in a sample from a local community with parameters θ, m, J is:

$$P[D \mid \theta, m, J] = \frac{J!}{\prod_{i=1}^{S} n_i \prod_{j=1}^{J} \Phi j!}\frac{\theta^S}{(I)_J}\sum_{A=S}^{J} K(D,A)\frac{I^A}{(\theta)_A} \qquad (2.5)$$

(Etienne 2005) where θ, m, J and S are as defined for the SNM and MFNM, n_i is the number of individuals of species i, Φ_j is the number of species with abundance j. The notation $(a)_b$ is the Pochhammer symbol, or rising factorial, also denoted as a^b which is defined as:

$$(a)_b = a(a+1)(a+2)...(a+b-1) = \frac{(a+b-1)!}{(a-1)!} \quad (2.6)$$

I is related to m in the same way as θ is related to ν'

$$I = \frac{m}{1-m}(J-1) \quad (2.7)$$

$K(D,A)$ is

$$K(D,A) = \sum_{a_i=1}^{n_i} \prod_{i=1}^{S} \frac{\bar{s}(n_i,a_i)(a_i-1)!}{(n_i-1)!} \quad (2.8)$$

where a_i is the number of ancestors of species i, and $\bar{s}(a,i)$ is the absolute value of a Stirling number of the first kind (or the ith coefficient of the falling factorial $a^{\underline{b}}$):

$$a^{\underline{b}} = a(a-1)(a-2)...(a-b+1) \quad (2.9)$$

The sum of all the a_i must equal A, the total number of ancestors in the local community. To fit the GNM, the combination of m and θ that minimizes expression (2.5) is found, rather than calculating log-likelihoods for each species abundance separately and then summing them. Although an analytical expression exists for the expected SAD in this model (Etienne and Alonso 2005), the expected SAD can also be obtained by using Hoppe urns to simulate samples from a distribution with a certain parameter combination (Etienne 2005). Hoppe urns work much like the metacommunity simulation algorithm in the SNM, where each individual in the sample is given a species label according to probabilities determined by θ and m. However, the GNM's Hoppe urn represents a dispersal-limited community at equilibrium, rather than a fully mixed metacommunity. One advantage of the Hoppe

urn approach (over the analytical formula) is that, as with the SNM, a number of simulations can be run to estimate the variance of the number of species (across simulations) in each abundance class, as well as the mean.

2.6. Comparison of the three main neutral models

Discrepancies between predictions of different neutral models are of great importance, because they can affect the outcome of empirical tests. In particular the differences in ancillary assumptions previously discussed can potentially generate variations in predictions that are independent of the fundamental assumption of neutrality. Apparent discrepancies between different neutral model versions have previously been reported between the SNM and MFNM (Chisholm and Burgman 2004) and between the MFNM and the GNM (Etienne and Alonso 2005). Here I review the extent of these discrepancies, and show they can be resolved by small adjustments in the models. For completeness, and to illustrate the patterns predicted by neutral models, I present also a comparison of SADs generated by the SNM and the GNM.

The SNM has been suggested to predict a lognormal-like distribution for parameter combinations (low m and medium to high θ) for which the MFNM predicts a low diversity flat SAD (Chisholm and Burgman 2004). This discrepancy is very important because the SNM's ability to generate lognormal-like distributions has repeatedly been invoked as one of the strengths of neutral models (Hubbell 1997b; Hubbell 1997a; Hubbell 2001; Volkov et al. 2003). However, it seems that this apparent discrepancy is an artefact of failing to allow time for simulated communities to equilibrate. When the simulations described above are run for sufficiently many

community turnovers, the SNM's SADs do converge to the MFNM's predictions of flat SADs (Chisholm and Burgman 2004). For extremely large J (~100,000), both the SNM and the MFNM do generate a lognormal-like SAD (Hubbell and Borda-De-Agua 2004). Hence, provided the distributions are allowed to stabilize, the SNM and MFNM seem to generate equivalent SADs.

The MFNM has also been suggested to differ from the GNM. The MFNM is based on an approximation, which assumes Jm to be infinitely large. SADs predicted by the GNM have been reported to have fewer species than those predicted by the MFNM, when $Jm < \infty$ (Etienne and Alonso 2005). However, SADs are indistinguishable if $Jm >> J$ (Etienne and Alonso 2005), which is the only realistic scenario, given that a metacommunity is composed of many local communities. Furthermore, Jm also does not affect within (Hubbell 2001) and between community patterns in the SNM (Chapter 5) as long as it is sufficiently larger than J. Thus, differences between the MFNM and the GNM caused by Jm are also easily resolved by adjusting the models to biologically meaningful scenarios.

SADs predicted by the SNM, and GNM vary similarly with spatial scale and parameter values. A log-series distribution is predicted for the metacommunity (Hubbell 2001; Etienne 2005). For the local community, SADs vary considerably with the three parameters: m, J, and θ (Hubbell 2001; Fig 2.1, 2.2 and 2.3). Isolation decreases species richness (the height of the bars in SADs) and the proportion of rare and abundant species in local communities (the shape of SADs, Fig 2.1). As m decreases rare species go locally extinct, and abundant species become more abundant. This is reflected by the modal class moving to the right on a SAD plot. In particular, extremely low immigration rates (m = 0.0001) lead to communities with a

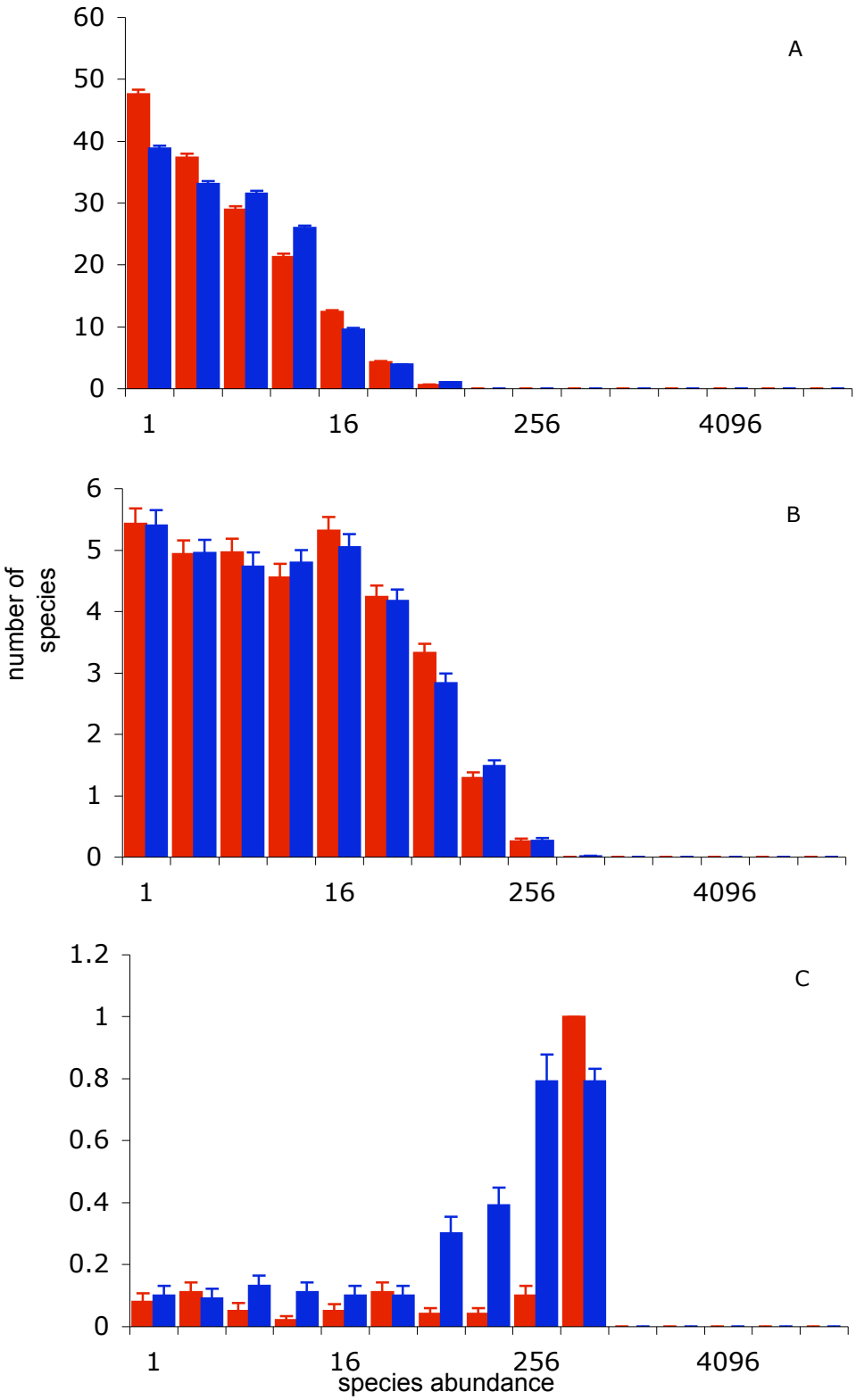


Fig. 2.1 - SADs predicted by the SNM (blue) and GNM (red) for a J of 1000, θ of 50 and m of 0.9 (A), 0.01 (B) and 0.0001 (C). Bars represent the mean of 100 simulations and error bars one standard error.

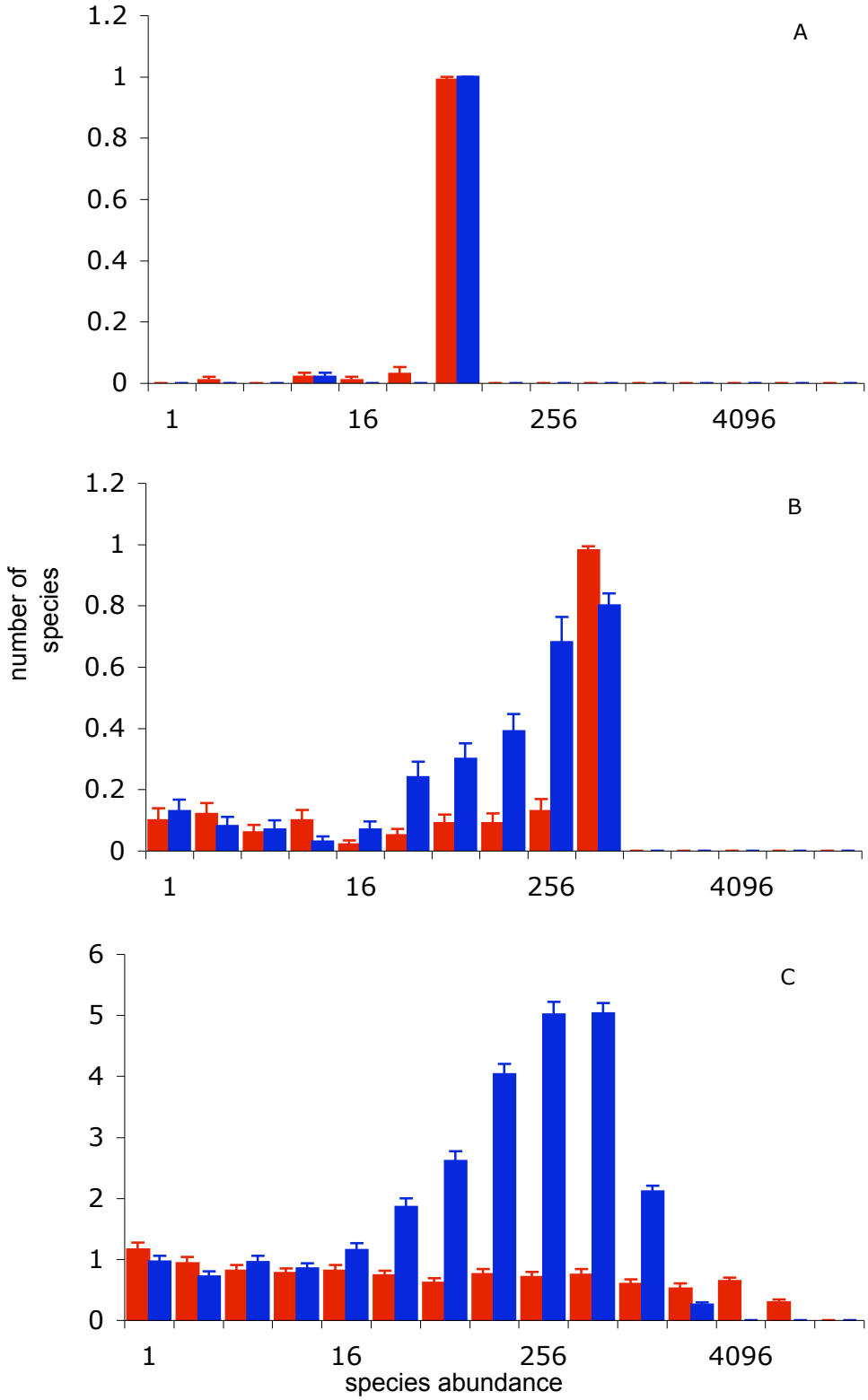


Fig. 2.2 - SADs predicted by the SNM (blue) and GNM (red) for a θ of 500, m of 0.01 and J of 100 (A), 1000 (B) and 10000 (C). Bars represent the mean of 100 simulations and error bars one standard error.

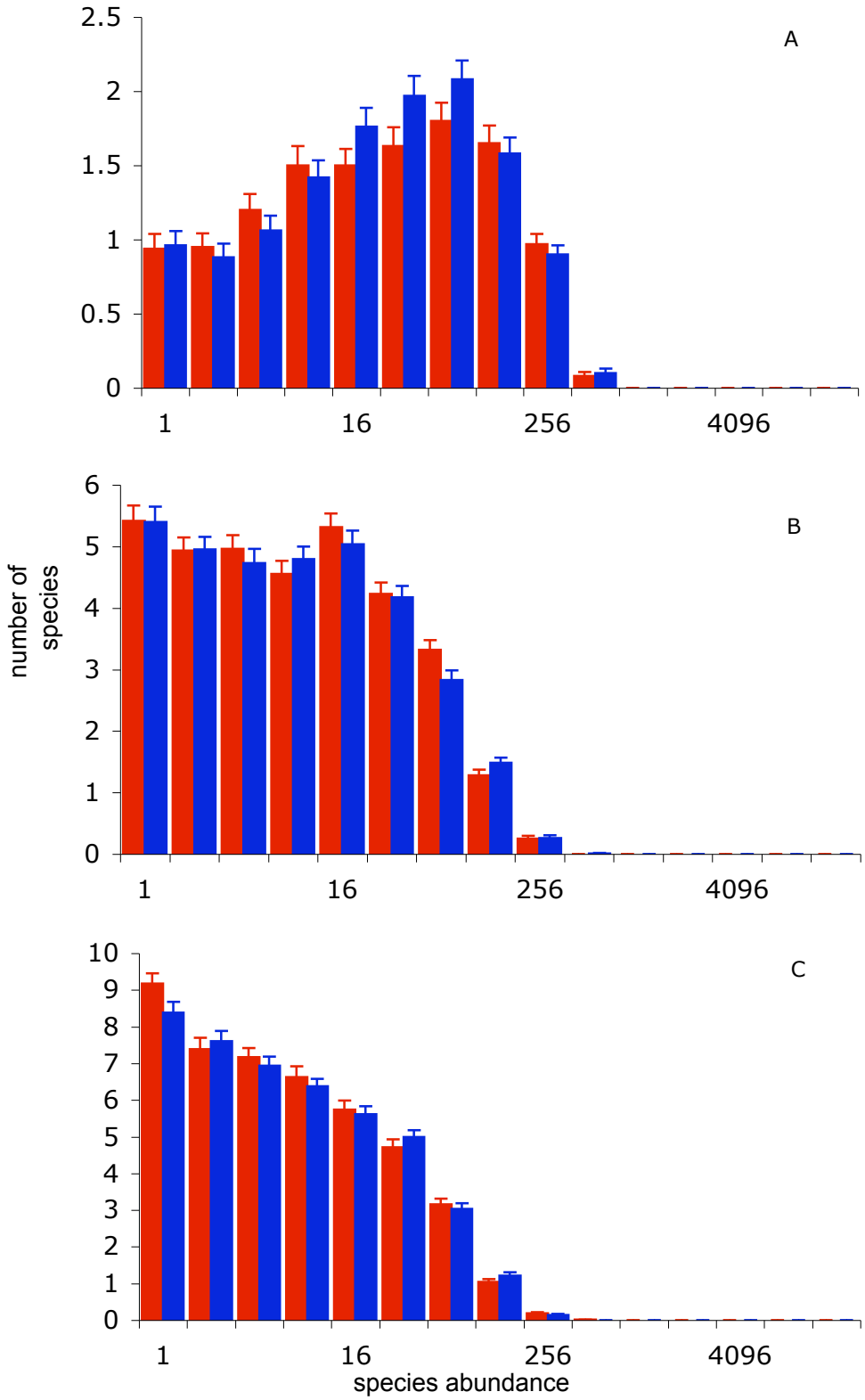


Fig. 2.3 - SADs predicted by the SNM (blue) and GNM (red) for a J of 1000, m of 0.01 and θ of 5 (A), 50 (B) and 500 (C). Bars represent the mean of 100 simulations and error bars one standard error.

single dominant species (Fig 2.1 C). However, the degree of isolation that leads to mono-dominance depends on J (Fig 2.2), as larger communities can sustain more species for a certain immigration rate. Additionally, higher diversity at the metacommunity scale (higher θ) leads to higher species richness at the local community scale (Fig 2.3). This effect is more pronounced for rare species, and thus, for constant m and J, θ affects not only the height of the distribution, but also the location of the internal mode.

The SNM and the GNM modelling approaches generate similar predictions for most parameter values (Fig 2.1 to 2.3). The only apparent exception is for very low m and relatively high J (Fig 2.2 B, C), for which the GNM predicts a flat SAD with very few species, and the SNM a left-skewed lognormal-like distribution. This is analogous to the previously-reported differences between the SNM and the MFNM (Chisholm and Burgman 2004). However, the predicted abundance distributions for the SNM, like those shown in Fig. 2.2 B and C, are transient. When the SNM is run for more community turnovers, the differences are resolved because the SNM loses its lognormal-like shape as it tends to monodominance (Chapter 5). Nevertheless, this discrepancy highlights the importance of verifying that simulations have stabilized before making comparisons with empirical data. Thus, although SADs vary considerably with model parameters, the variability is consistent within these two model versions, and any differences are well within the variance inherent to simulations. As both the SNM and the GNM have been shown to be equivalent to the MFNM (see above), tests of neutral models using any of the three versions should yield similar results.

2.7. Practical considerations

The SNM relies exclusively on simulations. Therefore, this model version is the most susceptible to uncertainty, both in terms of the variation between simulation runs, and the stabilization (through time) of the patterns generated. When this model is fitted to empirical data, to reduce uncertainty related to simulation noise, abundances are usually classified into octaves (log2 classes of abundance). Multiple simulations are run, so that the expected SAD used for fitting uses the mean number of species (across simulations) in each abundance class to reduce susceptibility to simulation variation. When fitting the SNM to empirical SADs J is assumed to be equal to the sample size, and for each SAD, θ and m are sequentially estimated by Maximum Likelihood methods using Hubbell's (2001) sequential estimation procedure. All of this, however, makes the SNM extremely computationally intensive. Nevertheless the SNM is the most flexible of all neutral models, as different ancillary assumptions can easily be incorporated with small changes to the simulation algorithm. Thus, it is extremely useful as a means to identify areas in which to invest analytical effort, and to examine combinations of assumptions that are not analytically tractable. It also facilitates quantifying the variance associated with neutral dynamics. Finally, some aspects of community structure can be readily examined by simulation (e.g. community similarity, as in Chapter 4), but cannot yet be examined with analytical versions of the theory.

The MFNM and GNM avoid uncertainty in parameter estimates related to the stochastic fluctuations inherent to simulations. This increased accuracy allows them to be fitted to un-binned species abundances, instead of octave-classified abundances. Thus the MFNM and GNM provide more sensitive tests of goodness-of-

fit, and allow more accurate analyses of deviations between model predictions and observed SADs. Because the MFNM is essentially a sampling theory (Alonso and McKane 2004), when fitting this model J is the sample size, and m is the only estimated parameter, which is estimated by Maximum Likelihood methods. For every value of m attempted in the fitting procedure, p_nS is solved for θ, by constraining the integral of the expected number of species with abundances between 0 and J so that it is equal to S. Thus, the MFNM has only one estimated parameter. However, because the integral in expression (2.4) must be solved numerically, this modelling approach is still extremely computer intensive and susceptible to numerical error (McGill et al. 2006), as well as difficult to implement. Furthermore, it is based on a approximation, and thus is not an exact analytical solution. In contrast, calculating the likelihood for the GNM, is far less computationally intensive than for the other neutral models. The GNM's numerical efficiency allows more comprehensive explorations of likelihood surfaces, as well as efficient estimation of confidence intervals for maximum likelihood parameter estimates.

2.8. Conclusions

Biodiversity neutral models all share the fundamental neutrality assumption, although ancillary assumptions vary to some extent between different model versions. However, previous findings of differences between predictions of the models appear to be due to failure to reach equilibrium in simulation models. Thus the "best" formulation of neutral theory to use can be based on its suitability for the aims of a particular study, rather than on the assumptions of the models. In Chapter 3, I used the SNM and the MFNM to test whether neutral models can predict coral SADs across

multiple scales, habitats and a biodiversity gradient. This study was completed before publication of the tractable form of the GNM (Etienne 2005). However, the results presented above strongly suggest that the findings in Chapter 3 would not have been different if I had used the GNM model instead. In Chapter 4, I examine patterns of community similarity, and I use the SNM because the analytical models currently do not make predictions regarding between community patterns. In Chapter 5, I exploit the computational efficiency of the GNM to examine parameter estimates for a coral community, and how these are affected by sample size. Again, the equivalence of the different neutral models shown in the present Chapter supports the robustness of the results in Chapter 5 to the choice of GNM.

Chapter 3: Coral species abundance distributions – a multi-scale test of neutral theory

3.1. Introduction

Neutral theory (Hubbell 2001) explains community structure based on the assumption that all individuals, regardless of species, are demographically identical. Hence, macro-ecological patterns are driven by speciation, demographic stochasticity, and dispersal limitation, with demographic and ecological differences between species having a comparatively negligible effect. The neutrality assumption contradicts much of classical ecological theory, which explains biodiversity and community structure based on local adaptation, and inter-specific differences in demographic rates, competitive abilities, and resource use (Hutchinson 1959; MacArthur 1960; Sugihara 1980; Tilman 1982; Tokeshi 1990). The assumption of neutrality has been branded as obviously "wrong" (Brown 2001; Mazancourt 2001; Baker 2002; Enquist et al. 2002; Norris 2003), because niche differences are likely to be essential to determine species coexistence (Chesson 1991). However, neutral theory should ultimately be judged based on its ability to explain empirical data.

Coral assemblages are particularly suited to test neutral theory predictions. Most ancillary assumptions of neutral theory are appropriate for the biology of corals. In fact, the theory was originally proposed for tropical forests (Hubbell 1979; Hubbell and Foster 1986), which are often compared to coral reefs in terms of their high diversity. Neutral communities are, by definition, composed of trophically equivalent species, as is generally the case for coral assemblages. The life cycle of individuals in the neutral model includes an adult stage fixed in the same local community, and a dispersive reproductive stage, connecting different local communities into a

metacommunity. This closely matches reef corals' life cycle. Space (and the associated access to light) is the primary limiting resource, and there is a strong competitive advantage to incumbent space occupants. Therefore it is not surprising that coral assemblages have been repeatedly postulated as ideal for testing neutral theory predictions (Hubbell 1997b; Hubbell 1997a; Whitfield 2002; Chave 2004; Williamson and Gaston 2005).

Neutrality, however, is a highly controversial premise, since there is little empirical evidence for competitive equivalence (Abrams 2001). However, equalising processes, such as priority effects and life-history trade-offs, have long been recognized as mechanisms of coexistence (Chesson 2000). From this perspective, it has been argued that neutrality approximates this kind of unpredictability in competitive interactions (Chave 2004), due, for instance, to intra-specific variability in competitive ability (Buss and Jackson 1979; Connolly and Muko 2003). Hence, neutrality has been argued to be a reasonable simplifying assumption when analysing community structure (Hubbell 2001). However, support for this proposition depends on the ability of neutral models to predict observed patterns.

Species abundance distributions (SADs) are one of the most important ecological predictions generated by neutral theory (Chapter 1). The classical statistical model for SADs is the lognormal distribution, and it has been shown to provide good fit to coral abundance distributions (Connolly et al. 2005). Initial support for neutral theory was drawn from its apparent ability to characterise community structure in empirical communities better than the lognormal (Hubbell 2001; Volkov et al. 2003), although subsequent tests have generated contradictory results (McGill 2003c). In this

Chapter I test the goodness-of-fit of neutral models to coral SADs, and compare it to the fit of a poisson lognormal (Connolly et al. 2005).

Support for ecological hypotheses typically varies with scale (Levin 1992). This is particularly true for neutral theory, where defining the two spatio-temporal scales (local community and metacommunity) in ecological and evolutionary time-frames is crucial to ensure appropriate testing (Hubbell 2001; McGill et al. 2006). Local adaptation and habitat heterogeneity may be important at some scales but not others, and hence support for neutral theory may vary with the scale at which it is tested (McGill et al. 2006). However, most empirical tests of neutral theory have focused on a single spatial scale. Here I analyse coral SADs across multiple scales: with local communities examined at the scale of sites (1-2 km), islands (10-100 km) and regions (>1,000 km).

In this Chapter I test whether neutral theory is consistent with observed coral SADs, and I compare neutral theory's fit to the data with that of the lognormal. The tests are done at three different spatial scales, and for coral assemblages from across a 10,000 km biodiversity gradient, and three reef habitats (reef flat, crest and slope). Previous empirical tests using a single forest dataset have reached opposite conclusions depending on which model version and which statistical tests are used (McGill 2003c; Volkov et al. 2003). Hence, both the SNM (simulation neutral model – Chapter 2 (Hubbell 2001)) and the MFNM (mean field neutral model – Chapter 2 (Volkov et al. 2003; McGill et al. 2006)) are fitted to the data. I conduct a comprehensive analysis of relative abundance patterns, to understand how the absolute and relative performance of neutral theory depends upon spatial scale. I develop a new goodness-of-fit test for neutral theory that is based on the actual unit of

ecological sampling (i.e., individuals), rather than on species (as conventional tests assume), and I apply it to these data. I also test the fit of the MFNM relative to the lognormal, using model selection statistics. The combination of fitting multiple neutral models, and of testing both absolute and relative fit at multiple spatial scales, and for assemblages from multiple habitats, and across a biodiversity gradient makes this study a particularly robust test of neutral theory.

3.2. Methods

3.2.1. Data collection

Coral species abundances were measured at several locations from across the Indo-Pacific to examine coral community structure patterns and to describe diversity and biogeographical trends. Sampling followed a hierarchical design with three spatial scales: regions, islands and sites. The five regions - Indonesia, Papua New Guinea, Solomon Islands, American Samoa and French Polynesia - are distributed along the Indo-Pacific gradient of coral biodiversity. Three high islands were selected in each region and four sites were chosen at each island (Karlson et al. 2004). Abundance of a species can be measured by the number of individuals, or by their biomass. In neutral models all individuals are implicitly assumed to be the same size, and hence the two measures are equivalent. However, in real communities, this is not the case, and different species can have dramatically different sizes. In fact the distribution of body sizes is one key macroecological pattern for which explanations are currently being sought (Brown 1999). This is particularly problematic in the case of colonial organisms, like corals, where individuals can be defined as a colony, or as the units that compose the colony. Furthemore, coral SADs using numbers of

individuals are strikingly different from SADs using colony cover (which is a proxy for biomass) (Connolly et al. 2003). In this study, each colony was counted as a single individual, so that all individuals originated from sexual reproduction, rather than by colony growth. This reflects more closely the type of lottery competition inherent to neutral models. At each site, all of the coral colonies intercepted by ten 10m long haphazardly placed transects were counted and identified to species. This was repeated in each of three reef habitats: slope, crest and flat using a total of 1800 transects. The different habitats were treated separately because their species composition is highly differentiated and neutral theory assumes homogeneous habitat. To generate island-level SADs, samples were pooled by summing the abundances of each species across the four sites on each island. Similarly, summing the abundances of each species across the 12 sites in each region created region-level SADs. A total of 60 site, 15 island, and 5 regional SADs were obtained for each habitat.

3.2.2. Testing the goodness-of-fit of the neutral model

To determine if the neutral model can accurately describe coral SADs we tested the goodness-of-fit of the SNM and the MFNM to the data at each of the three spatial scales. The MFNM was fitted to each of the SADs by finding the value of m that maximizes the log-likelihood:

$$LL = \sum_{i=1}^{J} O_i \log(E_i / S) \tag{3.1}$$

Where O_i is the observed number of species with abundance i, E_i is the expected number of species with abundance i, and S is the total number of species. Expected SADs were obtained using expression (2.4) in Chapter 2 (Volkov et al. 2003), using MATLAB code kindly provided by BJ McGill (McGill et al. 2006).

Algorithms for fitting the MFNM are notoriously prone to numerical error, and, for this code, some numerical errors were present that led, in some cases, to dramatic underestimates of the probability of observing a species with a certain abundance. Therefore, I modified the code to use a slower, but more accurate function, in order to be able to confidently fit the model to species abundances rather than octave classified abundances. For this reason both the SNM and the MFNM were extremely computationally intensive (requiring several months on a supercomputer composed of 16 500 MHz MIPS R14000 processors and 68 400 MHz MIPS R12000 processors, housed at the High Performance Computing Section of James Cook University).

I tested the goodness-of-fit of the neutral models by comparing the observed deviances with the corresponding expected deviances under the null hypothesis of neutral dynamics. Observed model deviance is calculated as:

$$d = -2(LL - LL_s) \tag{3.2}$$

where LL is the log-likelihood of the best fitted MFNM to the data, and LL_s is the log-likelihood of the saturated model (Burnham and Anderson 1998). The saturated model follows the observed distribution exactly, that is the probability of observing a species with abundance i in the saturated model is equal to the proportion of species with abundance i in the data. Total deviance is the sum of deviances for all replicates at each spatial scale and each habitat. Expected deviance is estimated by implementing a parametric bootstrap: species are randomly sampled from a theoretical distribution of abundances with the same parameters as the data, and the model is fitted to these simulated samples. This procedure yields a null distribution of deviances. The proportion of bootstrap replicates with a deviance higher than the observed deviance (p) is a measure of the probability that goodness-of-fit of the

model is acceptable. If p is low, then it is unlikely that the log-likelihood of the data is typical of a neutral community. Following convention, we take $p<0.05$ as our critical threshold value for rejecting the neutral model.

Traditional goodness-of-fit tests for SADs, such as chi-square tests, treat species as the units of sampling, when in fact individuals are being sampled, and this is also true of the procedure described above. The SNM allows me to avoid this problem by simulating communities with the same number of sampling units (individuals) as the data. The SNM was fitted to each of the octave-classified SADs by Maximum Likelihood sequential estimation of θ and m (Hubbell 2001). I developed a method based on parametric bootstrapping to test whether the goodness-of-fit of the model to the data is worse than would be expected if a community were undergoing truly neutral dynamics. 1000 SNM datasets were simulated with the parameters estimated for the data. For each of the simulated datasets and each spatial scale the global goodness-of-fit (GGOF) was calculated as:

$$GGOF = \sum_i LL_i \tag{3.3}$$

where LL_i is the log-likelihood of *ith* SAD. As a result, I obtain an expected distribution of GGOF statistics, under the null hypothesis that the data were generated by neutral dynamics. Thus, the proportion of simulated datasets with a log-likelihood more negative than the data's (p) is an estimate of the probability that the goodness-of-fit of the data is consistent with a community undergoing neutral dynamics. Note that the goodness-of-fit statistic for the simulated data sets was calculated relative to the best-fit distribution for the empirical data, rather than from a best-fit calculated by fitting the SNM to each simulated abundance distribution separately (which would

have been computationally prohibitive). The effect of this more exhaustive process would have been to increase the goodness-of-fit of the simulated data sets (i.e., the null distribution of fit statistics), without changing the fit of the actual data set. This would in turn increase the likelihood of rejecting the neutral model. Thus, this test is conservative with respect to rejecting the neutral model.

3.2.3. Assessing the level of parameterisation of the MFNM

To account for the effects of parameter uncertainty on the lack of fit, the MFNM was fitted separately for each site, and I also fit reduced parameter models. Specifically, the parameters were constrained to be constant for the entire dataset, for all sites within a region, and all sites within an island (Connolly et al. 2005). For each constrainment scale, the model was fitted by finding the value of m that maximized the sum of the log-likelihoods (expression 3.1) of all the sites included in that constrainment scale. Similarly, I fitted the MFNM for each island and for each region separately, and with constant parameters for the entire dataset and at the region scale, in the case of the islands. The relative fit of the different levels of parameterisation was compared using Akaike's Information Criterion (*AICc*) (Akaike 1985):

$$AICc = -2MLL + 2p + \frac{2p(p+1)}{n-p-1} \qquad (3.4)$$

Where MLL is the maximum log-likelihood, p is the number of parameters of the model, and n is the sum of the number of observed species abundances in each SAD. The estimated best model is the model with the lowest *AICc*. To quantify the uncertainty associated with model selection, Akaike weights were calculated:

$$W_i = \frac{e^{-\Delta_i/2}}{\sum e^{-\Delta_j/2}} \quad (3.5)$$

where Δ_i is the difference between the *AICc* of model *i* and the lowest *AICc* of the models being compared. *W* estimates the probability that a model is the best among the models being compared. Aggregate comparisons were made by summing the *MLL*s and computing *AICc* and *W* using the corresponding (total) number of parameters and observations.

3.2.4. Comparing the neutral model with the Poisson Lognormal

To test the relative goodness of fit of neutral models in comparison with other available models, the fit of the MFNM was compared with the Poisson lognormal. The best level of parameterisation of the two models was used in this comparison, as the Poisson lognormal is best parameterised with constant parameters at the region scale (Connolly et al. 2005). The two models were compared using *AICc*, as described above. Because the SNM was fit to binned abundances, its likelihoods were not comparable to those of the other models, so it was not used in this analysis. Although the use of *AICc* for selection between different models is superior to approaches that rely on arbitrary *P*-values, its use is nonetheless somewhat controversial (Boik 2004). Therefore, I also quantify and examine the deviations between the data and the predictions of each model. For this analysis, I use both MFNM and SNM, as well as the poisson lognormal.

3.3. Results

3.3.1. Testing the goodness-of-fit of the neutral model

Observed and fitted SNM and MFNM SADs are presented in Appendix I, Figs A.I.2-37. Parameter estimates for θ and m were similar for the SNM and the MFNM (Fig 3.1). Estimates of the diversity parameter θ increased predictably with increasing spatial scale, across the biodiversity gradient and the three reef habitats (Appendix I, Fig. A.I.1). θ varied between 1.3 (on a reef flat at a single site in Samoa) and 54.3 (on reef slopes at the regional scale in Indonesia). Estimates of θ with the MFNM were extremely well predicted by estimates with the SNM (Fig 3.1 A, R^2 of 0.9914, 0.9949 and 0.9887 for the site, island and region scales respectively), although the MFNM had slightly higher estimates (see figure legend for regression equations). Estimates of m were consistently high. Over 50% were above 0.8, over 90% above 0.7, and all estimates were above 0.25. However, regressions between the estimates of m with the MFNM and the SNM had lower R^2 (0.5599, 0.1692 and 0.695 for the site, island and region scales respectively), mostly because of the cases in which either one model or the other had m estimates below 0.8.

The species parametric bootstrap analysis shows the MFNM exhibits significant lack of fit at the site and island scales ($p < 0.001$ in both cases). In contrast, at the region scale, the goodness-of-fit of both neutral models is low, but within expected values for a neutral community (MFNM $p = 0.1625$). Results from the individual parametric bootstrap are entirely consistent with the species parametric bootstrap. At the site and island scales the SNM exhibits significant lack of fit, whereas at the region scale goodness-of-fit is low, but within expected values (site and island $p < 0.001$, region $p = 0.187$, Fig 3.2).

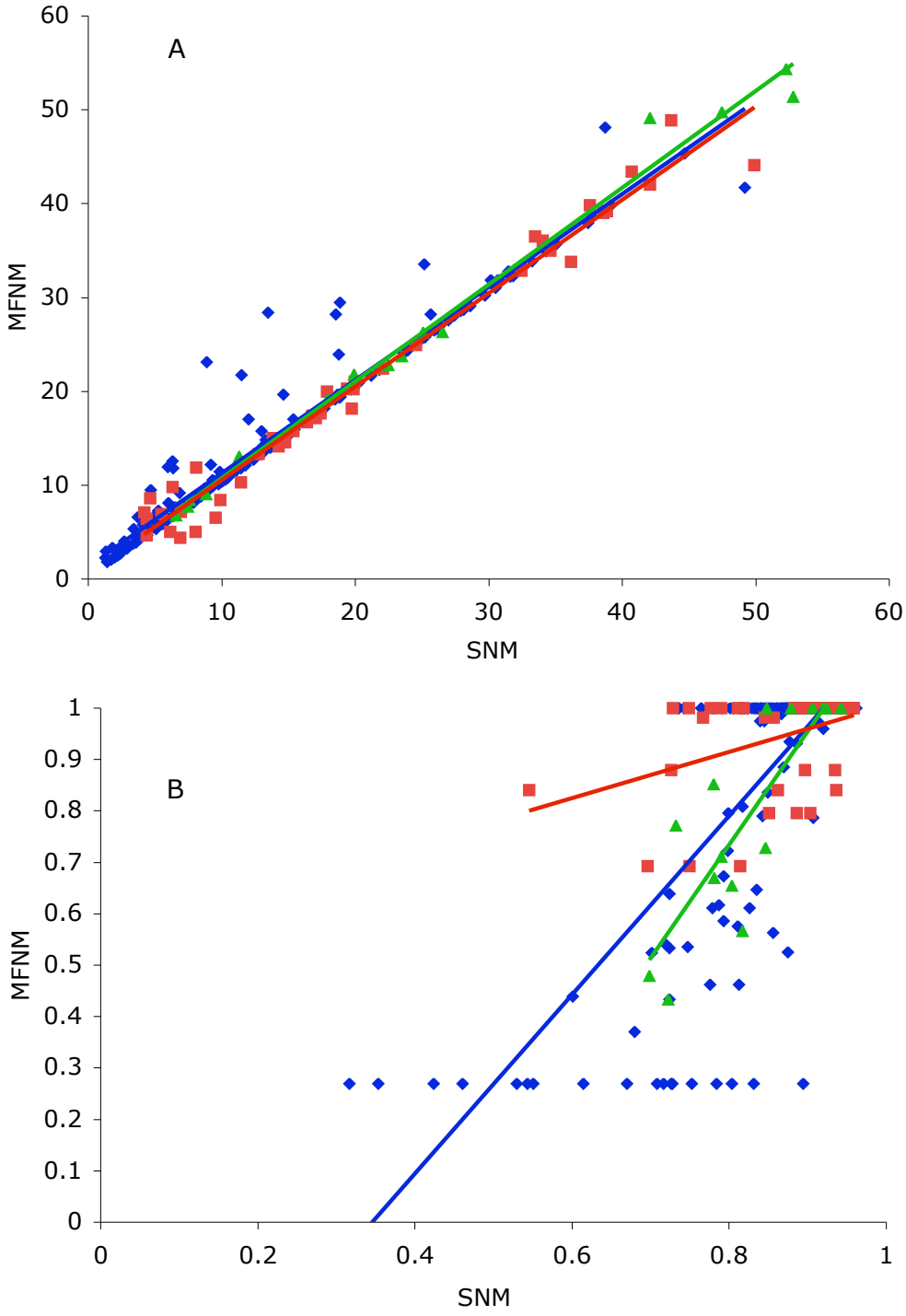


Fig 3.1 - Parameter estimates from the SNM and MFNMfor θ (A) and m (B). Blue diamonds represent sites, red squares islands, and green triangles regions. Best fitted regressions are respectively θ_{MFNM}= 0.9914*θ_{SNM}+1.3857 and m_{MFNM}=1.7402*m_{SNM}-0.6016 at the site scale, θ_{MFNM}= 0.9949*θ_{SNM}+0.0670 and m_{MFNM}=0.4502*m_{SNM}+0.5549 at the island scale and θ_{MFNM}= 1.0321*θ_{SNM}+0.3966 and m_{MFNM}=2.2053*m_{SNM}-1.0304 at the region scale. See text for R^2 values.

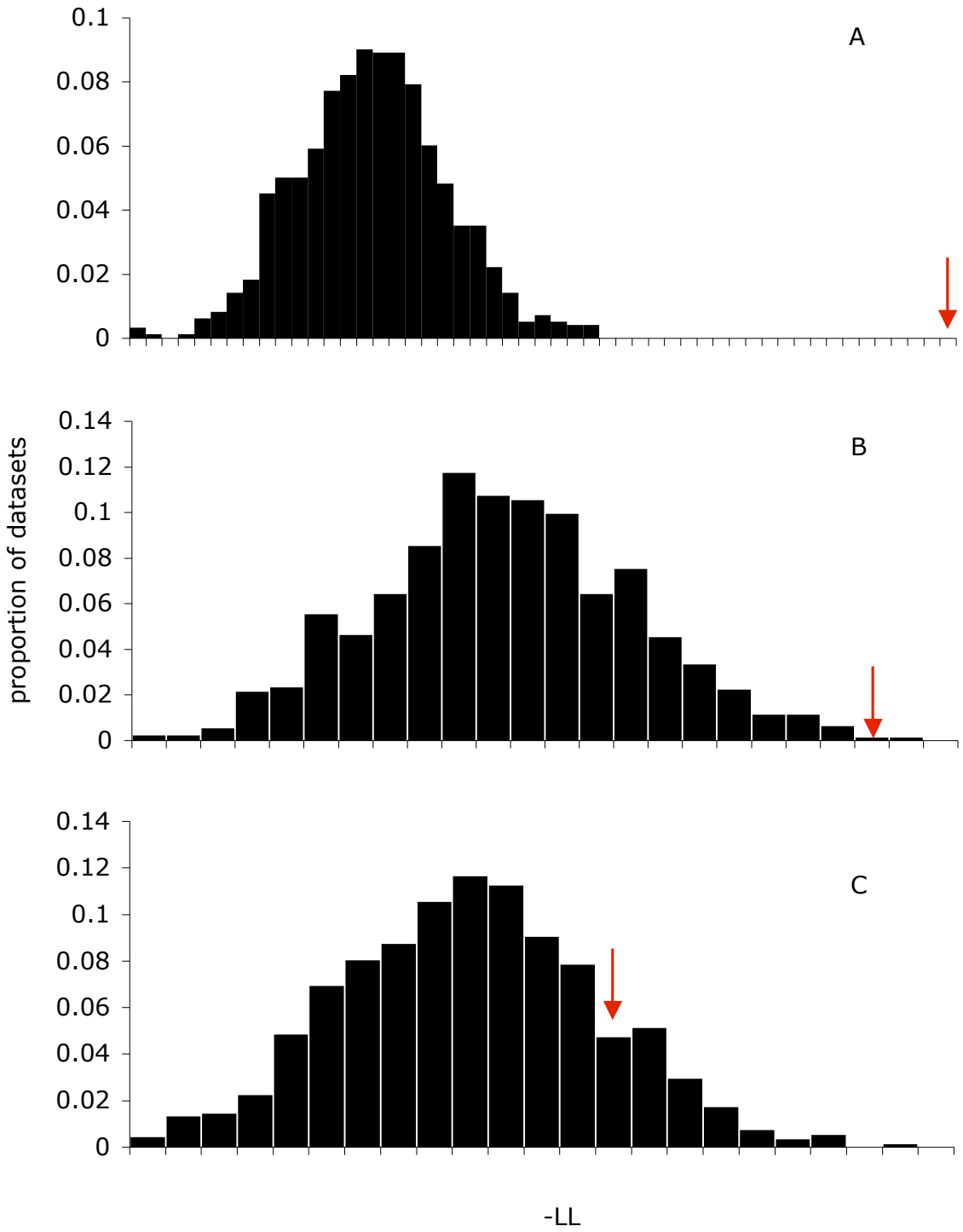


Fig 3.2 - Frequency distribution of negative Log Likelihood (-LL) of simulated neutral datasets at the site (A), island (B), and region (C) scales. The arrow marks the –LL of the empirical data. The –LL of the empirical data is much higher than any simulated datasets at the site and island scales, but not at the regional scale.

The cause of the lack of fit at site and island scales is apparent in a plot of the deviations between observed and best-fit neutral SADs (Fig 3.3). In the data there are more rare species, and fewer common species than predicted by the neutral model at the site and island scales, but the differences are smaller at the regional scale. The disparity between observations and predictions is observed both with the SNM and the MFNM (Fig 3.3).

3.3.2. Comparing the neutral model with the Poisson Lognormal

As for the Poisson lognormal (Connolly et al. 2005) the best parameterisation level for the MFNM is for parameters to be constant at the region scale (over 99.9% support in the *AICc* comparison). The Poisson lognormal has more support than the MFNM at all scales, as evidenced by the Akaike weights. At the site scale the Poisson lognormal has 99.7% vs. 0.3% for the MFNM. At the island scale the Poisson lognormal has over 99.9% support whereas the MFNM has less than 0.1%. Finally, at the regional scale the Poisson lognormal has 96.1% vs. 3.9% for the MFNM.

These results are consistent with visual inspection of the deviations between the data and each model's predictions. Deviations are smaller on average for the Poisson lognormal than either of the neutral models, at every scale (Fig 3.3). At the site scale, the SADs typically have a high number of singletons and a steeply decreasing number of increasingly abundant species. The Poisson lognormal captures the shape remarkably well, whereas the MFNM severely underestimates the number

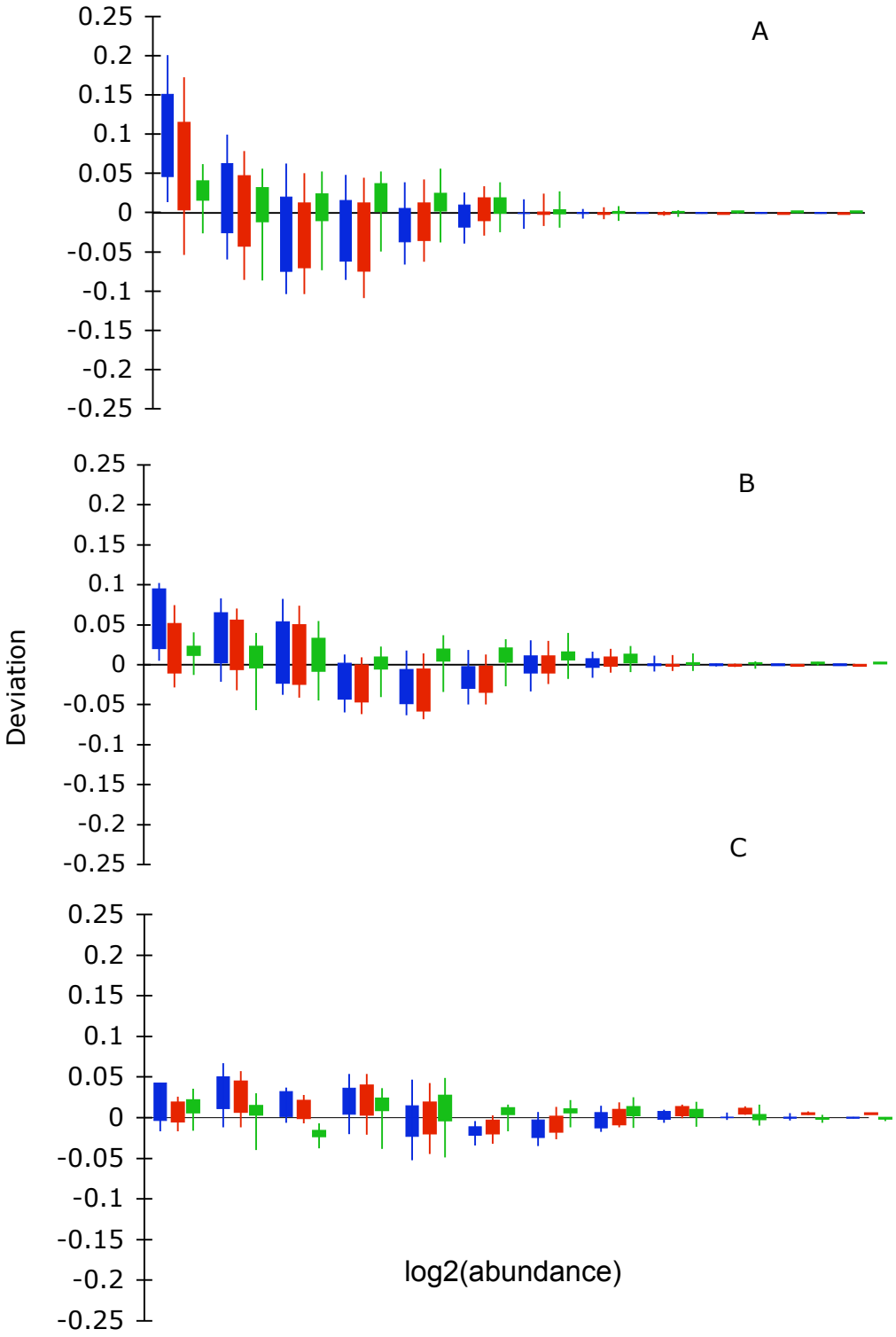


Fig 3.3 - Deviation between empirical species abundance distributions (SADs) and the best fit simulation neutral model (blue), analytical neutral model (red) and Poisson lognormal (green) for the site (A), island (B) and region scales (C). Deviations are calculated as the difference between the proportion of species observed and predicted by the model on each octave of abundance. Bars indicate quartiles and lines one standard deviation.

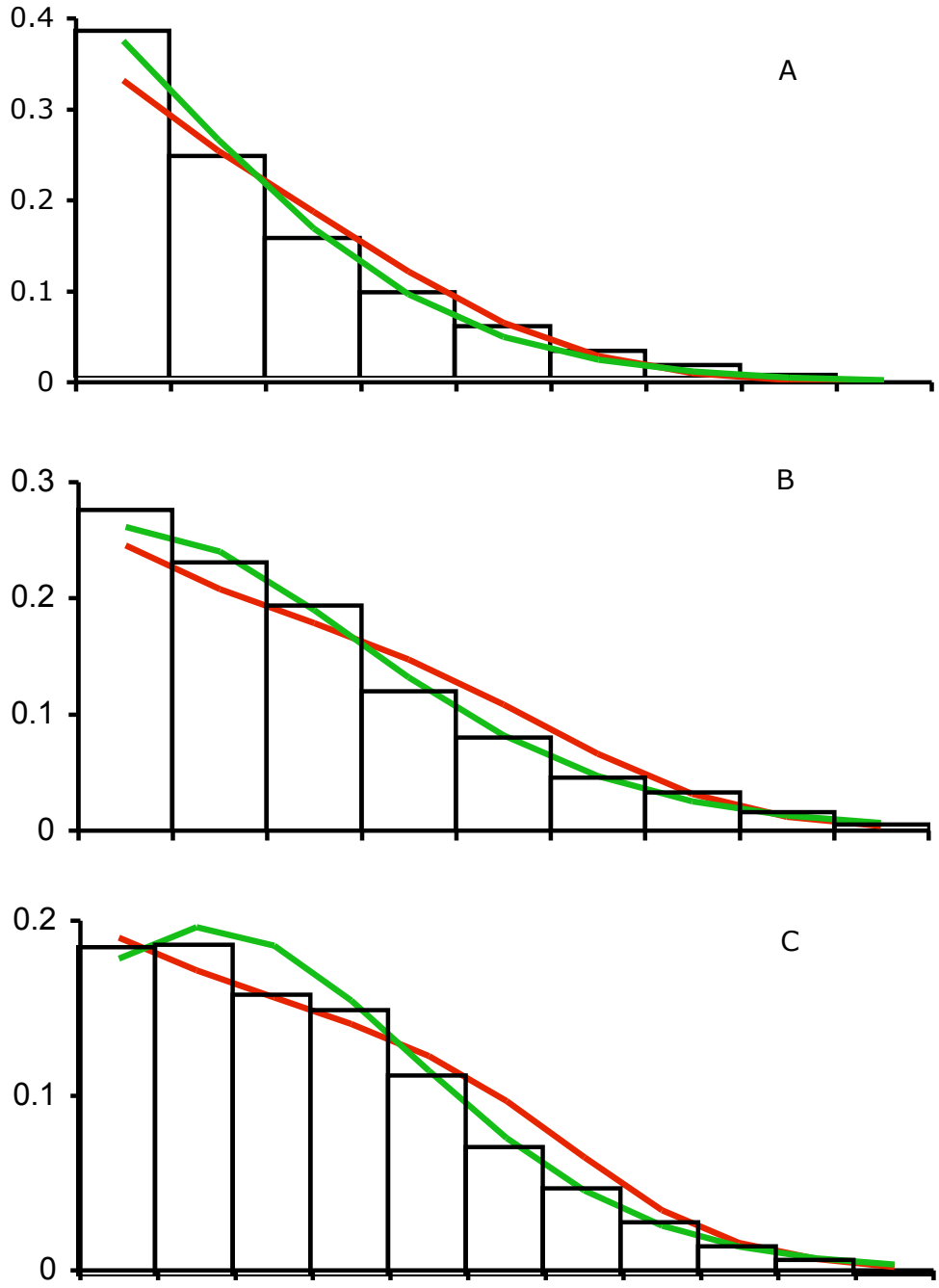


Fig 3.4 - Comparison between mean empirical SADs (bars), mean fitted MFNM (red line) and mean fitted lognormal (green line), for the site (A), island (B) and region (C) scales.

of singletons and overestimates the number of species with intermediate to high abundances (Fig 3.4). At the island scale, the pattern is similar, although the decrease in number of species is not as steep (Fig 3.4). At the region scale, the SADs usually have an internal mode and the Poisson lognormal follows the curve closest in most octaves (Fig 3.4). If SADs are examined individually it becomes apparent that the MFNM only outperforms the Poisson lognormal in cases when neither model fits well. In high diversity habitats and regions the Poisson lognormal follows observed SADs closely. The MFNM occasionally has better goodness of fit than the Poisson lognormal in depauperate regions and habitats, where neither model follows the observed pattern accurately (e.g. Appendix I, Fig. A.I.35 D and E).

3.4. Discussion

The results show substantial differences between neutral theory predictions and SADs of empirical coral assemblages. There is some debate about which groups of organisms are appropriate to test neutral theory (McGill et al. 2006). However, this Chapter presents empirical evidence against neutral theory from communities for which neutral theory has repeatedly been suggested to be most applicable (Hubbell 1997b; Hubbell 1997a; Whitfield 2002; Chave 2004; Williamson and Gaston 2005). Therefore, these results challenge neutral theory as a general explanation of biodiversity patterns in animal and marine communities.

The importance of scale has long been recognized in ecology, and it is possible that neutral dynamics are appropriate at some scales but not others. In fact, at the largest spatial scale examined (local communities pooled at the regional scale), coral SADs are not significantly different from neutral models (Fig 3.1). This is surprising, because at the regional scale empirical SADs should be approaching

metacommunity dynamics and deviating from local community SADs. This result was probably caused by the lack of power of the test at this scale, since the dataset includes only 5 regional SADs for each habitat, as opposed to 15 island and 60 site SADs for each habitat. However, sample sizes at the regional scale are much larger than at the site and island scales. Therefore, it is possible that sample size may influence the goodness-of-fit of neutral models to SADs. In Chapter 5, I investigate this by fitting a neutral model to SADs with sample sizes varying 300-fold. Nevertheless, the model selection process consistently supports the Poisson lognormal against the neutral model across spatial scales that vary 10^0 – 10^4 km across the Pacific Ocean. Hence, the conclusion that neutral theory does not provide a good model for SADs of coral assemblages is robust and independent of scale.

Empirical tests of neutral models have in some cases supported neutral theory (Volkov et al. 2003; Olszewski and Erwin 2004) and in others provided evidence against it (McGill 2003c; Etienne and Olff 2004). Because these conflicting results are sometimes obtained using the same dataset (McGill 2003c, Volkov et al. 2003), it is possible that such mixed results are in part caused by different testing methodologies. For example, using the SNM (McGill 2003c) or the MFNM (Volkov et al. 2003) can lead to opposite conclusions regarding a single dataset. In fact, in this Chapter the two models led to different estimates for the immigration parameter *m*. This was probably caused by uncertainty in the estimation process of the SNM related to, for example, stochastic variation in expected distributions, the sequential estimation procedure or fitting the model to octave classified data, instead of raw abundances. These discrepancies motivated the use of both models in the goodness-of-fit test, as well as the analysis regarding parameter estimates presented in Chapter 5. Additionally, the statistics used to test the goodness-of-fit, and whether octave

classes or actual abundances are used can also influence the results. Hence, multiple approaches are needed to ensure the robustness of tests of model fits to empirical SADs. The consistency of the results presented here when using two different versions of the model, two different tests of goodness–of-fit, and a model selection process against the Poisson lognormal, emphasises the robustness of this conclusion.

The most relevant example of a dataset where neutral theory also failed was a study of the beetles (Coleoptera) of Borneo. This dataset has the same qualitative deviations from the model as the coral data analysed here: more rare species and fewer common species than predicted by the model (Hubbell 2001). Hubbell (2001) suggests that these data are not intrinsically incompatible with the neutral model, and are to be expected under highly leptokurtic dispersal kernels with fat long tails. Under such dispersal, species should be very common in some locations and rare elsewhere. SADs in spatially explicit neutral models are indeed greatly influenced by the type of local within-community dispersal, and a stronger skew towards rare species occurs under such localised dispersal (Chave et al. 2002). However, it is not clear how a similar skew would arise from neutral dynamics with localized dispersal between local communities with high immigration, as seems to be the case both with the Indo-Pacific corals and, indeed, the Coleoptera of Borneo.

It is difficult to infer what causes the lack-of-fit of a model to SADs. Consequently, curve-fitting has been questioned as the best way of testing biodiversity models, despite being the mostly widely used approach. Additionally, interpreting the superior fit of the lognormal is not straightforward. The lognormal is unquestionably a valuable descriptive model, and arguably an appropriate null model for SADs (McGill 2003b; Williamson and Gaston 2005). A number of biological

explanations have been put forward for the lognormal (May 1975; Tokeshi 1990; Engen and Lande 1996; Sugihara et al. 2003). However the lognormal distribution is a statistical description of a pattern, not an ecological theory (Harte 2003). Because there are a variety of biological mechanisms that can give rise to lognormal distributions, its good fit is not sufficient, by itself, to identify the particular ecological processes that should be included in an improved theory of biodiversity.

In this Chapter I performed a comprehensive and robust test of neutral model goodness-of-fit to coral SADs, and presented evidence that neutral models are not consistent with coral community structure. However, because of the reasons highlighted above, testing the goodness-of-fit of biodiversity models to these curves is useful mostly as a first test. In the following Chapter I build on these results, and move beyond this approach to examine between-community patterns to try to understand why coral communities diverge from neutral theory predictions.

Chapter 4: Neutral Dynamics and Coral Reefs: patterns of community similarity

4.1. Introduction

Neutral theory (Hubbell 2001) proposes that species relative abundances in a community can be explained by the processes of birth, death, speciation and migration, under the assumption that all individuals are demographically identical. Niche assembly theory, in contrast, assumes that differences in species' resource use are the major factors determining community structure (MacArthur 1960; Tilman 1982; Tokeshi and Schmid 2002; Sugihara et al. 2003). Most empirical tests of neutral theory have focused on assessing the fit of a neutral model to relative abundance patterns within local communities, or for a metacommunity as a whole (Hubbell 2001; McGill 2003c; Volkov et al. 2003). However, several other models based on very different assumptions can generate similar relative abundance patterns (Engen et al. 2002; Mouquet and Loreau 2003). Hence, such approaches, used in isolation, have been argued to provide only weak tests of community structure models (McGill 2003b; Connolly et al. 2005). Instead, testing the predictions of a theory against multiple patterns (Adler 2004) is preferable, particularly where competing theories make different predictions. In Chapter 3, I discussed how coral species abundance distributions (SADs) from across the Indo-Pacific deviate significantly from neutral model predictions. Here, I analyse patterns of community similarity in the dataset used in Chapter 3, and compare them to neutral theory predictions.

High community similarity has been suggested as evidence for niche theory and against neutral theory. Niche assembly theory predicts that limiting similarity and complementarity of ecological niches should stabilize community structure and lead

to high community similarity (Clark and McLachlan 2003). By contrast, in neutral models each species undergoes independent random walks, and ecological drift should lead to increasing divergence between communities (Clark and McLachlan 2003; Maurer and McGill 2004), and thus to low community similarity. Because, the two competing theories have different expectations, examining patterns of community similarity should support either one or the other. High similarity in coral communities would provide support for niche theory, and low similarity for neutral theory.

Several examples of communities with high similarity in space or time have been suggested to provide evidence against neutral dynamics (McGill et al. 2006). One such example is tree composition of floodplain forests, which is constant in plots of similar habitat and disturbance history over scales of 1 to 40 km (Terborgh et al. 1996). Composition of *terra firma* forests is also dominated by a few species which are the same over scales of thousands of kms (Pitman et al. 2001). Another example are coral paleo-communities, for which constancy in composition has been suggested to support niche assembly driven community structure rather than dispersal limitation (Pandolfi 1999). Finally, pollen paleo-records trapped in lakes are also more constant in space and time than predicted by simulated neutral communities (Clark and McLachlan 2003). However, in these studies community similarity is generally assessed using only the most abundant species, which could lead to large differences in less common species being unnoticed. Furthermore, under neutral dynamics, it is precisely the most abundant species that are expected to contribute most to the immigrants pool, and thus be common to most local communities (Bell 2001). Most importantly, these studies do not take into account the homogenizing effects of immigration. Theoretical analyses of neutral models have shown that ecological drift leads to divergence between communities (Maurer and McGill 2004), but community

similarity is determined by the strength of dispersal limitation (Hubbell 2001; Maurer and McGill 2004; Volkov et al. 2004). Here I compare patterns of community similarity from coral communities across the Indo-Pacific, with those of neutral communities simulated with the parameters estimated in Chapter 1. This approach takes into consideration the effects of both dispersal limitation and diversity on similarity patterns. Similarity patterns under neutral dynamics are extensively characterized to ensure that the results are not affected by uncertainty in the parameter estimates obtained in Chapter 3.

4.2. Methods

4.2.1. Data

Sampling followed a hierarchical design with three spatial scales: site, island and region (Chapter 3, (Karlson et al. 2004)). At each site the coral colonies intercepted by ten 10-m long haphazardly placed transects were identified to species and counted, at each of three reef habitats (flat, crest and slope). Four sites were sampled for every island, and three islands were sampled at each of the five regions. The dataset is composed of 60 sites of each reef habitat.

4.2.2. Simulations

To characterize similarity patterns under neutral dynamics, I simulated metacommunities with θ of 1, 5, 10, 50, 100 and 500, and Jm of 10^7. I simulated 100 local communities with J of 10^4 individuals and m of 0.999, 0.1, 0.01, 0.001, 0.0001, 0.00001 and 0.000001, sampling immigrants from each of the metacommunities previously simulated. (I used large values of J, instead of the sample sizes, for this

particular analysis because small local communities with m of 0.01 or smaller cannot sustain a diverse community, and tend towards mono-dominance [Chapter 2]). These parameter values span a very broad range of parameter space, including highly unlikely values, and encompass all published empirical estimates to date. Each local community was simulated for 50,000 time steps with 1% of the individuals replaced at each time step (i.e., 500 turnovers of the local community). There is presently some dispute about the extent to which the community structure patterns reported for neutral models are stable (Chisholm and Burgman 2004; Hubbell and Borda-De-Agua 2004). In particular, analysis of neutral models over time scales much longer than those used in most neutral model analyses has shown a trend of increasing variance of abundances and decreasing local species richness (Maurer and McGill 2004). To verify if community composition had indeed stabilized, I ran additional simulations for up to 1500 turnovers. These simulations are exceptionally extensive: the local communities are one to two orders of magnitude greater and included six-fold more community turnover than previous studies.

The best-fitting estimates of m and q for both SNM and MFNM, obtained in Chapter 3, were used. Because the best-fit parameterisation of the ANM (with over 99.9% support) is for all local communities within the same habitat and region to have the same parameters, the metacommunity was defined at the region scale for the analyses of community similarity. To examine community similarity under neutral dynamics, I simulated the dynamics of 15 metacommunities: one for each habitat type in each region. The dynamics of the 12 local communities within each of the 15 metacommunities were simulated with J equal to the sample size for each corresponding local community in the coral dataset, and m and θ equal to the values estimated from the fits to the corresponding coral metacommunity. To verify that the

results were robust to metacommunity size, I simulated groups of local communities for fixed θ and m, with Jm varying from twice the size of the local communities to five orders of magnitude greater. The effect of metacommunity size was assessed by computing Spearman correlations of Jm with the mean and standard deviation of similarity.

4.2.3. Similarity.

I calculated Bray-Curtis community similarities between all pairs of sites in the data set, and for each set of 100 simulated neutral communities. Because neutral theory assumes homogeneous habitat, only within habitat similarities were calculated. To verify that the results were not sensitive to the choice of the Bray-Curtis statistic as a measure of community similarity, I repeated the analysis using Jaccard statistics, which use only presence-absence data, and thus contain less information but are more robust to sampling effects (Legendre and Legendre 1998).

The frequency distribution of coral community similarities was compared with that of the dataset simulated with the parameters estimated for the SADs. To verify that the results were not sensitive to uncertainty in the parameter estimates, I also compared the observed similarities with those of the range of parameter combinations simulated for the characterization of neutral similarity patterns. Finally, extensions of neutral theory that explicitly incorporate spatial structure predict decreasing similarity with distance (Hubbell 2001; Chave and Leigh 2002). This dataset allows assessing whether distance-decay in similarity is a likely cause of the differences between the data and neutral model's predictions. The most distant sites are more than 10,000 km apart and differ in their regional species pools. Therefore, to assess the effect of distance-decay on similarity patterns, I additionally analysed the

frequency distribution of similarity values obtained using only all the pairs of sites from each habitat, and only pairs of sites from the same island.

4.3. Results

The simulations showed that frequency distributions of community similarity under neutral dynamics are approximately normally distributed. Jaccard and Bray Curtis similarities showed similar trends (e.g. Fig. 4.1) so only Bray Curtis similarities are reported henceforth. Mean similarity decreases with increasing diversity, which is controlled by the parameter θ (Fig. 4.2.A). This is because in species-rich communities there are more species whose abundances can vary among communities. Isolation (low m) decreases mean similarity and increases its variance (Fig. 4.2.B). However, even moderate rates of immigration (0.1) generate local communities with very high similarity. The frequency distribution of community similarity exhibits no trend over the substantial range of metacommunity sizes (Jm) analysed for any of the parameter sets (Fig. 4.3, Spearman's r, all correlations $-0.454 < r < 0.115$, $P > 0.187$, $n = 10$ for mean similarity; $-0.358 < r < 0.430$, $P > 0.214$, $n = 10$ for standard deviation of similarity). Nevertheless, to be conservative, I used the largest metacommunities (five orders of magnitude greater than local community size) in all the analyses. Finally, the analysis of the 1500-turnover simulations reveals that community structure had stabilized by 500 turnovers for parameter values above very low migration rates (m~0.001 or greater), regardless of the value of θ (Fig. 4.4). For communities with immigration rates below 0.001, community structure does not equilibrate, even after 1500 turnovers (Fig. 4.4.A, B). In fact, with these extremely low migration rates communities slowly lose species as they approach mono-

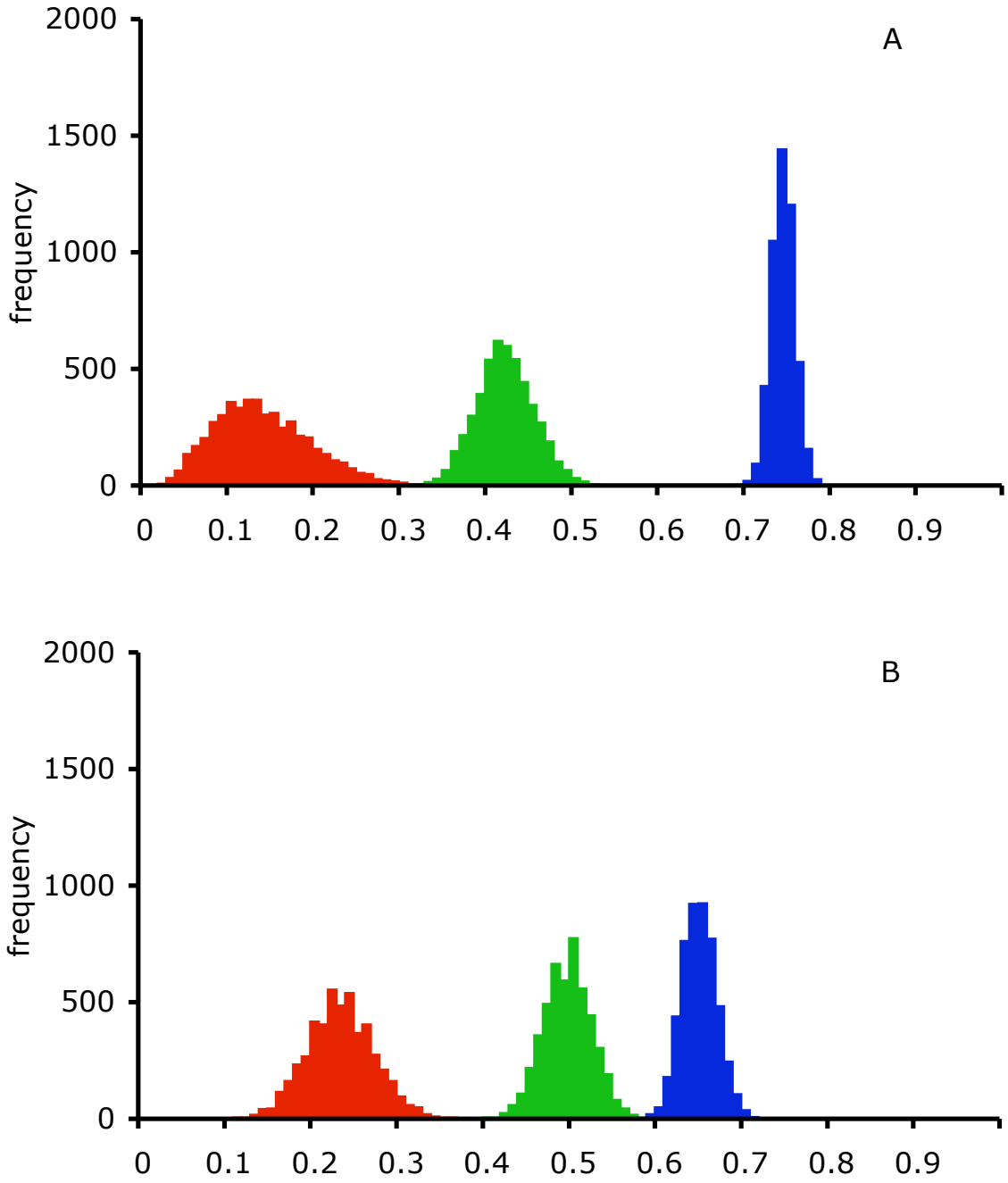


Fig 4.1 - Frequency distribution of Bray-Curtis (**A**) and Jaccard (**B**) similarities for neutral model simulations (*m* of 0.1 in blue, 0.01 in green and 0.001 in red, J of 10,000 individuals and θ of 50). Note that changes in the frequency distributions with parameter values is similar in both similarity measures.

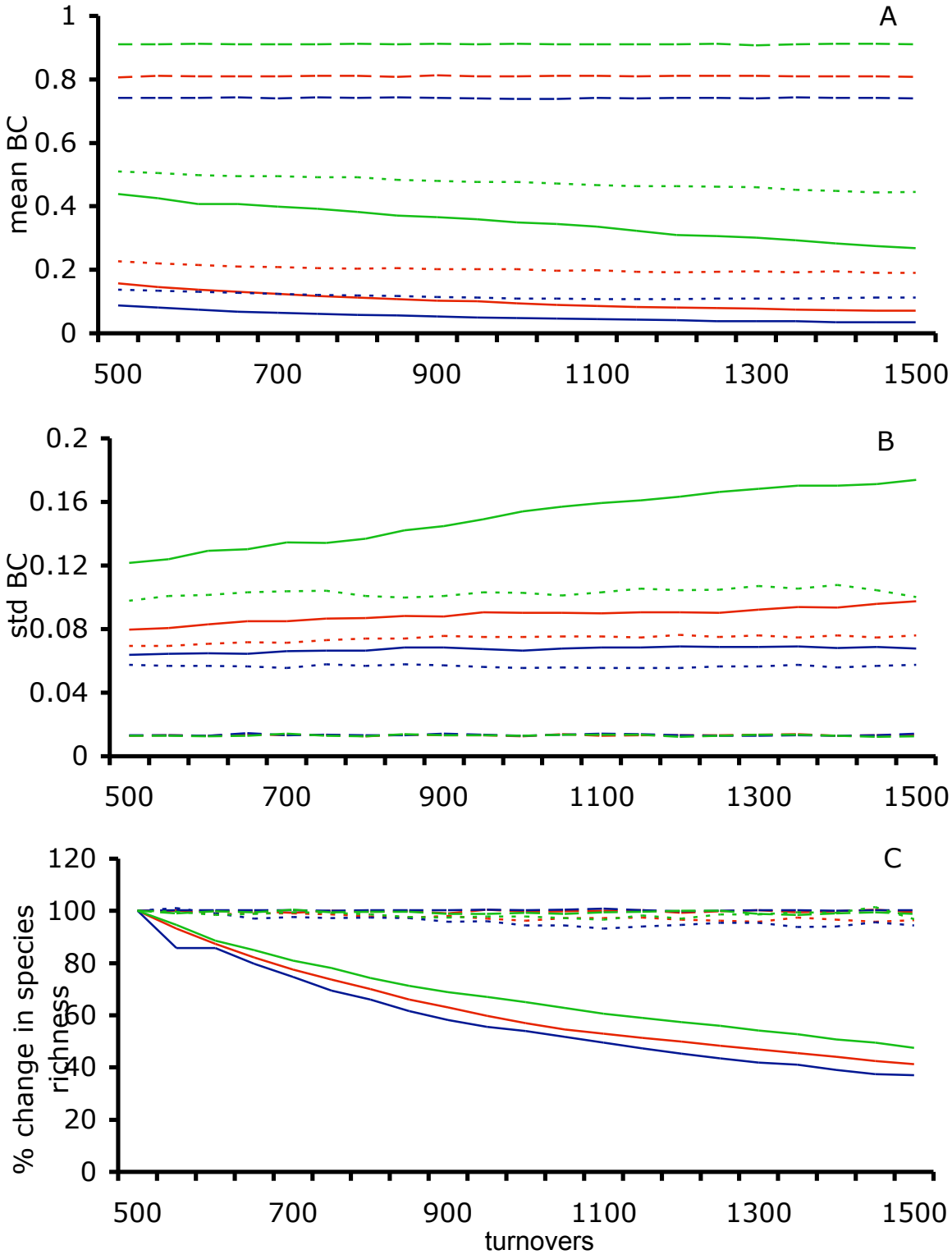


Fig 4.4 - Mean (**A**) and Standard Deviation (**B**) of Bray-Curtis similarity, and % change in species richness at 500 turnovers (**C**) through time at 50 turnover intervals from 500 to 1500 turnovers, for neutral simulations with J of 10 000 individuals, θ of 100 (blue), 50 (red) and 10 (green), and m of 0.1 (large dash), 0.001 (small dash), and 0.00001 (solid line). Note the trend towards mono-dominance in the communities with lowest migration rate.

dominance (Fig. 4.4.C). When local communities with smaller *J* are used, neutral theory predictions are nearly identical for high immigration rates (such as those estimated for our data), and the trend towards mono-dominance at lower immigration rates becomes even more pronounced (Chapter 2).

The similarities predicted for the parameter estimates are consistently high (Fig. 4.5.A). Mean similarity for simulations of each of the five regions varies between 0.77 and 0.88, with corresponding standard deviations varying between 0.019 and 0.023. In contrast, the observed similarities of coral communities are markedly lower, and more variable, than predicted by neutral theory (Fig. 4.5.B). Mean similarity of coral communities is between 0.25 and 0.60, and standard deviation varies between 0.080 and 0.131. Clearly, similarities predicted by neutral theory differ significantly from observed patterns. Figure 4.5 reports results for the slope, however, the pattern is common to the three habitats (Fig. 4.6). For the very broad range of parameter values examined in the characterization of similarity patterns under neutral dynamics, the community similarity distributions produced are markedly different from those exhibited by the data (Fig. 4.6). Therefore, the results are not caused by uncertainty in the parameter estimates (including uncertainty due to estimation procedure for the SNM, as suggested by Fig. 3.1). Finally, the distribution of similarity values calculated using only sites less than 10 km apart are as different from neutral model patterns as similarities calculated among sites up to 10,000 km apart. Indeed, while mean similarity is greater when the analysis is confined to nearby sites, there is no corresponding reduction in variance towards values consistent with neutral theory (Fig. 4.6).

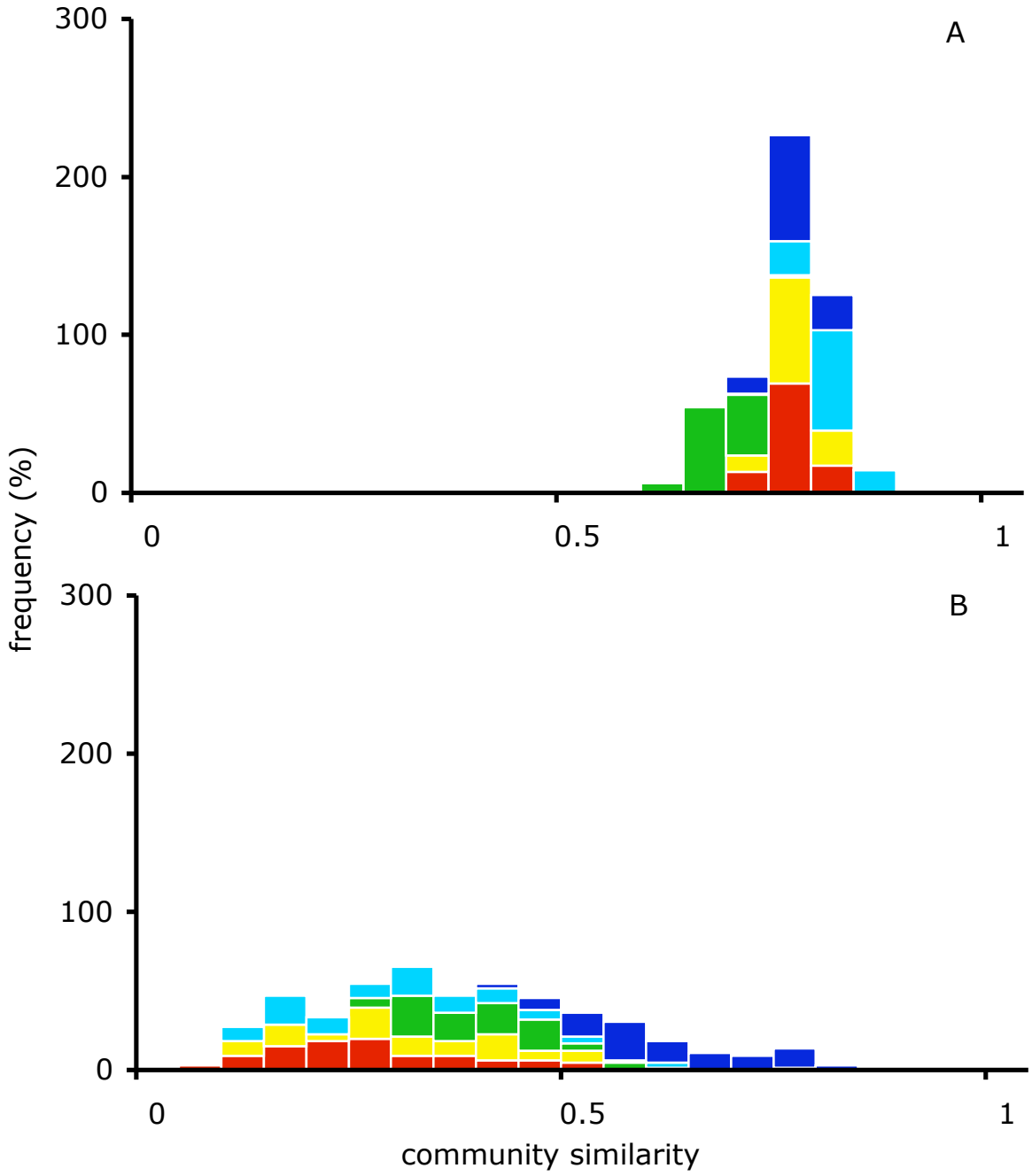


Fig 4.5 - Frequency distribution of Bray-Curtis similarities for (**A**) local communities simulated with parameters estimated from the data's species abundance distributions and (**B**) observed coral assemblages on reef slopes. Parameter estimates were: m= 0.905, 0.916, 0.915, 0.867, 0.748 and θ =25.4, 26.5, 30.6, 10.8, 6.3 respectively for Indonesia (red), Papua New Guinea (yellow), Solomon Islands (green), Samoa (light blue), and French Polynesia (dark blue). The heights of the bars sum to 100% separately for each region, so the stacked bars may exceed 100%. Note the low mean and high variance of the observed distributions in comparison to the simulations. Flat and crest assemblages differ from neutral model predictions in a very similar fashion (Fig 5.6).

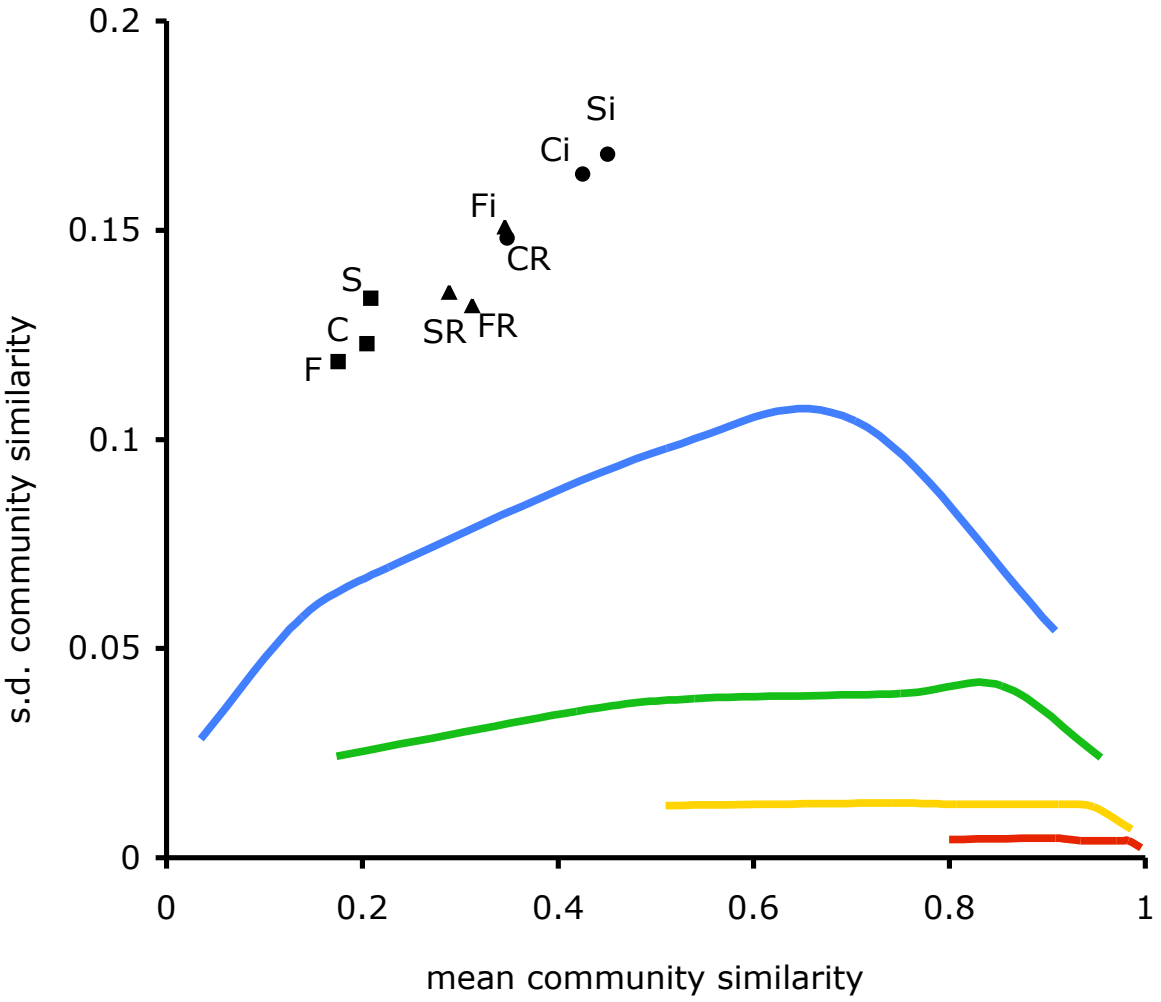


Fig 4.6 - Bray-Curtis similarity distributions for neutral simulations and coral communities. Because community similarity frequency distributions are approximately normal (Supplementary Information), they can be compared in terms of their mean and standard deviation. Each contour plots the mean and standard deviation of community similarity for $1< \theta < 500$ (from right to left on each contour) for a particular value of m (0.999 in red, 0.1 in yellow, 0.01 in green and 0.001 in blue). Coral assemblages Bray-Curtis similarity distributions of each habitat are plotted as black points: Si within island slope assemblages, Ci within island crest, assemblages, Fi within island flat assemblages, SR within region slope assemblages, CR within region crest assemblages, FR within region flat assemblages, S among all slope assemblages, C among all crest assemblages, F among all flat assemblages. Note how all data points, regardless of spatial scale or habitat, fall outside and comparably distant from the contours generated by neutral simulations.

4.4 Discussion

The finding that observed community similarities are lower than neutral model predictions is unexpected, and contrary to the widespread view that high similarity over space and time in empirical data constitutes evidence against neutral theory (Terborgh et al. 1996; Pitman et al. 2001; Clark and McLachlan 2003; Gilbert and Lechowicz 2004). However, others have argued that neutral theory can, in fact, produce high similarities (Hubbell 2001; Volkov et al. 2004). The results confirm that neutral theory can produce extremely high similarities, provided that immigration rates from the metacommunity are high, or diversity is very low. However, the analyses also reveal that variance in community similarity under neutral dynamics is consistently low, much lower than in the data, regardless of parameter values, thereby also confirming that this large discrepancy between neutral model predictions and the data is robust to uncertainty in parameter estimates.

For extremely low migration rates, community structure does not equilibrate, even after 1500 turnovers. Therefore, the possibility that distributions of community similarity for these parameter values could resemble those of the data cannot be ruled out. However, this superficial resemblance should not be interpreted as evidence for good fit of the neutral model, for several reasons. Firstly, for these parameter values, local species richness progressively decreased as communities tended towards mono-dominance. This is clearly not representative of scleractinian coral communities, which have sustained moderate to high biodiversity for tens of thousands of years. Moreover, the immigration rates necessary to produce these distributions (m<0.001, or less than one immigrant to a local community per ten thousand births) differ from

the estimated migration rates from the local community distributions ($0.75<m<0.92$) by over two orders of magnitude, and they are also inconsistent with population genetic data, which indicate moderate to high levels of immigration at spatial scales comparable to those of the local communities (Ayre et al. 1997; Ayre and Hughes 2000).

The higher variance of real coral communities compared to neutral theory predictions points to a strong role for environmental variability in determining patterns of community similarity on coral reefs. This high variance in community similarity indicates that the data contain a much broader mixture of sites that are very similar and very different, compared with neutral theory. Because habitat differences were minimized by the sampling design (see Methods and Chapter 3), the most likely source of this variability is spatio-temporal environmental stochasticity: the tendency for different local communities to experience fluctuations in environmental conditions differently. On coral reefs, even adjacent reefs can have markedly different environmental histories (Connell et al. 2004). Such historical differences would tend to decrease mean similarity if species differ in their responses to environmental fluctuations. These mechanisms would also be expected to increase the variability in community similarity values: communities that have experienced similar environmental histories would tend to have higher-than-average similarity (e.g., communities at similar successional stages), while communities with markedly different histories would tend to have lower-than-average similarity. In contrast, neutral models assume that metacommunities are environmentally homogeneous in space and time. Indeed, differences among species in responses to environmental fluctuations challenge a core assumption of neutral theory: that species are demographically identical.

In addition to the environmental homogeneity and neutrality assumptions, which are shared by all neutral models (McGill et al. 2006), there are other assumptions that differ among alternative formulations. However, these assumptions seem to be less plausible candidates as causes of the discrepancies between the neutral model and the data. Firstly, neutral models can have different assumptions about how speciation occurs (Hubbell 2001). However, because speciation occurs rarely compared with changes in relative abundance within species, it seems unlikely that different speciation mechanisms would have a marked effect on community similarity patterns, at least for realistic speciation rates (Ricklefs 2003). A second assumption is the "mainland-island" framework: metacommunities are held constant during local community dynamics (Hubbell 2001). An alternative formulation uses an archipelago framework with the metacommunity changing as local communities change (Bell 2000). While this might decrease community similarity over time, there is no obvious reason to expect it to increase variance in similarity among local communities, which is a major cause of the discrepancies between coral communities and neutral model predictions. Thirdly, spatially explicit neutral models predict decreasing similarity with distance (Hubbell 2001; Chave and Leigh 2002). Given that the most distant sites are more than 10,000 km apart, dispersal limitation is certainly occurring within the sampling scale. If dispersal limitation was the principal cause of the discrepancies between the data and the neutral model, the empirical data should converge towards the neutral model as the spatial scale at which community similarity is calculated decreases. However, the data do not converge to the model, because the high variance in coral community similarity does not decrease as the spatial scale decreases, even when similarity is calculated only for sites less than 10 km apart.

Because coral communities have been viewed as among the assemblages most likely to exhibit neutral dynamics (Hubbell 1997b; Whitfield 2002; Chave 2004; Williamson and Gaston 2005), the findings seriously challenge neutral theory's utility as a general theory of biodiversity and biogeography. Neutral theory was initially proposed as a null model for macroecology, predicting the diversity patterns that could arise from the action of demographic stochasticity and dispersal limitation alone (Hubbell 2001). However, early findings that neutral models can exhibit close fit to empirical species-abundance distributions rapidly shifted the focus towards whether or not neutral theory is sufficient, by itself, to explain macroecological regularities, such as species-abundance distributions and species-area relationships (McGill 2003c; Volkov et al. 2003; Adler 2004; Wootton 2005). Although the results do not support this latter hypothesis, they do show how using neutral theory as a null model can shed light on the mechanisms responsible for macroecological patterns.

In recent years, there has been a renewed focus on the use of ecological theory to inform conservation priorities. At the large ecological scales relevant to this endeavour, the prevailing view is that niche apportionment rules stabilize community structure over space and time (Pandolfi 1996; Terborgh et al. 1996; Pitman et al. 2001; Clark and McLachlan 2003; Gilbert and Lechowicz 2004). The results, however, support an alternative view in which species' different responses to spatial and temporal environmental fluctuations play a crucial role in the maintenance of biodiversity (Chesson and Huntly 1989). From this perspective, protecting biodiversity requires preserving the patterns of connectivity that allow species to find and exploit suitable environments that are patchy over both space and time. Fluctuation-mediated coexistence has received comparatively little attention in the ongoing controversy about neutral theory, despite the existence of tractable models of

species' relative abundance that explicitly incorporate effects of environmental stochasticity (Engen and Lande 1996). Given the accelerating pace of coral reef habitat loss (Hughes et al. 2003), a renewed focus on these mechanisms, and the body of ecological theory associated with them, is urgent.

Chapter 5: Unveiling a Coral Species Abundance Distribution

5.1. Introduction

So far I showed that coral communities of the Indo-Pacific deviate significantly from neutral model predictions both in their species abundance distributions (SADs, Chapter 3) and community similarity distributions (Chapter 4). However, these approaches raised questions that can only be answered with a different dataset. Firstly, in Chapter 3 goodness-of-fit of neutral models appeared to increase with spatial scale. However, in the dataset used, spatial scale cannot be unambiguously separated from sample size, as larger scale SADs are obtained by pooling samples from smaller spatial scales. Secondly, in Chapter 4 I showed that patterns of community similarity predicted by neutral models varied substantially with parameter values. Hence, confidence in the parameter estimates for the data is very important to ensure the reliability of the results. Specifically, it is important to verify whether parameter estimates are affected by sample size. To address these questions I collected and analysed a dataset of coral species abundances from a single location and habitat, with a total sample size approximately equal in magnitude to the total sample size of the Indo-Pacific dataset (when all regions and habitats are combined).

Diversity patterns are notoriously affected by sampling. Rare species are less likely to be observed in small samples than in large samples. Thus, sample size affects estimates of species richness, diversity and, most importantly for this study, species relative abundances (Magurran 2004). The effects of sampling have been present in the theoretical study of SADs from early on. The first statistical distribution proposed for SADs – the logseries – was defined as a distribution expected for a random sample

from populations with heterogeneous abundances (Fisher et al. 1943). The apparent concordance with a logseries distribution was later suggested to be the result of incomplete sampling, as the true underlying distribution for ecological SADs was proposed to be lognormal (Preston 1948). Because Preston suggested that incomplete sampling acted as a "veil" truncating part of the distribution, the process of exposing this underlying distribution with increasing sample size is known as "unveiling".

The logseries and lognormal differ most in the numbers of rare species. Most species in a logseries distribution are extremely rare. By contrast in a lognormal distribution most species are moderately abundant, and there is a small number of extremely rare, and extremely abundant species. The predicted change in shape of a SAD with increasing sample size is approximately as Preston (1948) had suggested. Poisson sampling from a lognormal distribution shows that as sample size increases, SADs are initially logseries-like, then an internal mode becomes apparent, and finally, at large sample sizes, they resemble a lognormal distribution (Lande et al. 2003b). Thus, if real SADs are indeed lognormally distributed, then increasing sample size should gradually unveil a lognormal distribution. In this Chapter I examine how coral SADs change with increasing sample size, as the true abundance distribution of a coral community is unveiled.

Empirical SADs usually vary from logseries-like to lognormal-like (see Chapter 1), and much of this variation is usually attributed to sample size. Small samples, relative to the size of the community being sampled, usually have the highest number of species represented by only one individual (singletons) (Pielou 1969). In marine communities in particular SADs tend to have many singletons (Hughes 1984; Hughes 1986). Therefore the use of the lognormal has been questioned when there is

no information to suggest that the true underlying SAD has an internal mode (Dewdney 1998; Dewdney 2000). However, numerous studies present lognormal-like SADs e.g.(Whittaker 1965; Wilson 1991), and the lognormal has become the classical distribution for SADs.

Large or exhaustive samples tend to have left-skewed, lognormal-like SADs (Nee et al. 1991; Hubbell 2001; Olszewski and Erwin 2004). This skewness, which has been the object of much recent attention, can be attributed to pooling samples (McGill 2003a). However, a left-skewed lognormal-like SAD can also be caused by the presence of occasional or migrant species (Magurran and Henderson 2003). Thus, the same pattern may be caused by an ecological process, or be the result of the sampling protocol. Discriminating between the effects of these two processes is essential to ensure that adequate patterns are used for testing ecological hypothesis. Specifically, it is important to assess whether the species that are apparent specialists in a certain habitat have a different abundance distribution than occasional species, typical of other habitats. In this Chapter I analyse the contribution of species not typical of a habitat to the habitat's SAD.

In the context of neutral theory, another potential confounding factor emerges: the effects of dispersal limitation and sample size are very similar. SADs predicted by neutral theory vary from logseries-like to lognormal-like as immigration rate decreases (Hubbell 2001; Chapter 2). These are the same changes expected by increasing sample size. In fact, the changes in SAD shape with immigration are somewhat related to the unveiling process: lower immigration rates correspond to a larger proportion of an independent community being sampled. However, sampling protocol can also potentially affect parameter estimates. A community may be

characterized by exhaustively identifying and counting all the individuals of a relatively small area, or, alternatively, by pooling several small samples distributed over a larger area. For the same sample size, these two approaches correspond to different levels of unveiling (sample size relative to the universe being sampled) and dispersal limitation. In the first case, the distribution is totally unveiled, and only effects of dispersal limitation are present. In the second case, dispersal limitation is smaller (a greater area is involved), but a smaller proportion of the universe is sampled. A sampling theory for a dispersal limited neutral community has been developed, which takes both factors into account (Etienne and Alonso 2005). However, the extent to which this sampling theory can discriminate between the effects of sample size and dispersal limitation has not yet been tested with empirical data. If sample size is truly accounted for, then parameter estimates for a community should not change with sample size.

In this Chapter I present an extensive characterization of a coral community in the Great Barrier Reef. This is, to my knowledge, the largest sample of a living coral community ever collected from a single site. It allows for an unprecedented characterization of the patterns of species relative abundances in a local coral community. Additionally, I separate species not typical of the habitat sampled, and examine whether excluding these species changes the shape of SADs. I examine how neutral model parameter estimates for a coral community vary as sample size increases nearly 300-fold. I examine changes in deviations between observed and predicted SADs with sample size, and I ask whether and how these change with increasing sample size. In summary, this Chapter focuses on three questions: 1) How does the SAD of a local coral community change as it is unveiled? 2) How do species

not typical of a habitat contribute to the habitat's SAD? and 3) How do neutral model parameter estimates and goodness-of-fit change with sample size?

5.2. Methods

5.2.1. Data Collection

Sampling took place between South and Palfrey Islands, in the Lizard Island group, Great Barrier Reef (Fig 5.1). Zonation patterns in reef communities are marked, and different habitats have different (albeit overlapping) species composition (Veron 2000). Because neutral models assume habitat homogeneity, this study focused on a single habitat: the exposed crest. The study site faces SE, which is the predominant wind direction, and therefore is often exposed to waves and currents. The exposed crest habitat is clearly defined in the field as a band between the reef slope (nearly vertical wall in this site) and the reef flat (horizontal surface that extends into the lagoon), where most waves break. It has approximately constant depth (~ 3 m) and a diverse coral assemblage with high cover, albeit small average colony size. The coral assemblage of the exposed crest habitat was characterized by identifying to species and counting the coral colonies within 276 belt-transects (10 x 1 m). The sample corresponds to the majority of the coral colonies on the exposed crest between South and Palfrey islands, within a total area of 2,760 m^2.

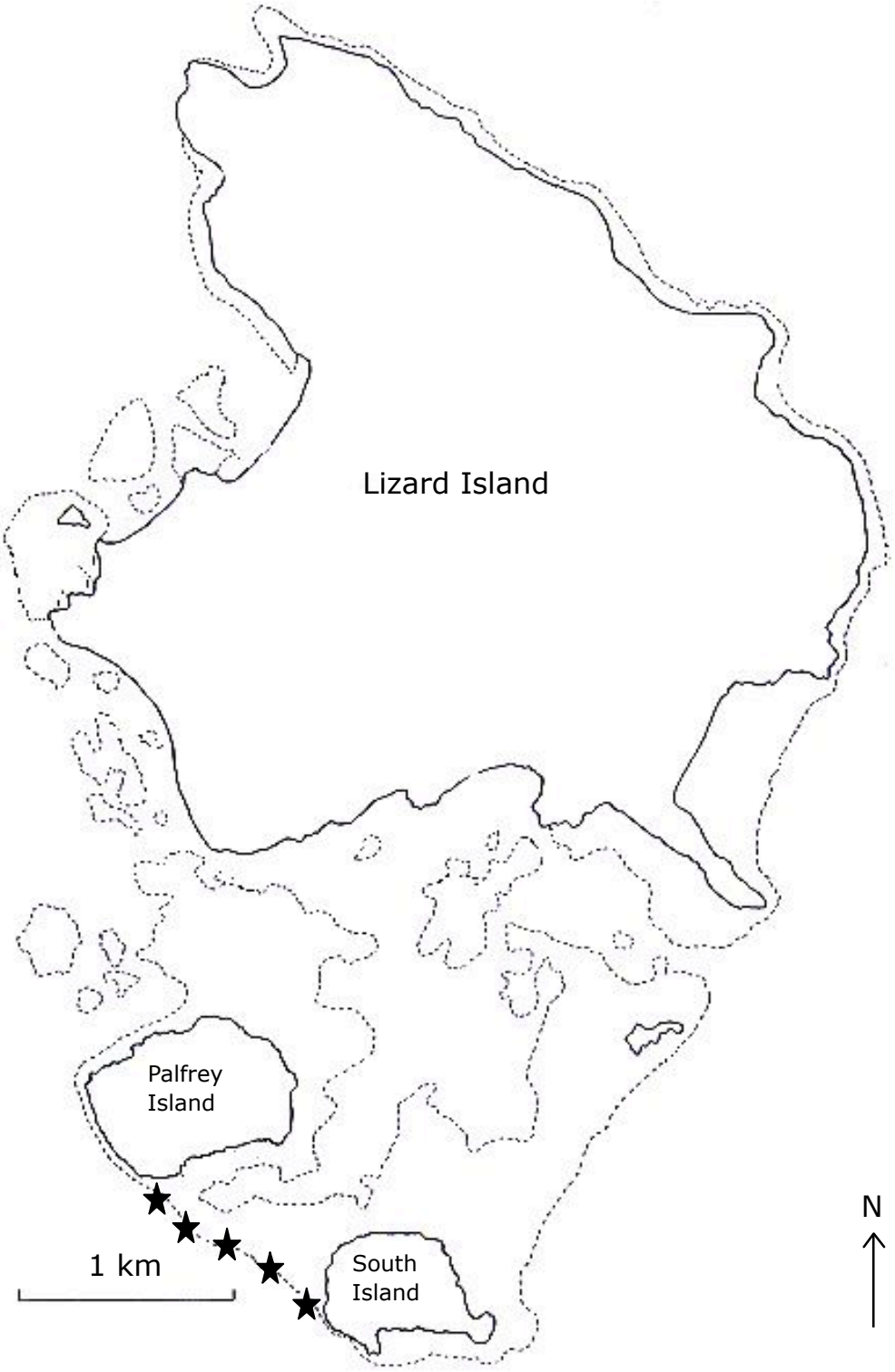


Fig 5.1 - Study site (marked by black stars): crest habitat between South and Palfrey Islands, in the Lizard Island Group, Great Barrier Reef, Australia.

5.2.2. Analysis

To obtain different sample sizes, an increasing number of transects (at intervals of 5: 1, 5, 10, 15…) were randomly selected to be pooled. This generated a total of 56 different samples with sample sizes varying over three orders of magnitude. I analysed the change with sample size of patterns of species relative abundance by plotting SADs and rank-abundance relationships for the different sample sizes.

To analyse the contribution of species not typical of the crest habitat to the SADs, I classified species according to their preferential habitat. Species habitat description in monographs are often not detailed enough to allow a consistent classification, therefore I used information from the CCRB dataset (Chapter 3). This dataset has species abundances for the three main reef habitats: flat, crest and slope. Species that were substantially more abundant at either the flat or the slope habitats were classified as non-crest, and examined separately. I used two different cut-off values: more than twice, and more than three times, as abundant on flat/slope relative to crest. SADs were then plotted separately for crest and non-crest species.

I analysed how parameter estimates and goodness-of-fit of the Genealogical Neutral Model (GNM) (Etienne 2005; Chapter 2) changed with increasing sample size. The GNM was fitted individually to each of these SADs, by calculating the log-likelihood (Etienne 2005; expression 2.5 in Chapter 2) for an extensive scope of parameter combinations ($10{,}000 > \theta > 1$ and $0.999 > m > 0.00008$). PARI/GP code for the *K(D,A)* component of expression 2.5 was provided by Etienne (2005), and I implemented the remaining steps in MATLAB. Two likelihood peaks were observed in every case. Only one of these peaks was biologically reasonable (see Results and

Discussion), so it was used in the remaining analyses. A second more detailed likelihood surface was then obtained for a more restricted scope of parameter combinations around this maximum ($50 > \theta > 1$ and $0.999 > m > 0.1$), in order to accurately estimate the best-fitted parameters for each SAD. 95% profile likelihood confidence intervals were calculated according to standard likelihood practice as the parameter values whose log-likelihood satisfied:

$$LL_{CL} = MLL - cl \quad (1)$$

where *MLL* is the maximum likelihood of the biologically reasonable peak and *cl* is half the critical value for 95% of a $\chi 2$ distribution for the appropriate number of degrees of freedom (Hilborn and Mangel 1997). For univariate confidence limits *cl* is 1.92, which is half of 3.84, the critical value for one degree of freedom. For bivariate confidence limits (confidence ellipses) *cl* is 3.00, which is half of 5.99, the critical value for 2 degrees of freedom.

Predicted SADs for each set of parameters (J, m and θ) were obtained by simulating 1,000 samples using the GNM urn scheme (code provided by Etienne (2005)). The simulation approach allowed me to calculate not only the mean but also the variance of each abundance class, and thus to estimate the expected stochastic variability in SAD shape under neutral dynamics. Observed and predicted SADs were then compared.

5.3. Results

In total, the 276 transects included 44,291 coral colonies and 156 species (see Appendix III for a species list). More than half the colonies (28,774) belonged to one of the ten most abundant species: *Porites* massive[1], *Acropora cuneata*, *Acropora hyacinthus*, *Acropora* dig-gem [2], *Goniastrea retiformis*, *Pocillopora damicornis*, *Acropora valida*, *Acropora nasuta*, *Pocillopora verrucosa* and *Favites halicora* (in order of abundance). A large proportion of species observed were extremely rare and, in particular, 19 species were represented by a single individual.

Species relative abundance patterns vary substantially with sample size (Fig 5.2). Small samples (1 transect or ~200 colonies) have a logseries-like SAD, with a large number of singletons, and a steeply decreasing number of increasingly abundant species. As sample size increases (20 transects or ~ 2,000 colonies) an internal mode appears, suggesting a lognormal distribution might be unveiling. However, as sample size increases further (140 transects or ~ 20,000 colonies, and 276 transects or 44,291 colonies), multiple modes become apparent. The apparent multimodality of the distribution does not seem simply the result of sampling noise, given its consistency, as it appears in SADs across a two-fold change in sample size and the modes can be seen travelling right in the graph as sample size increases.

---

[1] *Porites* massive cannot be resolved to species without microscopic observations, which would be impossible without destructive sampling. The number of colonies involved made this prohibitive both in terms of sampling time and the impact on the community contravening the Great Barrier Reef Marine Park Authority regulations for the Lizard Island Group.

[2] *Acropora* dig-gem has been shown, based on morphological, reproductive, and genetic criteria to be distinct from other species in the *Acropora humilis* species group (Wolstenhome 2004). Therefore, I treat it as a distinct species here, even though it has yet to be formally named.

The multiple modes are not caused by the presence of non-crest species (Fig 5.3). The two cut-off values used (species twice and three times more abundant in other habitats) led to classification of 48 and 67 species, respectively, as "non-crest". These two cut-off values produced similar results: the SAD distribution of crest species remained multimodal (and, indeed, the SAD of non-crest species is also multimodal). Therefore, the shape of SADs in this crest habitat appears not to be caused by the presence of transient species specialized for other habitats.

Across all sample sizes, there were consistently two peaks in the log-likelihood surface (Fig 5.4 and 5.5). The two peaks are present in approximately the same positions in every SAD examined, regardless of sample size (see Appendix II for all the Log-likelihood surfaces). The smaller peak is approximately for θ of 21 and m of 0.9. The largest peak is ridge shaped, and abuts the edge of the parameter space explored (up to $\theta < 500{,}000$), where it is still increasing. The m values corresponding to this peak decrease with sample size from 0.069 (1 transect or ~200 colonies – Fig 5.4 A) to 0.0004 (276 transects or 44,291 colonies – Fig 5.5 B).

The parameter estimates of this second peak are orders of magnitude higher than any previously estimated; therefore, I assessed whether this peak was biologically realistic by investigating the consequences of both θ and m values. I used the (under)estimate of θ (500,000) to generate a metacommunity of 10 million individuals (calculated with expression on page 165 of (Hubbell 2001). This metacommunity has over 1,500,000 species of corals, even though 10 million individuals corresponds only to less than 1 km^2 of reef. Furthermore, even if values of θ close to the lower confidence limits for this peak are used (θ~321), the metacommunity of

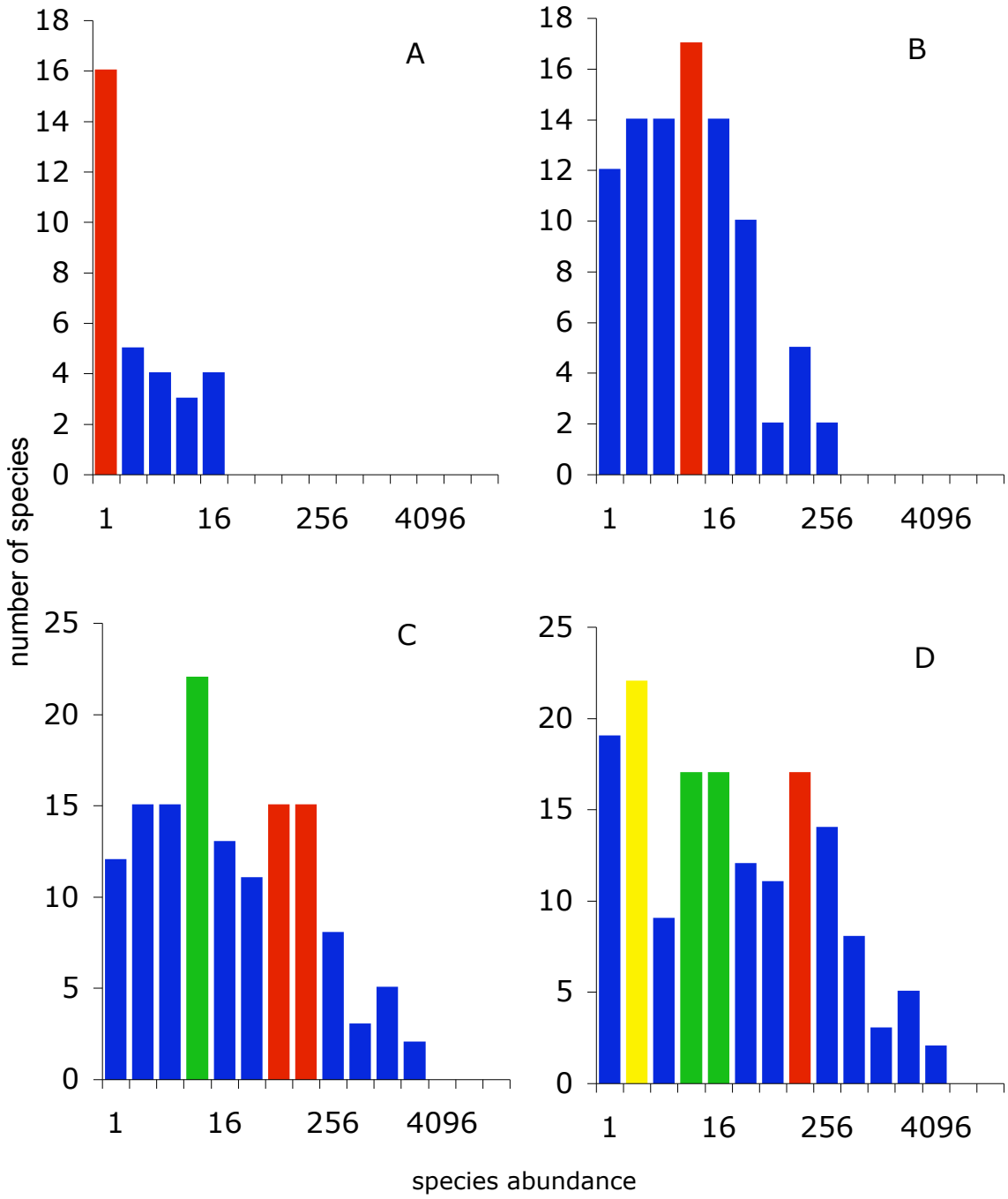


Fig 5.2 - Species abundance distribution for different sample sizes: 164 colonies or 1transect (**A**), 2755 colonies or 20 transects (**B**), 22939 colonies or 140 transects (**C**) and 44291 colonies or 276 transects (**D**). Note how modal classes (marked in red, green and yellow) travel to the right as sample size increases.

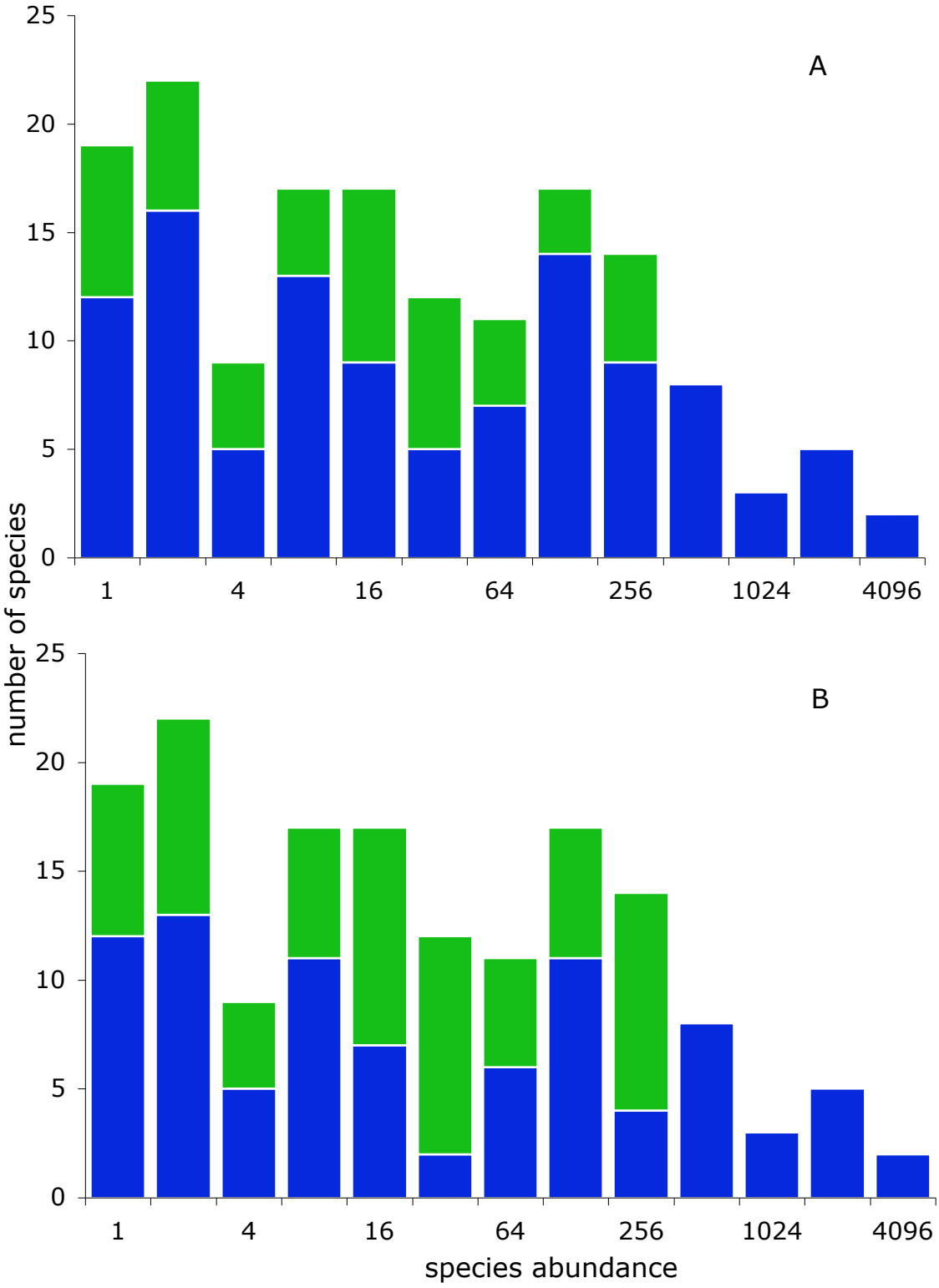
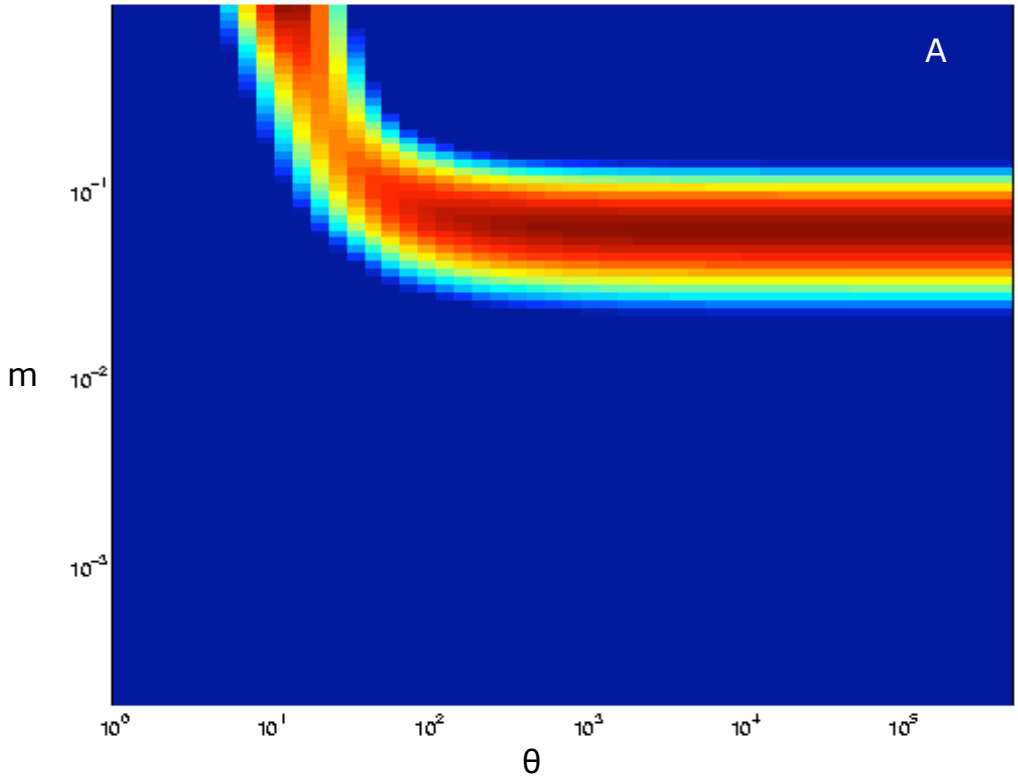


Fig 5.3 - Species abundance distributions for crest (blue) and non-crest species (green). Non-crest species were defined as being three times more abundant (A) or two times more abundant (B) in other habitats (see Methods).



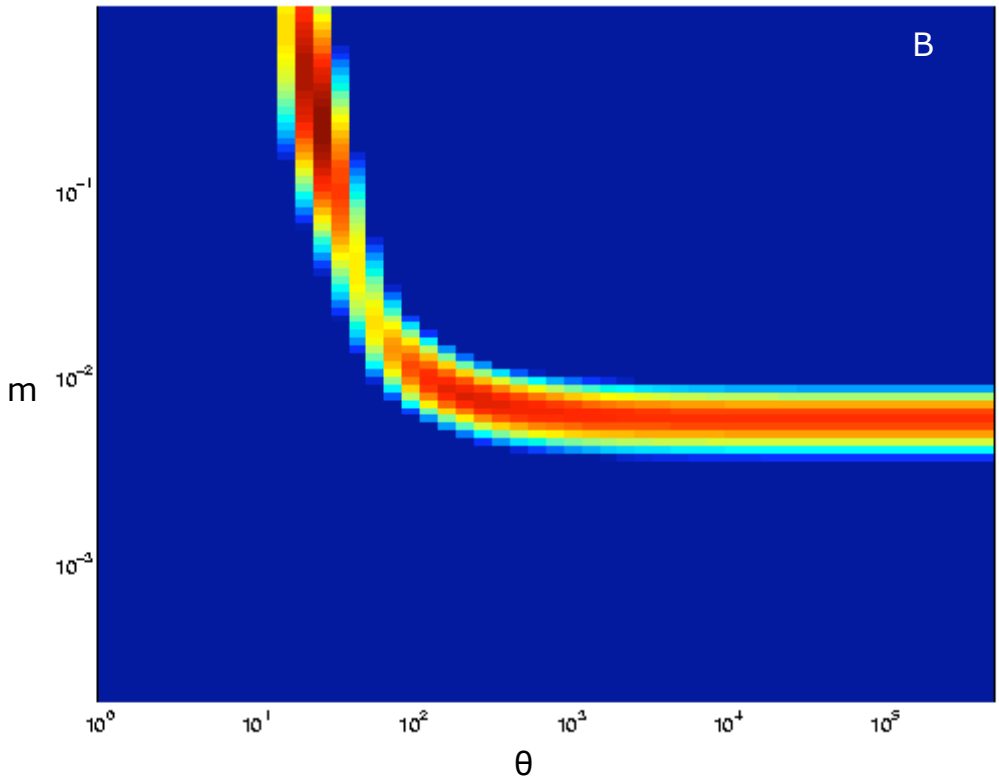


Fig 5.4 - Log-Likelihood surface for 164 colonies or 1transect (**A**) and 2755 colonies or 20 transects (**B**). Log-Likelihoods below 10 units of the likelihood maximum are truncated (represented in dark blue) to increase clarity. Colours warm as the likelihood increases

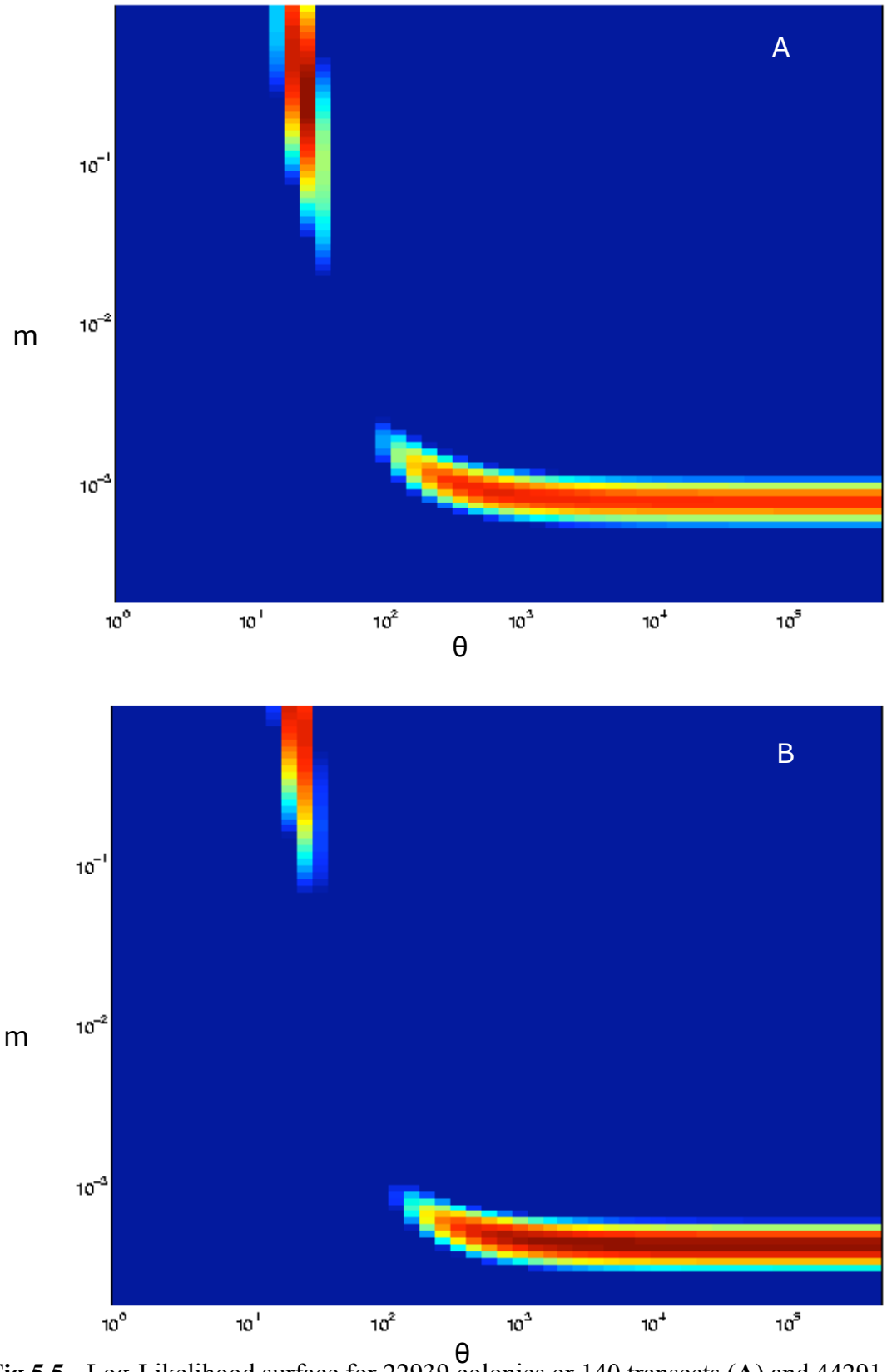


Fig 5.5 - Log-Likelihood surface for 22939 colonies or 140 transects (**A**) and 44291 colonies or 276 transects (**B**). Log-Likelihoods below 10 units of the likelihood maximum are truncated (represented in dark blue) to increase clarity. Colours warm as the likelihood increases

10 million individuals still contains over 3,300 species. Similarly, m~0.07 for single transects implies that 93% of the new colonies establishing on a 10m^2 area of reef were produced by adults within that same transect. Because both sets of values are wildly implausible (see Discussion), the high θ, low m peak was ignored in subsequent analysis and best-fitted parameters henceforth refer to the first maximum likelihood peak.

Profile likelihood confidence limits around the parameter estimates, using the fine-grained likelihood surfaces, reveal that the biologically realistic peak varies slightly in location with sample size (Fig 5.6-5.8). Estimates of θ increased from 13 to 20 as sample size reached 2,000 individuals, but remained constant thereafter (Fig 5.8 A). Estimates of m varied between 0.99 for small samples, to 0.3 for samples between 5,000 and 25,000 colonies, and back to 0.99 for larger samples (Fig 5.8 B). Most importantly, 95% confidence around maximum likelihood parameter estimates (MLE) overlap substantially for all sample sizes, suggesting that differences between MLEs are not statistically significant.

Predicted SADs deviate considerably from observed distributions for all sample sizes (Fig 5.9 and 5.10). For the smallest sample sizes deviations are similar to those described in Chapter 3 for the site and island scales: the model underestimates the number of rare species, and over-estimates the number of species with intermediate to high abundances (Fig 5.9 A). As sample size increases to ~2,000 colonies, the observed and predicted abundance distributions exhibit greater concordance, although the predicted distribution does not capture a small mode of very abundant species (Fig 5.9 B). As sample size increases further, several internal

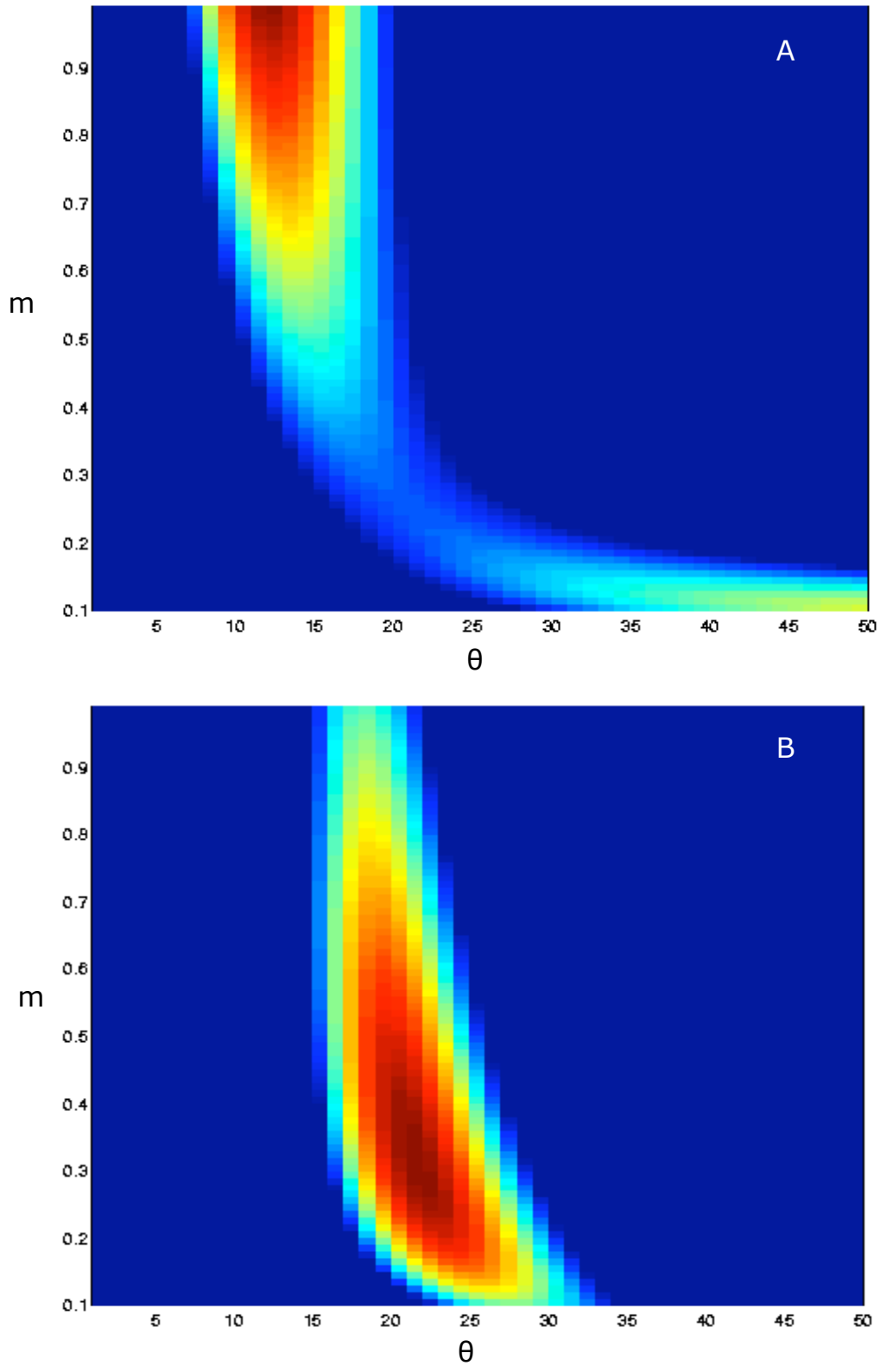
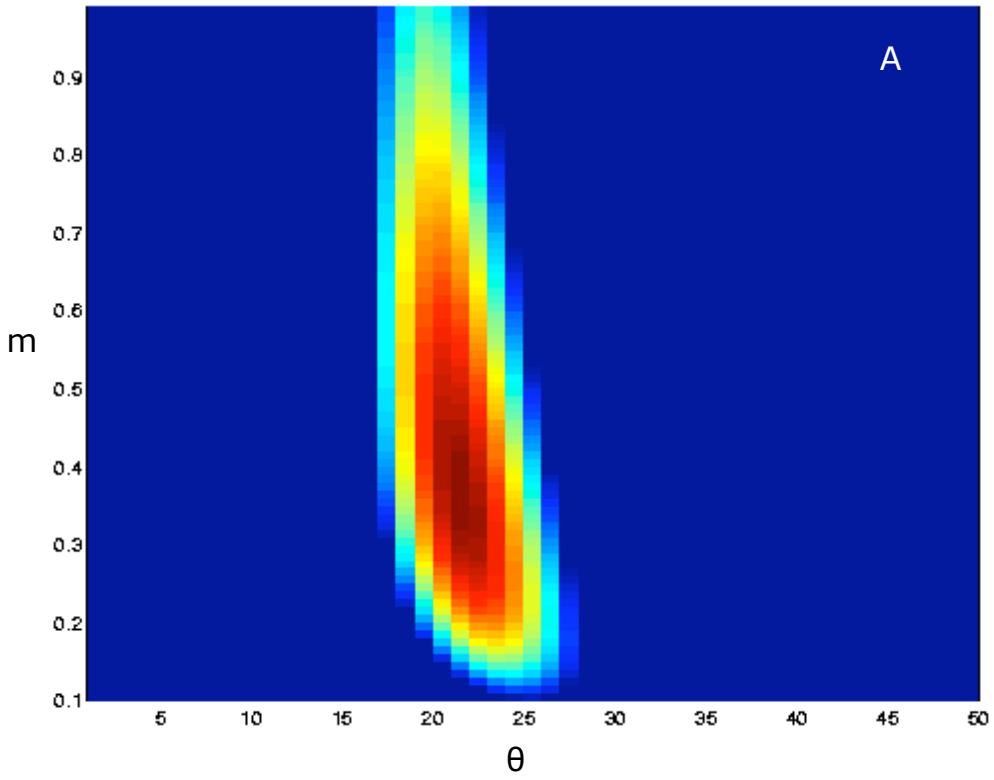


Fig 5.6 - Log-Likelihood surface for 164 colonies or 1transect (**A**) and 2755 colonies or 20 transects (**B**). Log-Likelihoods below 3 units of the likelihood maximum are truncated (represented in dark blue) to increase clarity. Colours warm as the likelihood increases. Note that all colors except dark blue correspond to the parameter confidence ellipse. 79



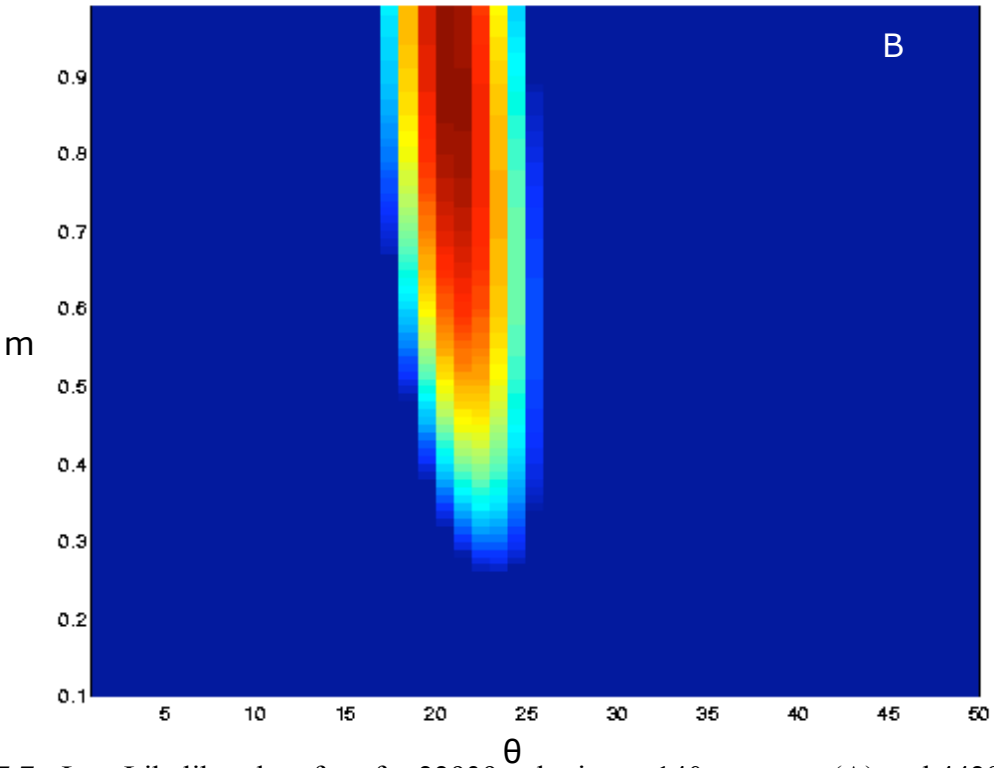


Fig 5.7 - Log-Likelihood surface for 22939 colonies or 140 transects (**A**) and 44291 colonies or 276 transects (**B**). Log-Likelihoods below 3 units of the likelihood maximum are truncated (represented in dark blue) to increase clarity. Colours warm as the likelihood increases. Note that all colors except dark blue correspond to the parameter confidence

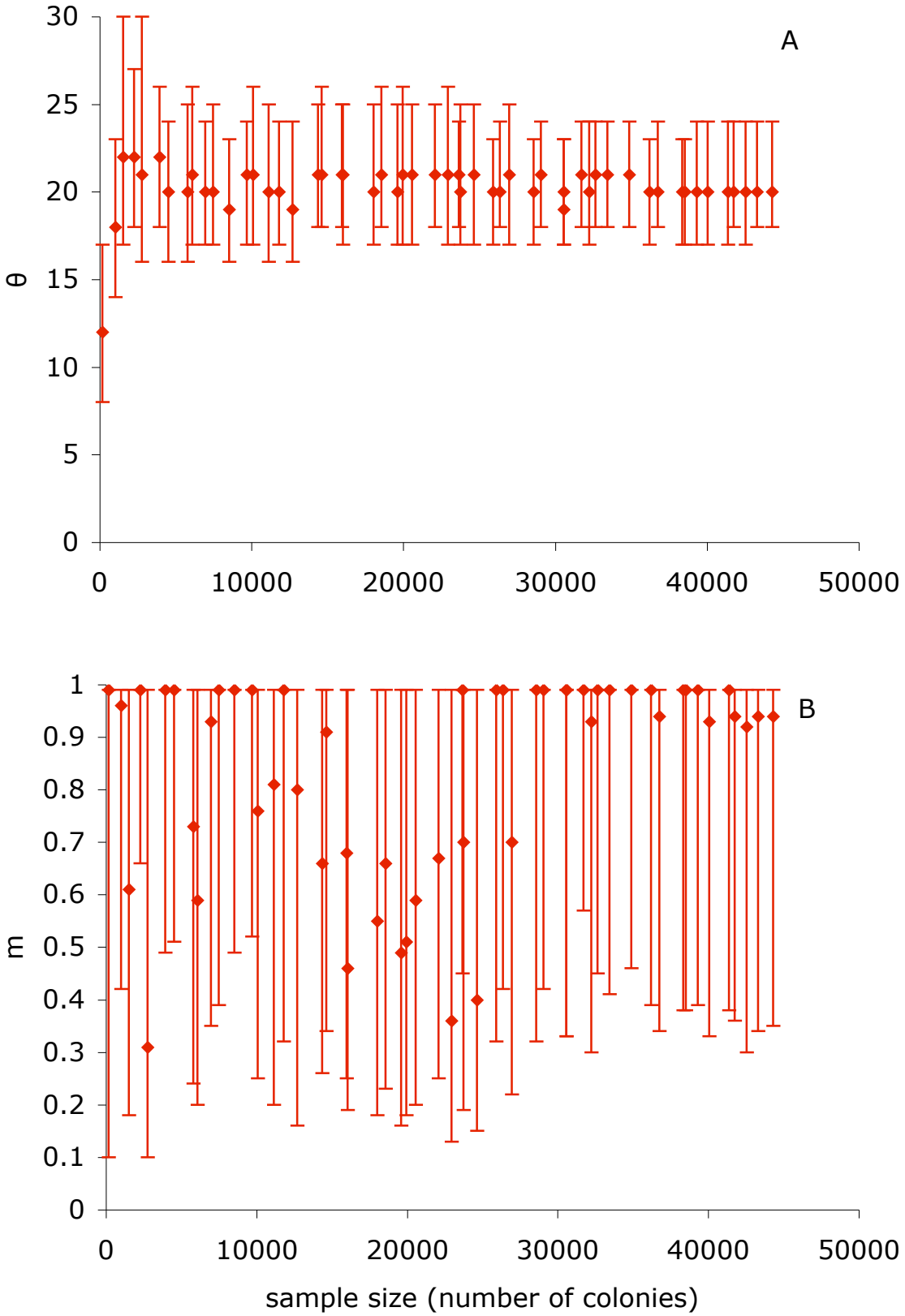


Fig 5.8 - Change in parameter estimates with sample size for m (**A**) and θ (**B**). Diamonds mark the maximum likelihood estimate and error bars mark 95% profile likelihood confidence limits.

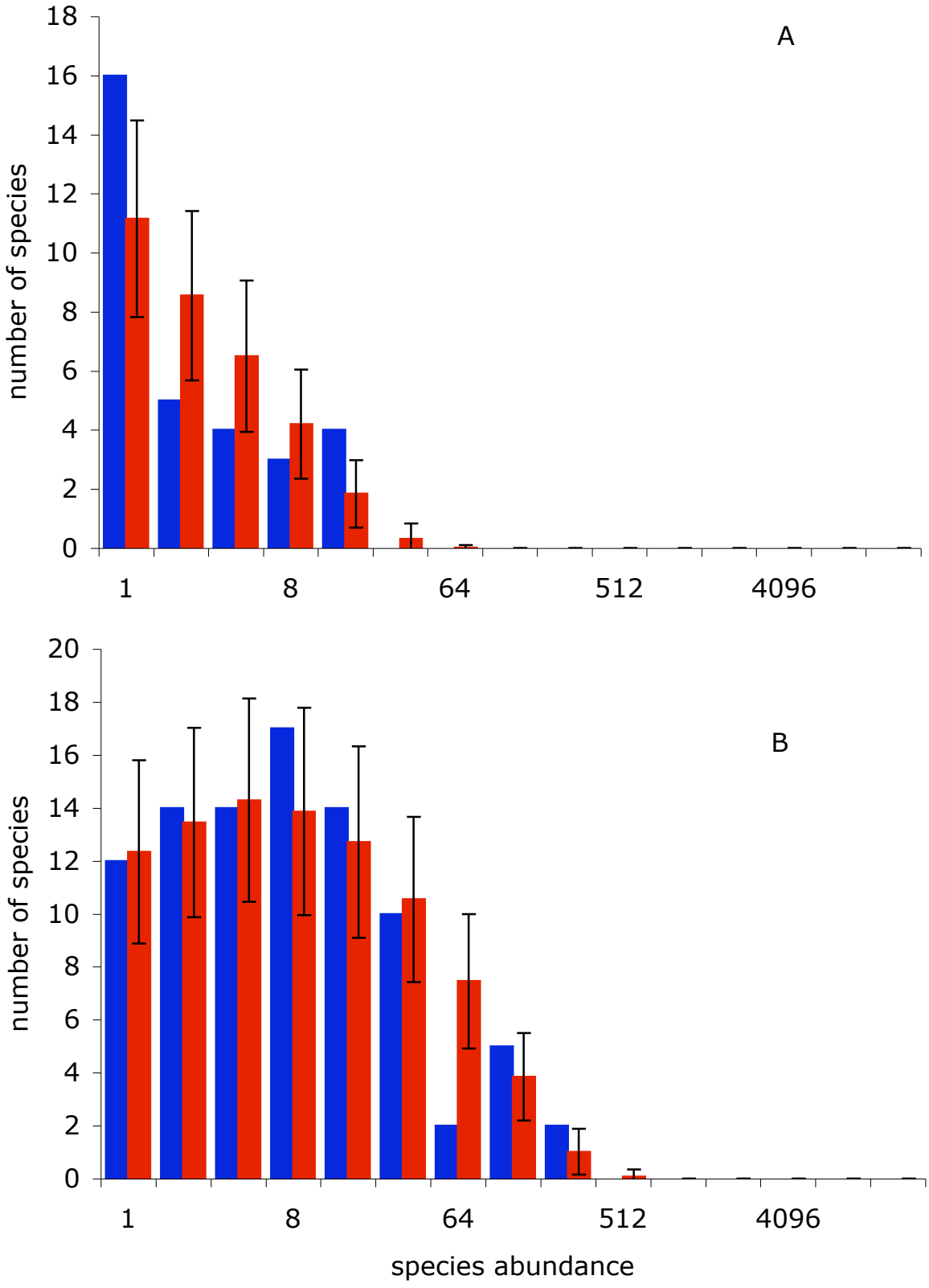


Fig 5.9 - Observed (blue) and best fitted (red) species abundance distributions (SADs) for 164 colonies or 1transect (**A**) and 2755 colonies or 20 transects (**B**). Best fitted SADs are mean of 1000 simulated samples (bars) +/- 1 standard deviation (error bars).

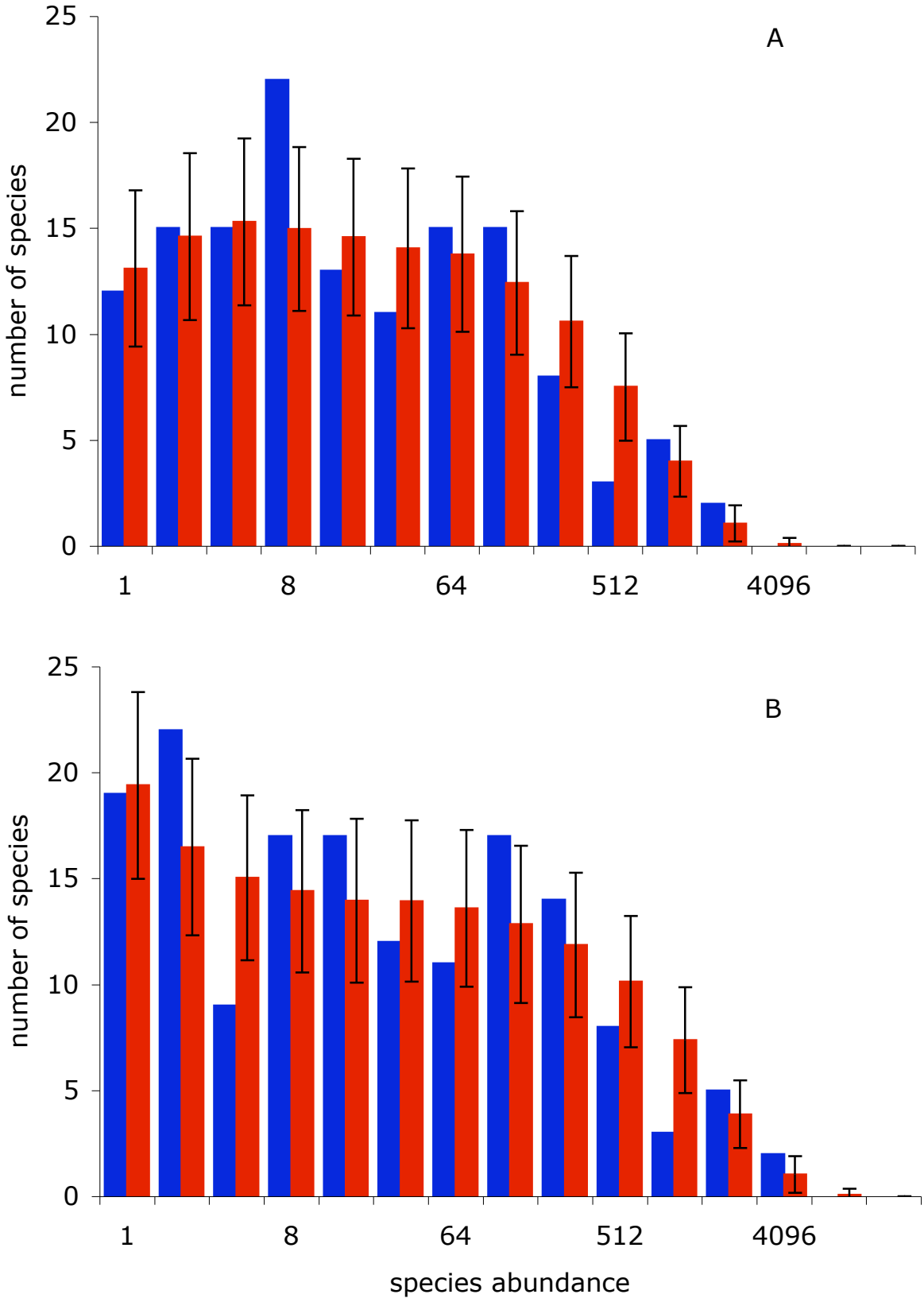


Fig 5.10 - Observed (blue) and best fitted (red) species abundance distributions (SADs) for 140 transects (**A**) and 44291 colonies or 276 transects (**B**). Best fitted SADs are mean of 1000 simulated samples (bars) +/- 1 standard deviation (error bars).

modes become apparent in the data, which the model does not capture (Fig 5.10 A and B).

5.4. Discussion

Increasing sample size does not seem to unveil a lognormal distribution for a crest coral community. This is particularly important because the lognormal has become the classical statistical distribution for SADs. Theoretical models of community structure are judged by their ability to generate lognormal distributions (e.g. Tokeshi 1990; Engen and Lande 1996), and the lognormal is the distribution against which other models are typically compared (Harte 2003; McGill 2003c; Nee and Stone 2003). Much of its support has come from the presence of an internal mode in SADs, if sample sizes are large enough (Whittaker 1965; Wilson 1991). However totally unveiled distributions are extremely rare in the literature, as the sampling effort involved is often enormous. Hence, the widespread use of the lognormal as a null expectation for SADs has been criticized (Dewdney 1998; Williamson and Gaston 2005), particularly in cases when an internal mode is not present (Hughes 1986). This study shows that the appearance of an internal mode with increasing sample size does not necessarily correspond to the partial unveiling of a lognormal distribution. As sample size is increased further SADs gained a multimodal appearance.

Multimodal SADs have previously been observed in empirical data, but are usually ignored. For example, the neutral model has been shown to provide a better fit than the lognormal to a brachiopod fossil community (Olszewski and Erwin 2004). However, neither model fits the data particularly well, in part because of what appears

to be a multimodal empirical abundance distribution. Others have sought to explain the multiple modes by partitioning species according to their preferential habitat and fitting different models to the SADs for the different partitions (Magurran and Henderson 2003). However, in the present study the sample collected was as ecologically homogenous as possible because it was collected at a single location in time and space. Furthermore, removing non-crest species does not reduce the multimodal appearance of the SADs, and hence the mixture of species adapted for different habitats is not likely to be causing this pattern.

Lognormally distributed SADs are often observed for heterogeneous communities. The Barro-Colorado dataset, for example, includes a variety of habitats in terms of altitude, soil and humidity (Hubbell et al. 1999). Another example is the British bird abundance dataset, which includes all the habitats in Great Britain (Nee et al. 1991). In fact, the coral dataset used in Chapters 3 and 4, which is of the same order of magnitude in terms of sample size, has a partially unveiled lognormal SAD when all habitats and regions are combined (Connolly et al. 2005). It is thus possible that habitat heterogeneity promotes lognormality. Lognormal SADs have been suggested to result from many variables influencing species abundances interacting multiplicatively, as a consequence of the Central Limit Theorem (CLT) (May 1975). Habitat heterogeneity should increase the number of variables involved, and consequently decrease the probability of observing departures from the CLT. Larger scales should have a similar effect, and thus metacommunities may well be more lognormal-like than local communities.

The fit of the GNM to these data is maximized by two parameter combinations, which correspond respectively to high immigration and low speciation,

or to low immigration and high speciation. (Etienne et al. 2006) report similar results for shrub communities in South Africa. The presence of two likelihood peaks is problematic because they correspond to two very different ecological hypotheses. Hence it is important to assess which peak is biologically realistic. In the case of the South African shrub communities, distinct isolation and high speciation rates are entirely consistent with the high level of endemism found in these communities (Latimer et al. 2005). In the current study, however, the high speciation peak leads to unrealistically high numbers of species at the metacommunity scale. Approximately 800 coral species have currently been described for the entire planet (Veron 2000). Even using the lower confidence limit of the high-θ peak, and an absurdly small metacommunity size, predicted a metacommunity richness over four times this number. Moreover, the values for the immigration parameter m in this peak (0.069 for an area of 10 m^2 to 0.0004 for an area of 2760 m^2) are inconsistent with the basic population biology of scleractinian corals, most of which have external fertilization and whose larvae spend days to weeks in the plankton before settling back on the reef.

Parameter estimates of the biologically realistic likelihood maximum do not vary significantly with sample size. Thus, the GNM can successfully separate the effects of dispersal limitation and incomplete sampling, which have similar qualitative effects on SADs. In particular, it is reassuring that fitting neutral models to small samples does not lead to over-estimates of immigration rates, as might be expected if incomplete sampling could not be separated from dispersal limitation. Furthermore, the confidence limits for the parameter estimates help explain the uncertainty regarding estimates of m in Chapter 3. The relative flatness of the likelihood surface around the maximum likelihood peak means that even small numerical errors in the

likelihood calculation can lead to differences in the best-estimated parameter. Thus, these results lend support to the parameter estimates in Chapter 3 since the parameter values of the two datasets are of comparable magnitude, and differences between SNM and MFNM estimates (Fig. 3.1) are within parameter confidence limits. Finally, confidence in the estimates of *m* reinforces the conclusions of Chapter 4, because differences between real and neutral communities were most pronounced for high immigration rates.

Understanding the processes that determine species relative abundances remains one of the central "unanswered questions in ecology" (May 1999). Fitting explanatory models to SADs is possibly the most common approach to this question. However it is increasingly criticized as providing only weak tests of ecological hypothesis (McGill et al. 2006). The main objection to using curve fitting as an empirical test of a model is that the correlation between theoretical and observed curves does not necessarily imply a correspondence between pattern and process (Magurran 2005). However, the reverse is not true: failing to fit empirical SADs clearly suggests that the processes included in neutral models cannot, by themselves, explain coral patterns of species relative abundances. In concordance with Chapter 3, neutral model predictions deviate markedly from observed SADs. However, in contrast to the Indo-Pacific dataset, the abundance distribution characterized here is apparently different from a lognormal as well.

Chapter 6: Conclusions

This thesis provides new and critical understanding of the processes that affect coral communities, achieved by examining exceptionally large datasets and comparing observed patterns with predictions from theoretical models of community structure. Firstly, relative abundance patterns of reef corals depart significantly from neutral theory predictions. This result is particularly important because corals were thought to be among the most likely communities to conform to neutral theory assumptions. This raises serious questions about using neutral theory as a general explanation for community structure. Secondly, patterns of community similarity differ from neutral theory predictions. Most importantly, they differ in the opposite direction of niche theory predictions, which suggest coral communities are strongly affected by environmental stochasticity. Thirdly, as sample size increases at a single location, the SAD that emerges appears to be multimodal, not lognormal. This suggests that that we do not yet have a good theory to explain patterns of species relative abundance.

Coral species abundance distributions (SADs) deviate from theoretical distributions. Observed patterns depart consistently and significantly from neutral theory predictions, across several habitats, bio-geographical regions, and spatial scales (Chapter 3). Neutral models have been compared to SADs of trees in tropical forests (Volkov et al. 2003; Volkov et al. 2005), fossil molluscs (Olszewski and Erwin 2004), birds (McGill 2003c), freshwater fish (Etienne and Alonso 2005), and intertidal communities (Wootton 2005) among others. Results have been inconsistent, but tended to reject neutral models (see McGill et al. 2006 for a review). Coral

communities add to the list of communities that do not conform to neutral theory predictions.

Since the publications that brought attention to neutral theory as a general explanation of diversity patterns appeared (Bell 2000; Bell 2001; Hubbell 2001), there has been considerable theoretical progress regarding neutral models. Several neutral model versions have been developed, with slightly different ancillary assumptions (Volkov et al. 2003; Etienne and Olff 2004; McKane et al. 2004; Etienne 2005; He 2005). In Chapter 2 I reviewed these models and compared the patterns generated by the three main modelling approaches to show that they generate similar predictions regarding species abundance distributions (SADs). The results from Chapters 3 and 5 combined also confirm that neutral models are essentially equivalent, and thus deviations from neutral model predictions must be a consequence of violation of the assumptions common to all the model versions.

In Chapter 5 I present a novel test of neutral models that uses parameters estimated from local community SADs, to test neutral model predictions regarding between community variability. I show that, for this pattern as well, coral communities deviate considerably from the model predictions. Hence, none of the patterns examined are consistent with neutral models, and these results add strength to the conclusion that the processes included in neutral models cannot explain coral community structure. Most importantly, coral communities are more variable than neutral model predictions, instead of less variable as many proponents of niche assembly have argued. This suggests that it is the violation of the spatial and temporal environmental homogeneity assumption that causes these results. The importance of environmental history for coral community structure has long been recognized

(Hughes 1989; Hughes 1994) and suggests that the state of a coral reef depends greatly on unpredictable but frequent disturbances.

The effects of environmental variability and disturbances on community structure patterns are best understood with a hypothetical example. Consider several hypothetical nearby reefs with similar communities. A bleaching event caused by a mass of unusually hot waters affects some of the reefs but not others. As a consequence, the reefs affected suffer mass mortality of bleaching sensitive species, whereas the unaffected reefs remain unchanged. Mean community similarity decreases because of differences between affected and unaffected reefs. However, resemblance between similarly affected reefs increases variance in similarity. Because disturbances on coral reefs are frequent, localized and selective spatio-temporal environmental variability is a likely explanation for the patterns observed in this thesis.

Coral SADs at a local level also seem to deviate from the classical lognormal distribution. For extremely large samples in a single habitat and local community, the distribution coral species abundances has several apparent modes (Chapter 5). Multimodal SADs have previously been reported (e.g. (Magurran and Henderson 2003; Olszewski and Erwin 2004)) but no theoretical model can currently explain these patterns. Given the results of Chapter 4, it will be most interesting to examine how environmental stochasticity and, in particular, frequent disturbances affect species relative abundances at a local scale.

In conclusion, the results of this thesis challenge neutral theory as a general explanation for community dynamics. The departures observed from neutral theory predictions clearly suggest that coral communities are strongly affected by processes

not included in neutral models. However, this thesis also highlights that neutral models are extremely useful as null models against which empirical patterns should be compared. This was the original purpose of neutral models (Caswell 1976; Hubbell 1979) and has the potential to dramatically advance our understanding of the processes that determine biodiversity.

References

Abrams, P. A. 2001. A world without competition. Nature 412:858-859.

Adler, P. B. 2004. Neutral models fail to reproduce observed species-area and species-time relationships in Kansas grasslands. Ecology 85:1265-1272.

Akaike, H. 1985. Prediction and Entropy, Pages 1-24 *in* A. C. Atkinson, and S. E. Fienberg, eds. A celebration of statistics, The ISI centenary volume. Berlin, Springer Verlag.

Alonso, D., and A. J. McKane. 2004. Sampling Hubbell's neutral theory of biodiversity. Ecology Letters 7:901-910.

Ayre, D. J., and T. P. Hughes. 2000. Genotypic diversity and gene flow in brooding and spawning corals along the Great Barrier Reef, Australia. Evolution 54:1590-1605.

Ayre, D. J., T. P. Hughes, and R. J. Standish. 1997. Genetic differentiation, reproductive mode, and gene flow in the brooding coral *Pocillopora damicornis* along the Great Barrier Reef, Australia. Marine Ecology Progress Series 159:175-187.

Baker, O. 2002. Law of the Jungle. New Scientist 173:28-31.

Bell, G. 2000. The distribution of abundance in neutral communities. The American Naturalist 155:606-617.

—. 2001. Neutral Macroecology. Science 293:2413-2418.

Bellwood, D. R., T. P. Hughes, C. Folke, and M. Nystrom. 2004. Confronting the coral reef crisis. Nature 429:827-833.

Bengtsson, J. 2002. The unified neutral theory of biodiversity and biogeography. Ecological Economics 42:497-498.

Boik, R. J. 2004. Commentary on: Why Likelihood? Pages 167-180 *in* M. L. Taper, and S. R. Lele, eds. The Nature of Scientific Evidence: Statistical, Philisophical, and Empirical considerations. Chicago, University of Chicago Press.

Brown, J. H. 1999. Macroecology: progress and prospect. Oikos 87:3-14.

—. 2001. Towards a general theory of biodiversity. Evolution 55:2137-2138.

Burnham, K. P., and D. R. Anderson. 1998, Model selection and inference, a practical information-theoretic approach. New York, Springer-Verlag.

Buss, L. W., and J. B. C. Jackson. 1979. Competitive networks nontransitive competitive relationships in cryptic coral reef environments. American Naturalist 113:223-234.

Caswell, H. 1976. Community structure: a neutral model analysis. Ecological Monographs 46:327-354.

Chapin, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper et al. 2000. Consequences of changing biodiversity. Nature 405:234-242.

Chase, J. M. 2005. Towards a really unified theory for metacommunities. Functional Ecology 19:182-186.

Chave, J. 2004. Neutral theory and community ecology. Ecology Letters 7:241-253.

Chave, J., and E. G. Leigh. 2002. A spatially explicit neutral model of beta-diversity in tropical forests. Theoretical Population Biology 62:153-168.

Chave, J., H. C. Muller-Landau, and S. A. Levin. 2002. Comparing classical community models: consequences for patterns of diversity. The American Naturalist 159:1-23.

Chesson, P. 1991. A need for niches? Trends in Ecology and Evolution 6:26-28.

—. 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31:343-366.

Chesson, P., and N. Huntly. 1989. Short-term instabilities and long-term community dynamics. Trends in Ecology and Evolution 4:293-298.

Chisholm, R. A., and M. A. Burgman. 2004. The unified neutral theory of biodiversity and biogeography: Comment. Ecology 85:3172-3174.

Clark, J. S., and J. S. McLachlan. 2003. Stability of forest biodiversity. Nature 423:635-638.

Connell, J. H., T. P. Hughes, C. C. Wallace, J. E. Tanner, K. E. Harms, and A. M. Kerr. 2004. A long term study of competition and diversity of corals. Ecological Monographs 74:179-210.

Connolly, S. R., D. R. Bellwood, and T. P. Hughes. 2003. Indo-Pacific biodiversity of coral reefs: Deviations from a mid-domain model. Ecology 84:2178-2190.

Connolly, S. R., T. P. Hughes, D. R. Bellwood, and R. H. Karlson. 2005. Community structure of corals and reef fishes at multiple scales. Science 309:1363-1365.

Connolly, S. R., and S. Muko. 2003. Space preemption, size-dependent competition, and the coexistence of clonal growth forms. Ecology 84:2979-2988.

Dewdney, A. K. 1998. A general theory of the sampling process with applications to the "veil line". Theoretical Population Biology 54:294-302.

—. 2000. A dynamical model of communities and a new species-abundance distribution. Biological Bulletin 198:152-165.

Dial, R. 2002. The unified theory of biodiversity and biogeography. Quarterly Review of Biology 77:50.

Engen, S., and R. Lande. 1996. Population dynamic models generating the lognormal species abundance distribution. Mathematical Biosciences 132:169-183.

Engen, S., R. Lande, T. Walla, and P. J. DeVries. 2002. Analyzing spatial structure of communities using the two- dimensional Poisson lognormal species abundance model. American Naturalist 160:60-73.

Enquist, B. J., J. Sanderson, and M. D. Weiser. 2002. Modeling macroscopic patterns in ecology. Science 295:1835-1836.

Etienne, R. S. 2005. A new sampling formula for neutral biodiversity. Ecology Letters 8:253-260.

Etienne, R. S., and D. Alonso. 2005. A dispersal-limited sampling theory for species and alleles. Ecology Letters 8:1147-1156.

Etienne, R. S., A. M. Latimer, J. A. Silander, and R. M. Cowling. 2006. Comment on "Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot". Science 311.

Etienne, R. S., and H. Olff. 2004. A novel genealogical approach to neutral biodiversity theory. Ecology Letters 7:170-175.

Ewens, W. J. 1972. The sampling theory of selectively neutral alleles. Theoretical Population Biology 3:87-112.

Fisher, R. A., A. S. Corbet, and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12:42-58.

Gaston, K. J., and S. L. Chown. 2005. Neutrality and the niche. Functional Ecology 19:1-6.

Gilbert, B., and M. J. Lechowicz. 2004. Neutrality, niches, and dispersal in a temperate forest understory. Proceedings of the National Academy of Sciences of the United States of America 101:7651-7656.

Harpole, W. S., and D. Tilman. 2006. Non-neutral patterns of species abundance in grassland communities. Ecology Letters 9:15-23.

Harte, J. 2003. Tail of death and resurrection. Nature 424:1006-1007.

He, F. 2005. Deriving a neutral model of species abundance from fundamental mechanisms of population dynamics. Functional Ecology 19:187-193.

Hilborn, R., and M. Mangel. 1997, The Ecological Detective, confronting models with data: Mongraphs in Population Biology, v. 28. New Jersey, Princeton University Press.

Hubbell, S. P. 1979. Tree dispersion, abundance, and diversity in a tropical dry forest. Science 203:1299-1309.

—. 1997a. Niche assembly, dispersal limitation, and the maintenance of diversity in tropical tree communities and coral reefs. Proceedings of the Eight International Coral Reef Symposium, Panama 1:387-396.

—. 1997b. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. Coral Reefs 16:s9-s21.

—. 2001, The Unified Neutral Theory of Biodiversity and Biogeography: Monographs in Population Biology, v. 32. Princeton, Princeton University Press.

Hubbell, S. P., and L. Borda-De-Agua. 2004. The unified neutral theory of biodiversity and biogeography: Reply. Ecology 85:3175-3178.

Hubbell, S. P., and R. B. Foster. 1986. Commoness and rarity in a neotropical forest: implications for tropical tree conservation, Pages 205-231 *in* M. E. Soule, ed. Conservation Biology: the Science of scarcity and diversity. Sunderland, Sinauer Associates Inc.

Hubbell, S. P., R. B. Foster, S. T. O'Brien, K. E. Harms, R. Condit, S. J. Wechsler, S. J. Wright et al. 1999. Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science 283:554-557.

Hughes, R. G. 1984. A model of the structure and dynamics of benthic marine invertebrate communities. Marine Ecology Progress Series 15:1-11.

—. 1986. Theories and Models of species abundance. The American Naturalist 128:879-899.

Hughes, T. P. 1989. Community structure and diversity of coral reefs: the role of history. Ecology 70:275-279.

—. 1994. Castatrophes, phase shifts, and large-scale degradation of a Caribbean Coral Reef. Science 265:1547-1551.

Hughes, T. P., A. H. Baird, D. R. Bellwood, M. Card, S. R. Connolly, C. Folke, R. Grosberg et al. 2003. Climate change, human impacts, and the resilience of coral reefs. Science 301:929-933.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577-586.

Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? American Naturalist 93:145-159.

Karlin, S., and J. McGregor. 1972. Addendum to a Paper by W. Ewens. Theoretical Population Biology 3:113-116.

Karlson, R. H., H. V. Cornell, and T. P. Hughes. 2004. Coral communities are regionally enriched along an oceanic biodiversity gradient. Nature 429:867-870.

Kimura, M. 1968. Evolutionary Rate at the Molecular Level. Nature 217:624-626.

King, J. L., and T. H. Jukes. 1969. Non-Darwinian evolution. Science 164:788-798.

Lande, R., S. Engen, and B. E. Saether. 2003a, Stochastic Population Dynamics in Ecology and Conservation. Oxford, Oxford University Press.

—. 2003b, Stochastic Population Dynamics in Ecology and Conservation: Oxford Series in Ecology and Conservation. Oxford, Oxford University Press.

Latimer, A. M., J. A. Silander, and R. M. Cowling. 2005. Neutral ecological theory reveals isolation and rapid speciation in a biodiversity hot spot. Science 309:1722-1725.

Legendre, P., and L. Legendre. 1998, Numerical ecology: Developments in environmental modelling, v. xv. Amsterdam; New York, Elsevier.

MacArthur, J. W. 1975. Environmental fluctuations and species diversity *in* M. L. Cody, and J. M. Diamond, eds. Ecology and Evolution of Communities. Cambridge, Belknap Press of Harvard University Press.

MacArthur, R. H. 1960. On the relative abundance of species. The American Naturalist 94:25-34.

MacArthur, R. H., and E. O. Wilson. 1967, The Theory of Island Biogeography. Princeton, Princeton University Press.

Magurran, A. E. 2004, Measuring Biological Diversity. Oxford, Blackwell Science.

—. 2005. Species abundance distributions: pattern or process? Functional Ecology 19:177-181.

Magurran, A. E., and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714-716.

Maurer, B. A., and B. J. McGill. 2004. Neutral and non-neutral macroecology. Basic and Applied Ecology 5:413-422.

May, R. 1999. Unanswered questions in ecology. Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences 354:1951-1959.

May, R. M. 1975. Patterns of species abundance and diversity, Pages 81-120 *in* M. L. Cody, and J. M. Diamond, eds. Ecology and Evolution of Communities. Cambridge, Belknap Press of Harvard University Press.

Mazancourt, C. 2001. Consequences of community drift. Science 293:1772.

McGill, B., B. A. Maurer, and M. D. Weiser. 2006. Empirical evaluation of neutral theory. Ecology 87: 1411-1423.

McGill, B. J. 2003a. Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? Ecology Letters 6:766-773.

—. 2003b. Strong and weak tests of macroecological theory. Oikos 102:679-685.

—. 2003c. A test of the unified neutral theory of biodiversity. Nature 422:881-885.

McKane, A., D. Alonso, and R. V. Sole. 2000. Mean-field stochastic theory for species-rich assembled communities. Physical Review E 62:8466-8484.

McKane, A. J., D. Alonso, and R. V. Sole. 2004. Analytic solution of Hubbell's model of local community dynamics. Theoretical Population Biology 65:67-73.

Mouquet, N., and M. Loreau. 2003. Community patterns in source-sink metacommunities. American Naturalist 162:544-557.

Nee, S., P. H. Harvey, and R. M. May. 1991. Lifting the veil on abundance patterns. Proceedings of the Royal Society of London, B 243:161-163.

Nee, S., and G. Stone. 2003. The end of the beginning for neutral theory. Trends in Ecology & Evolution 18:433-434.

Norris, S. 2003. Neutral theory: A new, unified model for ecology. Bioscience 53:124-129.

Olszewski, T. D., and D. H. Erwin. 2004. Dynamic response of Permian brachiopod communities to long-term environmental change. Nature 428:738-741.

Pandolfi, J. M. 1996. Limited membership in Pleistocene reef coral assemblages from the Huon Peninsula, Papua New Guinea: Constancy during global change. Palaeobiology 22:152-176.

—. 1999. Response of Pleistocene coral reefs to environmenatl change over long temporal scales. American Zoologist 39:113-.

Pielou, E. C. 1969, An Introduction to Mathematical Ecology. New York, Wiley-Interscience.

—. 1975. Species Abundance Distributions, Pages 19-31 Ecological Diversity. New York, Wiley Interscience.

Pitman, N. C. A., J. W. Terborgh, M. R. Silman, P. Nunez, D. A. Neill, C. E. Ceron, W. A. Palacios et al. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. Ecology 82:2101-2117.

Preston, F. W. 1948. The commonness, and rarity, of species. Ecology 29:254-283.

Ricklefs, R. E. 2003. A comment on Hubbell's zero-sum ecological drift model. Oikos 100:185-192.

Schwilk, D. W., and D. D. Ackerly. 2005. Limiting similarity and functional diversity along environmental gradients. Ecology Letters 8:272-281.

Silander, J. A. 2002. The ultimate answer is. 42? A critical review of S. P. Hubbell (2001) "The unified neutral theory of biodiversity and biogeography". Journal of Biogeography 29:299-301.

Sugihara, G. 1980. Minimal community structure: an explanaition of species abundance patterns. The American Naturalist 116:770-787.

Sugihara, G., L. F. Bersier, T. R. E. Southwood, S. L. Pimm, and R. M. May. 2003. Predicted correspondence between species abundances and dendrograms of niche similarities. Proceedings of the National Academy of Sciences of the United States of America 100:5246-5251.

Terborgh, J., R. B. Foster, and P. Nunez. 1996. Tropical tree communities: A test of the nonequilibrium hypothesis. Ecology 77:561-567.

Tilman, D. 1982, Resource competition and community structure. New Jersey, Princeton University Press.

—. 1990. Constraints and trade-offs: toward a predictive theory of compatition and succession. Oikos 58:3-15.

Tilman, D., and S. Pacala. 1993. The maintenance of species richness in plant communities, Pages 13-25 *in* D. Schluter, and R. E. Ricklefs, eds. Species diversity in ecological communities, historical and geographical perspectives. Chicago and London, Chicago University Press.

Tokeshi, M. 1990. Niche apportionment or random assortment: species abundance patterns revisited. Journal of Animal Ecology 59:1129-1146.

Tokeshi, M., and P. E. Schmid. 2002. Niche division and abundance: an evolutionary perspective. Population Ecology 44:189-200.

Vallade, M., and B. Houchmandzadeh. 2003. Analytical solution of a neutral model of biodiversity. Physical Review E 68.

Veron, J. E. N. 2000, Corals of the World. Townsville, Australian Institute of Marine Science.

Volkov, I., J. R. Banavar, F. He, S. P. Hubbell, and A. Maritan. 2005. Density dependence explains tree species abundance and diversity in tropical forests. Nature 438:658-661.

Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. Nature 424:1035-1037.

Volkov, I., J. R. Banavar, A. Maritan, and S. P. Hubbell. 2004. The stability of forest biodiversity. Nature 427:696-696.

Watterson, G. A. 1974. Models for the Logarithmic Species Abundance Distributions. Theoretical Population Biology 6:217-250.

Whitfield, J. 2002. Ecology: Neutrality versus the niche. Nature 417:480-481.

Whittaker, R. H. 1965. Dominance and Diversity in Land Plant Communities. Science 147:250-260.

—. 1972. Evolution and measurement of species diversity. Taxon 21:213-251.

Williamson, M., and K. J. Gaston. 2005. The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution. Journal of Animal Ecology 74:409-422.

Wilson, J. B. 1991. Methods for Fitting Dominance/Diversity Curves. Journal of Vegetation Science 2:35-46.

Wootton, J. T. 2005. Field parameterization and experimental test of the neutral theory of biodiversity. Nature 433:309-312.

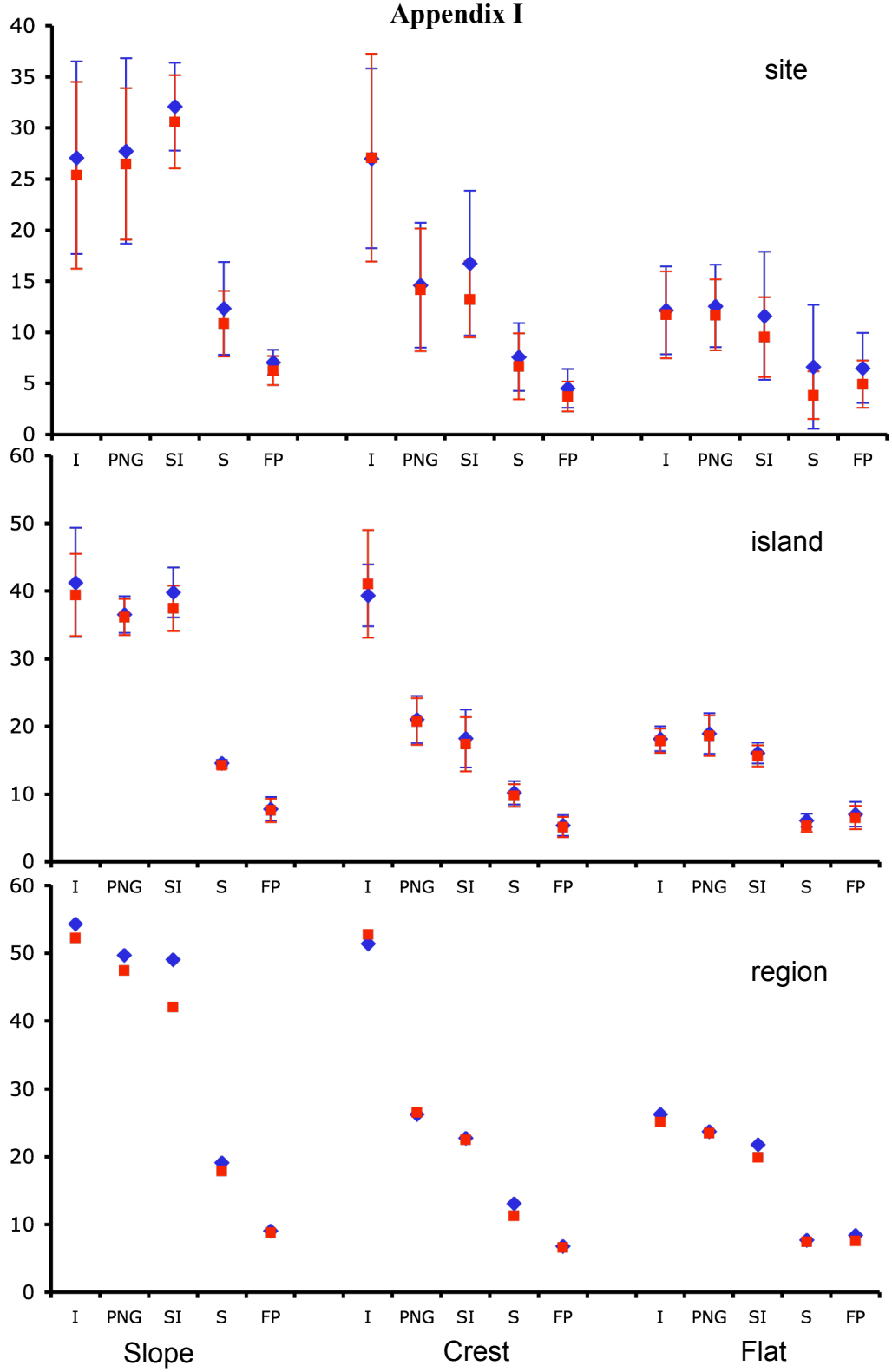


Figure AI.1 – Estimates of θ across the biodiversity gradient (Indonesia, Papua New Guinea, Solomon Islands, American Samoa and French Polynesia) and three habitats (slope, crest and flat) for the SNM (blue diamonds) and MFNM (red squares). Symbols represent the mean and error bars one standard deviation.

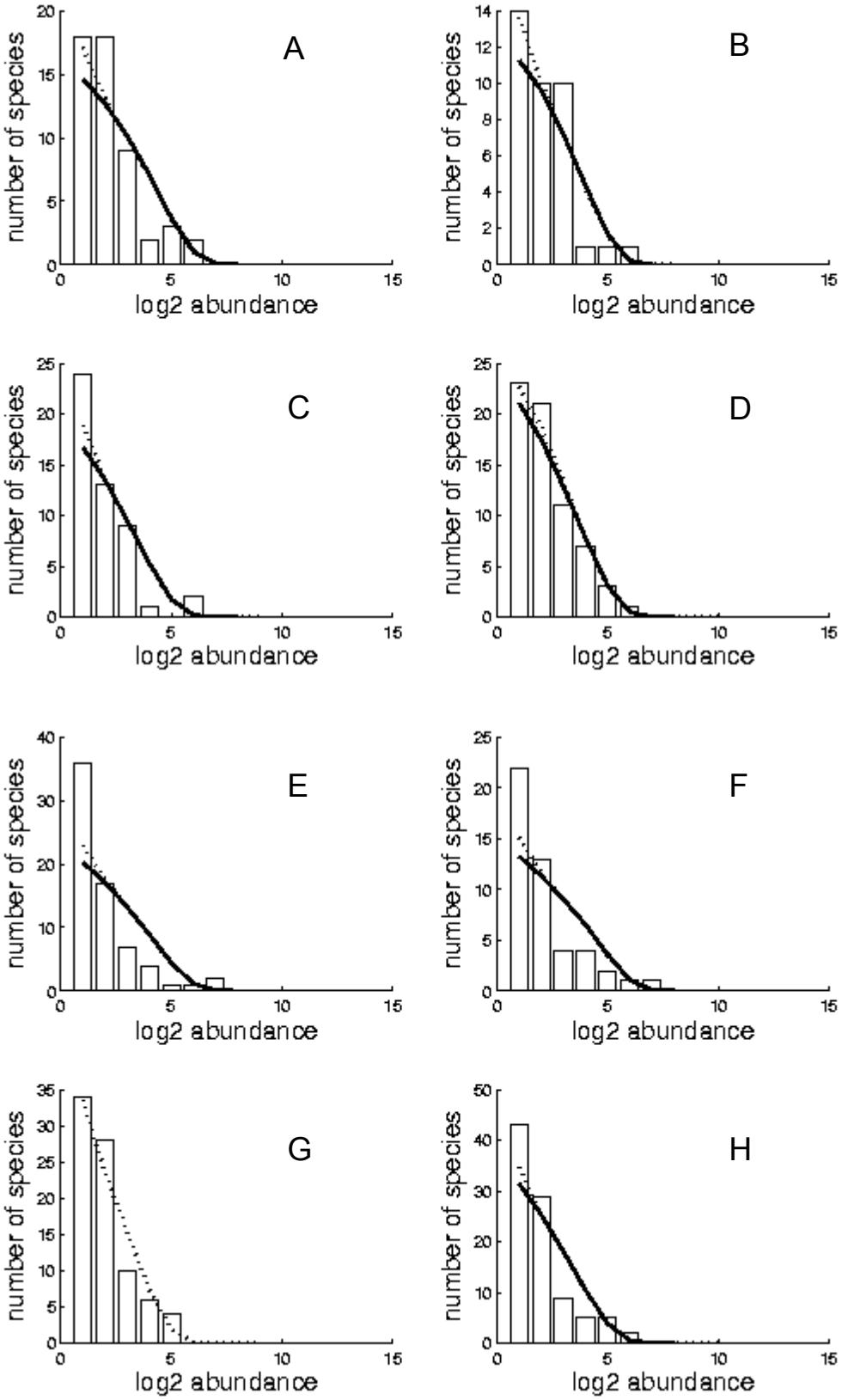


Figure AI.2 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Irian Jaya (A to D) and Manado (E to H).

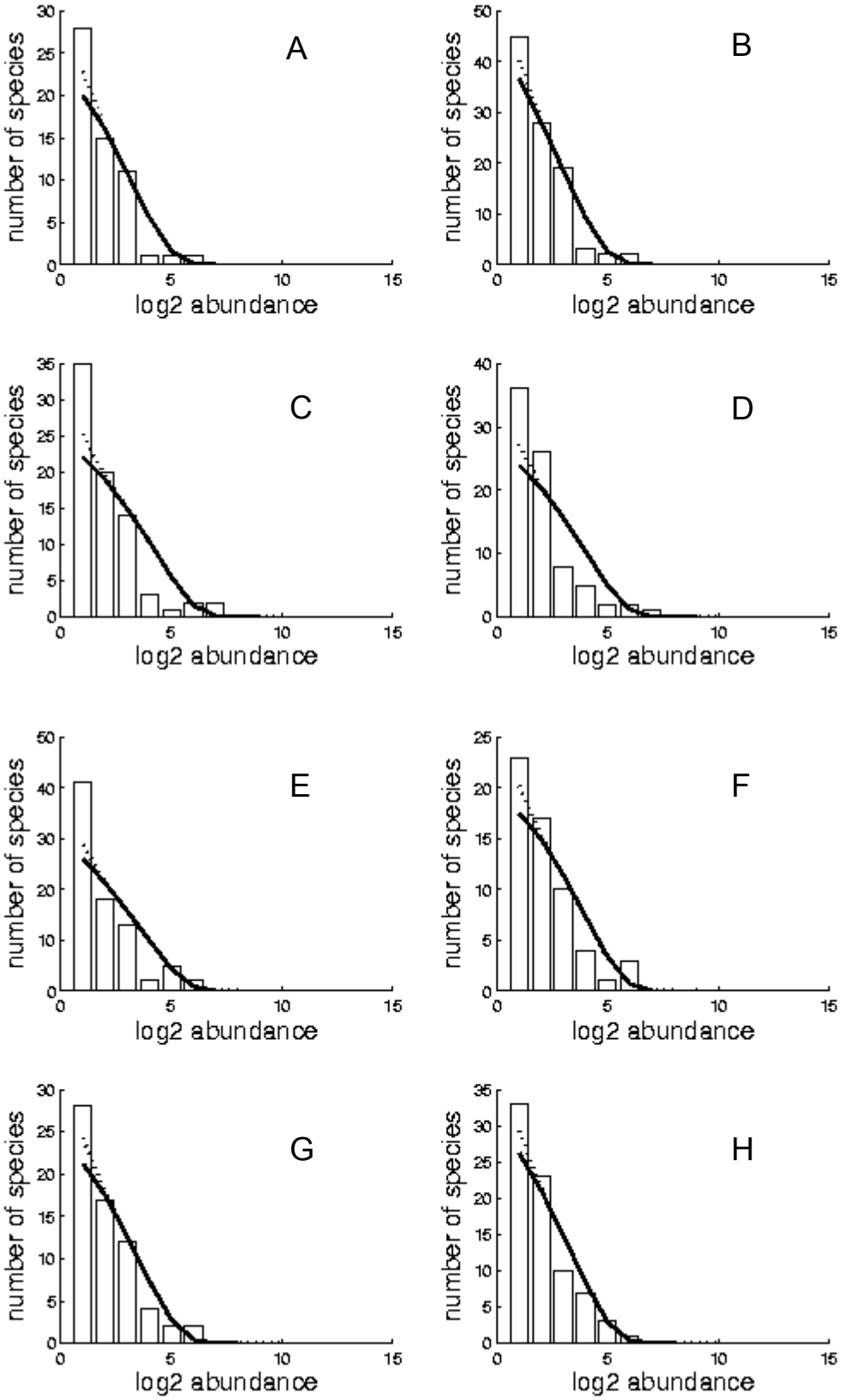


Figure AI.3 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Wakatobi (A to D) and Kimbe (E to H).

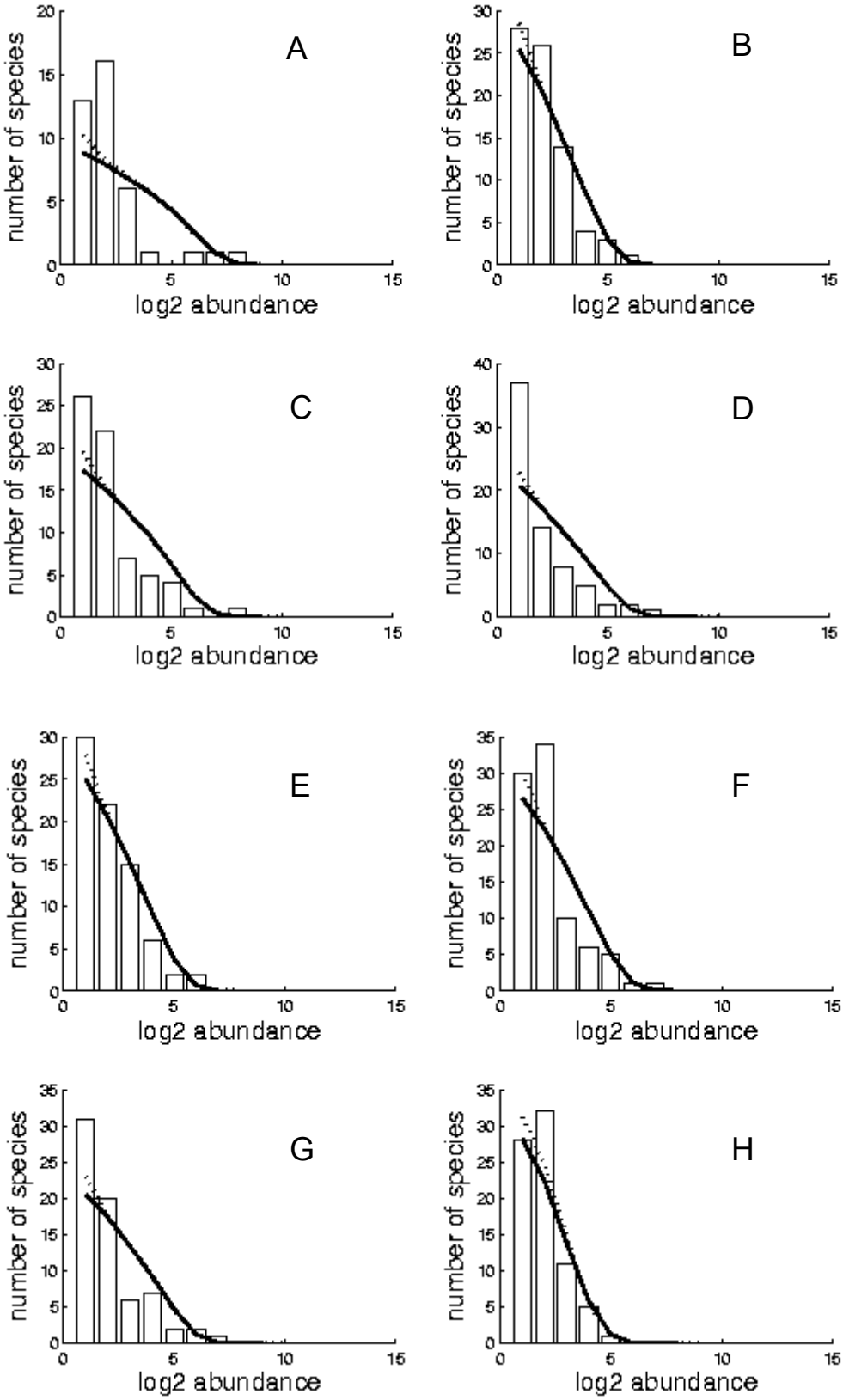


Figure AI.4 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Kavieng (A to D) and Madang (E to H).

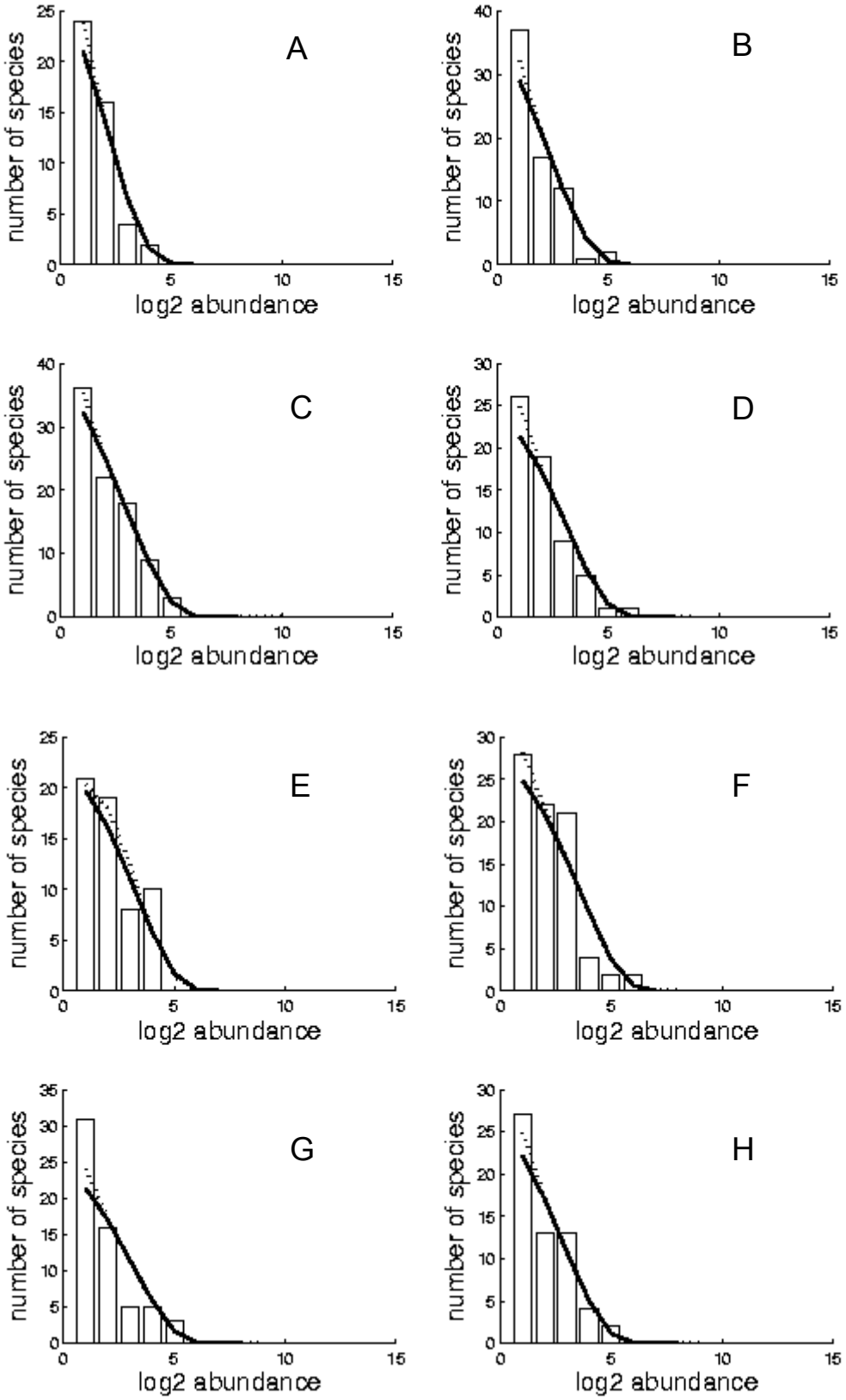


Figure AI.5 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Gizo (A to D) and Munda (E to H).

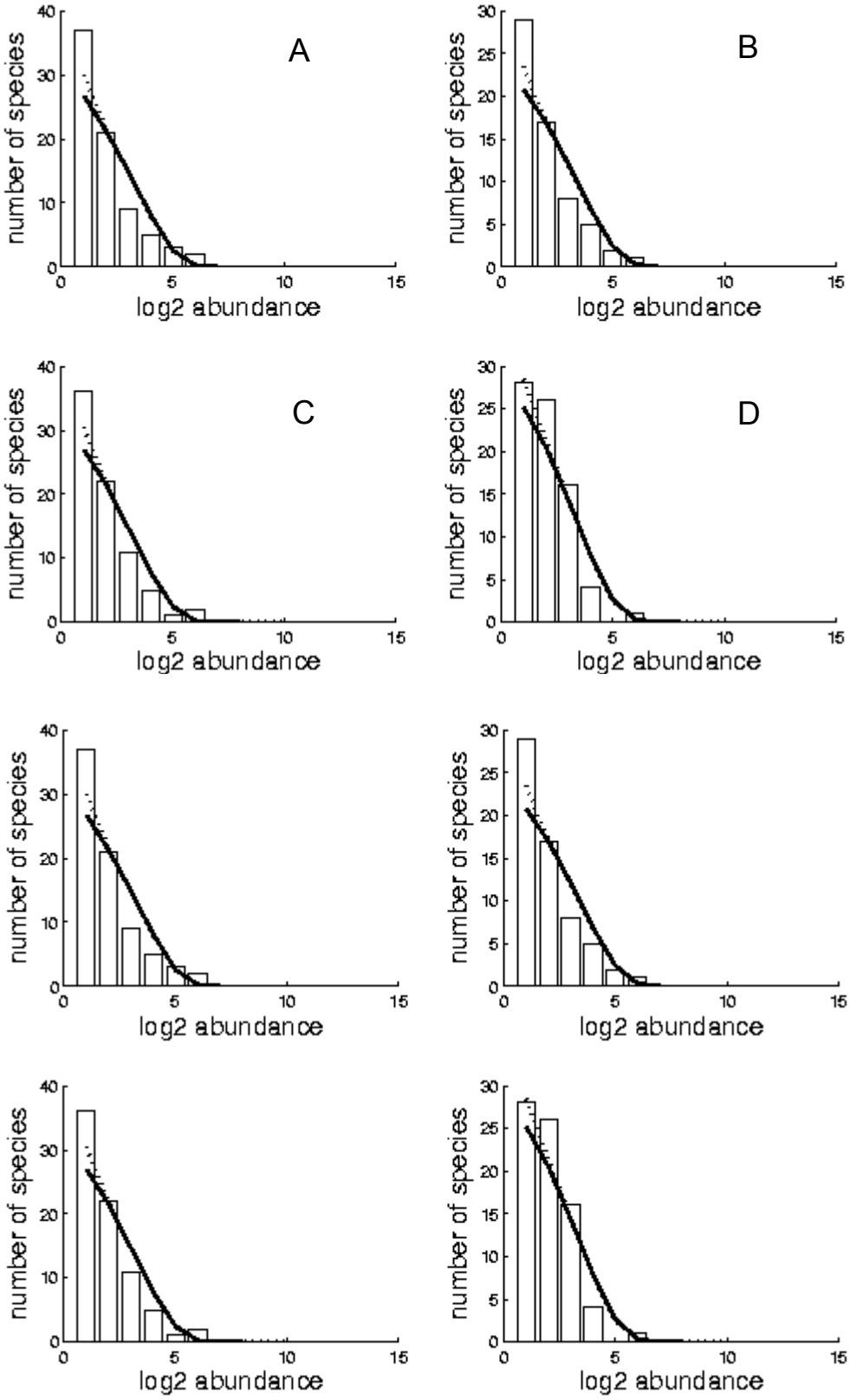


Figure AI.6 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Upei (A to D) and Tutuila (E to H).

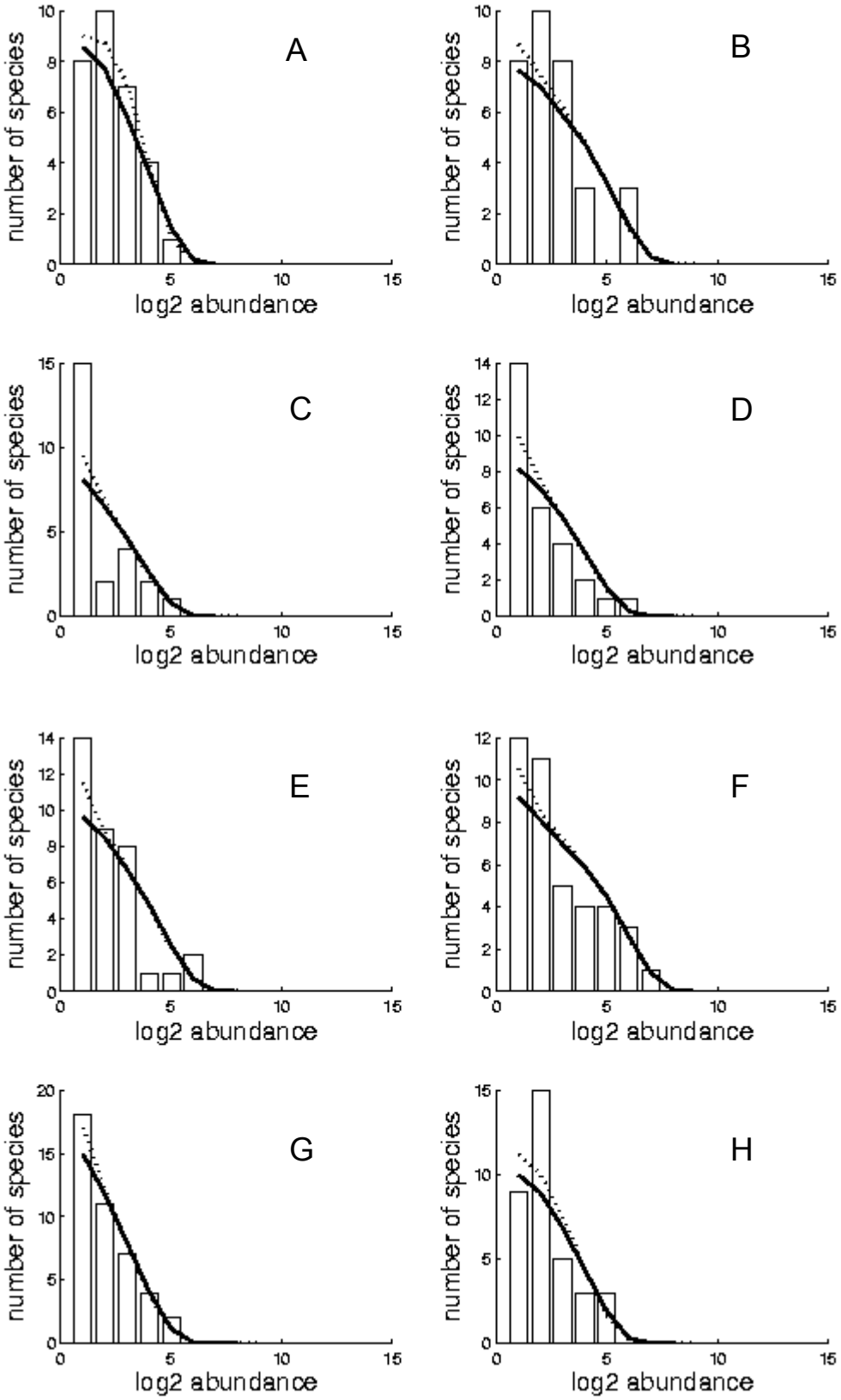


Figure AI.7 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Ofu (A to D) and Tau (E to H).

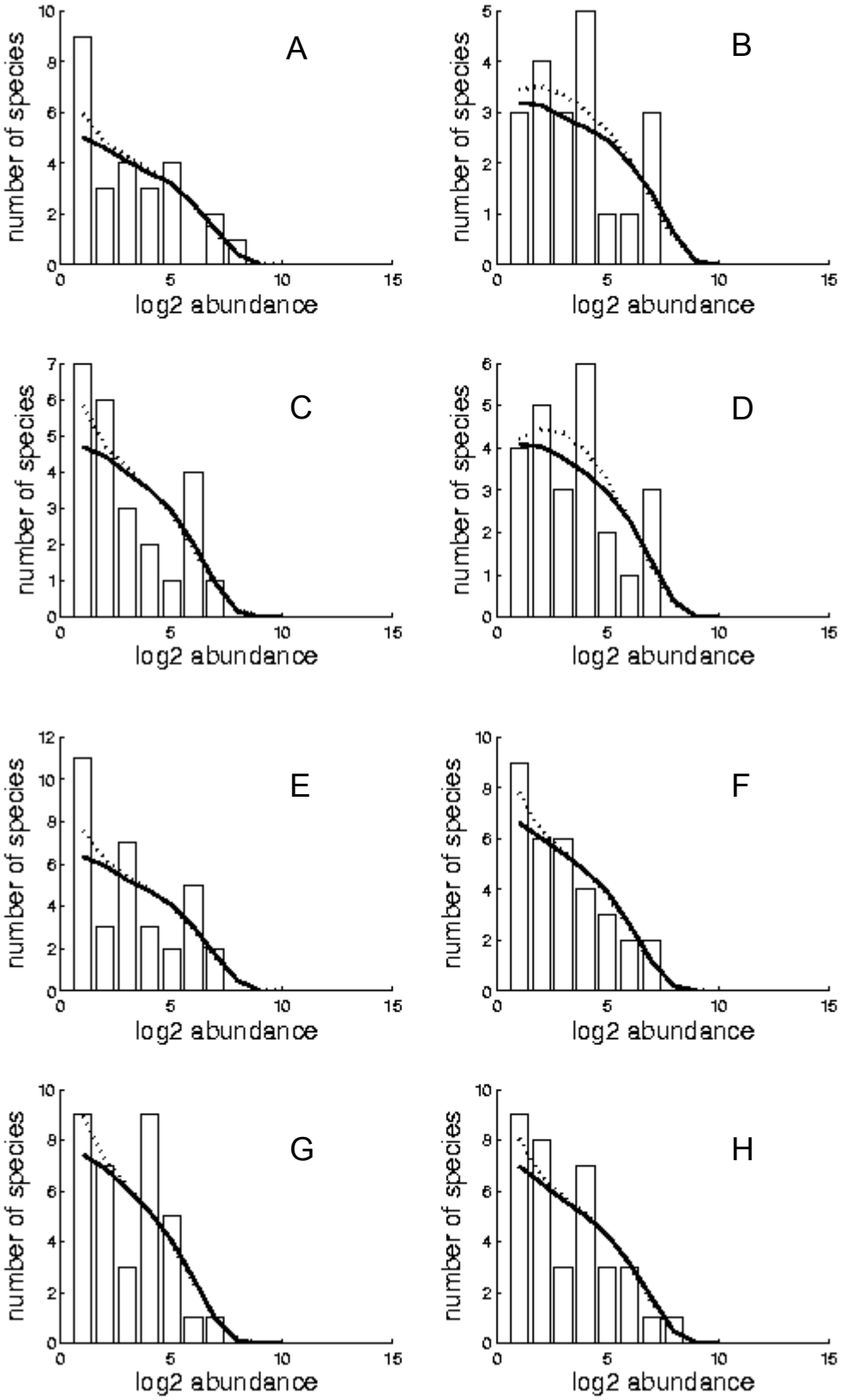


Figure AI.8 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Moorea (A to D) and Tahiti (E to H).

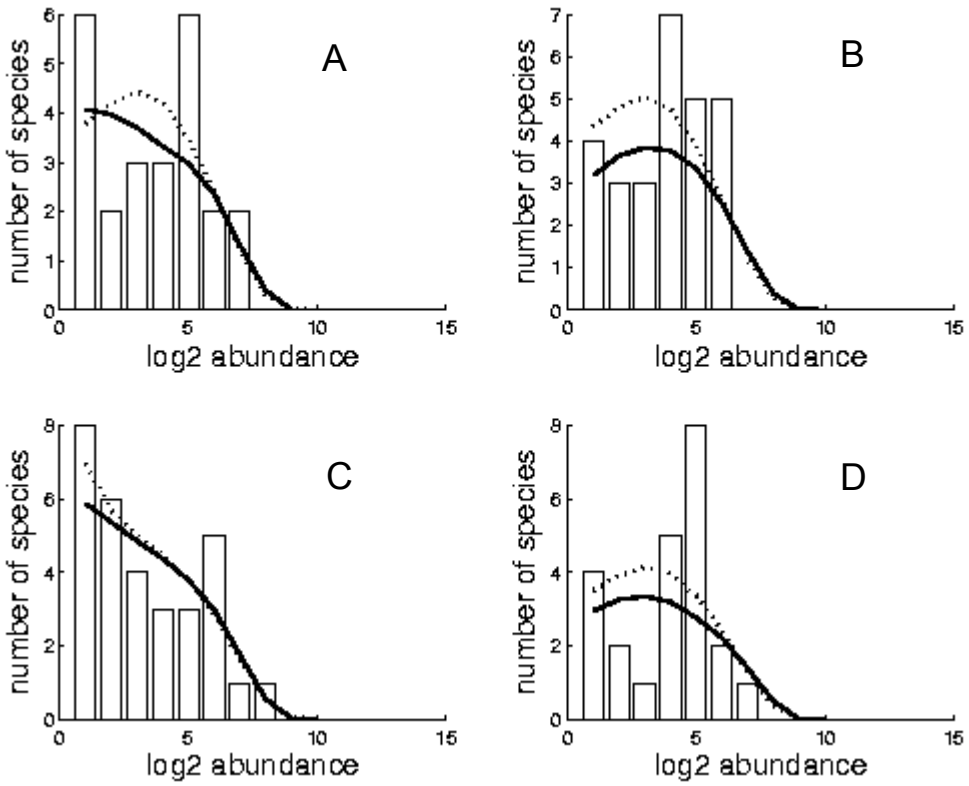


Figure AI.9 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Raiatea (A to D).

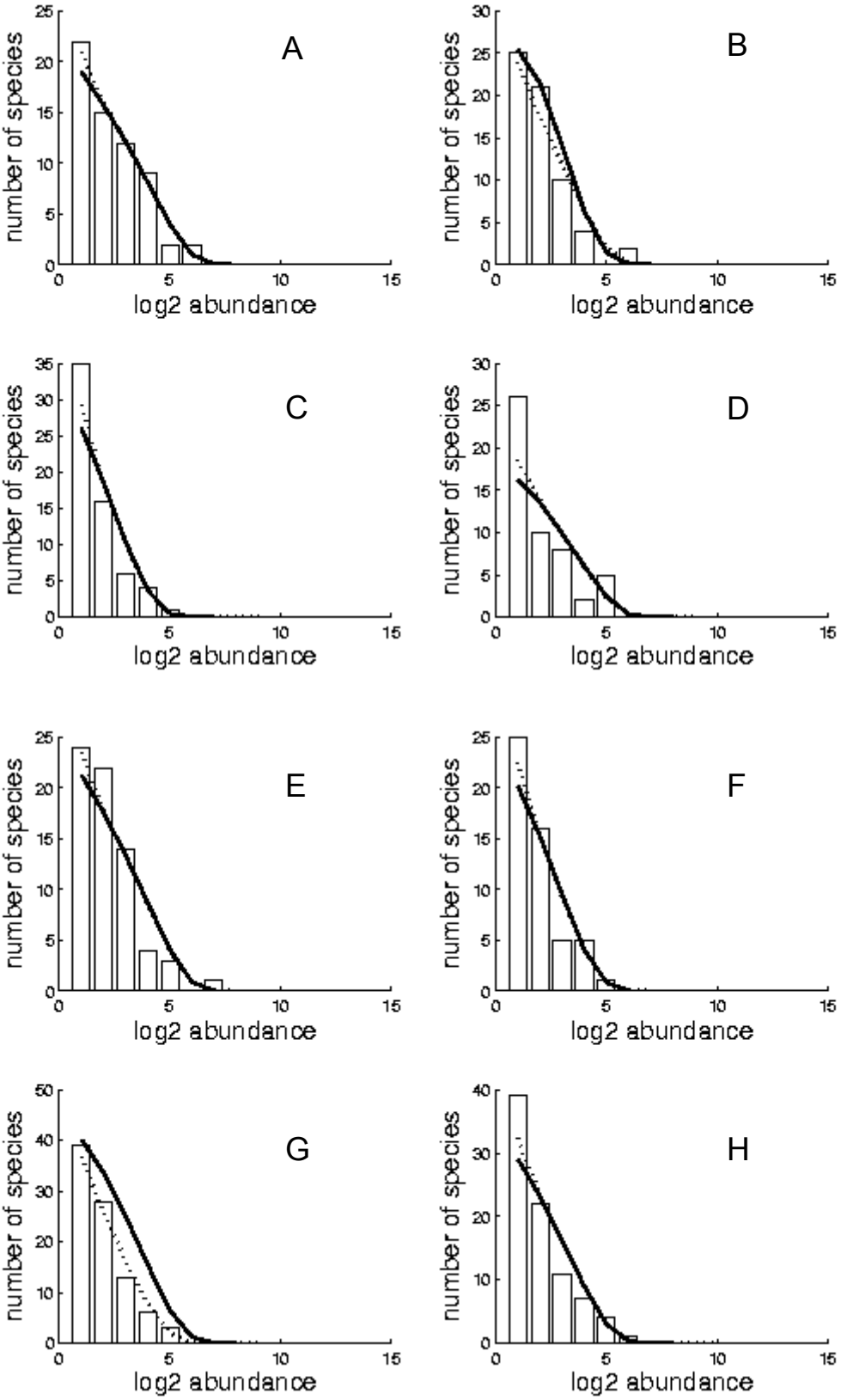


Figure AI.10 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Irian Jaya (A to D) and Manado (E to H).

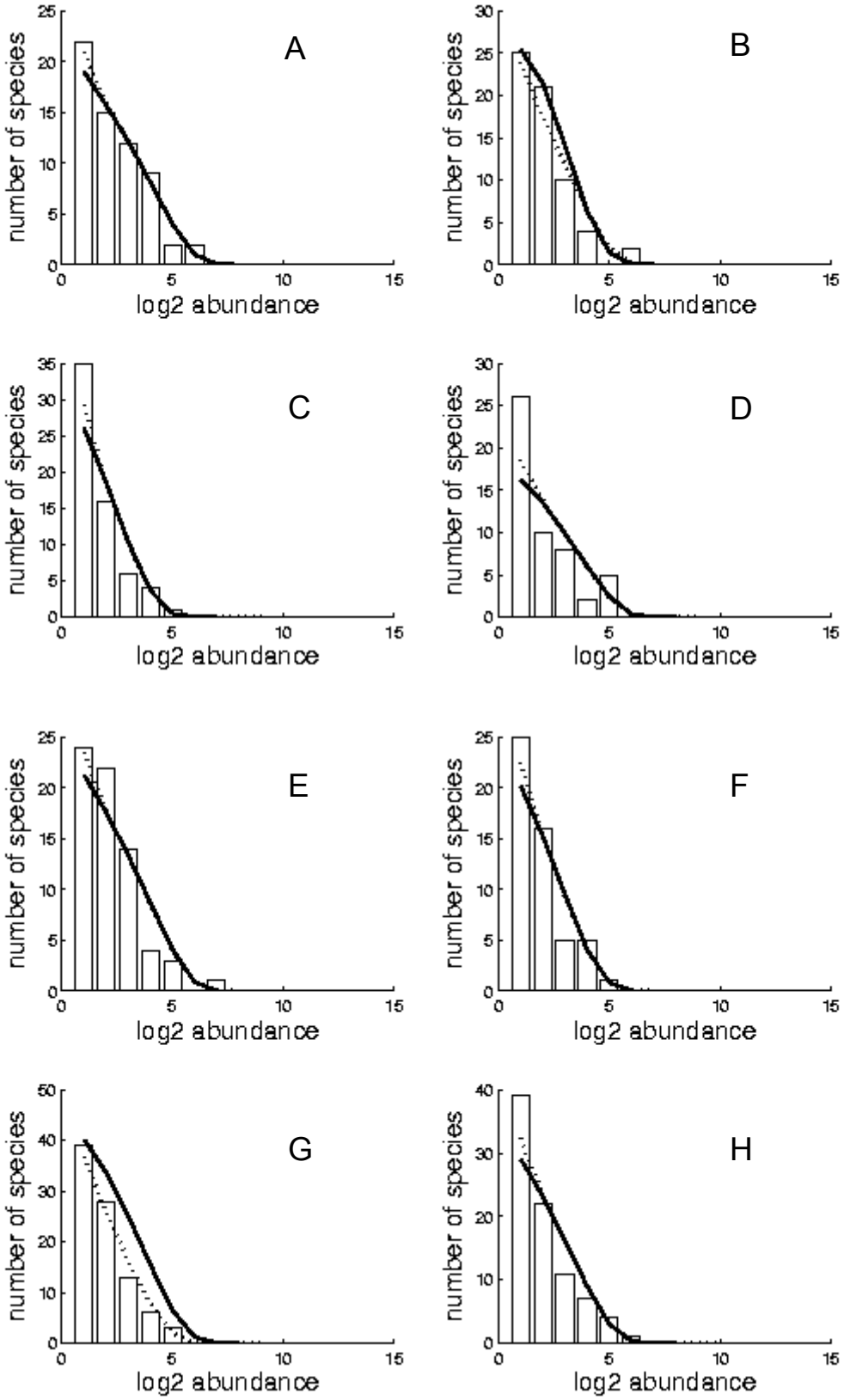


Figure AI.11 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Wakatobi (A to D) and Kimbe (E to H).

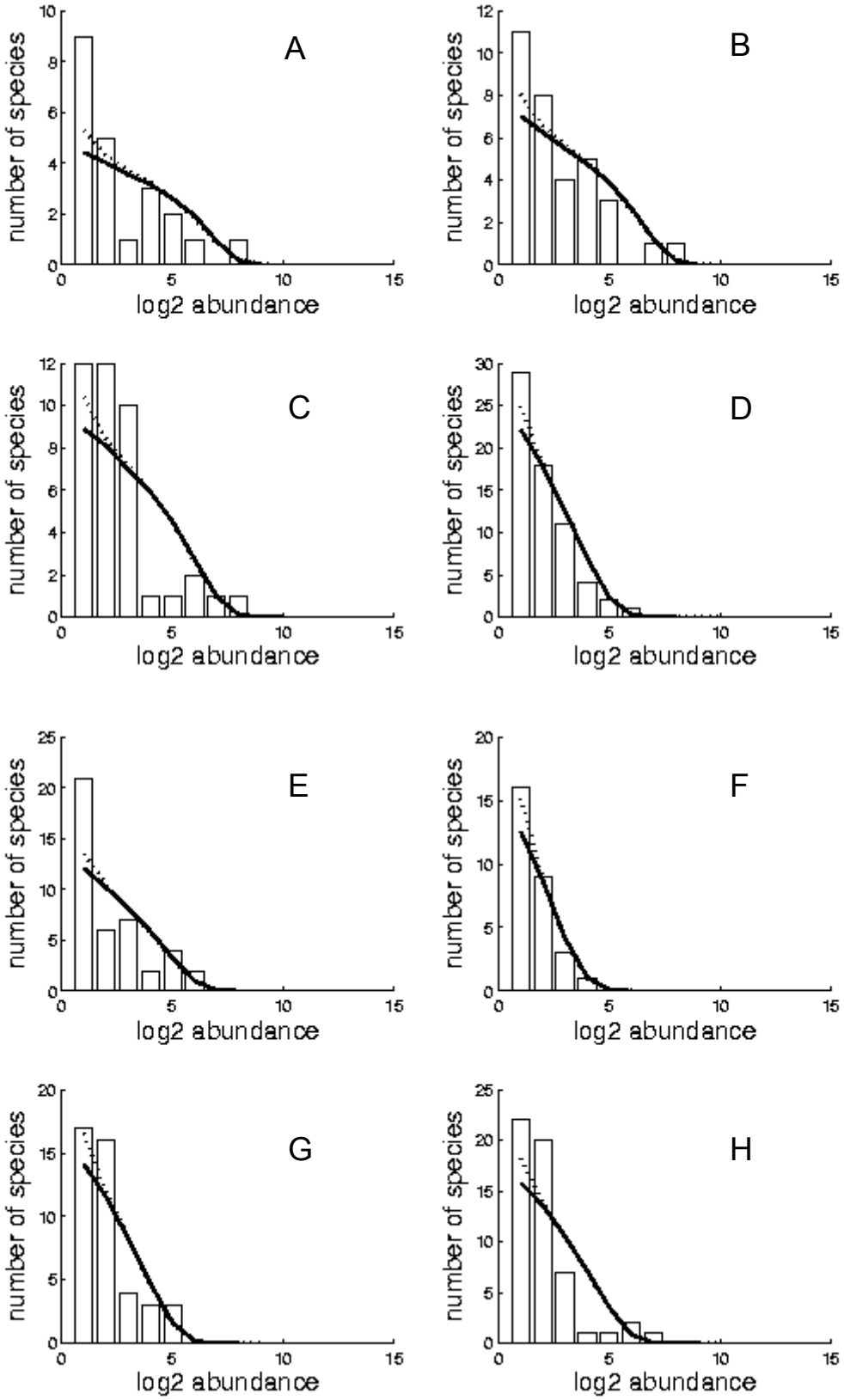


Figure AI.12 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Kavieng (A to D) and Madang (E to H).

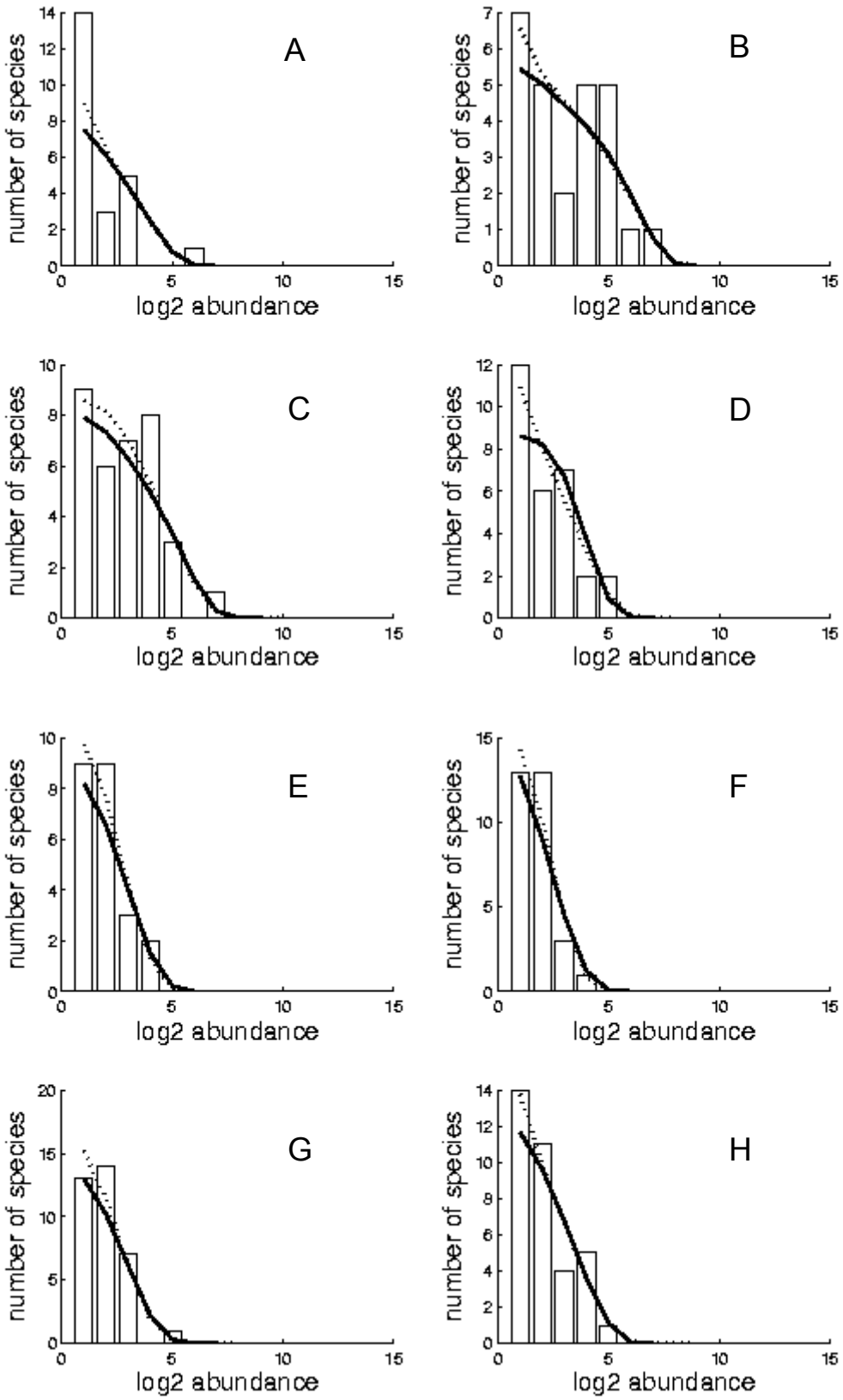


Figure AI.13 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Gizo (A to D) and Munda (E to H).

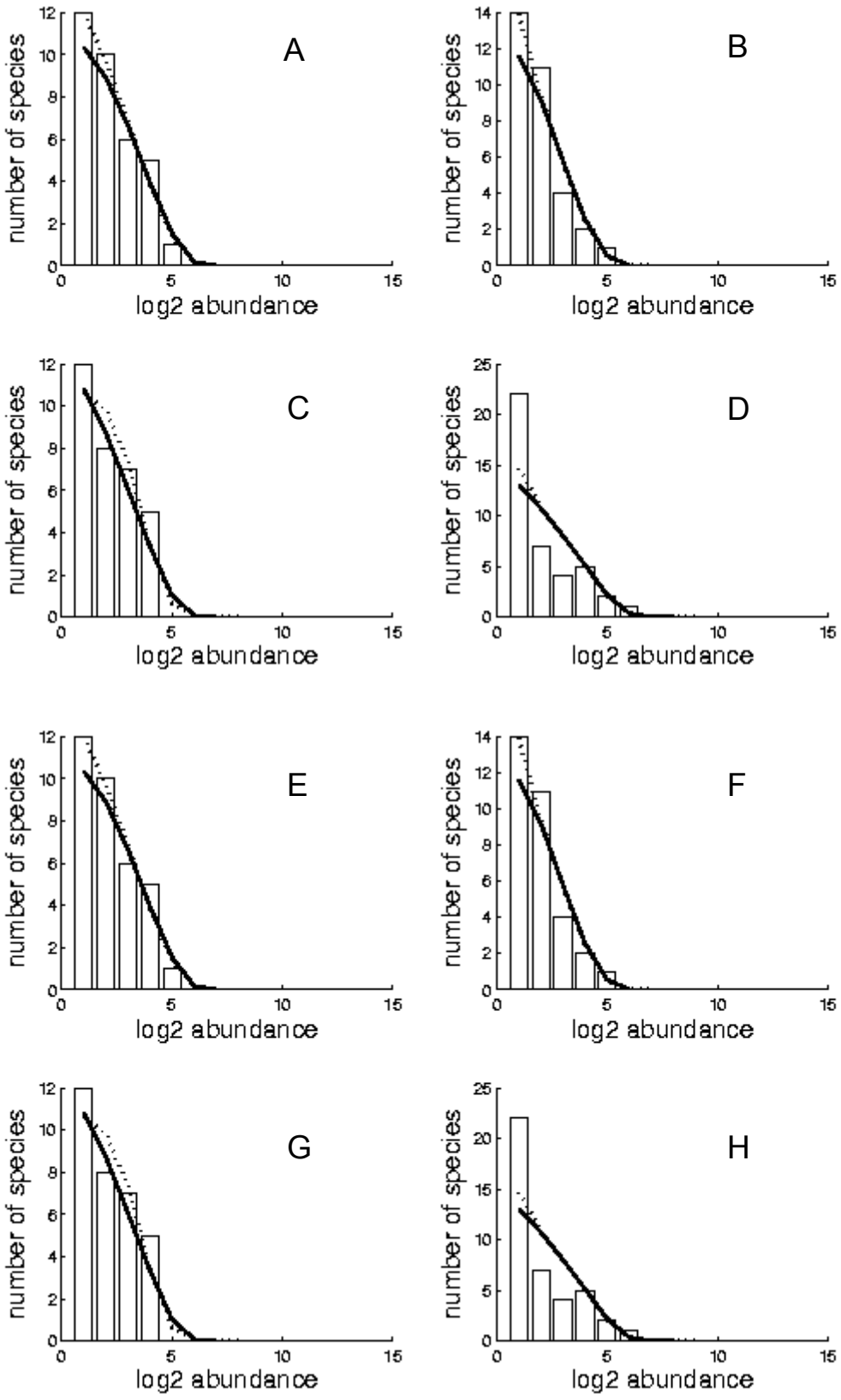


Figure AI.14 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Upei (A to D) and Tutuila (E to H).

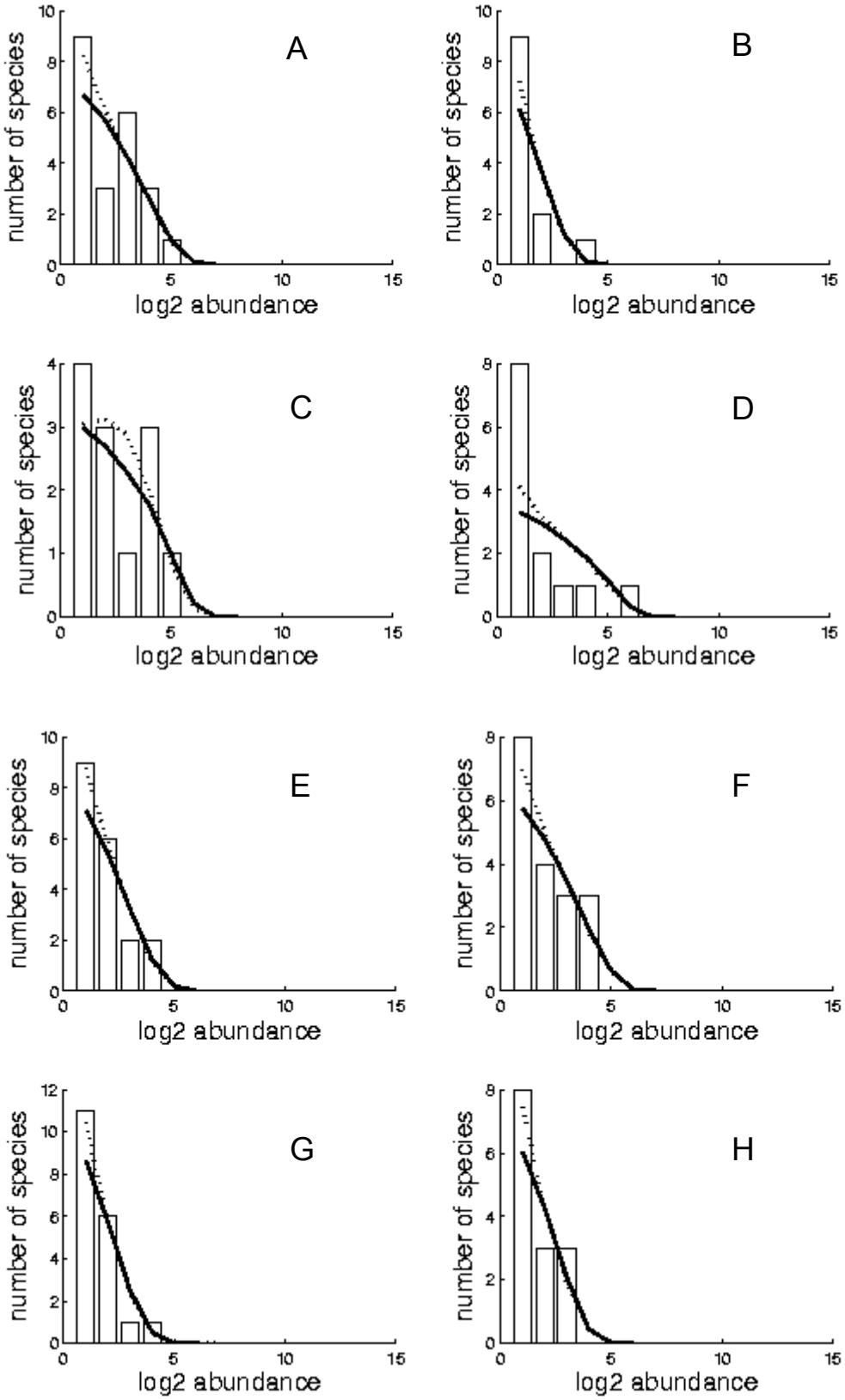


Figure AI.15 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Ofu (A to D) and Tau (E to H).

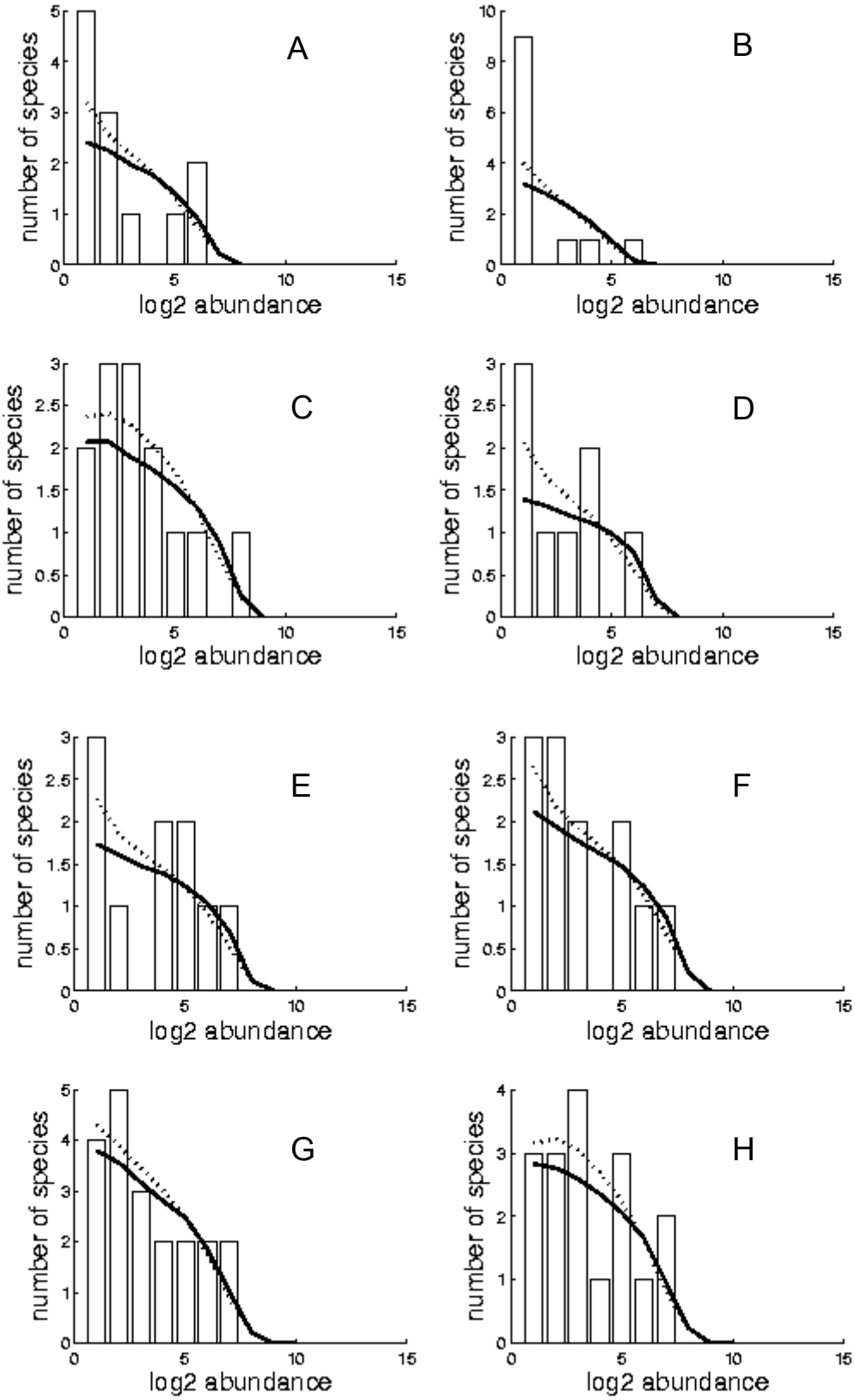


Figure AI.16 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Moorea (A to D) and Tahiti (E to H).

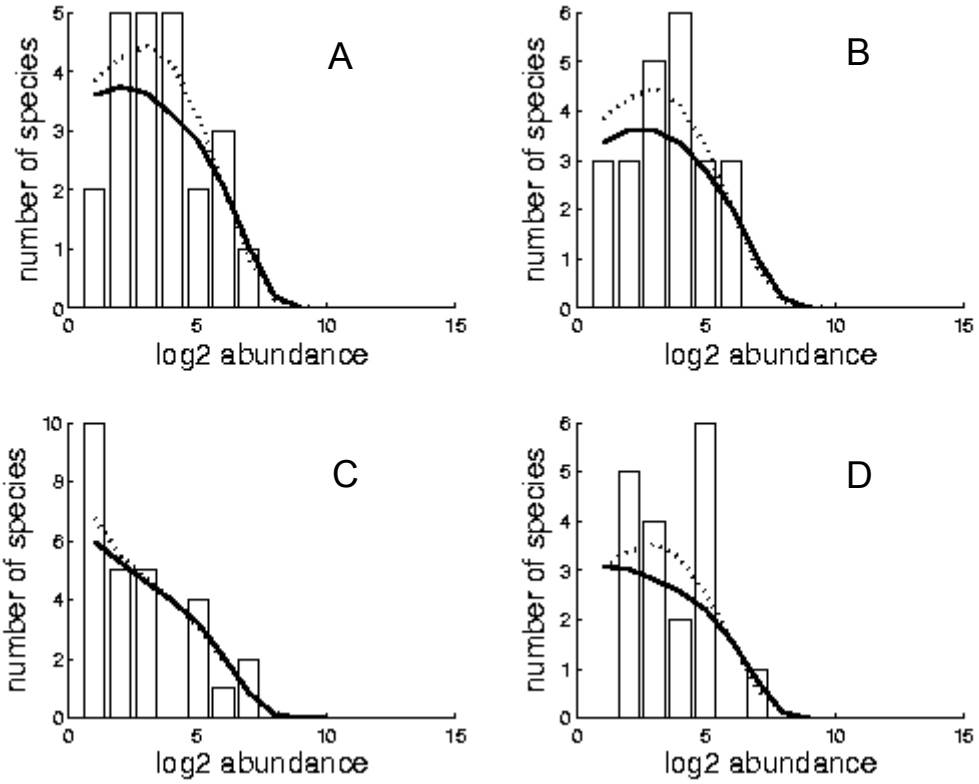


Figure AI.17 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Raiatea (A to D).

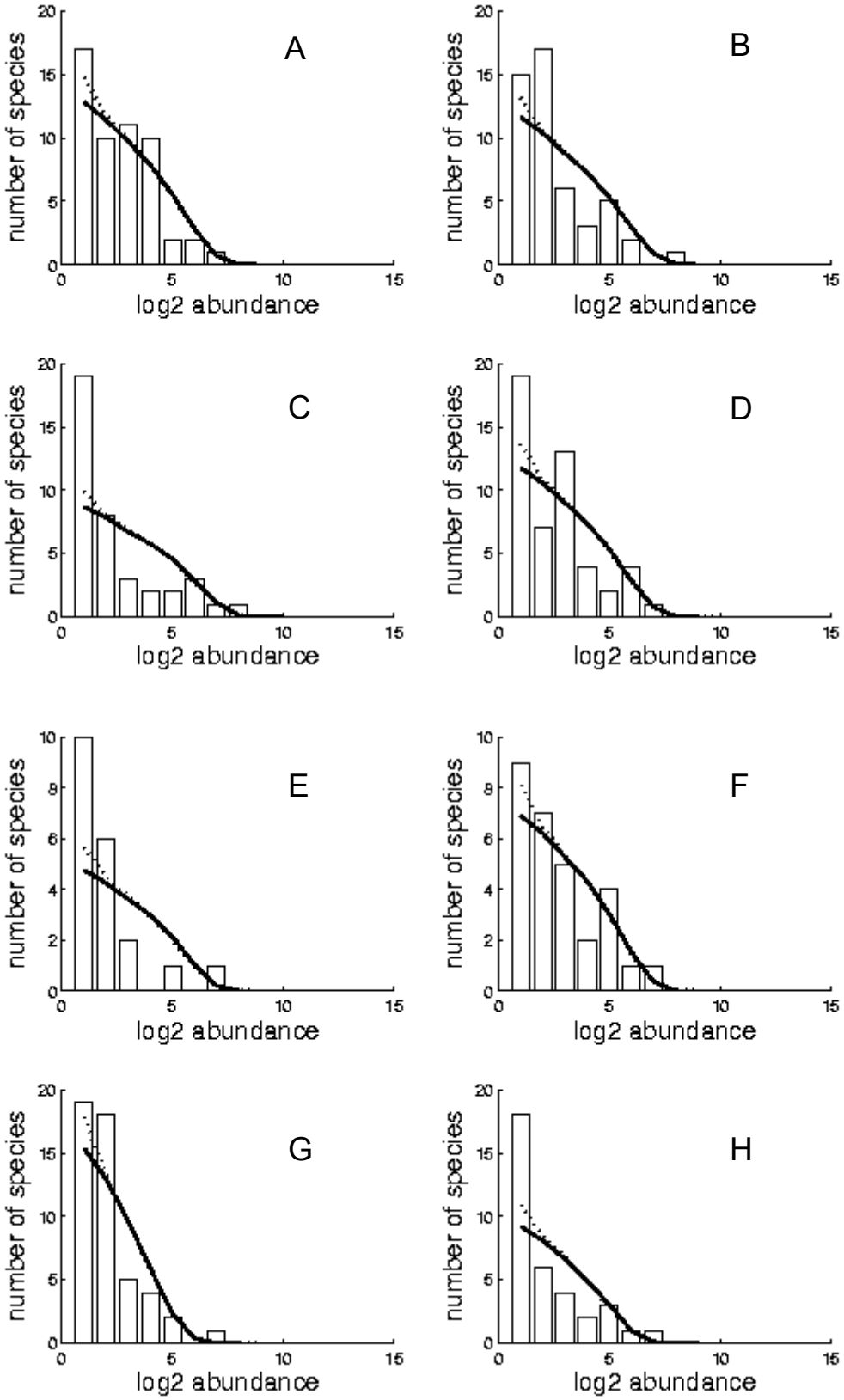


Figure AI.18 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Irian Jaya (A to D) and Manado (E to H).

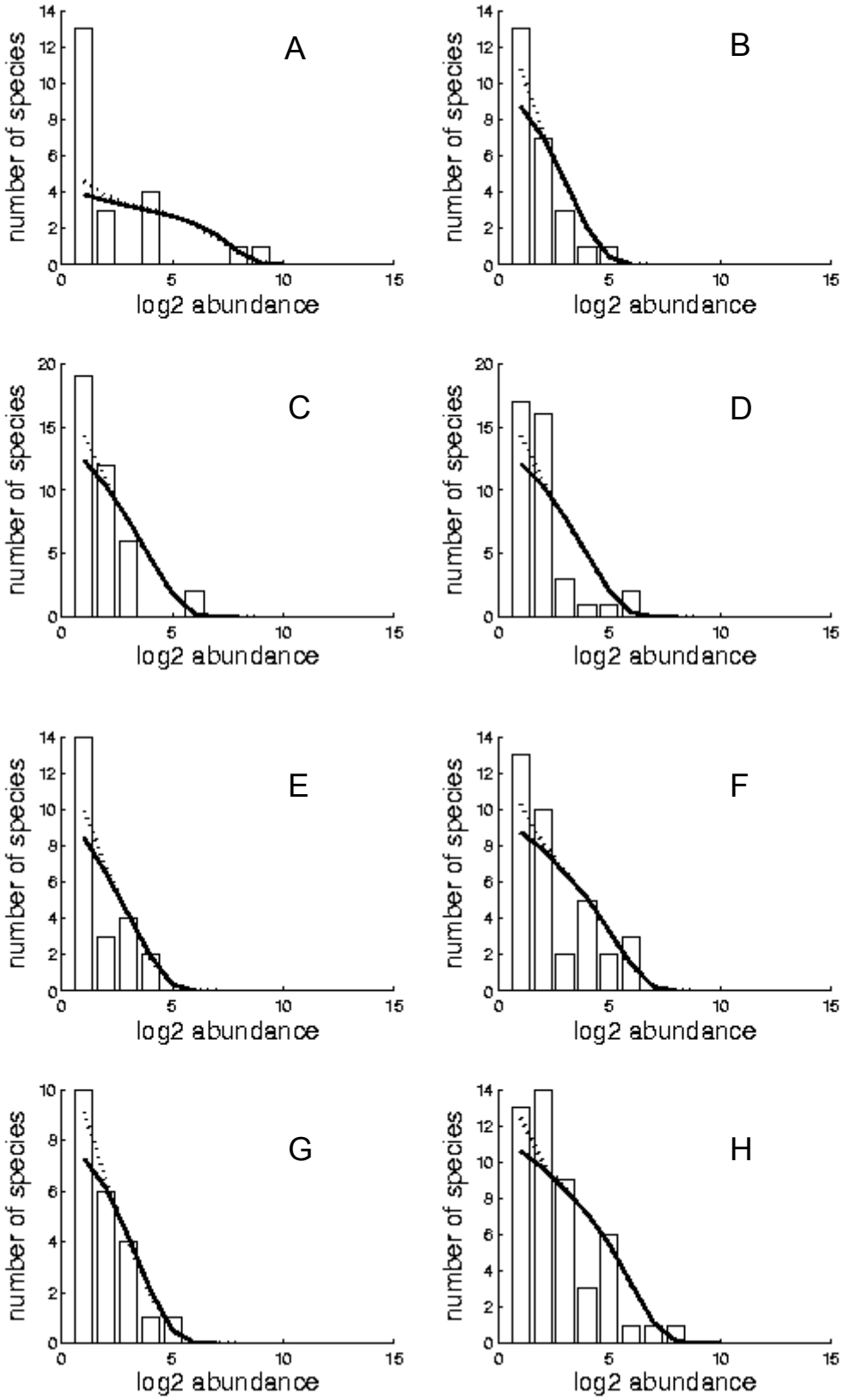


Figure AI.19 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Wakatobi (A to D) and Kimbe (E to H).

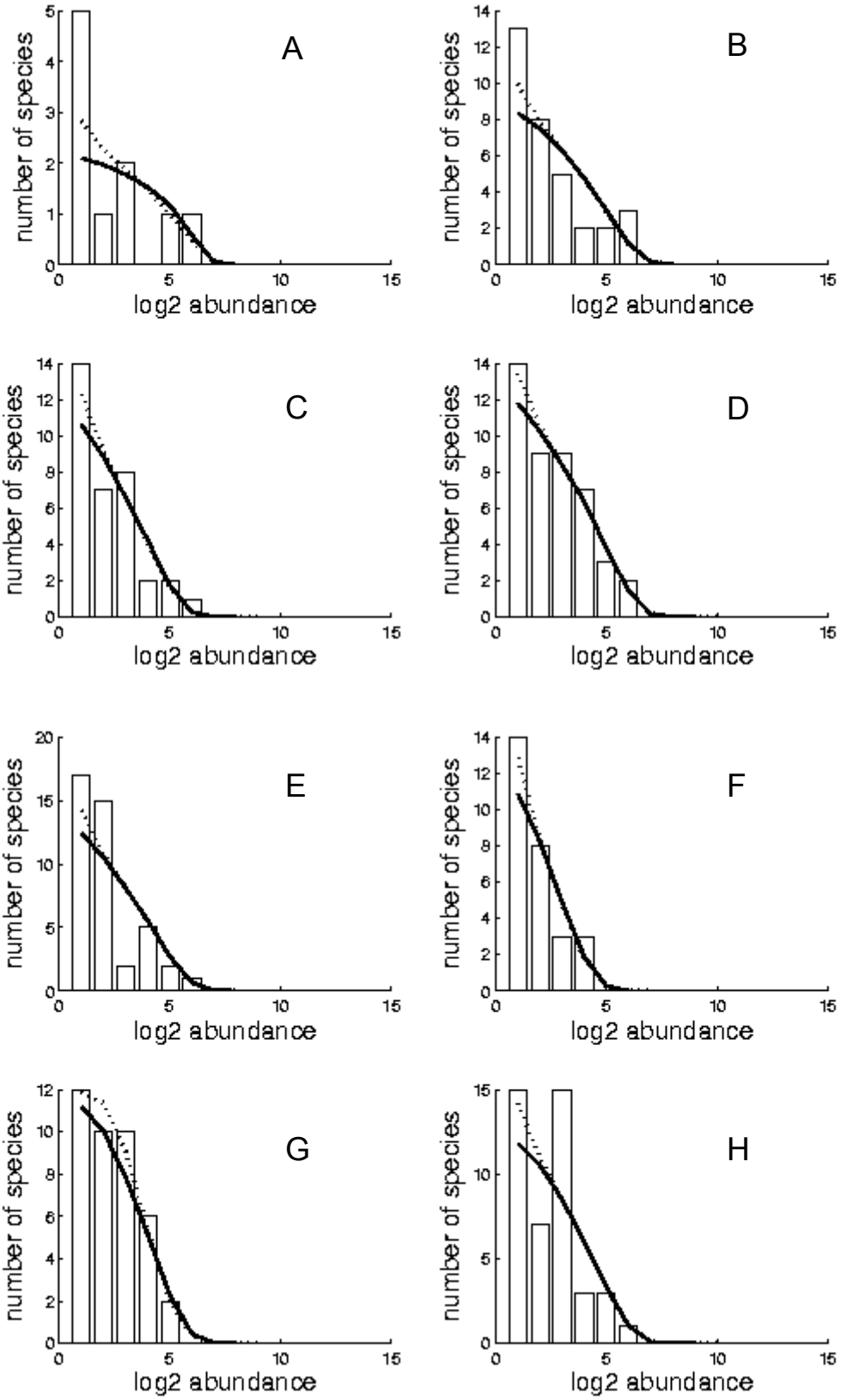


Figure AI.20 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Kavieng (A to D) and Madang (E to H).

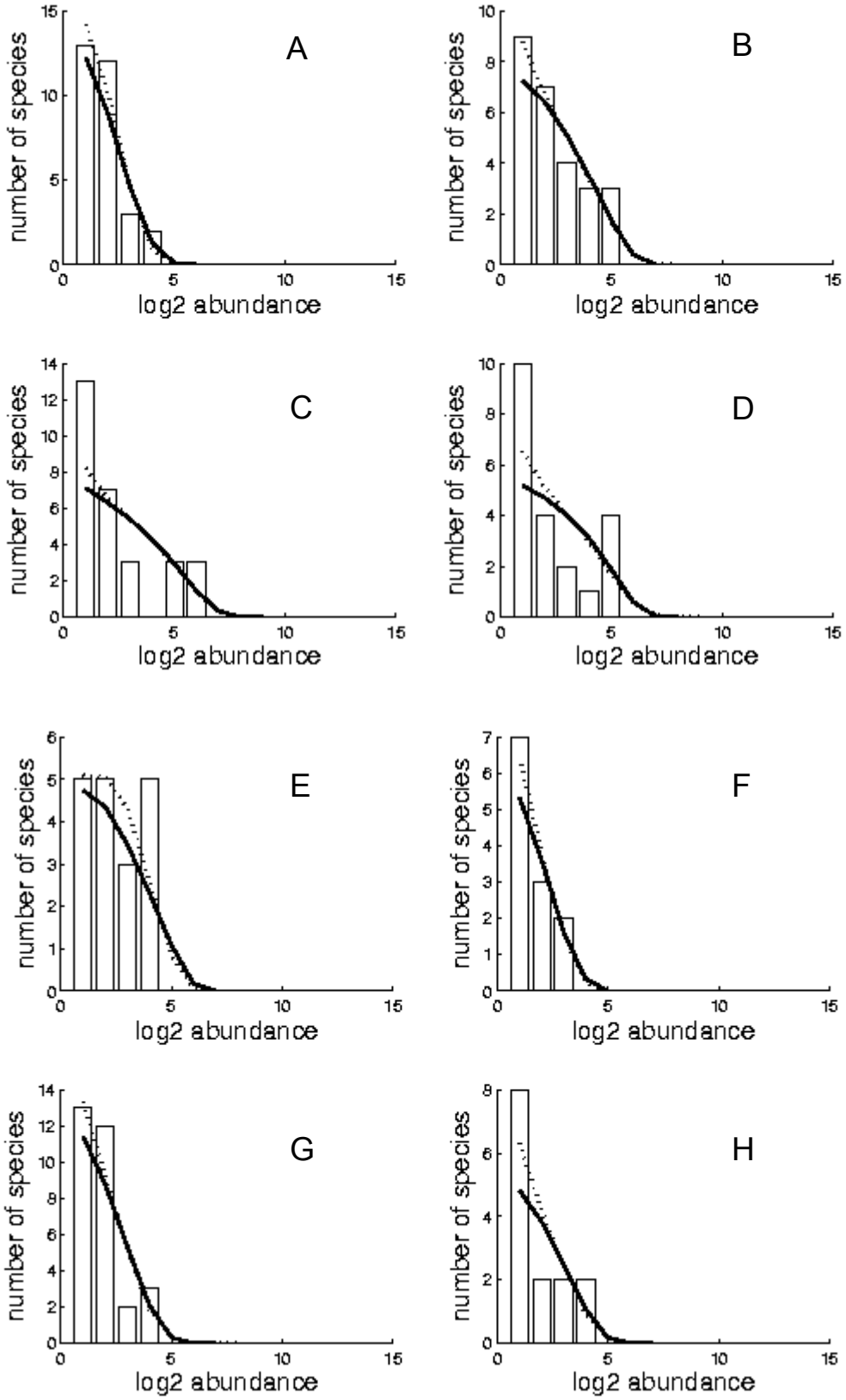


Figure AI.21 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Gizo (A to D) and Munda (E to H).

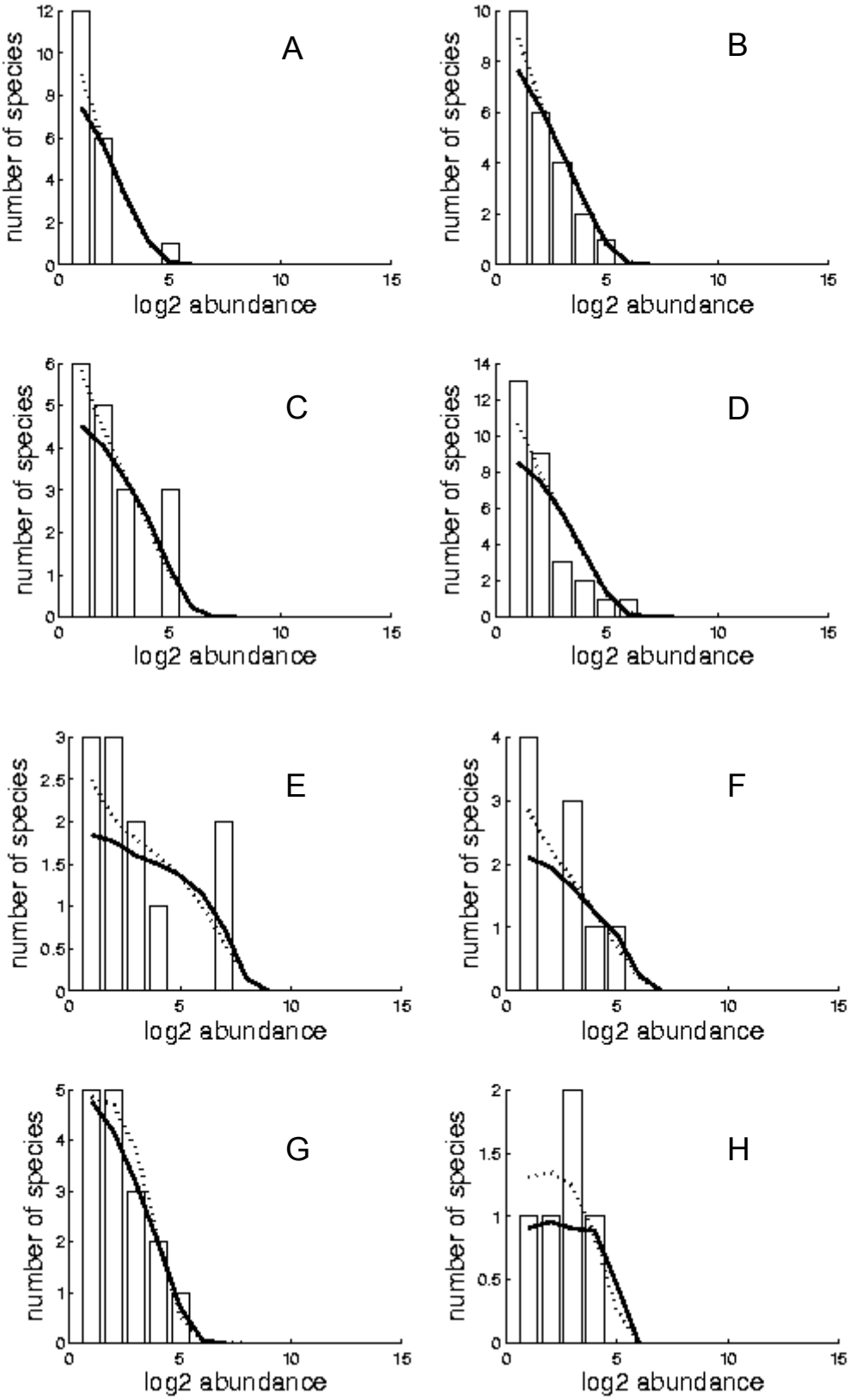


Figure AI.22 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Upei (A to D) and Tutuila (E to H).

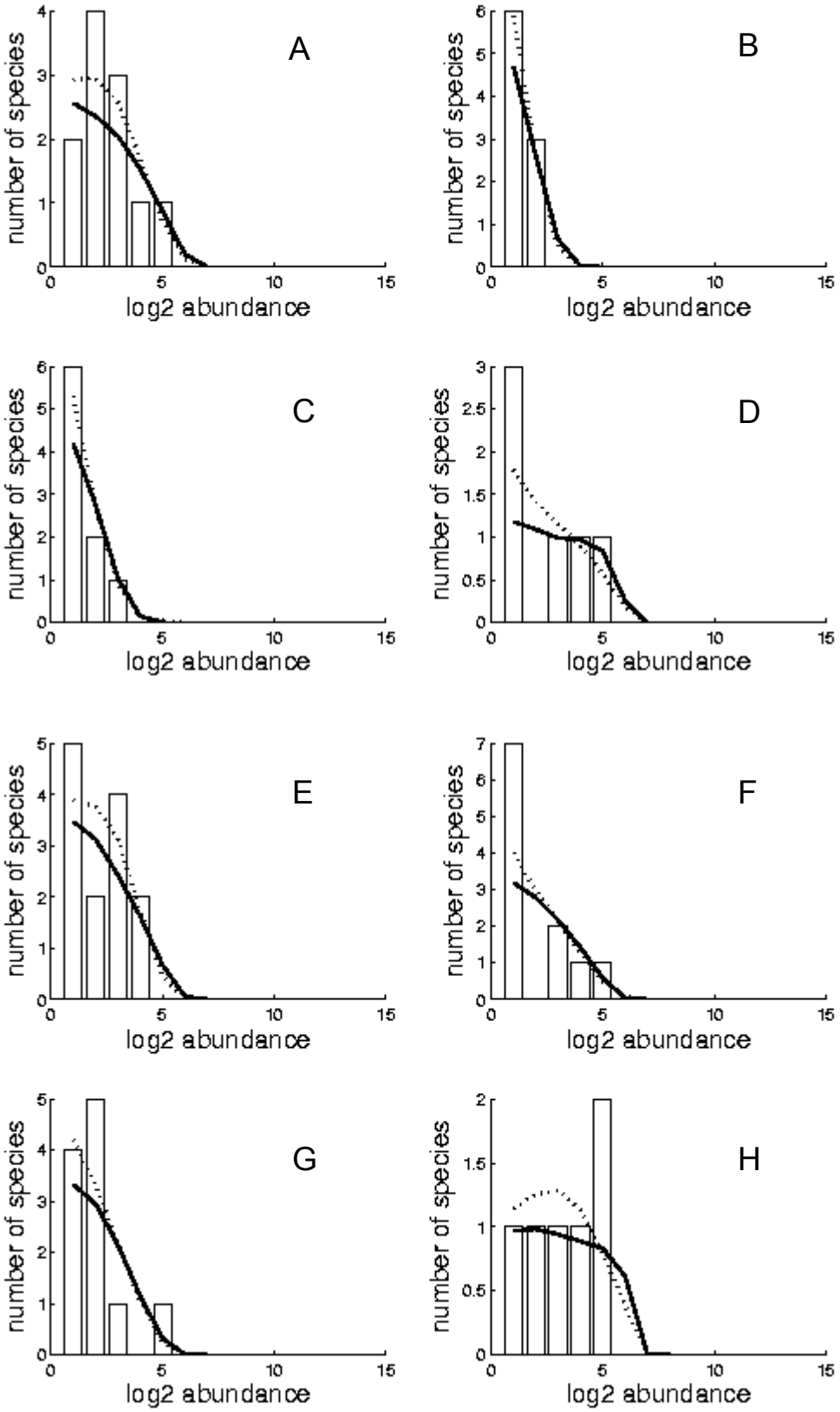


Figure AI.23 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Ofu (A to D) and Tau (E to H).

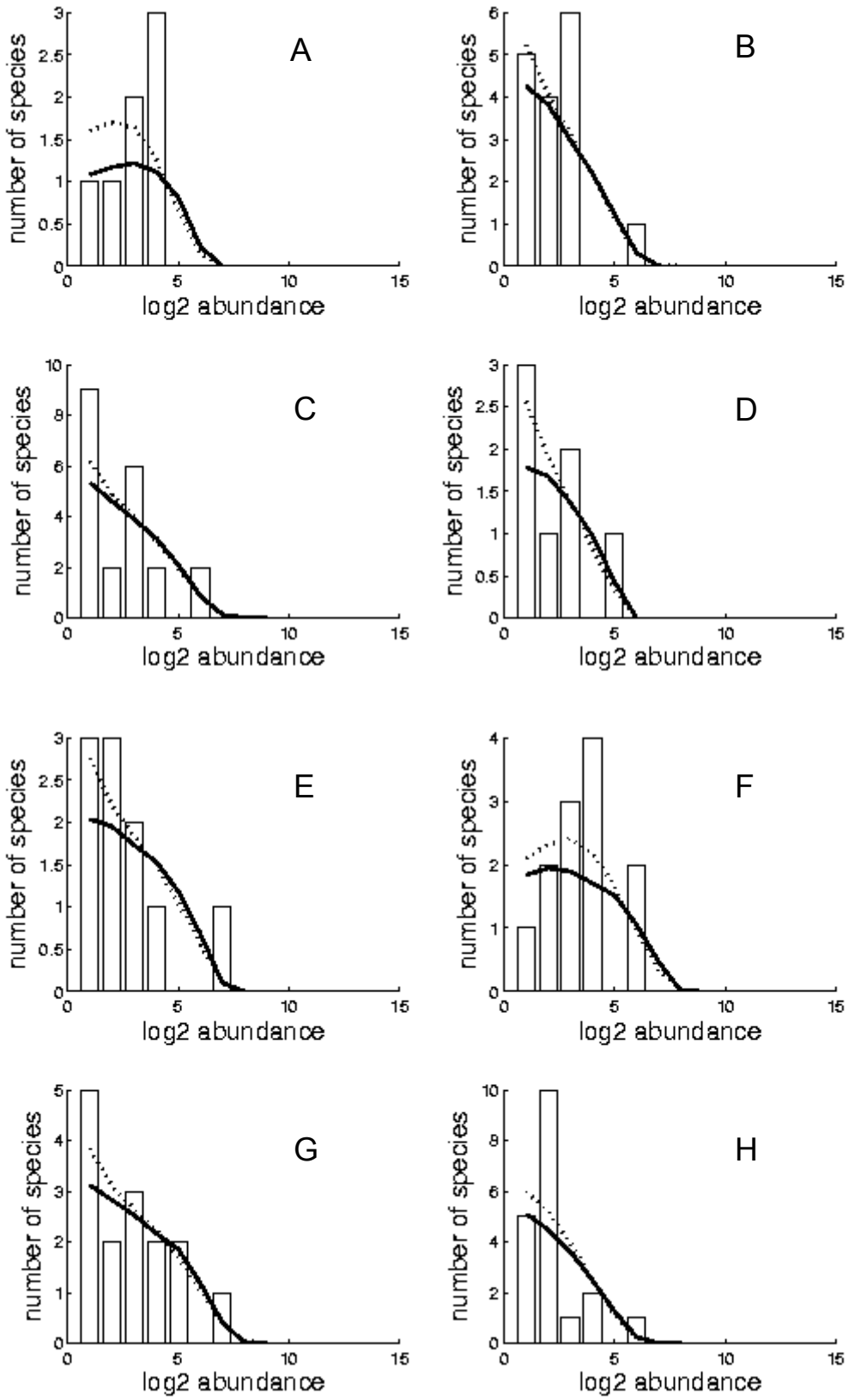


Figure AI.24 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Moorea (A to D) and Tahiti (E to H).

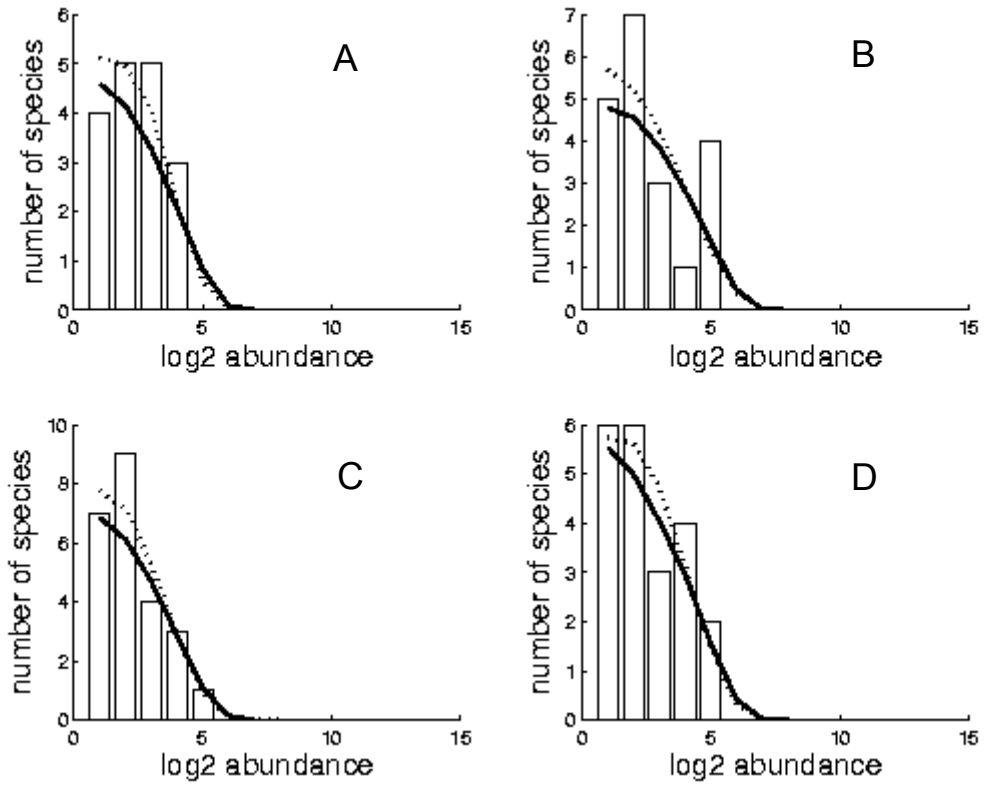


Figure AI.25 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for sites 1 to 4 in Raiatea (A to D).

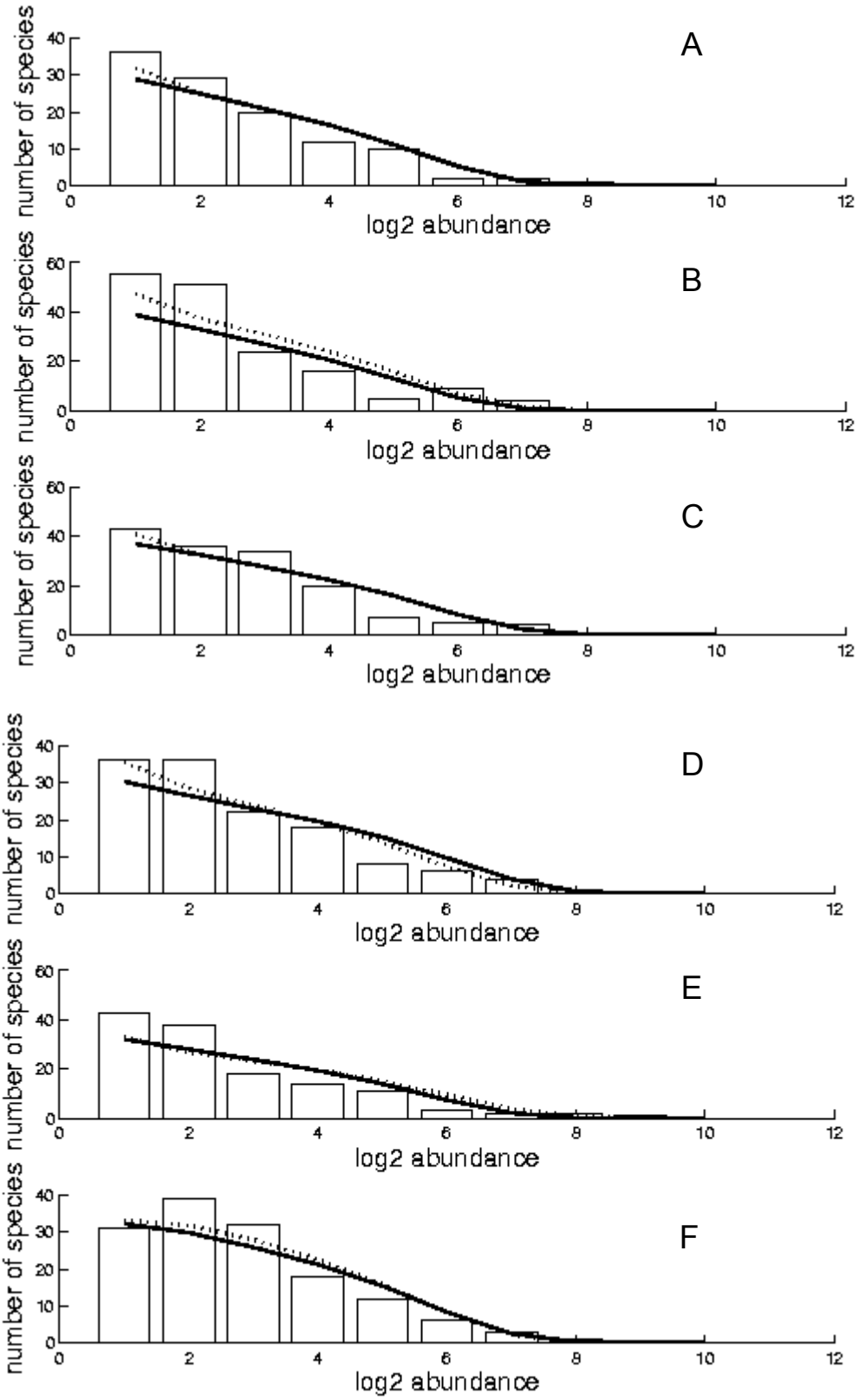


Figure AI.26 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Irian Jaya (A), Manado (B) and Wakatobi (C) in Indonesia, and Kimbe (D), Kavieng (E) and Madang (F) in Papua New Guinea.

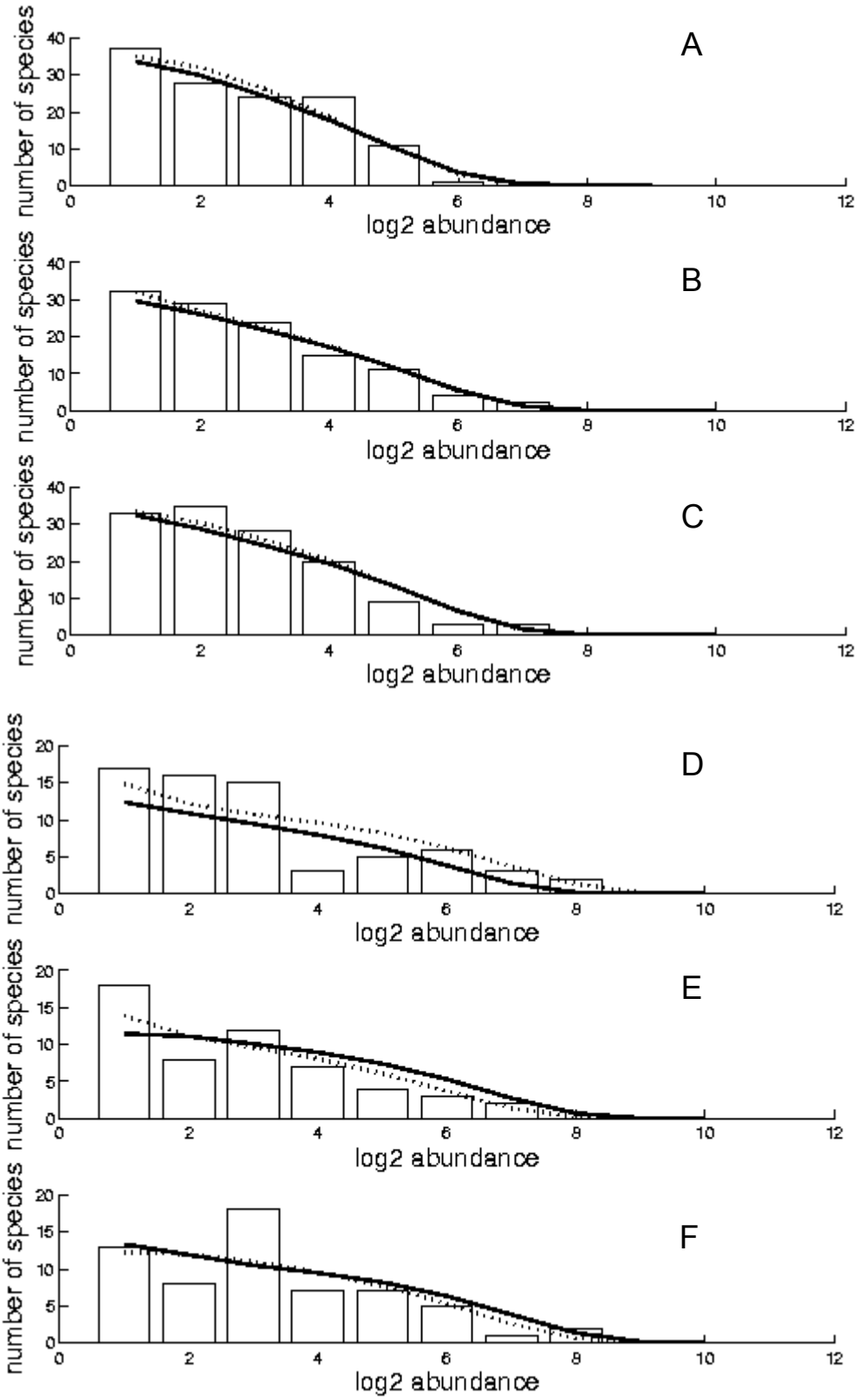


Figure AI.27 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Gizo (A), Munda (B) and Upei (C) in Solomon Islands, and Tutuila (D), Ofu (E) and Tau (F) in American Samoa.

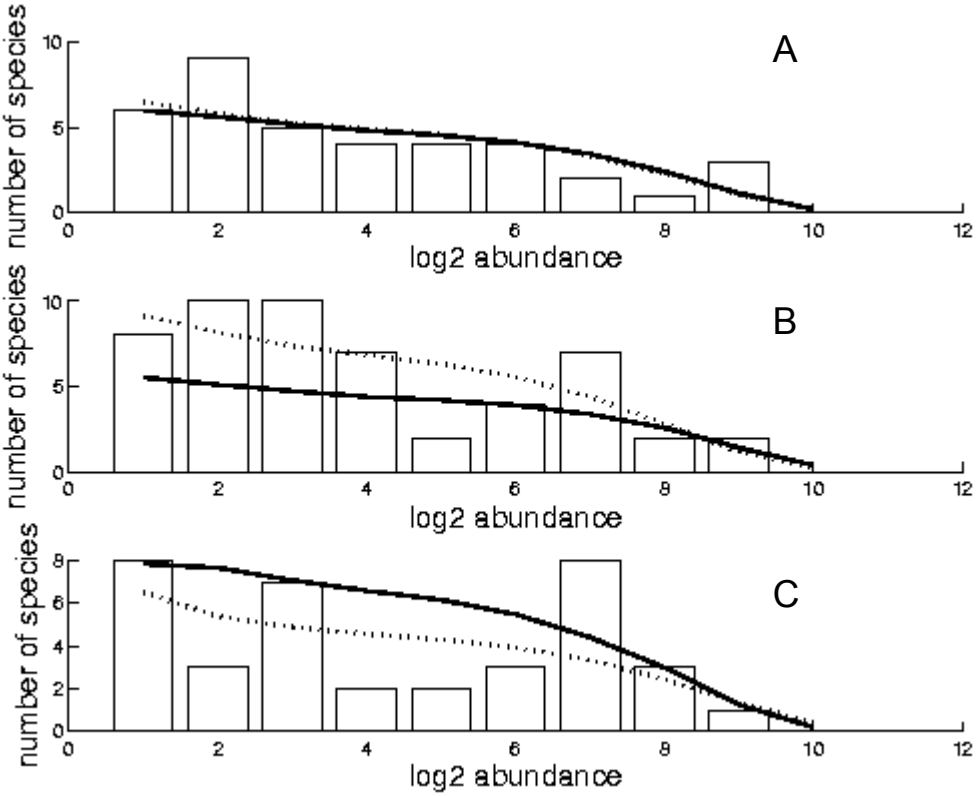


Figure AI.28 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Moorea (A), Tahiti (B) and Raiatea (C) in French Polynesia.

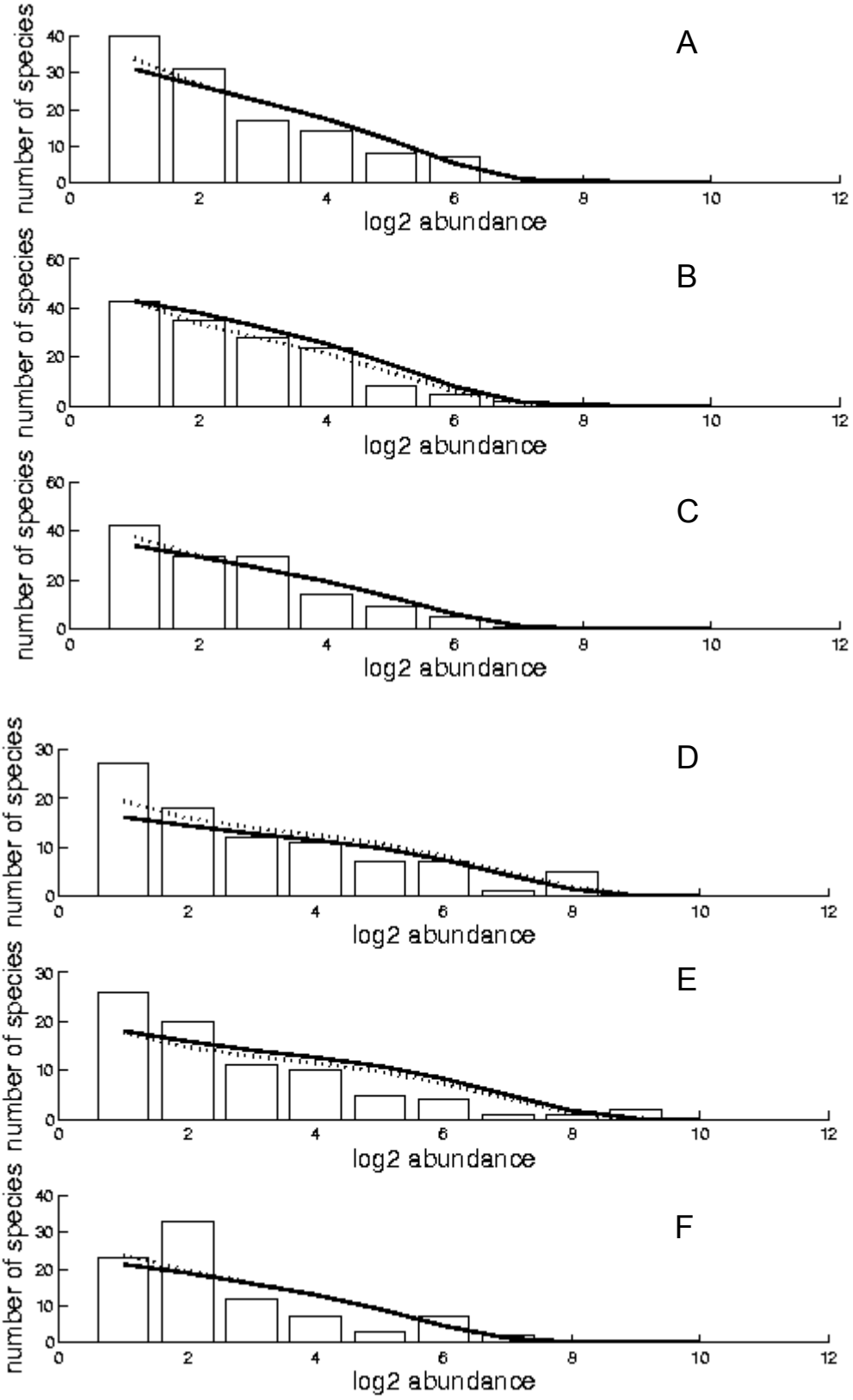


Figure AI.29 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Irian Jaya (A), Manado (B) and Wakatobi (C) in Indonesia, and Kimbe (D), Kavieng (E) and Madang (F) in Papua New Guinea.

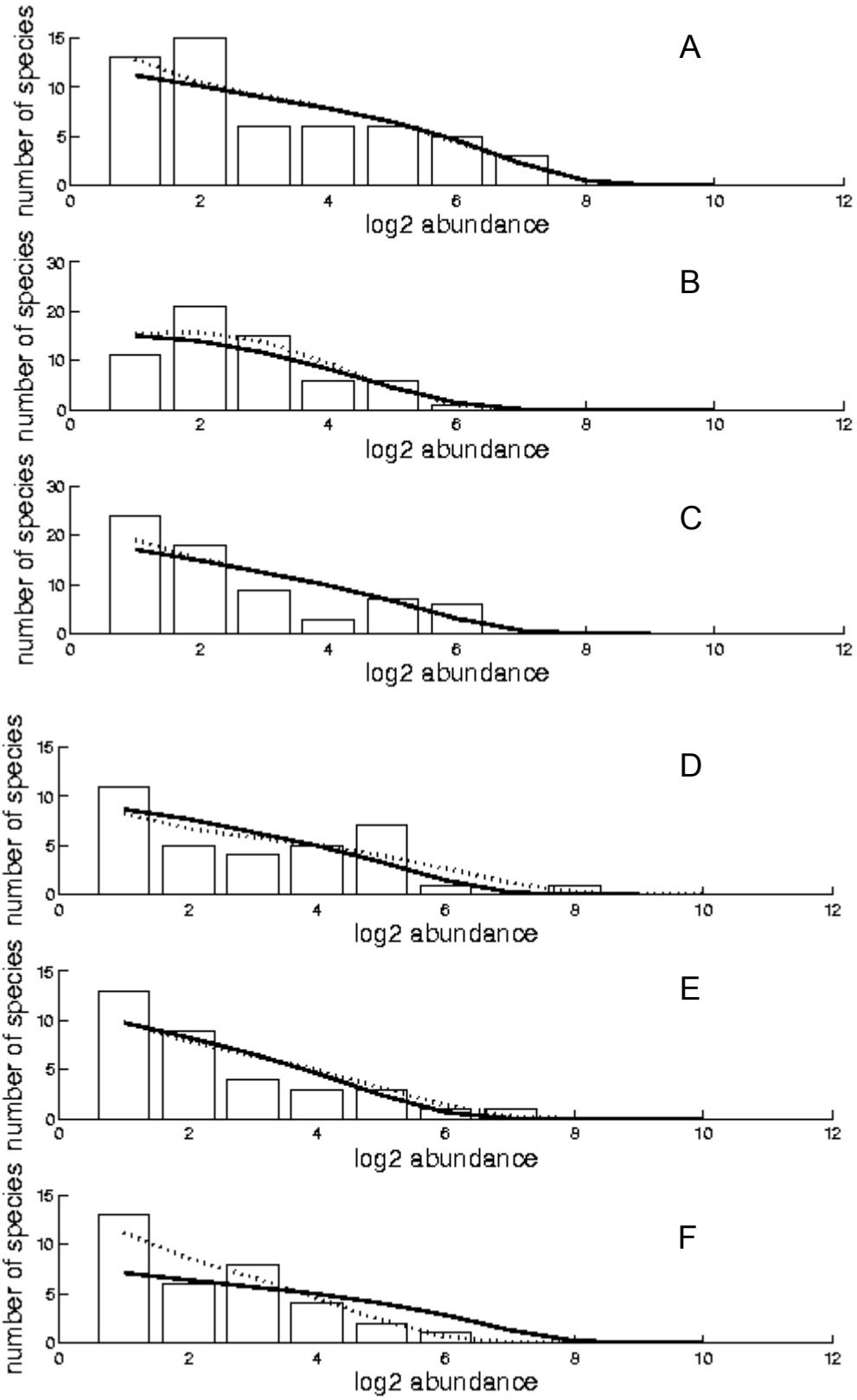


Figure AI.30 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Gizo (A), Munda (B) and Upei (C) in Solomon islands, and Tutuila (D), Ofu (E) and Tau (F) in American Samoa.

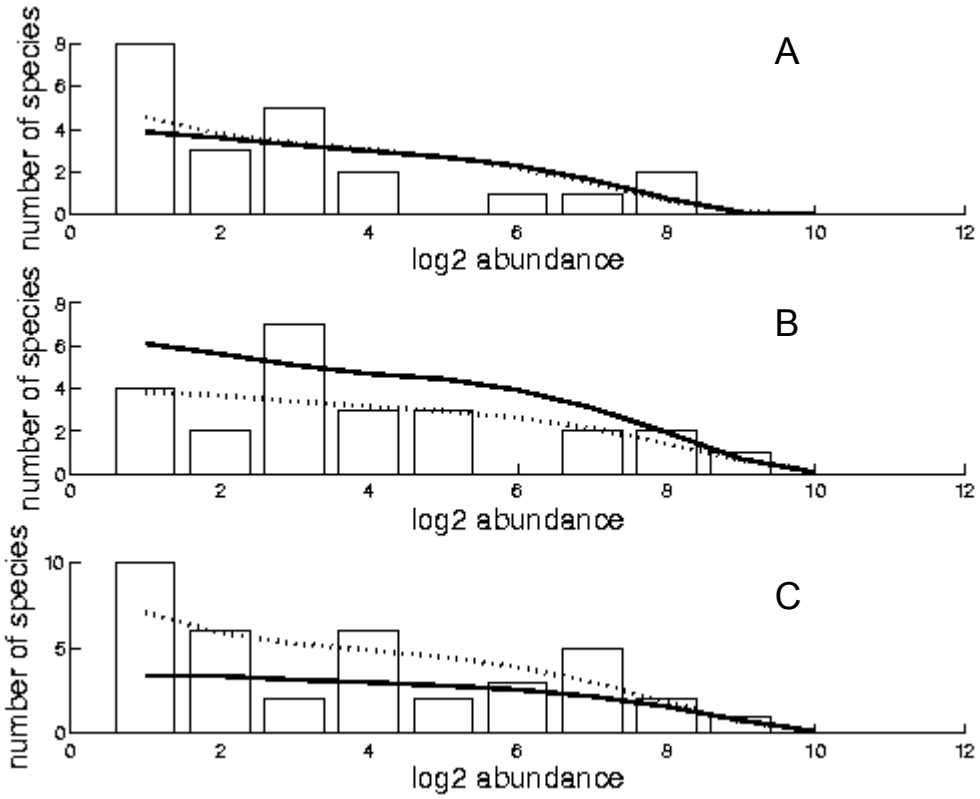


Figure AI.31 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Moorea (A), Tahiti (B) and Raiatea (C) in French Polynesia.

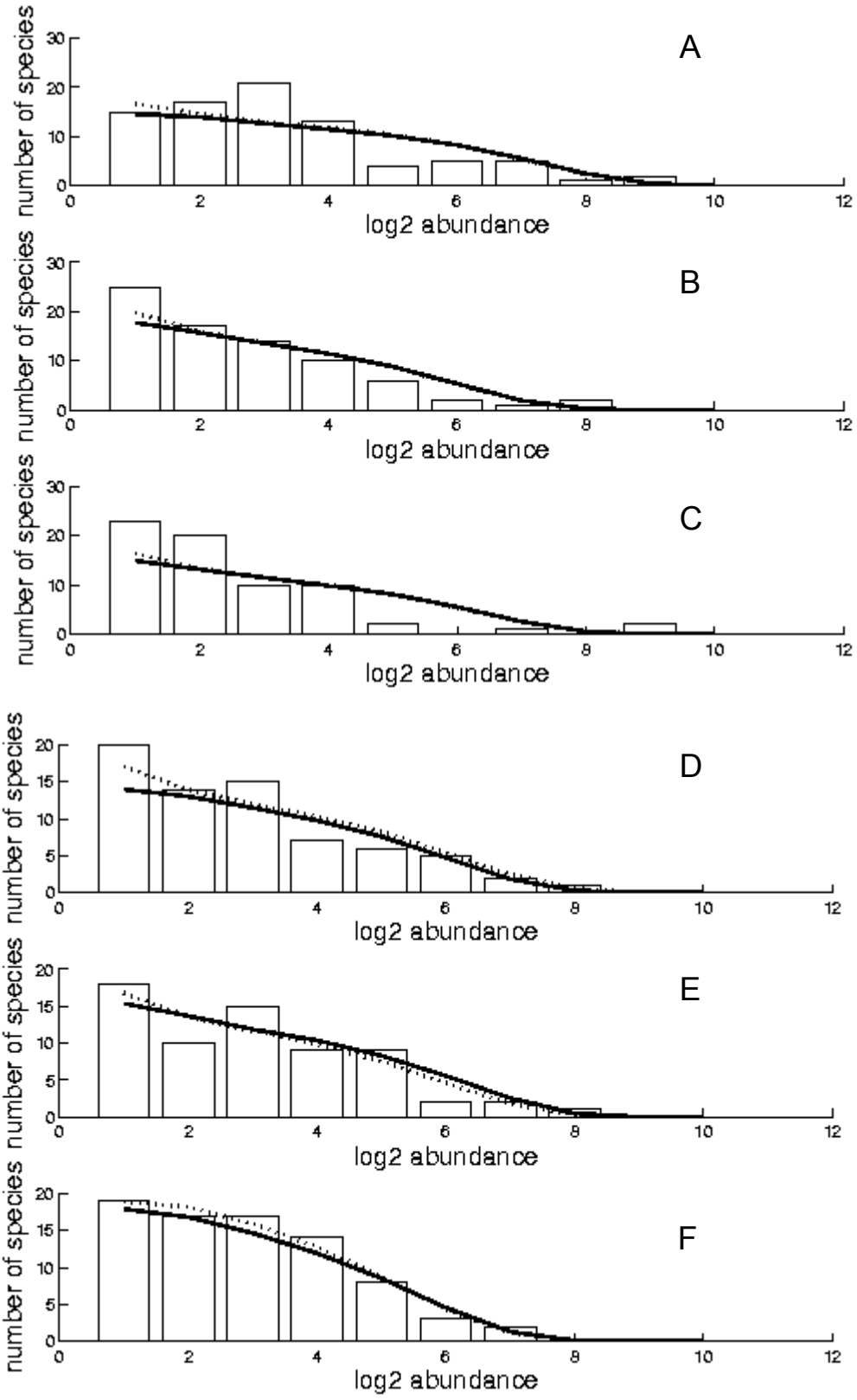


Figure AI.32 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Irian Jaya (A), Manado (B) and Wakatoby (C) in Indonesia, and Kimbe (D), Kavieng (E) and Madang (F) in Papua New Guinea.

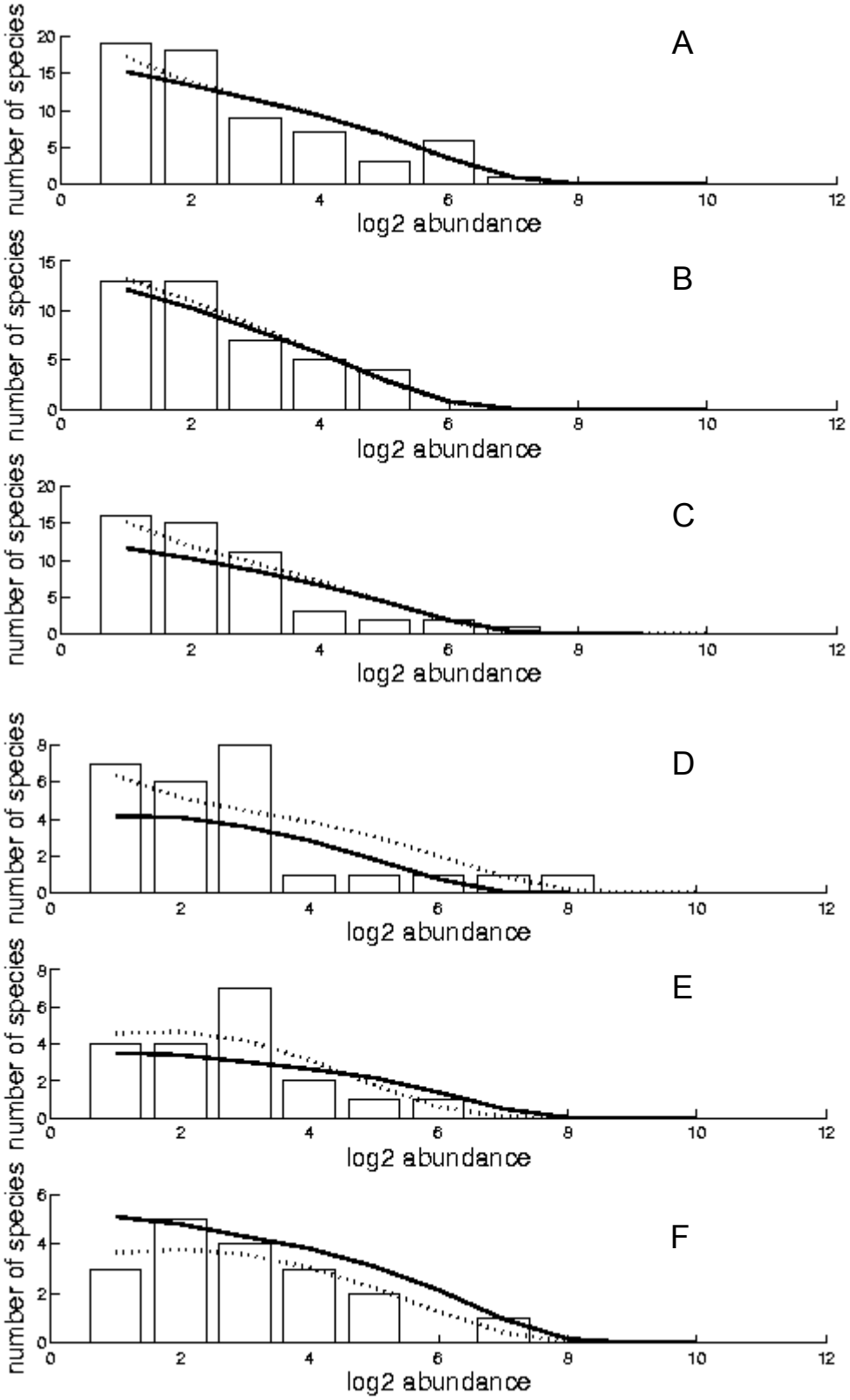


Figure AI.33 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Gizo (A), Munda (B) and Upei (C) in Solomon Islands, and Tutuila (D), Ofu (E) and Tau (F) in American Samoa.

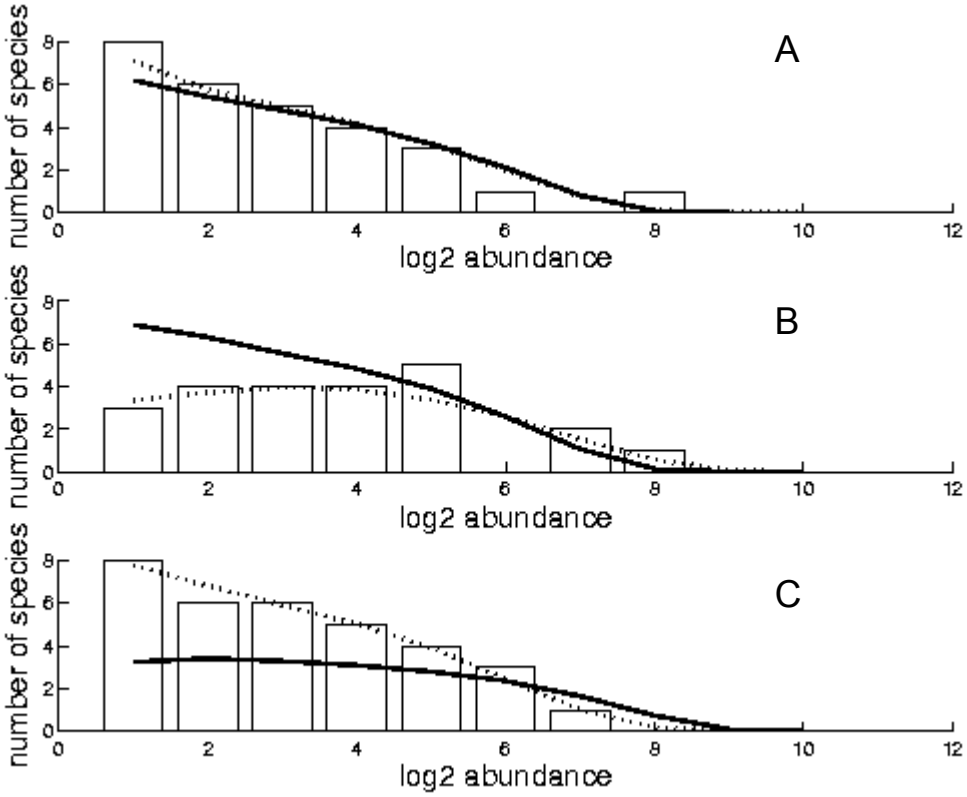


Figure AI.34 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Moorea (A), Tahiti (B) and Raiatea (C) in French Polynesia.

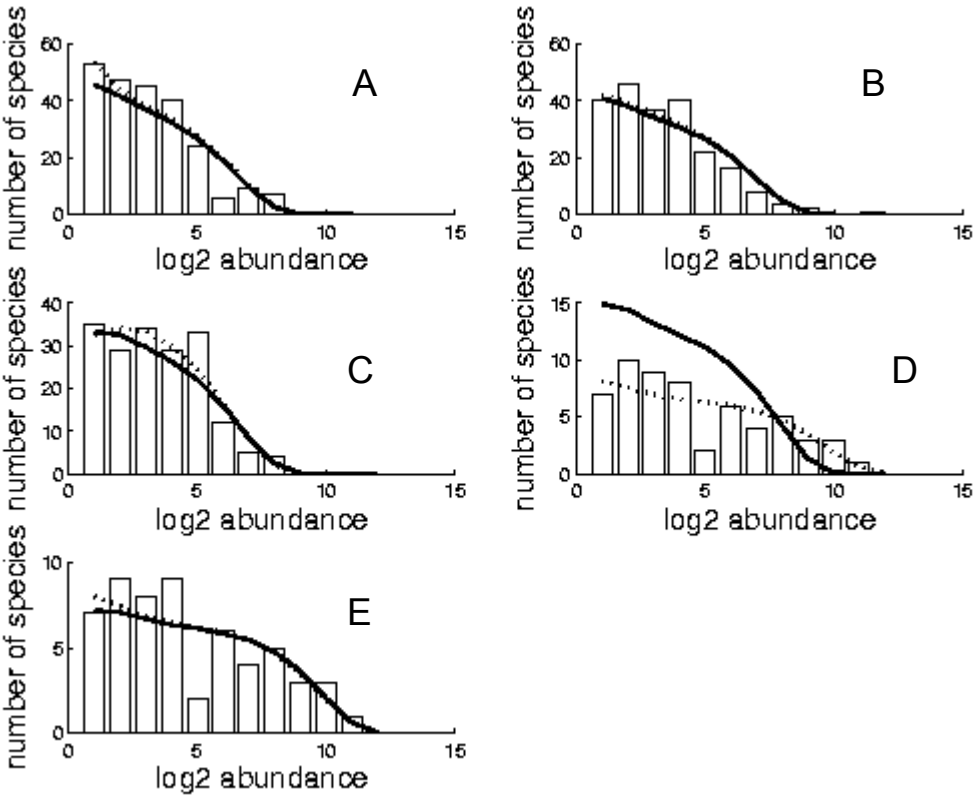


Figure AI.35 – Slope observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Indonesia (A), Papua New Guinea (B), Solomon Islands (C), American Samoa (D) and French Polynesia (E).

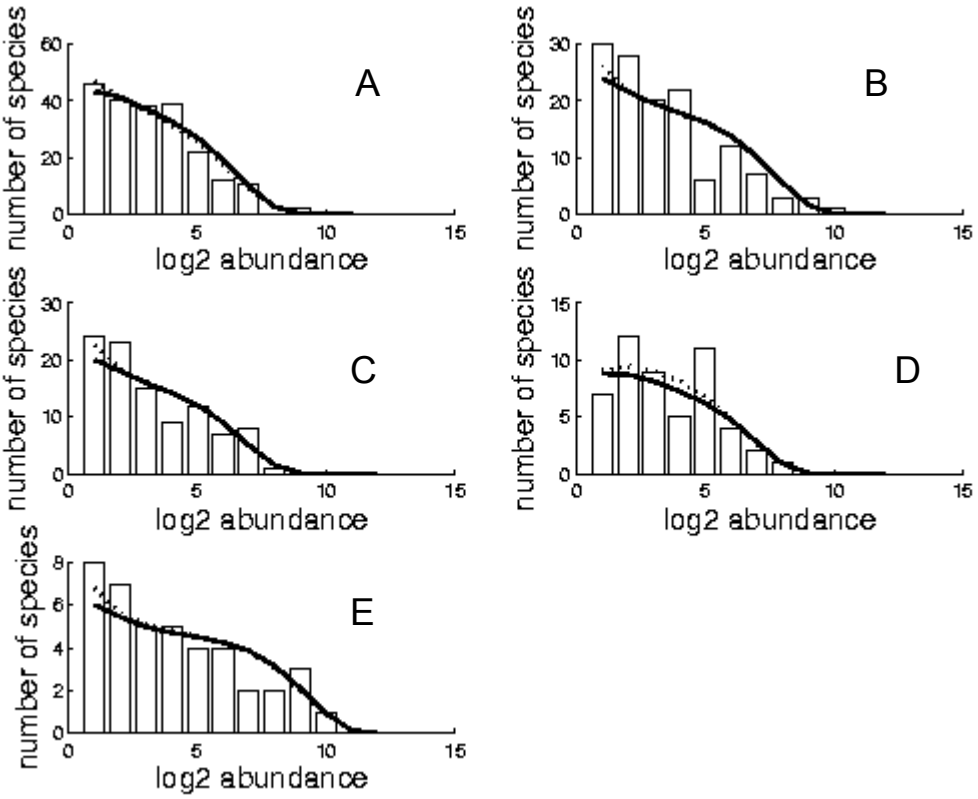


Figure AI.36 – Crest observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Indonesia (A), Papua New Guinea (B), Solomon Islands (C), American Samoa (D) and French Polynesia (E).

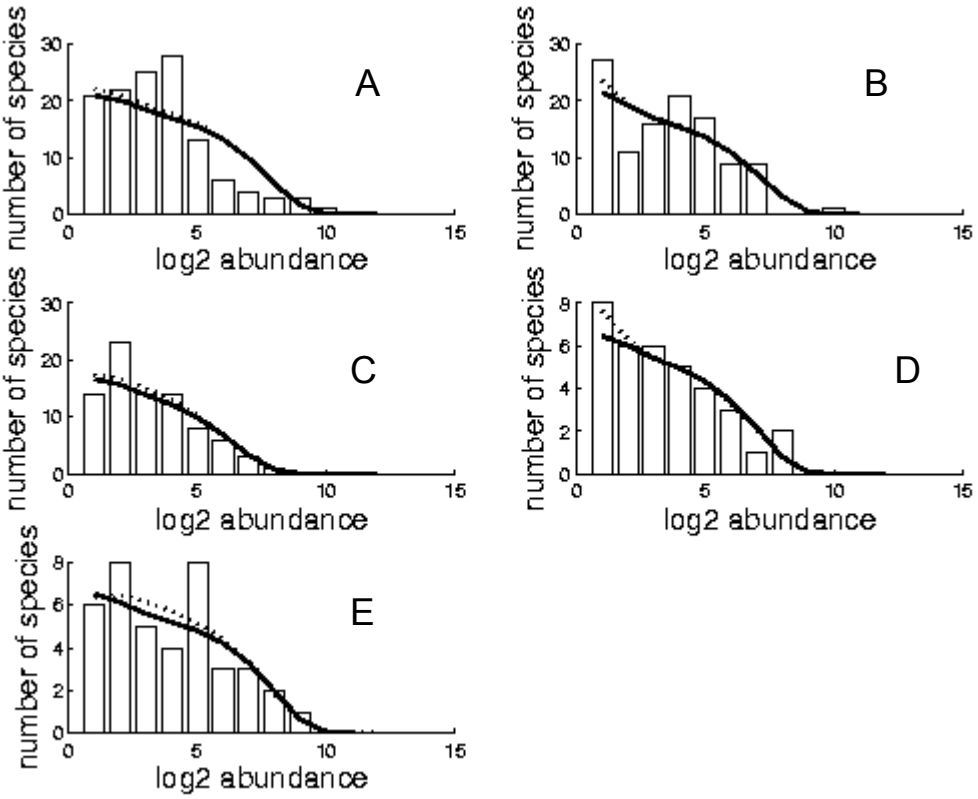


Figure AI.37 – Flat observed species abundance distributions (bars), best fitted SNM (full line), and MFNM (dashed line), for Indonesia (A), Papua New Guinea (B), Solomon Islands (C), American Samoa (D) and French Polynesia (E).

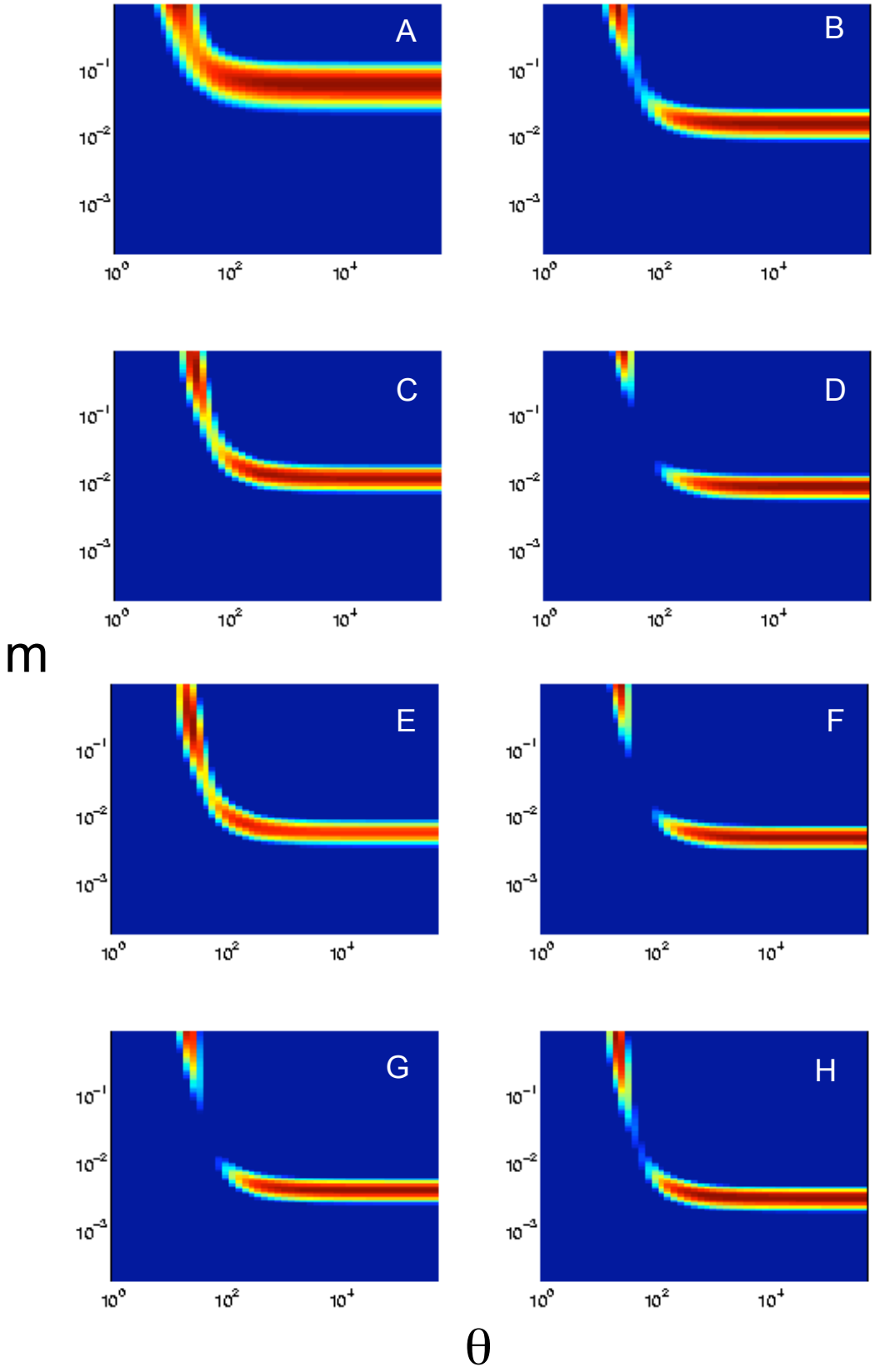


Figure A II.1 - Likelihood surfaces for 1 (A), 10 (B), 15 (C), 20 (D), 25 (E), 30 (F), 35 (G) and 40 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

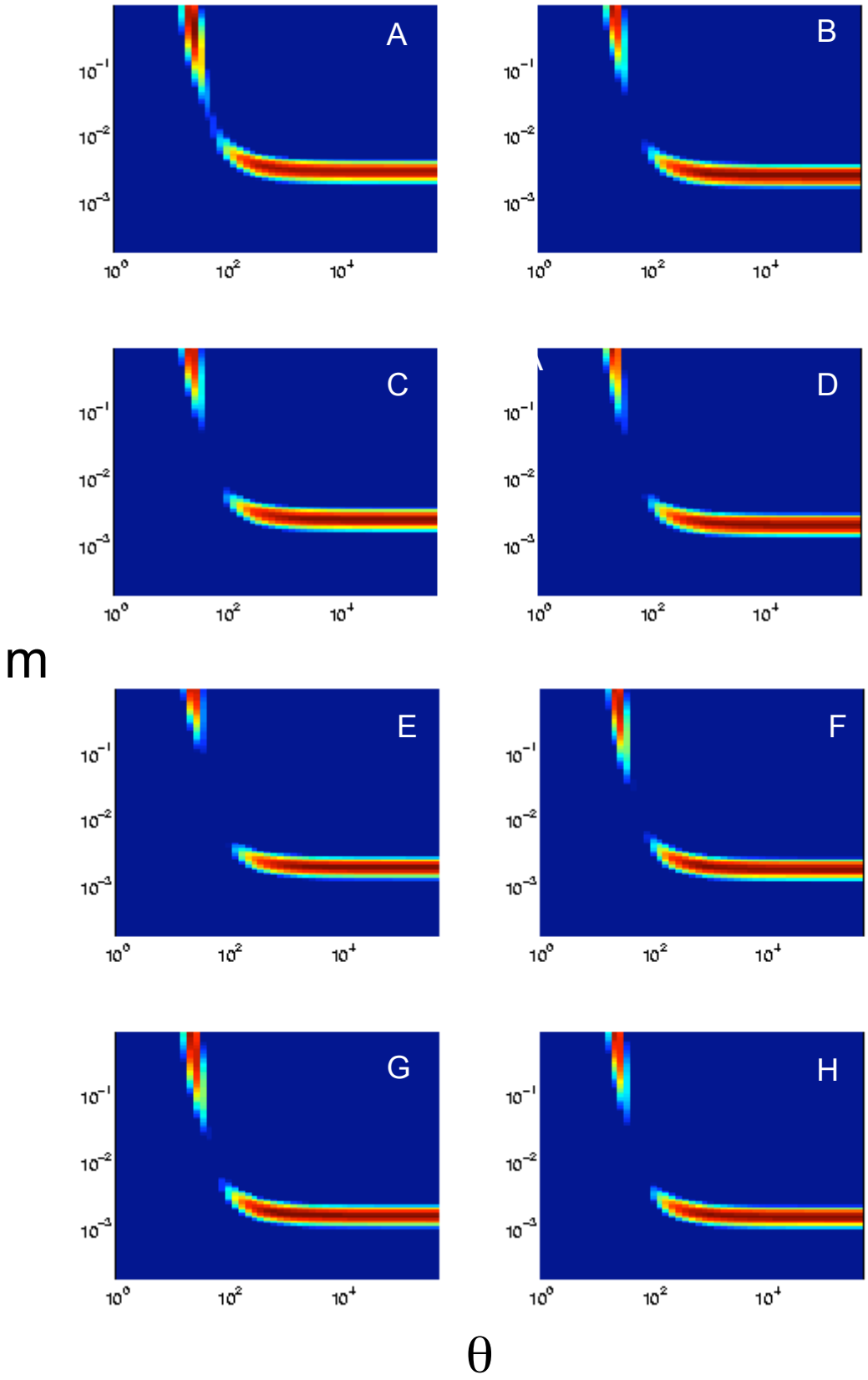


Figure A II.2 - Likelihood surfaces for 45 (A), 50 (B), 55 (C), 60 (D), 65 (E), 70 (F), 75 (G) and 80 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

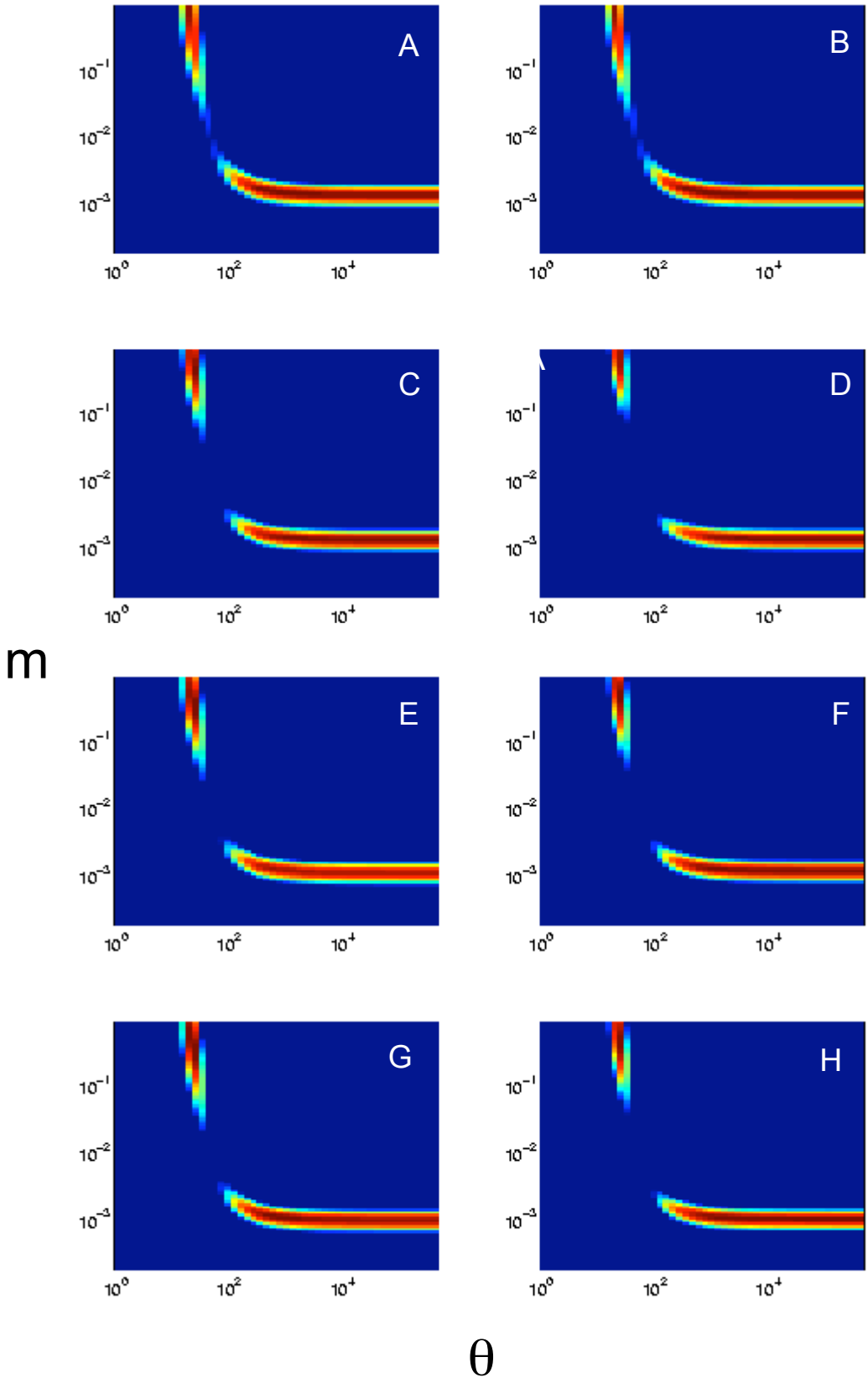


Figure A II.3 - Likelihood surfaces for 85 (A), 90 (B), 95 (C), 100 (D), 105 (E), 110 (F), 115 (G) and 120 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

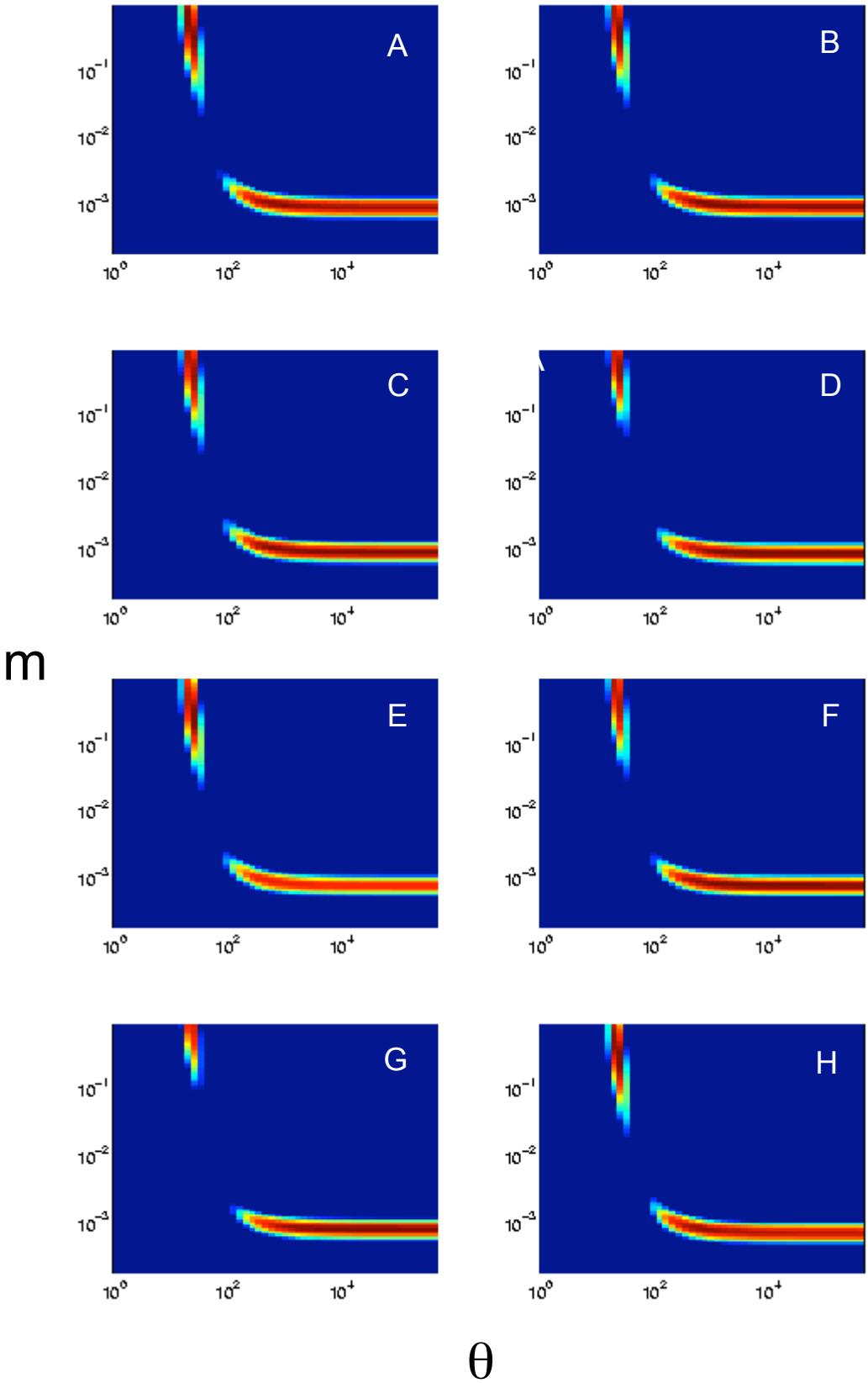


Figure A II.4 - Likelihood surfaces for 125 (A), 130 (B), 135 (C), 140 (D), 145 (E), 150 (F), 155 (G) and 160 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

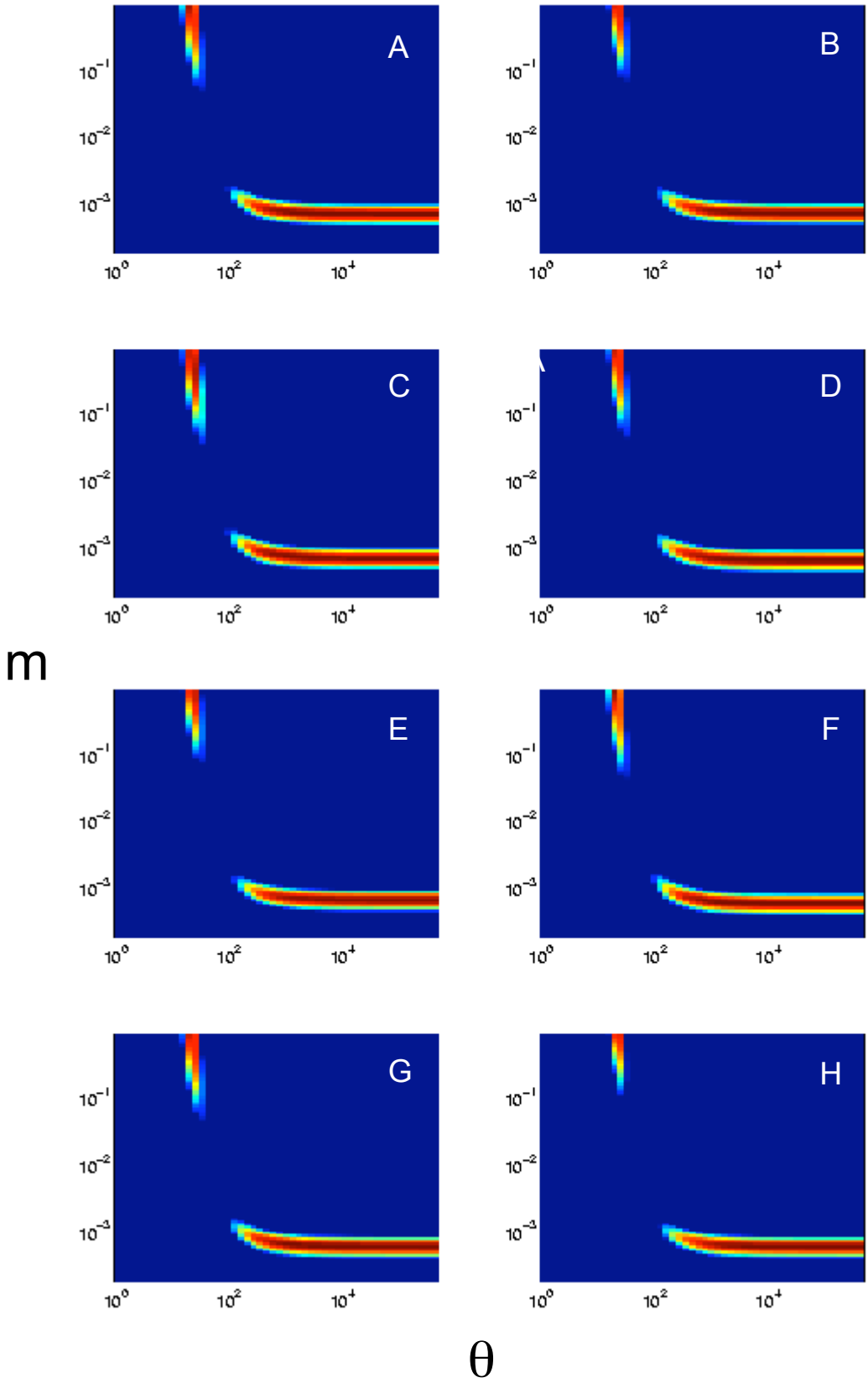


Figure A II.5 - Likelihood surfaces for 165 (A), 170 (B), 175 (C), 180 (D), 185 (E), 190 (F), 195 (G) and 200 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

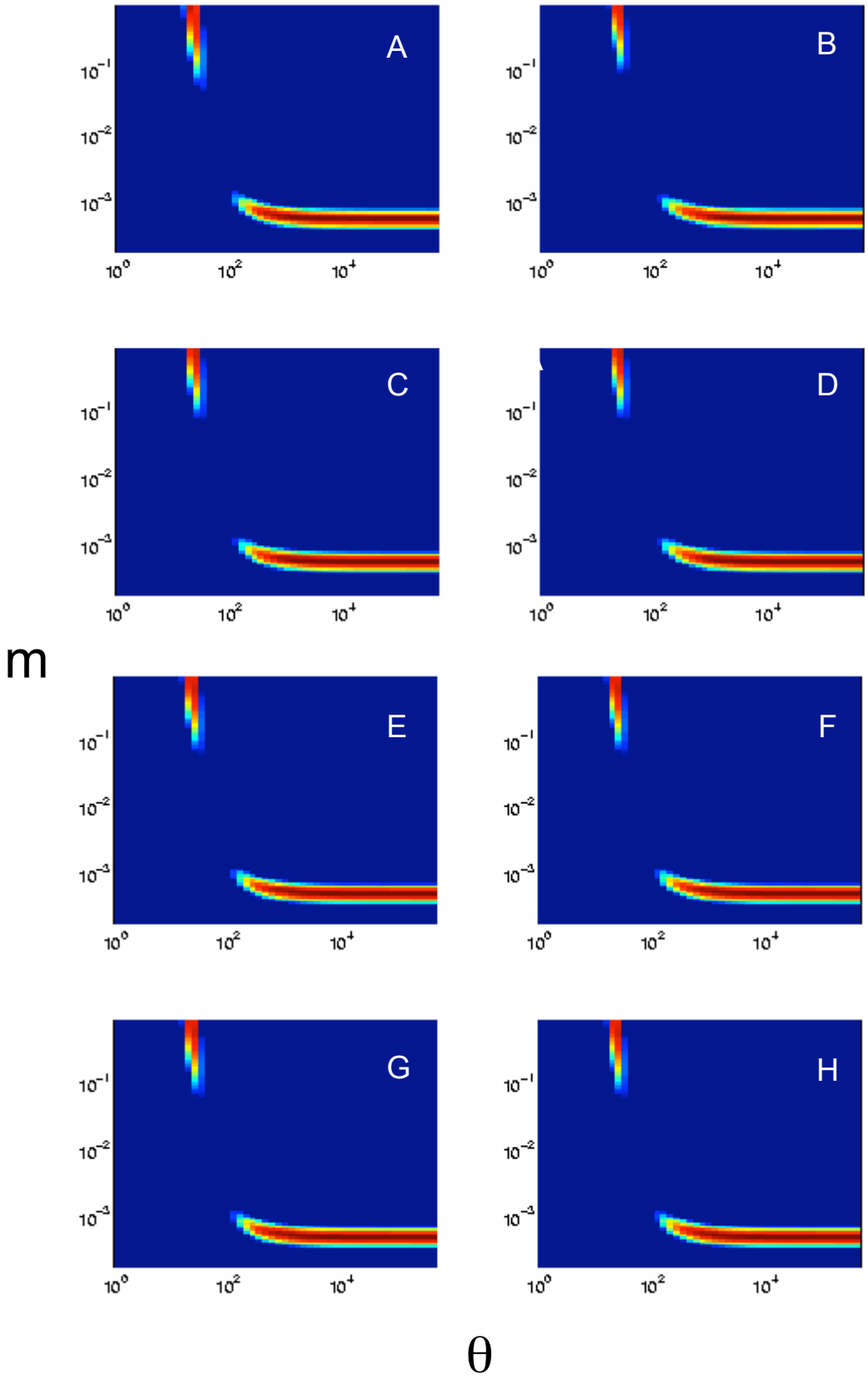


Figure A II.6 - Likelihood surfaces for 205 (A), 210 (B), 215 (C), 220 (D), 225 (E), 230 (F), 235 (G) and 240 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

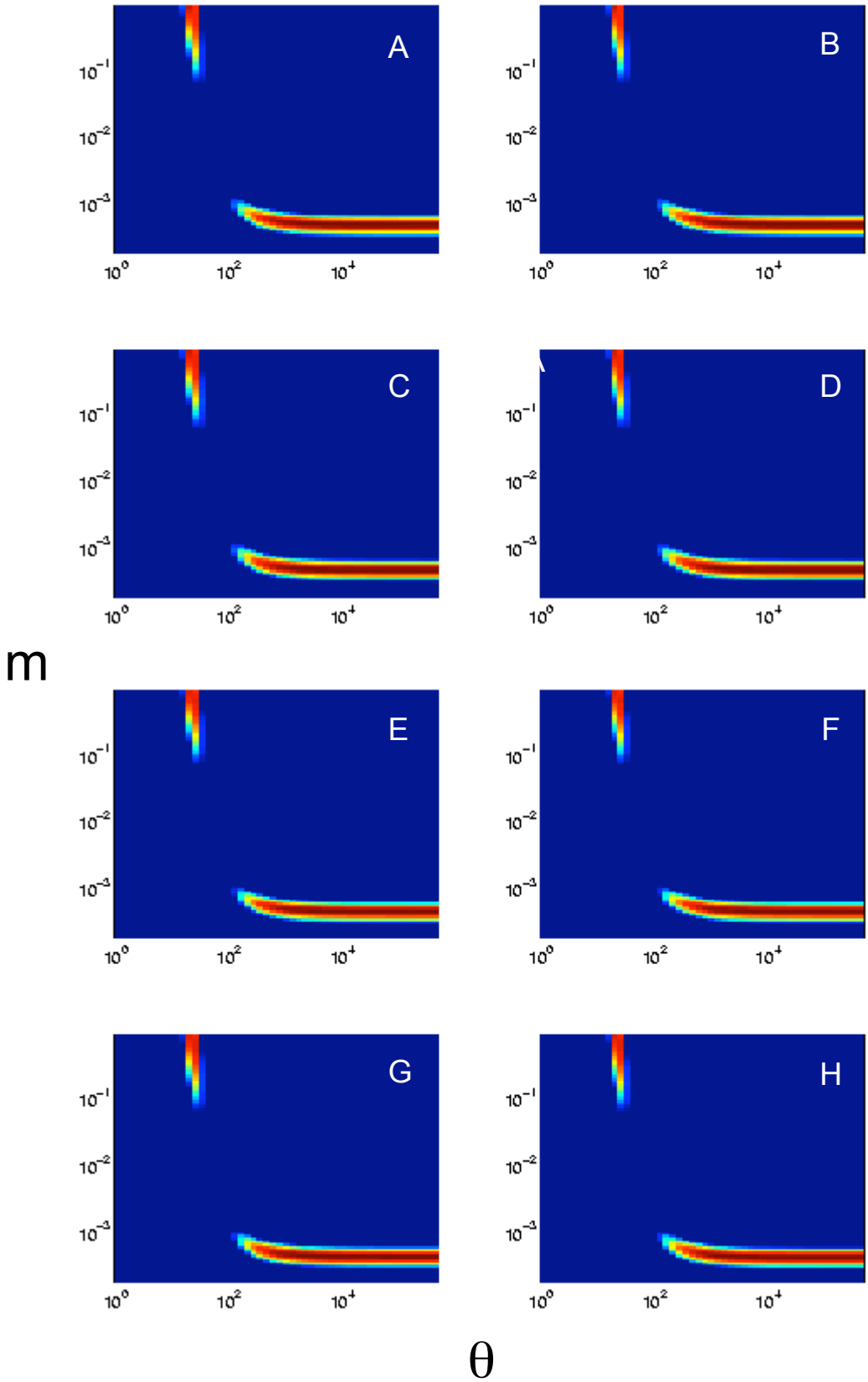


Figure A II.7 - Likelihood surfaces for 245 (A), 250 (B), 255 (C), 260 (D), 265 (E), 270 (F), 275 (G) and 280 (H) transects. Log-Likelihoods bellow 10 units of the likelihood maximum are truncated to increase clarity.

Appendix III

Species list for the crest between South and Palfrey Islands, Lizard Island Group, Great Barrier Reef

Acanthastrea echinata
Acanthastrea hemprichii
Acanthastrea lordhowensis
Acanthastrea regularis
Acropora abrotanoides
Acropora anthocercis
Acropora aspera
Acropora austera
Acropora brueggemanni
Acropora cerealis
Acropora clathrata
Acropora cuneata
Acropora cytherea
Acropora digitifera
Acropora divaricata
Acropora elseyi
Acropora dig-gem
Acropora florida
Acropora gemmifera
Acropora grandis
Acropora humilis
Acropora hyacinthus
Acropora intermedia
Acropora latistella
Acropora listeri
Acropora loripes
Acropora lutkeni
Acropora microclados
Acropora millepora
Acropora monticulosa
Acropora muricata
Acropora nana
Acropora nasuta
Acropora palifera
Acropora polystoma
Acropora pulchra
Acropora robusta
Acropora samoensis
Acropora sarmentosa
Acropora secale
Acropora selago
Acropora spathulata
Acropora tenuis
Acropora valida
Acropora verweyi
Acropora yongei
Alveopora sp.
Astreopora myriophthalma
Astreopora ocellata
Australogyra zelli
Coeloseris mayeri
Coscinaraea columna
Coscinaraea exesa
Ctenactis echinata
Cyphastrea chalcidicum
Cyphastrea microphthalma
Cyphastrea ocellina
Cyphastrea serailia
Diploastrea heliopora
Echinopora gemmacea
Echinopora horrida
Echinopora lamellosa
Favia favus
Favia laxa
Favia lizardensis
Favia matthaii
Favia pallida
Favia rotumana
Favia rotundata
Favia stelligera
Favites abdita
Favites complanata
Favites flexuosa
Favites halicora
Favites pentagona
Favites russelli
Fungia fungites
Fungia paumotensis
Fungia repanda
Fungia scutaria
Galaxea astreata
Galaxea fascicularis
Gardineroseris planulata
Goniastrea aspera
Goniastrea australensis
Goniastrea edwardsi
Goniastrea favulus
Goniastrea pectinata
Goniastrea retiformis
Goniopora pendulus

Goniopora tenuidens
Herpolitha limax
Hydnophora exesa
Hydnophora microconos
Hydnophora rigida
Leptastrea inaequalis
Leptastrea pruinosa
Leptastrea purpurea
Leptastrea transversa
Leptoria irregularis
Leptoria phrygia
Lobophyllia corymbosa
Lobophyllia hataii
Lobophyllia hemprichii
Merulina ampliata
Merulina scabricula
Montastrea annuligera
Montastrea curta
Montastrea magnistellata
Montastrea valenciennesi
Montipora aequituberculata
Montipora caliculata
Montipora crassituberculata
Montipora danae
Montipora efflorescens
Montipora floweri
Montipora foliosa
Montipora foveolata
Montipora grisea
Montipora hispida
Montipora monasteriata
Montipora spongodes
Montipora spumosa
Montipora tuberculosa
Montipora turgescens
Montipora undata
Montipora venosa
Montipora verrucosa
Pavona varians
Pavona venosa
Platygyra daedalea
Platygyra lamellina
Platygyra pini
Platygyra ryukyuensis
Platygyra sinensis
Platygyra verweyi
Plesiastrea versipora
Pocillopora damicornis
Pocillopora eydouxi
Pocillopora verrucosa
Porites massive
Psammocora digitata
Psammocora haimeana
Psammocora profundacella
Psammocora superficialis
Pseudosiderastrea tayami
Seriatopora caliendrum
Seriatopora hystrix
Stylocoeniella armata
Stylophora mordax
Stylophora pistillata
Symphyllia agaricia
Symphyllia radians
Symphyllia recta
Turbinaria mesenterina
Turbinaria reniformis
Turbinaria stellulata