THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Ecological systems in space: From islands to contiguous landscapes

Ankit Vikrant

Department of Space, Earth and Environment
Chalmers University of Technology
Gothenburg, Sweden, 2021

Ecological systems in space: From islands to contiguous landscapes

Ankit Vikrant

This thesis has been prepared using LaTeX.

Department of Space, Earth and Environment
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
`www.chalmers.se`

Printed by Chalmers Reproservice
Gothenburg, Sweden, November2021

Ecological systems in space: From islands to contiguous landscapes

ANKIT VIKRANT

Division of Physical Resource Theory
Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

Dispersal and interaction networks underlie many ecological systems, often dictating spatio-temporal dynamics as well as emergent patterns. Dispersal has implications for establishment for distinct subpopulations of individual species as well as for the emergence of universal spatial patterns and varied community level structures. Such implications and the search for mechanisms that drive them are the central themes of this thesis. The accompanying manuscripts discuss two very different approaches of analysing ecological systems.

The less complex of the two systems consists of sites along a coastal landscape that harbours the seagrass *Halodule uninervis.* The task is to partition the entire population into subpopulations that have very little dispersal across them. We use an algorithm to aggregates sites that are more strongly connected in the same subpopulation for a range of dispersal scenarios. These scenarios are useful since the dispersal processes and life-spans of the seagrass are not well-understood.

In subsequent work, the idea of complexity is more precise. We analyse ecosystems whose complexity is captured by the strength of interactions between species and the fraction of interaction links realized out of all possible ones. On one hand we study how ecological communities on islands are assembled from a large pool of species that can also immigrate from the mainland source, and if it can provide mechanisms to explain the scaling of species richness with island areas – namely the species-area relationship. A drastically different setting is large contiguous landscapes.

We posit that species interactions vary across space, which helps characterize habitat patches that would be connected by high dispersal of species. This simplified picture provides a range of situations to understand how habitat heterogeneity affects species richness.

An overarching theme in investigating these complex ecological systems is to first analyze the fully random case without assuming any structure. This null case admits description in terms of a few statistics that greatly simplify the study of system level properties and processes. One can then ask if there are parameter regimes where these properties break down. The difficulty in finding such regimes might indicate universal properties but even the sudden disappearance might unravel phase transitions or unexplored new properties.

Keywords: Metapopulation, complex ecological systems, species area relationships, metacommunity, community assembly, ecosystem stability

List of Publications

This thesis is based on the following publications:

Paper 1
R. D. Evans, K.M. McMahon, K. -J. van Dijk, K. Dawkins, M. Nilsson Jacobi, A. Vikrant , "Identification of dispersal barriers for a colonising seagrass using seascape genetic analysis". *Science of the Total Environment* (2021), 763, 143052..

Paper 2
A. Vikrant and M. Nilsson Jacobi , "Complex ecological communities and the emergence of island species area relationships". *Under review*, Preprint available at arXiv https://arxiv.org/abs/2103.07721.

Paper 3
A. Vikrant† , S. Pettersson† , M. Nilsson Jacobi , "Spatial coherence and the persistence of high diversity in spatially heterogeneous landscapes". *Under review.*

Contributions

Paper 1
RDE, KMM: Conceptualization, Methodology, Formal analysis, Writing - original draft, Project administration; KJvD, KD, MNJ, AV: Methodology, Formal analysis, Writing - review and editing.

Paper 2
AV conceived and designed the study, AV performed the simulations and wrote the paper, AV and MNJ analyzed the results.

Paper 3
Conceptualisation, interpretation of results and, review and editing: MNJ, SP, AV; Simulations, visualisation and original draft: SP, AV; Analytic results: SP

Acknowledgments

I want to thank Martin for his mesoscopic style of supervision – helping me keep stock of the larger picture but also enthralling me with forays into the finer details. Thank you Susanne for entertaining my questions and helping me wade through murky modelling possibilities. I am super grateful to my friends and family who at various times have sheltered me, fed me, and entertained me and my occasional oddities.

Acronyms

SAR: Species-area relationship

GLV: Generalized Lotka-Volterra

AHTO: Area-heterogeneity trade-off

Contents

CHAPTER 1

Introduction

Biological species interact in myriad ways with one another and their environment to generate the diversity of life that we see around us. In addition to ecology that concerns various aspects of organisms and their interactions with the environment, evolution is another cornerstone of organismal biology. Unlike physical systems – whose fundamental units are assumed to be immutable – biological entities could undergo changes through evolution, which has implications for species as well as the ecosystems they inhabit. Variability in the basic units of an ecological system could be further pruned and filtered by how individuals, species (and collections thereof) disperse in space. Ecology constrains how species can disperse in space, which results in the emergence of widespread spatial patterns and laws. Some of these patterns and properties are signatures of highly diverse and complex ecological communities that have received increasing attention lately [1, 2].

The oldest and the most prominent among such spatial laws is the species-area relationship (SAR) that describes how the number of species in a patch scale as the area of the patch increases. Two functional forms of the SAR have been most widely observed in empirical studies – the semi-log and the power law form [3]. The power-law form has especially been the subject of many studies trying to explain its origins. Preston used log-normal abundance distributions – that are widely observed – to argue for a power-law form of the SAR [4]. This work also demonstrated differences in studying nested samples in a contiguous landscape versus a system of isolates such as islands. A qualitatively different suggestion was that islands could be at internal equilibrium that constrained patterns of species diversity on them [4].

The simplicity of island systems prompted MacArthur and Wilson to propose an equilibrium theory of island biogeography [5]. They argued that the number of species found on any island depends on an equilibrium between immigration and extinction rates. Further, since increasing island size leads to higher immigration rates but lower extinction rates, one could also relate the number of species to island area. This simple model also inspired the study of species diversity patterns in other island-like ecosystems [6–8]. The fact that species diversity patterns could be reproduced without invoking species differences, has also been the basis of more successful recent theories (cite). However, the same simplifications prevent these theories from capturing temporal patterns and processes such as those related to population dynamics [9].

Island systems give way to the idea of having discrete spatial patches representing different populations that could be connected by migration or gene-flow. The concept of a metapopulation – meaning population of populations – had some primitive analogues already in 1950's [10], but the word was first used by Levins in a sense that is closer to its contemporary usage [11]. A metapopulation could witness local extinction of the species at one spatial patch that might be later recolonized through dispersal from other patches. Mainland-island systems are a special case of metapopulations where the mainland prevents local extinctions at the islands. Given a spatial distribution of many different populations, one could classify them into multiple metapopulations depending upon factors that impede migration or gene-flow across populations. Real ecosystems are much more complex in terms of the number of species they harbour and also the kinds of interactions between them. The idea that complex interactions varying in space can affect diversity patterns was the basis of the metacommunity idea.

Metacommunities can be studied in various paradigms that differ in the amount of dispersal and the variation in species traits resulting from environmental heterogeneity [12, 13]. The amount of dispersal between local communities dictates the influence of environmental stochasticity on community assembly at different scales. High dispersal not only acts to homogenize local communities across patches but it also increases the pace of community assembly such that stochastic effects from the environment have minimal bearing on the fate of these communities. This thesis considers metacommunities harbouring many interacting species such that these species cannot be reduced to their positions along one or few trait axes. This operational definition implies that habitat heterogeneity is captured through variation in interspecies interaction strengths across space, as in [1].

1.1 Motivation and Aim

This licentiate thesis broadly deals with spatial patterns and laws emerging from the ecology of species. One objective of Paper 1 is to identify clusters in a large network of seagrass populations that are established through underlying dispersal of vegetative fragments. Clustering in networks has been a very active research area over the past couple of decades, but the spatial and ecological constraints imposed on biological populations imply that the clusters identified might be sub-optimal using generic algorithms. My aim was to use an effective algorithm [14] that identifies different metapopulations such that sites within a metapopulation are more strongly connected than sites outside it. Also, if two clusterings have the same score – in that they are equally optimal – then how to best decide the better one? Combined with genetic analysis, the clustering results answer the question – what is the right spatial scale for management such that dispersal barriers do not hamper the re-establishment of seagrass populations following catastrophic events.

Many approaches explain species coexistence in a low-dimensional parameter space such as when species can be described using very few traits [15–17]. However, these methods have limited success when ecological communities are extremely high-dimensional [18], especially ones with a large number of species such that species interactions cannot be ignored and there is no clear niche-partitioning. One broad aim of my recent work is to describe the emergence of spatial laws and patterns in such high-dimensional ecosystems. Specifically, is to possible to understand the emergence of the widely observed SAR for high-dimensional communities? If yes, then how can one study this relationship using simple models of such communities? Further, what ecological systems are best suited to scrutiny using such simplified models? (Paper 2)

Ecosystems are better described as spatially-extended rather than well-mixed systems. This implies that species can move around across different locations and species interactions could vary spatially. For anyone who's worked on stability of large complex ecosystems, it is natural to ask whether a fully-feasible and stable ecosystem is possible for arbitrarily complex local communities? What kind of dispersal processes and spatial patterns of species interactions could enable highly complex yet stable ecosystems, theoretically? Further, do large contiguous landscapes entail high species diversity when understood as collections of many local communities connected by dispersal? (Paper 3)

1.2 Thesis outline

This thesis introduces some key concepts and existing theories in Chapter 2 before providing a summary of appended papers in Chapter 3. The chronological order of papers corresponds to increasingly complex ecological communities. Paper 1 identifies the most likely scenarios to explain the partition of a colonizing seagrass into distinct metapopulations along the northern and western coasts of Australia. This study uses two complementary approaches, one utilizing ocean current data to estimate transition probabilities between different sites and the other based on genetic data. Paper 2 uses a modification of Generalized Lotka-Volterra equations to study factors that could result in different island SAR forms. Paper 3 studies large contiguous landscapes as habitat patches connected by high-dispersal. We use a spatially-extended version of the GLV equations where dispersal is modelled through diffusion on discretized space. Chapter 4 reflects upon the ecological systems and results discussed in this thesis to motivate possible future directions.

CHAPTER 2

Background

The spatial aspects of single-species metapopulations versus highly-diverse ecosystems differ a great deal conceptually. My contribution in paper 1 is closest in spirit to the problem of finding clusters in a large network. I therefore provide a brief background on what this means for generic large networks and specifically for spatio-ecological systems. Section 2.2 describes complex high-dimensional ecosystems that I study using random matrices where pairwise interactions between species are randomly drawn from some distribution. The dynamics of many interacting species could give rise to certain emergent patterns and laws that are typical of large ecosystems for a large part of the parameter space, even when more structure in the species interactions is incorporated. Such highly disordered ecosystems (and systems in general) admit description in terms of a few statistics, which provides some baseline expectations about properties and processes that might be otherwise intractable for such complex systems.

2.1 Community detection in networks

Spatially distributed populations of a species can be partitioned into distinct metapopulations. Such metapopulations would be expected to have two major but related properties. The dispersal across metapopulations is much lower than within them. In an evolutionary sense, there is negligible gene flow between different metapopulations. This partitioning problem is analogous to finding clusters within a large network where individual populations are nodes, and the transition probabilities characterize the edges between them. One class of methods for clustering in networks is based on spectral analysis that was originally used for graph partitioning [19].

The central object of study in such methods is usually the graph Laplacian whose second eigenvector – corresponding to the second smallest eigenvalue called Fiedler value shows distinct clusters of nodes such that there are much fewer edges across these clusters [19]. The graph Laplacian is also related to the transition matrix of a random walker on a graph. Therefore, the spectral methods can be extended to identify 'lumpings' of states of a Markov process. The problem of graph partitioning can be posed as a problem of minimizing the cut size – defined as the number of edges between two distinct groups of nodes – that could be written in terms of the graph Laplacian. Using a modularity matrix that is analogous to the graph Laplacian, one can similarly maximize the modularity of a given network to find constituent modules [20].

A broad class of methods rely on minimizing (or maximizing) certain measures that quantify the degree of good or bad division of a network into communities. Network modularity [21] is a measure that takes high values if there are many edges within the communities than across them. Such maximization methods are much faster than most other methods in dividing very large networks, especially when the networks are sparse. A highly efficient algorithm proposed in [22] utilized sophisticated data structures that considerably speed up the search and insert operations in such maximization problems.

A majority of spectral as well as hierarchical partitioning methods cannot be extended directly to ecological networks. This is especially true when the underlying networks are directed and asymmetric, or if the transition matrix is not symmetric with respect to some scalar product [23]. For example, the dispersal of a species between geographical sites can be encoded in a transition matrix that contains probabilities of transition from one site to another. An efficient hierarchical partitioning algorithm for this problem was introduced in [14] to minimize a function that characterizes the connectivity between different subpopulations. Ecological constraints on dispersal could be introduced in a tuneable penalty term that allowed for different comparable partitions. The ecological system studied in paper 1 uses this algorithm, and a heuristic is suggested to compare partitions that are equally optimal (defined by a score).

Algorithm for finding subpopulations using connectivity matrices

Paper 1 uses the algorithm proposed in [14] to aggregate geographical sites into distinct subpopulations. The analysis is based on fragment dispersal matrices – termed as connectivity matrices C here – whose elements (i.e. C_{ij}) define the probability of individuals dispersing from site i ending up at site j. A minimization problem is defined using a preliminary function that quantifies the connectivity between two subpopulations:

$$G(\beta) = 2\sum_{ij}(C_{ij} - \beta^{-1}E_{ij})\Omega_{ij} \tag{2.1}$$

where Ω is matrix with $\Omega_{ij} = 1$ if site i and j are in different subpopulations else $\Omega_{ij} = 0$. $C_{ij}\Omega_{ij}$ is the naïve connectivity that equals zero if all sites are placed in a single subpopulation. Sites across subpopulations with transition probability below β^{-1} is rewarded with a term $\beta^{-1}E_{ij}$ where the matrix E has all elements equal to 1. This term therefore prevent a trivial solution with only one subpopulation.

It is often useful to redefine Ω in terms of a vector s that takes the values ± 1. Different sites can be placed into two subpopulations depending on the sign of the respective element [20]. The relationship $\Omega_{ij} = (1 - s_i s_j)/2$ then allows the redefinition:

$$G(\beta) = -s^T C s + \beta^{-1}(\sum_i s_i)^2 \tag{2.2}$$

The equation 2.2 is written in a matrix form where terms independent of s have been removed since they do not affect the minimization. s^T is the transpose of the s matrix.

Monte-Carlo or simulated annealing techniques are computationally very expensive in trying to find the minimum for this problem. An efficient alternative is to allow s_i to take continuous values such that the final solution can be read off from the signs of s_i. To ensure that the minimization problem is bounded, the following function is more suitable:

$$G' = G + \frac{1}{\gamma + 1} \sum_i |s_i|^{\gamma+1} \tag{2.3}$$

where $\gamma \geq 1$ prevents the function from diverging. The minimum can be found by using an iterative scheme that solves the derivative of 2.3 starting from a random guess for s. The recursion reads as:

$$s_{t+1,i} = as_{t,i} + (1-a)\text{sign}(s_{t,i}) \times |\sum_j (C_{ij} + C_{ji})s_{t,j} - 2\beta^{-1} \sum_j s_{t,j}|^{1/\gamma} \tag{2.4}$$

where the derivative was $-\sum_j (C_{ij} + C_{ji})s_j + 2\beta^{-1} \sum_j s_j + |s_i|^\gamma \text{sign}(s_i)$. This also explains the choice of having $1/(\gamma+1)$ in the minimization function. Small but positive values of a guide the solution towards the minimum.

The identification of subpopulations can then be summarized as follows. Starting with some low value of $1/\beta$, a single subpopulation is split into two using the recipe described above. Then each subpopulation is recursively split as long as the global value of G decreases. Random combinations of the subpopulations are merged, and a possible merge is only accepted if G goes down. This gives the best possible partition for a given β. Sweeping over different β values gives many possible partitions that can be compared using a leakage measure:

$$Q = \langle \sum_{l \neq k} \tilde{C}_{kl} / \tilde{C}_{kk} \rangle_k \tag{2.5}$$

where $\tilde{C}_{kl} = \langle \sum_{i \in L_k} C_{ij} \rangle_{j \in L_l}$ is the mean connectivity from subpopulation l to k defined using set of sites L_k and L_l that lie in these subpopulations.

2.2 Stability and complexity of large ecological systems

Large ecosystems can be studied by stripping away the complexity of species interactions or the environment. At one end lies the niche paradigm that describes community properties by characterizing differences between species in terms of a few traits. In contrast to this selection dominated approach, the neutral view disregards inter-species differences such that species are identical on a per capita basis in their birth, death, dispersal or speciation rates [24]. A radically different approach is to retain the complexity of species interactions but assume that any structure in this network of interactions is washed away. This is the approach that I employ to study ecosystems where the complexity of species interactions cannot be ignored.

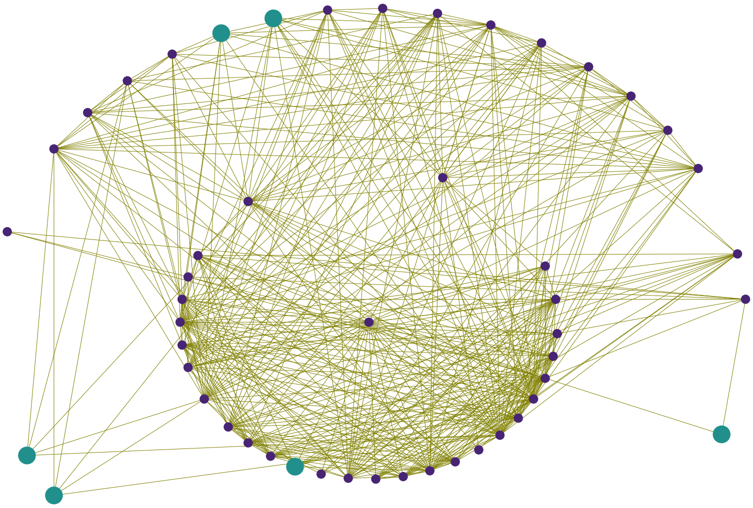

Figure 2.1: A food web with 50 species generated using the niche model in [25]. The larger nodes represent predators in a higher trophic level. The other species constitute an ecological community and lie in the same trophic level. In this sense, the community is embedded within a larger ecosystem represented by the entire food web. The given layout represents the most connected species closer to the centre. Not all possible interaction links are realized. The fraction of realized links is called the connectance.

An early example of describing such complex but random ecosystems is May's seminal work [26]. He used random matrix theory to analyse the stability of random ecosystems whose Jacobian can be written as:

$$\mathbf{J} = \mathbf{A} - \mathbf{I} \tag{2.6}$$

where $\mathbf{I}$ is an identity matrix with diagonal elements equal to 1. $\mathbf{A}$ is the interaction matrix, i.e., a random matrix where non-zero pairwise interaction strengths are drawn from a distribution symmetric around 0. Any interaction pair could have a non-zero interaction strength with probability c (i.e. the connectance). Equivalently, c is the fraction of links between species realized out of all the possible links. The eigenvalue spectrum of $\mathbf{J}$ is central to studying local stability. This form of stability relates to the behaviour of the system in response to small perturbations around an equilibrium point (or fixed point). If the largest eigenvalue of $\mathbf{J}$ is negative, then the system is locally stable and returns to the equilibrium point if slightly perturbed from it.

The matrix $\mathbf{A}$ can be parameterized in terms of the average interaction strength σ that is also proportional to the standard deviation of the underlying distribution. This parameter can be viewed as a scaling factor that scales all the interaction strengths by a constant factor. Equation 2.6 can be written more explicitly as:

$$\mathbf{J} = \sigma\mathbf{A} - \mathbf{I} \tag{2.7}$$

The above form helps tune the mean and standard deviation of $\mathbf{A}$ using just the σ parameter. May showed that stability of an ecosystem is limited by its complexity which is defined in terms of σ and the connectance c of the interaction matrix. This major result for random interaction matrices also provided intuition for possible aspects of stability in real ecosystems.

A simple implication is that communities with a large number of species are dominated by weak interactions between species, and strong interactions are more likely in communities with fewer connections between species. Also, for some given average interaction strength and connectance, interaction matrices with a modular structure would be more stable. This paper naturally raises the question: What systems have Jacobians of the form used by May, which are independent of the equilibrium abundances?

The population dynamics of many interacting species can be described by a system of ordinary differential equations (ODEs). One of the simplest such equations is the Generalized Lotka-Volterra (GLV) equations.

$$\frac{dy_i}{dt} = r_i y_i + y_i \sum_j B_{ij} y_j \tag{2.8}$$

where r_i is the intrinsic growth rate of i^{th} species. B is the interaction matrix whose elements are pairwise interaction strengths between species. Other general models of population dynamics differ in the functional form used for the growth and interaction terms. Equation 2.8 could be written in another form by explicitly including the carrying capacity K_i in a self-interaction term.

$$\frac{dy_i}{dt} = r_i y_i (1 - \frac{y_i}{K_i}) + \sigma y_i \sum_{j \neq i} B_{ij} y_j \tag{2.9}$$

Note that equation 2.9 could have up to 2^N equilibria for a system of N species since this system admits solutions with $x_i^* = 0$. The corresponding Jacobian depends on the equilibrium abundances x_i^*. The Jacobian does have a more elaborate structure than the one studied by May, but one can draw some interesting comparisons by analysing the case without any $x_i^* = 0$. For this case, if we set all r_i and K_i equal to 1, then the Jacobian can be written as:

$$\mathbf{J} = \mathbf{X}^* (\sigma \mathbf{A} - \mathbf{I}) \tag{2.10}$$

where $\mathbf{X}^*$ is a diagonal matrix with equilibrium abundances on the diagonal. This does resemble the Jacobian that May analysed but that case did not have the $\mathbf{X}^*$ part. If the interaction strengths are weak such that all species end up with equilibrium abundances close to 1 (i.e. the carrying capacities), then the GLV equations have similar properties to the kind of systems May analysed. This also explains the fact that May's framework implicitly excludes stable equilibria with $x_i^* = 0$. There is a rather sharp transition towards instability.

Equilibrium solutions with all positive abundances are called feasible solutions. The parameter regime corresponding to feasible solutions is also structurally stable, in that a small change in parameters does not change the qualitative dynamical behaviour of the system [27]. The GLV equations allow for a locally stable but structurally unstable phase characterized by single species extinctions. This phase facilitates the study of community assembly where not all species from the larger pool survive (Figure 2.2). This is an important

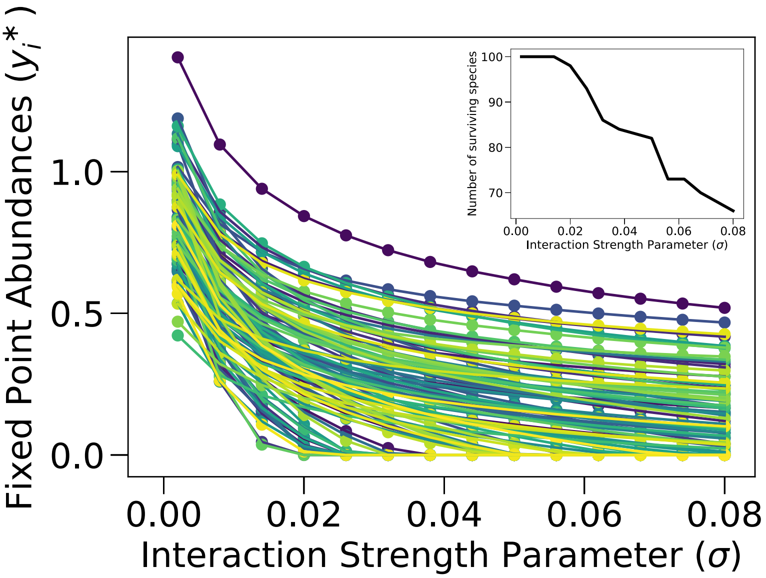

Figure 2.2: Equilibrium abundances of a competitive community of 100 species for increasing values of the interaction strength parameter (σ) (See equation 2.9). Each set of vertical dots represents an assembled community corresponding to a given value of σ. The bold black line in the inset traces the corresponding number of surviving species. Note that higher values of σ correspond to more extinctions. The interaction strengths, growth rates and carrying capacities are chosen from normal distributions with means -1, +1 and +1 respectively. The standard deviation is set to 0.2 for each of these.

aspect of papers 2 and 3.

2.3 Species-area relationships

It is debatable whether any universal quantitative laws exist in ecology. The species-area relationship (SAR) comes close to this distinction, but more than one functional forms of this relationship have been reported in empirical studies [3]. The two simplest forms of the SAR are the power law $S \sim A^z$ [28] and the semi-log $S \sim z \log A$ [29] – both forms suggest that species richness S increases monotonically with patch area A.

The power-law form has received much attention in terms of studies investigating its incidence. Its origin has been explained using empirical patterns such as lognormal species-abundance distributions [4] and clustering of conspecific (i.e. of same species) individuals [30]. The SAR could be studied for island-like systems or for nested sampling areas [4]. Island SARs differ in terms of processes such as immigration from a mainland pool, which is a key aspect of the analysis in Paper 2 (Section 3.2).

2.4 High-dimensional ecosystems in space

The past few decades have witnessed many approaches to study ecological communities distributed over spatially connected regions. A prominent approach is the metacommunity framework that considers local communities connected by dispersal of species. The spatial patches could be heterogeneous in the environmental conditions that further affect species traits [12, 13]. Paper 3 does focus on metacommunities but the term has a slightly different meaning operationally as in [1]. The focus is again on highly diverse complex communities but now these are spatially distributed over different patches that allow migration between one another.

The problem of studying complex ecosystems connected by migration draws from the tradition of embracing complexity over structure, which allows for understanding system-wide properties using a few statistics. Studies in this direction have uncovered novel results ranging from the stability-complexity paradox to newer types of dynamics that facilitate higher diversity of species [1, 2, 31]. These studies use similar approaches but we present the general paradigm using the GLV type equations discussed in [1]. Consider M patches connected by dispersal, each of which represents a well-mixed ecosystem with unstructured inter-species interactions:

$$\frac{d}{dt}\phi_{i,u} = \phi_{i,u}\left[K_{i,u} - \phi_{i,u} - \sum_{j\neq i} A_{ij,u}\phi_{j,u}\right] + \sum_{v} D_{i,uv}(\phi_{i,v} - \phi_{i,u}) \quad (2.11)$$

The dynamics of abundance $\phi_{i,u}$ of a species i in patch u depends in a familiar way on growth and interaction terms. For simplicity, the carrying capacities $K_{i,u}$ can be set to 1. $A_{ij,u}$ denotes the interaction matrix at a patch u, which could differ from other patches. The dispersal rate between patches u and v is $D_{i,uv}$.

It was shown in [31] that increasing the number of ecologically distinct patches – characterized by variation in species interactions – increases the stability of the meta-ecosystem. In [1], it was found that such spatially-extended systems can enter a chaotic phase for intermediate dispersal rates. A key feature of this phase are abundances fluctuations whose higher strength corresponds to higher fraction of species surviving. A cut-off is set below which species are considered extinct, but the results are unchanged qualitatively for different possible values. The authors of [2] found a similar chaotic regime for interaction matrices with an anti-symmetric structure.

Paper 3 considers only stable equilibria but the ideas of metacommunity and surviving fraction of species is used in the same vein as [1] (Section 3.3).

CHAPTER 3

Present Work

3.1 Paper 1

This study uncovers the spatial patterns in genetic connectivity of the seagrass *Halodule uninervis* and compares it to the oceanographic dispersal barriers along the northwestern coast of Australia. One key objective is to understand the resilience of the contemporary genetic structure to catastrophic disturbances. To counter the limited knowledge of dispersal properties of this seagrass, the oceanographic dispersal barriers are investigated for a range of fragment dispersal durations and for many different timescales over which the genetic structure could persist. Seagrasses are important ecosystems that have a range of functions from being a food-source for threatened species such as dugongs [33] to sequestering carbon [34]. Interestingly, the study landscape encompasses a large area north of Shark Bay, which is a global stronghold of dugongs.

My contribution pertains to oceanographic dispersal barrier (ODB) analysis that is described below.

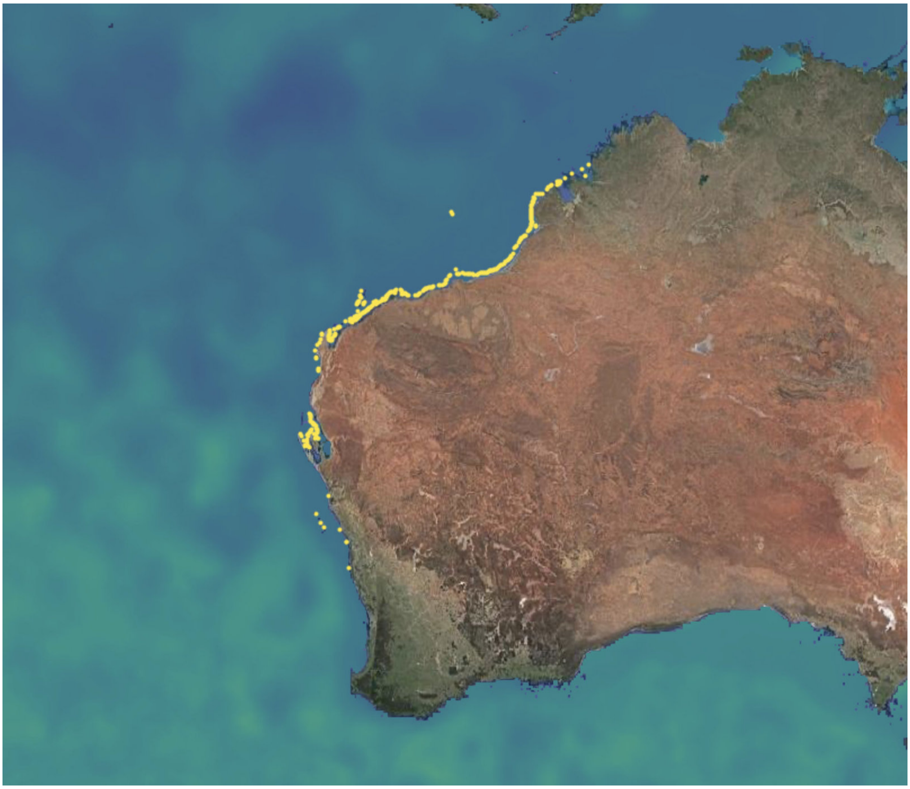

Figure 3.1: Seagrass population sites (yellow dots) along the north-western coast of Australia, which were analysed for dispersal barriers under different scenarios. A naive clustering of sites into metapopulations is possible based on geographical breaks in the distribution. However, this clustering is sub-optimal and might not capture the actual barriers to dispersal that is mediated by ocean currents and other biophysical factors. The shades of the ocean represent varying sea surface elevations that increase from lighter to darker shades. The map was generated using the HYCOM sea surface elevation data [32]

Methods

ODB analysis requires fragment dispersal matrices that are generated by simulating passive dispersal of vegetative fragments. Dispersal of biological species is sometimes modelled as a diffusion process but it is a much better approximation of passive dispersal by media such as ocean currents. The generation process uses a biophysical oceanographic model [35] combining data on ocean currents [36], fragment viability and mortality rate [37]. Preliminary dispersal matrices are first obtained by calculating likelihood of dispersal from source to destination populations – i.e. between all pairs of sites – for fragment dispersal duration (FDD) of 2, 7, 14 and 28 days. These matrices were then projected forward in time (for 10, 50 and 100 years) using a recursive recipe that combines matrix operations incorporating birth, subsequent migration and growth to carrying capacity [38]. The resulting fragment dispersal matrices represent combinations of different FDDs and life-history based temporal mixing scenarios. Dispersal barriers that create almost disconnected subpopulations (i.e. clusters of sites) are identified using the algorithm defined in [14].

Oceanographic dispersal barrier (ODB) analysis

Identifying subpopulations broadly amounts to dividing sites into clusters such that sites across different clusters have much lower transition probabilities than sites within. We use a threshold parameter to determine whether two sites are sufficiently likely to allow dispersal between them and hence be included in the same subpopulation. As in [14], by sweeping over the threshold parameter and varying the penalty for clustering weakly connected sites together, many alternative sets of subpopulations can be found. These solutions are compared by calculating the overall 'leakage' between constituent subpopulations, to find the optimal aggregation.

Some additional heuristics are used to find optimal solutions for the study dataset. There is a critical transition probability corresponding to each solution, which equals the inverse of the threshold parameter. This is useful when two solutions have the same number of clusters and equal leakage – the better solution should have a higher critical transition probability, which amounts to stronger connections within the clusters. If the best dispersal scenario corresponds to zero leakage between clusters, then we try to identify barriers within the largest of those clusters by allowing 'minimal non-zero leakage' between potential sub-clusters. This places ODB analysis in line with the genetic methods since different subpopulations are never entirely disjunct genetically.

Results

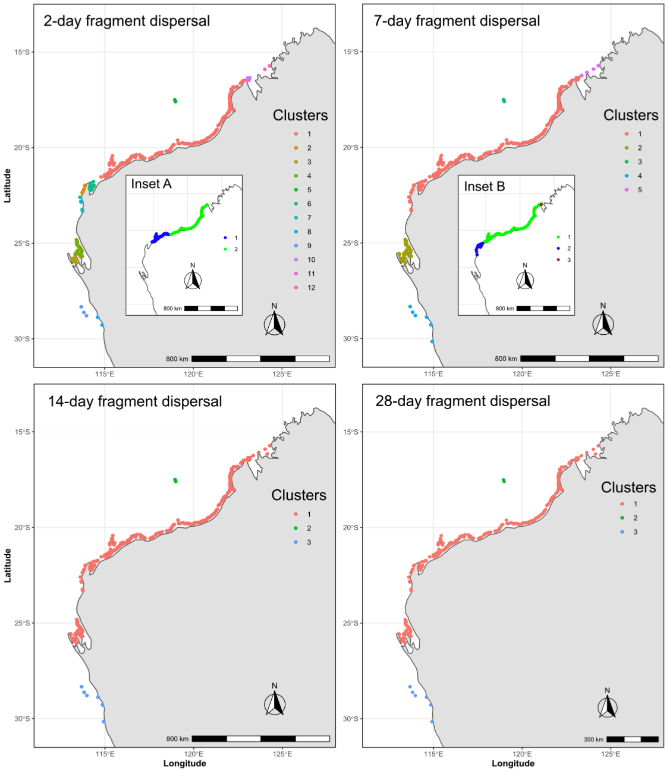

Figure 3.2: Results of the oceanographic barrier dispersal analysis for *H. uninervis* fragment dispersal based on basal modelling outputs from oceanographic modelling of 2-, 7-, 14- and 28-day fragment dispersal duration. Clusters indicate near zero leakage between identified groups. Inset A and B show clustering with minimal non-zero leakage for the base 2-day and the 7-day, respectively.

The genetic structure was best supported by the 2-day FDD model that identified 12 clusters with zero leakage that persisted up to the 100-year projection. The genetic results indicated an additional barrier (at Balla Balla) that was not captured by the ODB analysis. However, the largest cluster in the 2-day FDD split into two sub-clusters at this barrier when minimal non-zero leakage was allowed. The 7-day FDD could also identify some important dispersal barriers but the number of predicted clusters declined greatly for the 14- and 28-day scenarios – 3 clusters each.

Discussion

The ODB analysis used a variation of the algorithm in [14], which is particularly suited for the case of zero leakage in the predicted clusterings. The genetic clustering best matched the results of the 2-day FDD especially with minimal non-zero leakage – but the 7-day model also managed to capture some key dispersal barriers. The likely fragment dispersal duration might lie somewhere between 2 to 7 days. Since dispersal barriers exist at smaller spatial scales at the edge of the distribution, conservation should be managed at these scales to prevent loss of genetic diversity. The results from two complementary approaches show good agreement despite utilizing very different types of data and analysis approaches. ODB analysis could therefore serve as a cheap alternative to collecting genetic samples, especially in large landscapes where extensive sampling requires considerable resources.

3.2 Paper 2

This paper draws from the effects of the interaction strength parameter (σ) as in 2.9 to understand how assembled communities could be related to the patch area. The GLV equations can generate assembled communities of different sizes from a source pool through σ or equivalently the standard deviation of the interaction matrix [39]. The number of species decrease monotonically as σ increases in the regime where stable equilibria exist. This can be likened to the well-studied increase in species richness for increasing areas. It is interesting to ask if a modified version of the GLV equations could yield known forms of the species-area relationship without invoking any structure in species interactions? We address this question through a simple modification of the GLV equations that combines ecological intuition with observations from empirical studies.

Methods

The interaction strength parameter (σ) in equation 2.9 scales the strengths of all inter-species interactions. This effect is analogous to shrinking the area of a patch such that species encounter each other more frequently. Before analysing the consequences of spatial scaling, we make a few modifications to our existing system of equations. If the abundance densities are replaced by absolute abundances x_i , then equation 2.9 transforms to:

$$\frac{dx_i}{dt} = r_i x_i (1 - \frac{x_i}{K_i}) + \frac{x_i}{A} \sum_{j \neq i} B_{ij} x_j \tag{3.1}$$

where we do not use the σ parameter which was effectively the proxy for the standard deviation of the interaction matrix. Since K_i and x_i have the same units, additional factors of A do not appear in that term. Next, we account for the fact that absolute carrying capacities should change with changing areas. We modify equation 3.1 as:

$$\frac{dx_i}{dt} = r_i x_i (1 - \frac{x_i}{K_i (\frac{A}{A_{init}})^\gamma}) + \frac{x_i}{A} \sum_{j \neq i} B_{ij} x_j \tag{3.2}$$

where A_{init} is the minimum area that supports all species from the pool in the assembled community – this can be fixed by running the model numerically. We fix $\gamma = 0.25$, but in general $\gamma < 0.5$ is consistent with the results that are reported. The modifications we made translate to a weaker scaling of the carrying capacities than the scaling of inter-species interaction strengths. This constraint is equivalent to saying that self-interactions scale weakly with area. One would expect this situation if individuals of the same species cluster, which is not unusual in studies explaining the power-law SAR [30, 40]. Clustering of individuals translates to high self-interactions that do not scale drastically with changes in area.

Island SARs are particularly amenable to analysis through the model described. For a given A, the dynamics results in the assembly of a community of species from the species pool. By progressively sweeping over different values of A, we can get assembled communities representing different islands where species richness varies monotonically with A.

Community assembly with immigration

Immigration is another key aspect of island systems that could influence diversity patterns through introduction of new species [5], or the introduction of individuals of the existing species (demographic immigration) [41]. Demographic immigration is incorporated into our model using an additional term that captures incoming individuals from the mainland source pool:

$$\frac{dx_i}{dt} = r_i x_i (1 - \frac{x_i}{K_i(\frac{A}{A_{init}})^{\gamma}}) + \frac{x_i}{A} \sum_{j \neq i} B_{ij} x_j + \lambda e^{-\beta/\sqrt{A}} \tag{3.3}$$

The immigration term – $\lambda e^{-\beta/\sqrt{A}}$ – depends on A such that smaller islands receive almost negligible contributions through immigration. The exponential function further contains the parameter β – we fix β=1000 – that appears analogously as a characteristic length scale in [42]. We set the extinction limit to 10^{-5} even though this framework never really drives abundances to zero. The effective immigration rate equals λ for large A, which means that λ imposes limits on the immigration rates possible for an island system. We can compare results for different levels of immigration using this parameter.

Structured communities and empirical comparisons

We introduced many parameters to reflect the ecological setting of the Island SARs but this exercise helps avert some restrictive assumptions. In any case, our model still builds on the framework of studying random ecosystems especially in terms of how species interactions are structured. We also study ecological communities with more realistic connectances and distribution of interactions. This allows us to compare community level patterns that are insensitive to addition of further structure. The differences prompt scrutiny of additional mechanisms that drive the spatial patterns.

The findings are discussed in the context of empirical studies that benefit from SAR data over a large range of island areas [43–47]. Of particular interest are [43, 44], where two different Island SAR forms are discussed in relation to a common mainland.

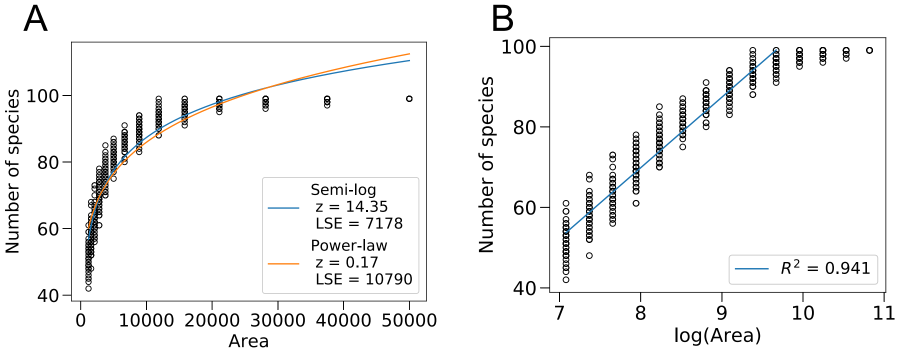

Figure 3.3: Species-area plots generated through 50 realizations of interaction matrix with mean = -1 and variance = 0.2. A_{init} = 50000. (A) The semi-log form shows a better fit. (B) The corresponding linear regression on a semi-log plot that shows an obvious upper asymptote.

Results

For no or low immigration rates, we get semi-log SARs (Figure 3.3). However, there exists an intermediate range of immigration rates that supports power-law SARs (Figure 3.4).

If interactions are structured such that the community contains many weak interactions and a fewer strong ones, then a higher skew towards weaker interactions shifts the SAR from a power-law to a semi-log relationship. Lower connectances – such as those found in real communities result in lower SAR slopes as in most natural communities.

Discussion

Our study identifies two major factors that differentiate semi-log and power-law SARs: immigration rates and skewness towards weak interactions. It is also evident that both functional forms show similar fits to the data except for very small and very large areas.

The case of remote archipelagos

A key result we report is that low immigration rates translate to semi-log SARs. As a crude test, empirical studies from remote archipelagos should

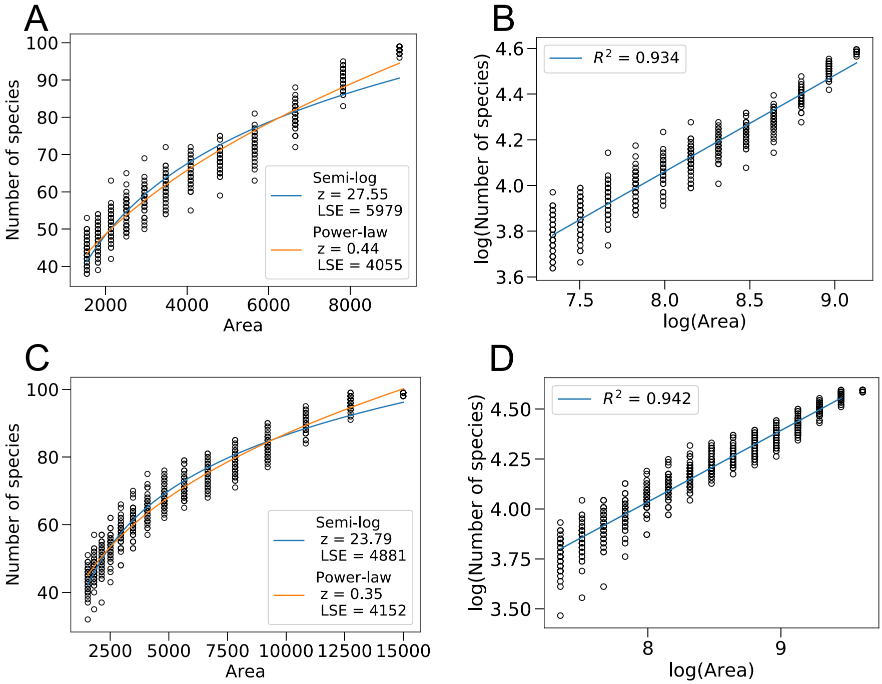

Figure 3.4: Species area plots demonstrating the better fit of power law SAR for intermediate values of immigration rates. Panels A and C show the fits for $\lambda = 0.1$ and $\lambda = 0.01$ respectively for 50 instances of the interaction matrix. Panels B and D correspond to the respective linear regressions on log-log plots. The interaction strength mean and variance are -1 and 0.2 respectively. $A_{init} = 15000$.

concur with this result. Two extensive studies from the Southwest Pacific are extremely well-suited to this end [43, 44]. The Solomon archipelago studied in [44] is more than 600 km away from the ‘source island’ of New Guinea. The authors find conclusive support for the semi-log form for islands that span over six orders of magnitude. The archipelago consists of three clusters of islands – each of which show very similar SAR slopes. In fact, the authors state that with increasing isolation of an archipelago, the SAR may shift in form from a power function to an exponential (semi-log). This observation is complemented by another study [43] (by the same authors), where they

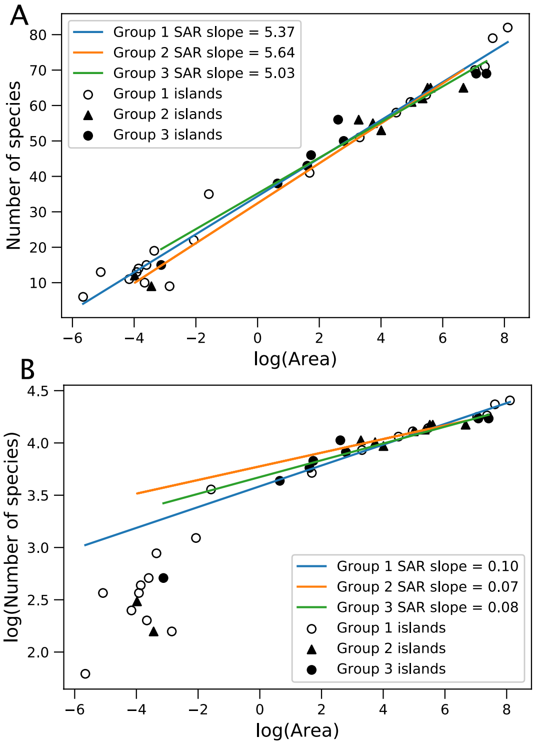

Figure 3.5: SAR plots for three groups of non-isolated islands within the Solomon Archipelago. These groups differ in how the islands within them were connected during the Pleistocene period. The islands in Group 3 did not have any history of connections. The semi-log relationship shows a good fit to data (A). The R^2 values for the regression lines are 0.978, 0.982 and 0.955 for Group 1, 2 and 3 respectively. The slopes for the different groups are very similar. Panel B shows a clear departure from a power-law relationship for smaller areas. The linear regression lines indicate a good fit for islands larger than one square mile. In particular, the R^2 value for such islands in group 1 is 0.976 from the power-law SAR. This demonstrates that a naive inference could support a power law, in spite of the islands spanning over four orders of magnitude in area ($>$ 1 square mile).

study islands that lie 5 to 500 km from New Guinea. That dataset shows a clear power-law for the SAR. Our finding on semi-log SARs is also supported by data from distant archipelagos in the Atlantic [45] as well as the Indian Ocean [46, 47] (Supplementary Appendix S2 in Paper 2).

Conclusion

This study disentangles factors that might result in the two most widely observed SAR forms for islands. We conclude that the semi-log functional form is observed for islands systems with low immigration rates. On remote archipelagos, the difference between the two functional forms might only be evident on the smallest islands where the power-law form overestimates species richness. This calls for a systematic estimation of the Island SARs in field studies, given that islands systems have witnessed disproportionately large number of extinctions [48, 49].

3.3 Paper 3

Consider a large contiguous landscape characterized by a many species interacting over different spatial locations. What diversity patterns emerge if the landscape is understood as a collection of many habitat patches connected by high dispersal of species? Of particular interest is the limit of very high dispersal resulting in spatial coherence of species densities across patches. When studied for the theoretical extreme of total randomness and no spatial correlation in inter-species interaction strengths, spatial coherence simplifies the analytic investigation of complex ecological systems. We further extend the emerging simple picture to more structured landscapes so as to explore various aspects of habitat heterogeneity and its relationship to diversity patterns.

Methods

We use a spatially-extended version of the GLV equations where dispersal between well-mixed patches is captured through diffusion. The equations read as:

$$
\begin{aligned}
\frac{\partial \phi_i(\mathbf{x},t)}{\partial t} &= r_i \phi_i(\mathbf{x},t)\left(1-\frac{\phi_i(\mathbf{x},t)}{K_i}\right) \\
&+ \phi_i(\mathbf{x},t)\sum_{j=1}^{N} A_{ij}(\mathbf{x})\phi_j(\mathbf{x},t) \\
&+ D\nabla^2\phi_i(\mathbf{x},t),
\end{aligned}
\tag{3.4}
$$

where $\phi_i(\mathbf{x},t)$ are species abundance densities for species i at position $\mathbf{x}$ and at time t. As previously, r_i and K_i represent intrinsic growth rates and carrying capacities respectively. These equations resemble reaction-diffusion equations, but we discretize the spatial dimension such that a given point in space represents a well-mixed patch that can undergo local GLV dynamics in the absence of dispersal. The Laplacian operator $\nabla^2 = \sum_{\alpha=1}^{d} \partial^2/\partial x_\alpha^2$ (where d is the dimension of space) is discretized on a two-dimensional lattice as:

$$
\nabla^2\phi_{i,\alpha\beta} = \left(\phi_{i,\alpha+1\beta} + \phi_{i,\alpha-1\beta} + \phi_{i,\alpha\beta+1} + \phi_{i,\alpha\beta-1} - 4\phi_{i,\alpha\beta}\right)/h^2. \tag{3.5}
$$

which implies that the Laplacian operator acting on abundance density $\phi_{i,\alpha\beta}$ at spatial patch (α,β) transforms it based on the abundance densities at adjoining patches and the given patch. Beyond this cumbersome notation, the diffusion term has a simple function – to enable dispersal between patches such that the abundance densities are ‘equalized’ or homogenized. The equation 3.4 also contains interaction and growth terms that dictate the dynamics of the system in conjunction with the diffusion term.

One primary objective of this analysis is to understand how spatial heterogeneity in interactions affects diversity patterns. The interaction strengths $A_{ij}(\mathbf{x})$ between any pairs of species are allowed to vary with the patch location $\mathbf{x}$, although the local means and variances of the interaction matrices are kept constant. In effect, interaction heterogeneity is used as a placeholder for habitat heterogeneity.

High-dispersal and spatial coherence

In the high-dispersal limit, abundance densities become coherent across spatial patches. An observation greatly simplifies the analysis of these spatially-extended systems – in the coherent limit, the system can be described by regular GLV equations with an effective interaction matrix $\bar{A}$ (cross ref fig). That is:

$$\frac{d\phi_i(t)}{dt} = r_i\phi_i(t)\left(1 - \frac{\phi_i(t)}{K_i}\right) + \phi_i(t)\sum_{j=1}^{N}\bar{A}_{ij}\phi_j(t), \tag{3.6}$$

Each entry within $\bar{A}_{ij}$ is an average of the corresponding pairwise interaction over all spatial patches. The standard deviation of $\bar{A}_{ij}$ could be used to study the stability and species richness of the spatially-extended ecosystem.

Spatial heterogeneity and species richness

It is unlikely that the interaction strength between two species is entirely uncorrelated across all spatial locations. A more reasonable setting is to consider some correlation between interaction strengths across space. We investigate how spatial heterogeneity affects the relationship between the number of habitats and species richness. This relationship has been actively studied in the context of the Area-Heterogeneity Trade-off (AHTO) hypothesis. The earliest proposals championed a unimodal relationship between the number of habitats (located in a fixed area) and species richness [50]. However, a recent experimental study reported a monotonic functional form [51].

The authors of [51] expected the positive relationship to be strengthened by increasing dispersal. The high dispersal coherent limit we described is conducive to theoretically understand these results. This is partly because the effect of environmental stochasticity is subdued when inter-species competition shows high variation [52]. High dispersal also increases the rate at which community assembly takes place, which means that stochastic effects are less dominant.

The simulation setup consists of nine patches, each of which could have a habitat characterized by an interaction matrix. The interaction strength for each species pair is correlated by a constant amount across habitats. We run our simulations for different spatial correlations ρ and plot the species richness resulting from an increasing number of habitats that could be randomly distributed over the nine patches.

It is unlikely that constant spatial correlations exist for very distant spatial patches. If a fixed correlation is assumed between adjacent spatial patches only, then how do species richness patterns compare against those from uniformly correlated cases? We address this question by analysing the situation in one spatial dimension.

Results

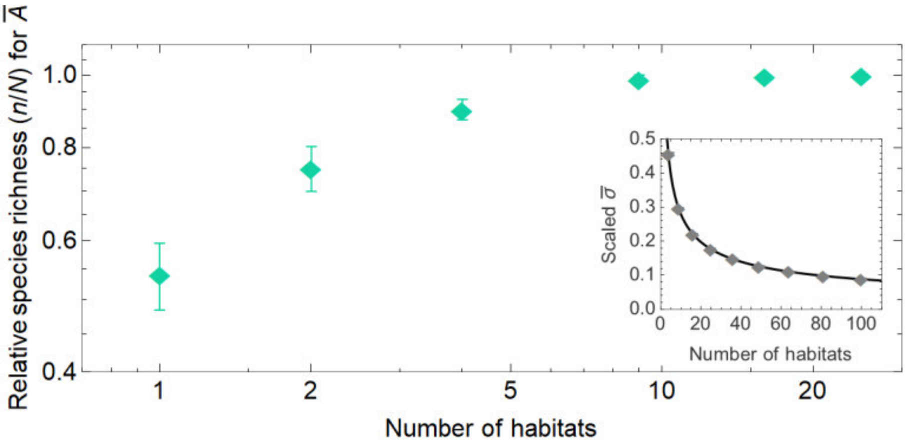

Figure 3.6: The plot shows how the relative species richness (n/N) increases with heterogeneity, in effect number of habitats. The data points are species richness averages from 20 runs of systems with $\mu_{A_{\alpha\beta}} = 0$ and $\sigma_{A_{\alpha\beta}} = 5/4\sqrt{cn}$ and one standard deviation errorbars. The species richness saturates at relative species richness 1, which corresponds to complete coexistence of all species in the original species pool. The inset shows the decrease of the standard deviation of the entries of the effective interaction matrix $\bar{\sigma}$. Since the effective system captures the dynamics and stability properties of the underlying metacommunity, this demonstrates (as in the framework of May) how the proxy for effective complexity of the metacommunity decreases, thereby allowing for a higher species richness.

If species interactions are totally uncorrelated in space, then the variance $\bar{\sigma}^2$ and mean $\bar{\mu}$ of the effective interaction matrix have exact analytical expressions. If each local interaction matrix has mean μ and standard deviation σ, then we find:

$$\begin{aligned} E[\bar{A}_{ij}] &= c\mu \\ &= \bar{\mu} \\ \\ Var[\bar{A}_{ij}] &= \frac{c}{G}(\sigma^2 + \mu^2(1-c)) \\ &= \bar{\sigma}^2, \end{aligned} \tag{3.7}$$

where c is the connectance. Each non-zero entry corresponding to species i and j at spatial patch g is considered a stochastic variable X_{ijg} where $i = j = 1, 2, ..N$, $g = 1, 2...G$.

Equation 3.7 implies that $\bar{\sigma} \to 0$ in the limit of infinitely many habitats (limit $G \to \infty$). Equivalently, a fully-feasible and stable ecosystem is theoretically possible for indefinitely large local σ if there exist an infinite number of habitats connected by high dispersal.

Spatial correlation mediates heterogeneity-richness relationships

Low values of ρ result in a significantly positive relationship between species richness and the number of habitats, irrespective of whether the interaction matrix mean is zero or negative (Figure 3.7) . For the case of zero interaction mean and spatial correlation ρ, the analytical expression for the variance $\bar{\sigma}_{\rho}^2$ of the effective interaction matrix can be calculated. This reads as:

$$\begin{aligned} Var[\bar{A}_{ij\rho}] &= \frac{\sigma^2}{G}\left(1 + (G-1)\rho\right) \\ &= \bar{\sigma}_{\rho}^2. \end{aligned} \tag{3.8}$$

where σ is the standard deviation of the local interaction matrices, as previously. In the limit of very large number of habitats G, we get

$$Var[\bar{A}_{ij\rho}] \sim \rho\sigma^2 \tag{3.9}$$

If one assumes spatial correlation ρ_{nn} between adjacent patches only, then the following result holds (for one-dimensional space):

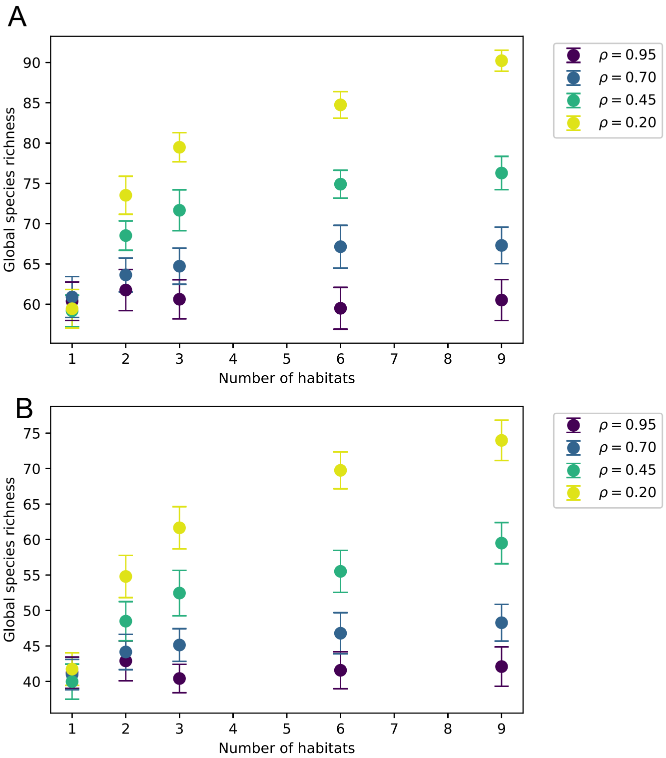

Figure 3.7: Relationship between habitat heterogeneity and species richness. For a given value of ρ, the plots show the mean species richness (with one standard deviation error bars) resulting from different number of habitats distributed over nine spatial patches. A given number of habitats corresponds to an equivalent number of correlated random interaction matrices such that the means are computed over 50 realizations of these matrices. The two panels correspond to (A) Mean interaction strength = 0. (B) Mean interaction strength = -0.5.

$$\begin{aligned} Var[\bar{A}_{ij\rho_{nn}}] &= \frac{\sigma^2\left(G + 2\sum_{\eta=1}^{G-1}\rho_{nn}^{\eta}(G-\eta)\right)}{G^2} \\ &= \bar{\sigma}^2_{\rho_{nn}}. \end{aligned} \tag{3.10}$$

If two habitat patches are separated by k patches between them, then the spatial correlation between them is equal to $(\rho_{nn})^{k+1}$. In other words, correlations fall off exponentially as the separation between patches increases. Given infinitely many spatial patches, $\bar{\sigma}_{\rho_{nn}} \to 0$. High species richness could still result from high spatial correlation between adjacent patches. The exponential decay of correlations with distance, paves way for large spatial variation in species interactions.

Discussion

Spatial coherence through high dispersal offers a unique perspective to understand diversity patterns in spatially-extended ecosystems. All the results we presented were obtained in this limit. We consistently found that spatial heterogeneity and high dispersal promote species richness for complex local communities.

The theoretical extreme of spatially uncorrelated interaction matrices provides simple yet rich insights. We show analytically that such a setting could allow for fully-feasible and stable metacommunities even if the local communities are extremely complex. Dispersal and heterogeneity have been shown to promote stability earlier as well [31].

The standard deviation of the effective interaction matrix could also be used to study diversity patterns in cases where species interactions are correlated uniformly everywhere or only between adjacent patches. We found that increasing the number of habitat patches in a fixed area results in significantly higher species richness, as opposed to the unimodal relationship advocated by the AHTO hypothesis. Since authors of [51] reported that increasing the number of habitats had a significantly positive effect on species richness, we surmise that this could be a consequence of high spatial variation in inter-species interactions.

A sufficiently large contiguous landscape would always have variation in species interaction strengths over space, which motivates the picture of many habitat patches connected to one another. The different cases we analysed consistently showed that spatial heterogeneity in interactions promotes high species richness. Factors that promote such heterogeneity could therefore explain how high diversity persists in large contiguous landscapes.

CHAPTER 4

Discussion and Future Work

Ecological processes in space can generate a range of patterns at species as well as the community level. With regards to dispersal, the papers discussed used this concept in very different ecological contexts.

Single species metapopulations have high dispersal probabilities for sites that lie within than across them. The coastal landscape we studied allowed for dispersal barriers that also indicate reduced gene-flow in an evolutionary sense. Paper 2 looked at complex ecological systems to understand the incidence of the island SAR. As discussed previously, a key ingredient that tipped the functional form of the relationship was dispersal of individuals from a mainland pool, i.e., demographic immigration. Mainland-island systems provided a simplified setting to study a spatial law (i.e. SAR) without explicitly invoking space. Paper 3 looked at the diametrically opposite case of contiguous landscapes by considering them as a patchwork of many habitats connected by high dispersal.

Interestingly, this drastic departure from an island-like configuration again facilitated a simplification with surprising results. High dispersal results in spatial coherence of species abundance densities, guaranteeing description of the system in terms of an effective interaction matrix. We used the properties of this matrix to get analytically exact results on stability and global species richness. The different cases yielded one consistent result: spatial heterogeneity in interactions has a positive effect on species richness, and this effect becomes stronger as the number of habitats increase. Although spatial coherence might be an idealized setting, but it does provide a framework to study processes that could drive high diversity in large contiguous landscapes.

Examples of such landscapes include tropical lowland forests that support astoundingly many species of various types. Communities of species such as trees show very high local (i.e. α) diversity in these forests, which is not well understood [53, 54]. Evolutionary aspects might be key to solving these puzzles but the temporal sequence of dispersal and speciation is only recently being uncovered.

A good example is the Amazon, where the absence of dispersal barriers means that dispersal and speciation are not trivially related. A recent phylogenetic study found evidence for widespread dispersal assembly of some diverse tree lineages [55]. Their proposal is to consider the entire Amazon basin as a metacommunity on evolutionary timescales.

Some interesting future avenues could be explored to understand these empirical findings. In particular, are speciation and widespread dispersal continuously in operation to generate high diversity? Or does speciation have a role only after local communities have been shaped by widespread dispersal? This does seem like an additional layer of complexity but stripping it away at the onset might deprive us of new patterns and simplifications that we constantly seek.

References

[1] F. Roy, M. Barbier, G. Biroli, and G. Bunin, “Complex interactions can create persistent fluctuations in high-diversity ecosystems,” *PLoS computational biology*, vol. 16, no. 5, e1007827, 2020.

[2] M. T. Pearce, A. Agarwala, and D. S. Fisher, “Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 25, pp. 14 572–14 583, 2020.

[3] S. Drakare, J. J. Lennon, and H. Hillebrand, “The imprint of the geographical, evolutionary and ecological context on species–area relationships,” *Ecology letters*, vol. 9, no. 2, pp. 215–227, 2006.

[4] F. W. Preston, “The canonical distribution of commonness and rarity: Part i,” *Ecology*, vol. 43, no. 2, pp. 185–215, 1962.

[5] R. H. MacArthur and E. O. Wilson, “An equilibrium theory of insular zoogeography,” *Evolution*, pp. 373–387, 1963.

[6] W. Dritschilo, H. Cornell, D. Nafus, and B. O’Connor, “Insular biogeography: Of mice and mites,” *Science*, vol. 190, no. 4213, pp. 467–469, 1975.

[7] T. Bell, D. Ager, J.-I. Song, J. A. Newman, I. P. Thompson, A. K. Lilley, and C. J. Van der Gast, “Larger islands house more bacterial taxa,” *Science*, vol. 308, no. 5730, pp. 1884–1884, 2005.

[8] S. I. Glassman, K. C. Lubetkin, J. A. Chung, and T. D. Bruns, "The theory of island biogeography applies to ectomycorrhizal fungi in subalpine tree islands at a fine scale," *Ecosphere*, vol. 8, no. 2, e01677, 2017.

[9] J. P. O'Dwyer and R. Chisholm, "A mean field model for competition: From neutral ecology to the red queen," *Ecology letters*, vol. 17, no. 8, pp. 961–969, 2014.

[10] H. G. Andrewartha, L. C. Birch, *et al.*, *The distribution and abundance of animals.* Edn 1. University of Chicago press, 1954.

[11] R. Levins, "Some demographic and genetic consequences of environmental heterogeneity for biological control," *American Entomologist*, vol. 15, no. 3, pp. 237–240, 1969.

[12] M. A. Leibold, M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, *et al.*, "The metacommunity concept: A framework for multi-scale community ecology," *Ecology letters*, vol. 7, no. 7, pp. 601–613, 2004.

[13] M. Holyoak, M. A. Leibold, and R. D. Holt, *Spatial dynamics and ecological communities*, 2005.

[14] M. N. Jacobi, C. André, K. Döös, and P. R. Jonsson, "Identification of subpopulations from connectivity matrices," *Ecography*, vol. 35, no. 11, pp. 1004–1016, 2012.

[15] R. Mac Arthur, "Species packing, and what competition minimizes," *Proceedings of the National Academy of Sciences*, vol. 64, no. 4, pp. 1369–1371, 1969.

[16] S. A. Levin, "Community equilibria and stability, and an extension of the competitive exclusion principle," *The American Naturalist*, vol. 104, no. 939, pp. 413–423, 1970.

[17] P. Chesson, "Quantifying and testing coexistence mechanisms arising from recruitment fluctuations," *Theoretical population biology*, vol. 64, no. 3, pp. 345–357, 2003.

[18] G. Barabás, R. D'Andrea, and S. M. Stump, "Chesson's coexistence theory," *Ecological Monographs*, vol. 88, no. 3, pp. 277–303, 2018.

[19] M. Fiedler, "Algebraic connectivity of graphs," *Czechoslovak mathematical journal*, vol. 23, no. 2, pp. 298–305, 1973.

[20] M. E. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Physical review E*, vol. 74, no. 3, p. 036 104, 2006.

[21] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026 113, 2004.

[22] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Physical review E*, vol. 70, no. 6, p. 066 111, 2004.

[23] M. N. Jacobi, "A robust spectral method for finding lumpings and meta stable states of non-reversible markov chains," *arXiv preprint arXiv:0810.1127*, 2008.

[24] S. P. Hubbell, "Neutral theory in community ecology and the hypothesis of functional equivalence," *Functional ecology*, vol. 19, no. 1, pp. 166–172, 2005.

[25] R. J. Williams and N. D. Martinez, "Simple rules yield complex food webs," *Nature*, vol. 404, no. 6774, pp. 180–183, 2000.

[26] R. M. May, "Will a large complex system be stable?" *Nature*, vol. 238, no. 5364, pp. 413–414, 1972.

[27] R. P. Rohr, S. Saavedra, and J. Bascompte, "On the structural stability of mutualistic systems," *Science*, vol. 345, no. 6195, 2014.

[28] O. Arrhenius, "Species and area," *Journal of Ecology*, vol. 9, no. 1, pp. 95–99, 1921.

[29] H. A. Gleason, "On the relation between species and area," *Ecology*, vol. 3, no. 2, pp. 158–162, 1922.

[30] H. G. Martín and N. Goldenfeld, "On the origin and robustness of power-law species–area relationships in ecology," *Proceedings of the National Academy of Sciences*, vol. 103, no. 27, pp. 10 310–10 315, 2006.

[31] D. Gravel, F. Massol, and M. A. Leibold, "Stability and complexity in model meta-ecosystems," *Nature communications*, vol. 7, no. 1, pp. 1–8, 2016.

[32] J. A. Cummings and O. M. Smedstad, "Variational data assimilation for the global ocean," in *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*, Springer, 2013, pp. 303–343.

[33] H. Marsh, A. Grech, and K. McMahon, "Dugongs: Seagrass community specialists," in *Seagrasses of Australia*, Springer, 2018, pp. 629–661.

[34] J. W. Fourqurean, C. M. Duarte, H. Kennedy, N. Marbà, M. Holmer, M. A. Mateo, E. T. Apostolaki, G. A. Kendrick, D. Krause-Jensen, K. J. McGlathery, *et al.*, "Seagrass ecosystems as a globally significant carbon stock," *Nature geoscience*, vol. 5, no. 7, pp. 505–509, 2012.

[35] J. T. Kool and S. L. Nichol, "Four-dimensional connectivity modelling with application to australia's north and northwest marine environments," *Environmental Modelling & Software*, vol. 65, pp. 67–78, 2015.

[36] E. P. Chassignet, H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, R. Baraille, and R. Bleck, "The hycom (hybrid coordinate ocean model) data assimilative system," *Journal of Marine Systems*, vol. 65, no. 1-4, pp. 60–83, 2007.

[37] L. M. Hall, M. D. Hanisak, and R. W. Virnstein, "Fragments of the seagrasses halodule wrightii and halophila johnsonii as potential recruits in indian river lagoon, florida," *Marine Ecology Progress Series*, vol. 310, pp. 109–117, 2006.

[38] J. T. Kool, C. B. Paris, S. Andréfouët, and R. K. Cowen, "Complex migration and the development of genetic structure in subdivided populations: An example from caribbean coral reef ecosystems," *Ecography*, vol. 33, no. 3, pp. 597–606, 2010.

[39] G. Bunin, "Ecological communities with lotka-volterra dynamics," *Physical Review E*, vol. 95, no. 4, p. 042 414, 2017.

[40] J. B. Plotkin, M. D. Potts, N. Leslie, N. Manokaran, J. LaFrankie, and P. S. Ashton, "Species-area curves, spatial aggregation, and habitat specialization in tropical forests," *Journal of theoretical biology*, vol. 207, no. 1, pp. 81–99, 2000.

[41] J. H. Brown and A. Kodric-Brown, "Turnover rates in insular biogeography: Effect of immigration on extinction," *Ecology*, vol. 58, no. 2, pp. 445–449, 1977.

[42] J. D. OSullivan, R. J. Knell, and A. G. Rossberg, "Metacommunity-scale biodiversity regulation and the self-organised emergence of macroecological patterns," *Ecology letters*, vol. 22, no. 9, pp. 1428–1438, 2019.

[43] J. M. Diamond, “Biogeographic kinetics: Estimation of relaxation times for avifaunas of southwest pacific islands,” *Proceedings of the National Academy of Sciences*, vol. 69, no. 11, pp. 3199–3203, 1972.

[44] J. M. Diamond and E. Mayr, “Species-area relation for birds of the solomon archipelago,” *Proceedings of the National Academy of Sciences*, vol. 73, no. 1, pp. 262–266, 1976.

[45] R. J. Whittaker, F. Rigal, P. A. Borges, P. Cardoso, S. Terzopoulou, F. Casanoves, L. Pla, F. Guilhaumon, R. J. Ladle, and K. A. Triantis, “Functional biogeography of oceanic islands and the scaling of functional diversity in the azores,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 38, pp. 13 709–13 714, 2014.

[46] L. D. Gooriah, P. Davidar, and J. M. Chase, “Species–area relationships in the andaman and nicobar islands emerge because rarer species are disproportionately favored on larger islands,” *Ecology and Evolution*, vol. 10, no. 14, pp. 7551–7559, 2020.

[47] ——, *Data from: Species-area relationships in the andaman and nicobar islands emerge because rarer species are disproportionately favored on larger islands*, Dryad, Dataset, 2020.

[48] C. Loehle and W. Eschenbach, “Historical bird and terrestrial mammal extinction rates and causes,” *Diversity and Distributions*, vol. 18, no. 1, pp. 84–91, 2012.

[49] D. R. Spatz, K. M. Zilliacus, N. D. Holmes, S. H. Butchart, P. Genovesi, G. Ceballos, B. R. Tershy, and D. A. Croll, “Globally threatened vertebrates on islands with invasive species,” *Science advances*, vol. 3, no. 10, e1603080, 2017.

[50] O. Allouche, M. Kalyuzhny, G. Moreno-Rueda, M. Pizarro, and R. Kadmon, “Area–heterogeneity tradeoff and the diversity of ecological communities,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 43, pp. 17 495–17 500, 2012.

[51] E. Ben-Hur and R. Kadmon, “An experimental test of the area–heterogeneity tradeoff,” *Proceedings of the National Academy of Sciences*, vol. 117, no. 9, pp. 4815–4822, 2020.

[52] A. M. Kramer and J. M. Drake, “Time to competitive exclusion,” *Ecosphere*, vol. 5, no. 5, pp. 1–16, 2014.

[53] R. Valencia, H. Balslev, and G. P. Y. Miño, "High tree alpha-diversity in amazonian ecuador," *Biodiversity & Conservation*, vol. 3, no. 1, pp. 21–28, 1994.

[54] E. G. Leigh Jr, P. Davidar, C. W. Dick, J. Terborgh, J.-P. Puyravaud, H. ter Steege, and S. J. Wright, "Why do some tropical forests have so many species of trees?" *Biotropica*, vol. 36, no. 4, pp. 447–473, 2004.

[55] K. G. Dexter, M. Lavin, B. M. Torke, A. D. Twyford, T. A. Kursar, P. D. Coley, C. Drake, R. Hollands, and R. T. Pennington, "Dispersal assembly of rain forest tree communities across the amazon basin," *Proceedings of the National Academy of Sciences*, vol. 114, no. 10, pp. 2645–2650, 2017.