Modelos ecológicos em redes complexas

Livia Akemi Hotta

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Livia Akemi Hotta

Modelos ecológicos em redes complexas

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Ciências – Ciências de Computação e Matemática Computacional. *EXEMPLAR DE DEFESA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Francisco Aparecido Rodrigues

USP – São Carlos
Julho de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

H834m Hotta, Livia Akemi
 Modelos ecológicos em redes complexas
 / Livia Akemi Hotta; orientador Francisco
Aparecido Rodrigues. - São Carlos - SP, 2017.
 58 p.

 Dissertação (Mestrado - Programa de
Pós-Graduação em Ciências de Computação
e Matemática Computacional) - Instituto
de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2017.

 1. Redes complexas. 2. relação espécie-área.
3. padrões de biodiversidade. I. Rodrigues,
Francisco Aparecido, orient. II. Título.

Livia Akemi Hotta

Ecological models in complex networks

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *EXAMINATION BOARD PRESENTATION COPY*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Francisco Aparecido Rodrigues

USP – São Carlos
July 2017

RESUMO

Hotta, L. A. **Modelos ecológicos em redes complexas**. 2017. 58p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Um dos padrões mais importantes que ocorrem em ecossistemas é a relação espécie-área, que relaciona o número de espécies em um ecossistema com a sua área disponível. O estudo dessa relação é fundamental para entender-se a biodiversidade e o impacto de políticas ambientais de preservação de espécies, de modo que é possível analisar desde os tamanhos das reservas necessários para a conservação das espécies e até verificar o impacto da intervenção humana em *habitats* naturais. Assim sendo, várias estratégias matemáticas e computacionais foram desenvolvidas para prever e entender esse padrão ecológico em modelos ecológicos. Todavia, muitas abordagens são simuladas em ambientes homogêneos e regulares, porém, sabe-se que, em cada ecossistema, há regiões com acidentes geográficos, variações de altitudes, vegetação e clima. Dessa forma, nesse trabalho, estamos interessados em estudar a influência de diferentes ambientes no processo de evolução das espécies. Para isso, consideramos modelos ecológicos que utilizam características geográficas para colonização e, comportamentos individuais como dispersão, mutação, acasalamento. Com isso, foi possível simular a propagação das espécies em diferentes topologias e analisar como ocorreu a dinâmica em cada uma delas. Assim, verificamos que a topologia regular e a dispersão homogênea dos indivíduos são duas características que maximizam a diversidade de espécies. E por outro lado, a formação de regiões mais densas e interações heterogêneas, contribuem para a diminuição da quantidade de espécies, apesar de em alguns casos, ajudarem na velocidade de propagação e colonização.

Palavras-chave: Redes complexas, relação espécie-área, padrões de biodiversidade.

ABSTRACT

Hotta, L. A. Ecological models in complex networks. 2017. 58 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

One of the most important patterns that occur in ecosystems is the species-area relationship, which says that the number of species increases with the sampled area. There is a great interest among ecologists about this pattern, since it is possible to verify the human impact on the environment and the area of reserves necessary to maintain species. Thus, motivated by the explanation of such behavior, some mathematical and computational strategies have been developed over the years. However, most approaches are simulated in homogeneous and regular scenarios, however, in the ecosystem, there are regions with landforms, different climates and vegetation. Thus, in this work, we are interested in studying the influence of different environments in the evolution process of the species. We consider ecological models that use geographical characteristics for colonization and individual behaviors such as dispersion, mutation, and mating. Thereby, it was possible to simulate the propagation of the species in different topologies and to analyze how the dynamics occurred in each case. Therefore, we verified that the regular topology and the homogeneous dispersion of the individuals are two characteristics that maximize the diversity of species. On the other hand, denser regions and heterogeneous interactions, contribute to the decrease the number of species, even when in some cases, they help in the speed of propagation and colonization.

Keywords: Complex network, species-area relationship, patterns of biodiversity.

LISTA DE FIGURAS

- Figura 1 – Redes complexas podem ser representadas por matrizes ou listas de adjacência. Considerando (a) uma rede não-dirigida. Em (b), tem-se a representação por matriz, sendo os elementos a_{ij} iguais a 1 se existir uma ligação entre os vértices i e j ou iguais a zero, caso contrário. Já no caso (c), tem-se a representação por lista de adjacência. Para uma aresta (i, j) , inclui-se uma lista ligada o elemento i na posição j do vetor, e o elemento j na posição i 20
- Figura 2 – Representação de uma rede para ilustrar o grau (k) e o grau médio dos vizinhos (knn). No caso, para o vértice 1 em vermelho, tem-se $k_1 = 3$. Já para knn_1 , o vértice 1 possui como vizinhos $\{2, 3, 4\}$, com graus $\{2, 2, 3\}$. Sendo assim, $knn_1 = \frac{2+2+3}{3} = 7/3$ 20
- Figura 3 – Ilustração de três redes com diferentes situações para o coeficiente de aglomeração do vértice 1, em vermelho. Em (a) tem-se todos os possíveis triângulos, sendo que todos os vizinhos de 1 são totalmente conectados. Nesse caso, o coeficiente é máximo, $cc_i = 1$. Já em (b), não se tem todos os possíveis triângulos, então, $cc_i = 1/3$. Por fim, em (c), o coeficiente é mínimo, $cc_i = 0$, quando não ocorre nenhum triângulo. 21
- Figura 4 – Ilustração do modelo *small-world* de Watts e Strogatz, para $c = 4$. O modelo varia entre a total regularidade à aleatoriedade de acordo com a probabilidade de reconexão. Em vermelho, estão as arestas selecionadas para a reconexão, e em preto as novas arestas. 23
- Figura 5 – Ilustrações do modelo livre de escala de Barabási-Albert. Em (a) Exemplo de uma rede construída a partir de 40 nós e $m = 2$. Em (b) A probabilidade de conexões para uma rede com 10.000 nós e $m = 6$ 23
- Figura 6 – Ilustrações do modelo livre de escala de Barabási-Albert. Em (a) Exemplo de uma rede construída a partir de 40 nós e $m = 2$. Em (b) A probabilidade de conexões para uma rede com 10.000 nós e $m = 6$ 24
- Figura 7 – Exemplo do modelo não linear de Barabási-Albert. O tamanho e a cor dos nós variam de acordo com a quantidade de conexões. (a) Quando $\alpha < 1$, o modelo se aproxima de uma rede aleatória, com isso, os nós possuem aproximadamente a mesma quantidade de conexões. (b) Já para $\alpha = 1$, o modelo é o original. (c) Para $\alpha > 1$, o modelo possui uma distribuição de conexão superlinear, sendo assim, há o aparecimento de nós super conectados. 25

Figura 8 – Ilustração da curva trifásica da relação espécie área para diferentes escalas. Figura criada com base em HUBBELL (HUBBELL, 2001).	28
Figura 9 – Ilustração de uma matriz na representação de sítios em uma rede, com nós e arestas. Assim, no caso, uma interação entre as áreas ocorre quando elas compartilham a mesma borda	29
Figura 10 – Fases do Modelo I	32
Figura 11 – Resultado das simulações do Modelo I na rede <i>small-world</i> com diferentes probabilidades de reconexão (ϕ) e mutação (μ). As figuras representam a relação espécie área em escala Log-Log. Em azul, estão as simulação para taxa de mutação $\mu = 0.1$. Em laranja, $\mu = 0.3$. Em roxo, $\mu = 0.5$. Em verde, $\mu = 0.7$. Em vermelho, $\mu = 0.9$	33
Figura 12 – Resultado das simulações do Modelo I para verificar qual propriedade da rede <i>small-world</i> está influenciando o processo de colonização. Fixamos $N = 10000$ e $\mu = 0.3$, variamos a probabilidade de reconexão ϕ entre 0.0 e 1.0.	34
Figura 13 – Comparação entre a média e mediana de 1000 simulações. Considerando uma rede <i>small-world</i> com 1000 vértices, taxa de mutação $\mu = 0.5$. Para cada coluna, tem-se gráficos com uma probabilidades de reconexão ϕ . Em vermelho representa-se a mediana e em azul, a média. No topo, tem-se a relação espécie-área. Embaixo, mostra-se a distribuição de probabilidade da quantidade de espécies.	35
Figura 14 – Comparação entre a média e mediana de 1000 simulações. Considerando uma rede não linear <i>scale-free</i> com 1000 vértices, taxa de mutação $\mu = 0.5$. Para cada coluna, tem-se um α . Em vermelho representa-se a mediana e em azul, a média. No topo, mostra-se a relação espécie-área. Embaixo, tem-se a distribuição de probabilidade da quantidade de espécies ¹	35
Figura 15 – Gráfico de comparação do Modelo I da relação espécie-área para diferentes topologia e taxas de mutação μ	36
Figura 16 – Gráfico em escala de cor para o valor de z estimado pelo método dos mínimos quadrados da curva $S = cA^z$. Pra cada linha é indicado um tipo de rede, e para cada coluna uma variação da taxa de mutação μ . Valores pequenos de z se aproximam da cor azul, e valores alto da cor vermelha.	36
Figura 17 – Representação matricial da colonização de espécies para diferentes topologias, sendo taxa de mutação $\mu = 0.3$ e área total $L \times L = 625$. Sendo (a), (b) e (c) redes <i>small-world</i> ; (d), (e) e (f) redes livre de escala Barabási-Albert; (g), (h) e (i) redes geográficas livres de escala. Os círculos representam as áreas, as arestas são as ligações e cada cor significa uma espécie.	38
Figura 18 – Representação das redes com a mesma dinâmica e parâmetros da Figura 17.	39

Figura 19 – Dados reais da curva espécie-área (BARBOUR, 1974) e curvas estimadas a partir dos resultados obtidos para taxa de mutação $\mu = 0.3$ dos diferentes tipos de redes. Cada curva estimada a partir das redes é indicado por uma letra e linhas tracejadas. A linha azul é a curva original estimada (STEIN; JURITZ, 1988) com $z = 0.15$. E em vermelho tem-se as duas curvas que mais se aproximam da original.	40
Figura 20 – Curvas obtidas para diferentes valores de μ	41
Figura 21 – Comparação entre os diferentes tipos de raio das sub-áreas na iteração = 1, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e o expoente da lei de potência $x = 0$	47
Figura 22 – Comparação entre os diferentes tipos de expoentes da lei de potência ($P(x) \approx x^\beta$) para a distribuição do raio das regiões, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$	48
Figura 23 – Comparação entre os diferentes tipos de raio das sub-áreas na iteração = 4000, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e o expoente da lei de potência $x = 0$	49
Figura 24 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração = 4000, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$	50
Figura 25 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração = 4000, variando-se a área e fixando-se a quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$	50
Figura 26 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração 4000, variando-se o raio das sub-áreas (r) e fixando-se a área 256x256, quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$	51
Figura 27 – Comparação entre os diferentes tipos de expoentes da lei de potência, fixando-se a área 256x256, a quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$	52

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Motivação e Objetivos	16
1.2	Organização	17
2	REDES COMPLEXAS	19
2.1	Caracterização de Redes Complexas	20
2.2	Modelos de Redes Complexas	22
2.2.1	<i>Modelo de Erdős-Rényi</i>	22
2.2.2	<i>Modelo de Watts e Strogatz</i>	22
2.2.3	<i>Modelo de Barabási-Albert</i>	24
2.2.4	<i>Modelo Espacial</i>	26
3	RELAÇÃO ESPÉCIE-ÁREA	27
4	MODELO BASEADO EM ESPÉCIES	29
4.1	Representação em redes complexas	29
4.2	Modelo I	30
5	MODELO BASEADO EM AGENTES	43
5.1	Modelo II	43
6	CONCLUSÃO	53
	REFERÊNCIAS	55

INTRODUÇÃO

A diversidade das espécies é um dos assuntos centrais na área de ecologia (MAYR, 2001). Um dos estudos fundamentais nesse campo é a teoria da evolução sugerida por Chales Darwin em 1859 no livro *On the Origin of Species*. Darwin apresenta, como ideia principal, a seleção natural, a qual os organismos com características mais adaptadas ao ambiente possuem maiores chances de sobrevivência, gerando também a diversidade das espécies. Ao longo dos anos, outras teorias surgiram e novas descobertas aprimoraram a teoria da evolução, como a variabilidade e mutação genética. Entretanto, ainda hoje, não estão listados todos os fatores envolvidos no processo evolutivo (COLLEY; FISCHER, 2013).

As interações entre espécies define a complexidade dos ecossistemas, onde há a emergência de padrões ecológicos bem conhecidos (CRAWLEY, 1997), tais como a relação espécie-área (em inglês, *species-area relationship*, SAR) (HOLT; LAWTON; ET, 1999) e a dominância e riqueza das espécies. Particularmente, o padrão SAR permite entender a diversidade e estimar o número de espécies em diferentes regiões (ROSENZWEIG, 1995). Desse modo, é possível quantificar qual o impacto da interferência humana nos *habitats* naturais, já que tal interferência pode provocar a diminuição da área de muitos ecossistemas e a extinção e isolamento de algumas espécies, além do desequilíbrio ecológico. Apesar de ser um padrão estudado desde meados do século XIX com o artigo publicado por Watson (WATSON, 1847-59), ainda não há um consenso sobre a causa da SAR. Dessa forma, diferentes modelos matemáticos foram propostos, baseado em abordagem estatística (CONCEIÇÃO *et al.*, 2014), curvas determinísticas (TJORVE, 2009), amostragem (KITZES; HARTE, 2014) e indivíduos (GRAS *et al.*, 2009). O modelo estatístico mais discutido da SAR é o apresentado por Arrhenius (ARRHENIUS, 1921), o qual generalizou a relação por uma lei de potência:

$$S(A) = cA^z, \quad (1.1)$$

sendo S o número de espécies e A o tamanho da área. As constantes c e z são usadas para o ajuste da curva. Variações nos parâmetros e formato da curva da SAR podem ocorrer dependendo de diferentes propriedades do ecossistema (MASHAYEKH; MACPHERSON; GRAS, 2014).

Além dessa preocupação em determinar o formato exato da curva, diversos modelos estocásticos foram propostos para verificar se é possível obter a relação espécie-área através de regras simples (AZAELE *et al.*, 2016). A maioria das estratégias propostas, na literatura, utilizam a modelagem baseada em indivíduos para a simulação do processo de especiação, e consequentemente, obtenção da curva espécie-área, como em (GRAS *et al.*, 2009; AGUIAR *et al.*, 2009; CONCEIÇÃO; OLIVEIRA, 2010). Esses modelos foram simulados em estruturas de grafos regulares (*grids*), sendo cada unidade de área um vértice, e o acesso aos vértices determinado pelas ligações entre eles.

Apesar da modelagem usando tais grafos fornecer resultados interessantes em muitas situações, ela é uma aproximação pouco vista na natureza, ocorrendo apenas entre animais rasteiros e planícies sem acidentes geográficos. Para uma modelagem mais real, é necessário considerar topologias mais heterogêneas, como as estudadas pela teoria das redes complexas (NEWMAN, 2003). Essa teoria vem sendo empregada em ecologia, mas apenas na análise de cadeias alimentares (COSTA *et al.*, 2011). Nesse caso, diversos estudos verificaram que as interações entre os indivíduos possuem alguns padrões e não ocorrem de maneira totalmente aleatória ou regular. Por exemplo em (MONTROYA; SOLÉ, 2002), os autores analisaram três cadeias alimentares e observaram que há uma grande ocorrência de triângulos em grupos de espécies com dietas similares. Verificaram também que a distribuição de conexões entre os vértices (espécies) segue uma lei de potência, mostrando que tais redes são altamente heterogêneas.

Nesse trabalho, estamos interessados em modelar a relação espécie-área em topologias heterogêneas, usando conceitos e métodos da teoria das redes complexas. Vamos estudar qual a importância da estrutura topológica na predição da biodiversidade, uma vez que o processo dinâmico pode ser influenciado se o meio, em que ele é simulado, for modificado. Para isso, utilizaremos o modelo baseado em espécies de (CONCEIÇÃO; OLIVEIRA, 2010), o qual considera que cada unidade de área possui uma capacidade de ser habitada e cada espécie pode sofrer mutações para gerar novas espécies. Além disso, estudamos o modelo proposto em (AGUIAR *et al.*, 2009), que considera fatores genéticos na dinâmica de populações. Verificamos que a estrutura da rede tem uma influência fundamental na diversidade das espécies.

1.1 Motivação e Objetivos

No estudo da biodiversidade, a relação espécie-área é um padrão muito observado nos ecossistemas. Vários modelos matemáticos e computacionais foram propostos para explicar essa propriedade, porém, devido à complexidade das espécies e de suas interações com fatores externos, até hoje, não há uma explicação completa sobre esse fenômeno.

Os modelos ecológicos publicados até o momento, são geralmente simulados em grades de duas dimensões, sendo que cada vértice tem o mesmo número de conexões, sendo todas locais (e.g. (CONCEIÇÃO; OLIVEIRA, 2010; AGUIAR *et al.*, 2009)). No mundo real, esse tipo de

abordagem representaria apenas áreas que possuem a mesma vegetação por grandes extensões, como as planícies. Porém, nem todos os *habitats* possuem essa natureza, principalmente, quando pensamos em lagos, rios, desertos e regiões montanhosas. Logo, cenários que consideram apenas interações locais conseguem reproduzir o comportamento apenas de animais terrestres, uma vez que eles só conseguem se locomover em curtas distâncias, o que não ocorre, por exemplo, com aves que podem migrar para regiões distantes. Uma maneira de aprimorar esses estudos, é incluir ambientes com estruturas heterogêneas, possibilitando, assim, inclusão de espécies que se dispersam por longas distâncias.

Nesse projeto estamos interessados em analisar a influência da estrutura topológica na evolução da população de espécies, considerando redes com ligações de longo alcance. Vamos considerar os modelos descritos em (CONCEIÇÃO; OLIVEIRA, 2010) e (AGUIAR *et al.*, 2009), sendo o primeiro baseado em espécies e o segundo, em indivíduos. Esses modelos consideram as principais características da evolução de uma espécie: mutação, genoma, competição e capacidade de uma área ser ocupada.

1.2 Organização

Os capítulos a seguir estão organizados da seguinte maneira: no Capítulo 2 descrevemos os conceitos sobre redes complexas, incluindo desde a sua representação até as principais medidas de caracterização e os modelos. No Capítulo 3, apresentamos a relação espécie-área, a qual iremos utilizar para verificar a influência da topologia na biodiversidade. No Capítulo 4, apresentamos os resultados da simulação do modelo baseado em espécies. Já no capítulo 5, são descritos os resultados para o modelo que considera indivíduos. Finalmente, no Capítulo 6 realizamos as considerações finais.

REDES COMPLEXAS

Uma rede complexa representa a estrutura de um sistema complexo, o qual possui elementos que interagem entre si de acordo com algum critério. Uma rede pode ser representada matematicamente por um grafo, isto é, um par $G(\mathcal{N}, \mathcal{E})$, o qual $\mathcal{N}$ e $\mathcal{E}$ representam a quantidade de elementos (nós ou vértices) e o conjunto de interações (arestas ou *links*), respectivamente.

Os pares de elementos podem estar conectados por meio de arestas com ou sem direção. Por exemplo, no contexto de uma cadeia alimentar, a ordem dos níveis tróficos devem ser representados por interações dirigidas. Porém, em uma rede social, pode-se adotar a estratégia sem direção, já que as relações de amizade são bidirecionais. Além disso, as arestas também possuem um peso, se necessário, como no caso de uma rede de transações comerciais, onde os pesos representam o montante das transações.

Computacionalmente, uma rede complexa pode ser implementada por meio de uma matriz ou uma lista de adjacência. Uma lista de adjacência, é uma estrutura de lista ligada que armazena apenas as arestas (i, j) . Já a matriz de adjacência A , cada elemento a_{ij} representa a conexão entre os vértices $i \rightarrow j$, isto é,

$$a_{ij} = \begin{cases} 1, & \text{se existe uma aresta entre } i \text{ e } j, \text{ ou seja, } (i, j) \in \mathcal{E}. \\ 0, & \text{caso contrário.} \end{cases}$$

A Figura 1 mostra um exemplo de representação da organização de uma rede não-dirigida em uma matriz e em uma lista de adjacência. Note que para uma rede não-dirigida, cada elemento $a_{ij} = a_{ji}$, isto é, a matriz é simétrica.

Nesse trabalho, utilizaremos a representação por matrizes, pois, apesar de utilizar mais memória para armazenar a rede, o tempo de execução dos algoritmos e simulações dos modelos estocásticos é mais rápido.

Figura 1 – Redes complexas podem ser representadas por matrizes ou listas de adjacência. Considerando (a) uma rede não-dirigida. Em (b), tem-se a representação por matriz, sendo os elementos a_{ij} iguais a 1 se existir uma ligação entre os vértices i e j ou iguais a zero, caso contrário. Já no caso (c), tem-se a representação por lista de adjacência. Para uma aresta (i, j) , inclui-se uma lista ligada o elemento i na posição j do vetor, e o elemento j na posição i .

2.1 Caracterização de Redes Complexas

Como a visualização é uma ferramenta que limita a extração de informações sobre a organização de redes, faz-se necessário o uso das medidas de caracterização topológica para descreveras suas propriedades. Uma das medidas para quantificar a centralidade de um nó i é o grau, o qual é definido pelo número de conexões de cada vértice, isto é,

$$k_i = \sum_{j=1}^N A_{ij}. \quad (2.1)$$

Essa medida é fundamental para a identificação de nós muito conectados, denominados de *hubs*.

O grau também permite quantificar outras propriedades importantes da rede. A probabilidade de selecionar um vértice com grau k , escolhido de maneira uniforme, permite quantificar o nível de heterogeneidade de uma rede. Além disso, a média do grau médio dos vizinhos de um nó j é calculada por

$$knn_j = \frac{1}{k_j} \sum_{i=1}^N k_i A_{ij}. \quad (2.2)$$

A Figura ?? exemplifica os conceitos de grau e grau médio dos vizinhos.

Figura 2 – Representação de uma rede para ilustrar o grau (k) e o grau médio dos vizinhos (knn). No caso, para o vértice 1 em vermelho, tem-se $k_1 = 3$. Já para knn_1 , o vértice 1 possui como vizinhos $\{2, 3, 4\}$, com graus $\{2, 2, 3\}$. Sendo assim, $knn_1 = \frac{2+2+3}{3} = 7/3$

Figura 3 – Ilustração de três redes com diferentes situações para o coeficiente de aglomeração do vértice 1, em vermelho. Em (a) tem-se todos os possíveis triângulos, sendo que todos os vizinhos de 1 são totalmente conectados. Nesse caso, o coeficiente é máximo, $cc_i = 1$. Já em (b), não se tem todos os possíveis triângulos, então, $cc_i = 1/3$. Por fim, em (c), o coeficiente é mínimo, $cc_i = 0$, quando não ocorre nenhum triângulo.

Redes complexas são caracterizadas também por curta distância entre pares de vértices. Sendo assim, seja d_{ij} a distância do menor caminho entre dois pares de vértices i e j . O caminho característico da rede é calculado como:

$$\ell = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}. \quad (2.3)$$

É importante ressaltar que por meio dessa medida é possível verificar o efeito pequeno-mundo (*small-world*). Essa propriedade pode ser observada quando o caminho característico da rede é relacionado com o N e $\langle k \rangle$ como $\ell \sim \frac{\log N}{\log \langle k \rangle}$. Em outras palavras, em redes que apresentam esse fenômeno, a distância entre quaisquer pares de nós cresce com o logaritmo do tamanho da rede, e decresce com a quantidade de conexões.

O número de menores caminhos que cada vértice participa pode ser entendida como uma medida de fluxo na rede. Dessa forma, o grau de intermediação (*Betweenness Centrality*) (FREEMAN, 1977) é calculada por:

$$B_u = \sum_{i,j} \frac{\sigma(i,u,j)}{\sigma(i,j)}, \quad (2.4)$$

sendo $\sigma(i,u,j)$ a quantidade de menores caminhos entre i e j passando por u e $\sigma(i,j)$ o total de menores caminhos possíveis entre i e j .

Em redes complexas observa-se também o aparecimento de subgrafos formados por três vértices totalmente conectados, ou seja, *loops* de ordem três (WATTS; STROGATZ, 1998). A ocorrência desse padrão de conexão é calculado pelo coeficiente de aglomeração (*clustering coefficient*), o qual mede a partir de um nó i qual a fração de possíveis triângulos formados por i e seus vizinhos. Matematicamente,

$$cc_i = \frac{\sum_{j=1}^N \sum_{m=1}^N a_{ij} a_{jm} a_{mi}}{k_i(k_i-1)} = \frac{2e_i}{k_i(k_i-1)}, \quad (2.5)$$

sendo e_i igual ao número de arestas entre os vizinhos de i . Vale ressaltar que a maioria das medidas apresentadas são calculadas para cada nó da rede, entretanto, como caracterização global utiliza-se a sua média amostral.

2.2 Modelos de Redes Complexas

No estudo de sistemas reais, como redes sociais, biológicas e transporte, é interessante entender qual as suas principais características e evolução. Nessa seção, iremos apresentar os modelos principais de redes complexas que utilizamos no projeto.

2.2.1 Modelo de Erdős-Rényi

Erdős e Rényi apresentaram o modelo conhecido como grafo aleatório $G_{N,p}$ (ERDŐS; RÉNYI, 1959; ERDŐS; RÉNYI, 1960; ERDŐS; RÉNYI, 1961), o qual considera uma probabilidade fixa para a ocorrência de cada aresta. Em outras palavras, a construção desse modelo se inicia com N vértices totalmente desconectados, e a cada passo, sorteia-se dois deles, os quais serão conectados com uma probabilidade p , sendo que cada par de vértices é escolhido apenas uma vez.

Várias propriedades foram demonstradas para esse modelo. Como o processo de construção da rede se trata de um processo de Bernoulli, o grau médio é dado por $\langle k \rangle = p(N - 1)$. Além disso, tem-se que a distribuição do grau é do tipo Binomial. No caso limite, quando N grande o grau médio fixo, a distribuição de conexão converge para distribuição de Poisson, ou seja,

$$P(K = k) = p_k = \binom{n}{k} p^k (1 - p)^{n-k} \simeq \frac{z^k e^{-z}}{k!}, \quad (2.6)$$

sendo o grau médio representado por z .

Apesar de ser matematicamente tratável, o grafo aleatório dificilmente é utilizado em aplicações práticas, uma vez que, dentre outras propriedades, em redes reais não se observa a distribuição de conexões na forma de uma Poisson. Dessa forma, geralmente, o modelo é utilizado apenas para efeito comparativo como um modelo nulo. Com o propósito de criar modelos mais próximos dos sistemas reais, surgiram os modelos pequeno-mundo e livre de escala.

2.2.2 Modelo de Watts e Strogatz

Duncan Watts e Steven Strogatz propuseram um modelo para reproduzir dois fenômenos observados em redes reais: a presença de *loops* de ordem três e o efeito pequeno-mundo (WATTS; STROGATZ, 1998). Para isso, combinaram duas características a regularidade e aleatoriedade.

Na rede regular nota-se uma grande quantidade de triângulos, mas sem a presença do efeito pequeno-mundo, uma vez que há uma grande distância entre os pares de vértices. Já a rede aleatória, observa-se o efeito *small-world*, entretanto com pouca ocorrência de triângulos. Assim, o modelo proposto por Watts e Strogatz, denominado pequeno-mundo (*small-world*), está situado entre essas duas redes, ver Figura 4.

Figura 4 – Ilustração do modelo *small-world* de Watts e Strogatz, para $c = 4$. O modelo varia entre a total regularidade à aleatoriedade de acordo com a probabilidade de reconexão. Em vermelho, estão as arestas selecionadas para a reconexão, e em preto as novas arestas.

No modelo original de Watts-Strogatz a construção inicia-se com uma rede regular, a qual cada vértice é conectado com c vizinhos mais próximos, sendo c um número par. Para cada aresta, selecionada apenas uma vez, tem-se a probabilidade p de reconectar apenas de um dos seus vértices terminais. Essas reconexões são importantes pois criam “atalhos” na rede, que diminuem a distância entre os vértices e geram o efeito *small-world*. Entretanto, essa definição de reconexão pode gerar componentes desconexos e resulta em redes não totalmente aleatórias.

Sendo assim, utilizaremos a sua variação proposta por Newman-Watts (NEWMAN; WATTS, 1999), o qual a reconexão ocorre com quaisquer outros pares de vértices escolhidos aleatoriamente. A partir disso, pode-se afirmar que a probabilidade de reconexão p determina a variação entre a regularidade e aleatoriedade. Quando $p = 0$ a probabilidade de reconexão é nula, logo, a rede será regular. Porém, para $p = 1$ todas as arestas são reconectadas, gerando uma rede totalmente aleatória, porém não no sentido $G_{n,p}$.

Figura 5 – Ilustrações do modelo livre de escala de Barabási-Albert. Em (a) Exemplo de uma rede construída a partir de 40 nós e $m = 2$. Em (b) A probabilidade de conexões para uma rede com 10.000 nós e $m = 6$.

A distribuição do grau para esse modelo é definida de acordo com a probabilidade p de reconexão. Assim, inicialmente, tem-se em média $(\frac{1}{2}nc)p$ arestas que serão reconectadas e, ncp vértices que participarão dessas ligações. Isto significa que em média cada vértice possui np

arestas de reconexão. Consequentemente, a distribuição do número de reconexões s é descrita como uma distribuição de Poisson,

$$p_s = e^{-cp} \frac{(cp)^s}{s!}. \quad (2.7)$$

Porém, o grau k de um vértice é dado pela quantidade de arestas c dos vizinhos, mais s arestas de reconexão, isto é, $k = c + s$. Substituindo $s = k - c$ na Equação 2.7, obtem-se a distribuição do grau,

$$p_k = e^{-cp} \frac{(cp)^{k-c}}{(k-c)!}, \quad (2.8)$$

para $k \geq c$ e $p_k = 0$ se $k < c$.

Dessa forma, assim como na rede de Erdős-Rényi, a distribuição do grau é em forma de Poisson, entretanto, o modelo de Watts-Strogatz consegue simular a ocorrência de triângulos e o efeito *small-world*.

2.2.3 Modelo de Barabási-Albert

Em 1999, motivados ainda pelo efeito *small-world*, os pesquisadores, Albert-László Barabási e Reka Albert, analisaram a estrutura topológica da rede WWW (*World Wide Web*).

Figura 6 – Ilustrações do modelo livre de escala de Barabási-Albert. Em (a) Exemplo de uma rede construída a partir de 40 nós e $m = 2$. Em (b) A probabilidade de conexões para uma rede com 10.000 nós e $m = 6$.

Nesse estudo, além do efeito *small-world*, eles verificaram que a distribuição de conexões não se aproximava dos modelos propostos até o momento, os quais possuíam uma escala típica com média e desvio padrão bem definidos. Pelo contrário, o aparecimento de alguns nós com muitas conexões (*hubs*) e muitos nós pouco conectados criava uma distribuição livre de escala, $P(k) \propto k^{-\gamma}$, com isso, eles propuseram um modelo para gerar redes livre de escala (BARABÁSI; ALBERT, 1999).

Figura 7 – Exemplo do modelo não linear de Barabási-Albert. O tamanho e a cor dos nós variam de acordo com a quantidade de conexões. (a) Quando $\alpha < 1$, o modelo se aproxima de uma rede aleatória, com isso, os nós possuem aproximadamente a mesma quantidade de conexões. (b) Já para $\alpha = 1$, o modelo é o original. (c) Para $\alpha > 1$, o modelo possui uma distribuição de conexão superlinear, sendo assim, há o aparecimento de nós super conectados.

Para a construção desse modelo livre de escala (*scale-free*) é necessário dois passos: crescimento e ligação preferencial. Assim, para construir uma rede com N vértices, inicia-se selecionando um subconjunto N_0 de vértices totalmente conectados. Em seguida, a cada passo de tempo t na fase do *crescimento*, adiciona-se um novo vértice com m ($m \leq N_0$) arestas conectadas com vértices já presentes na rede. Sendo que essas ligações ocorrem segundo a probabilidade de *ligação preferencial*, a qual é dada como:

$$\mathcal{P}_{i \rightarrow j}(t+1) = \frac{k_j(t)}{\sum_u k_u(t)}, \quad (2.9)$$

Sendo $\mathcal{P}_{i \rightarrow j}(t+1)$ a probabilidade do vértice i se conectar ao j no tempo $t+1$ e k_j o grau do vértice j . Para o modelo de Barabási-Albert, tem-se que para $N \rightarrow \infty$, a distribuição do grau segue uma lei de potência com $\gamma = 3$, ver Figura 6

Entretanto, é importante notar que em redes reais esse valor não é fixo, geralmente, ele varia entre dois e três, como a rede *World Wide Web* que possui $\gamma = 2.1$ (ALBERT; JEONG; BARABÁSI, 1999; BARABÁSI; ALBERT; JEONG, 2000), e a rede de citações com $\gamma = 3$. (REDNER, 1998).

O modelo livre de escala BA foi generalizado por Krapivsky *et al.* (KRAPIVSKY; REDNER; LEYVRAZ, 2000), Figura 7, a qual a ligação preferencial é dada de forma não-linear:

$$\mathcal{P}_{i \rightarrow j}(t+1) = \frac{k_j(t)^\alpha}{\sum_u k_u(t)^\alpha}, \quad \alpha > 0. \quad (2.10)$$

Dessa forma, para $\alpha = 1$, obtêm-se o modelo original de Barabási e Albert. Quando $\alpha < 1$, há pouca influência da probabilidade de conexão preferencial e a rede construída se aproxima de uma rede aleatória. Por outro lado, quando $\alpha > 1$, a probabilidade de conexão é

superlinear, isto é, há grande influência do grau na probabilidade. Como resultado, a rede gerada possui *hubs* muito conectados, os quais se ligam com quase todas as partes da rede.

2.2.4 Modelo Espacial

Além de características como quantidade ou probabilidade de conexões entre os nós ou ainda formação de estruturas na rede, por exemplo triângulos, há também redes que consideram a distância espacial entre os nós.

Dentre esses modelos, Bathélemy estudou e propôs o modelo que transita entre a rede livre de escala e a rede espacial (BARTHÉLEMY, 2003), aproximando-se mais de modelos de redes reais as quais, muitas vezes, possuem distâncias geodesicas porém não utilizam em seus modelos.

Assim, o modelo é descrito como: dado uma rede de N nós, em um espaço d – dimensional com tamanho linear L , a probabilidade de conexão entre um subconjunto de nós inativos e ativos é dado como:

$$P_{i \rightarrow j} \propto \frac{Z(k_j)}{\Delta(d_{ij})}. \quad (2.11)$$

sendo Δ e Z as funções consideradas, k_j a quantidade de interações do nó j , d_{ij} a distância entre os nós i e j . Caso a função $\Delta = \text{const}$, apenas a função Z irá influenciar essa probabilidade de conexão. Consequentemente, sendo $Z = k + 1$, recupera-se o modelo original de Barbási-Albert do modelo de ligações preferenciais, o qual obtem um expoente de $\gamma = 3$ para a lei de potência da distribuição de conexões.

Para determinar a função de distância, os autores consideraram a seguinte função de distância:

$$\Delta_{ij}(d) = e^{d/r_c}, \quad (2.12)$$

sendo r_c o raio de influência da distância na conexão, d a distância entre os nós i e j . Com isso, tem-se que devido a função exponencial, quando $r_c \gg d$, o expoente se aproximará de zero e a função $\Delta \approx \text{const}$, por outro lado para $r_c < d$ haverá influência da distância e a combinação de propriedade de uma rede espacial e livre de escala.

RELAÇÃO ESPÉCIE-ÁREA

Um dos padrões mais conhecidos na ecologia é a relação espécie-área (em inglês, *species-area relationship*, SAR). Basicamente, a SAR caracteriza o aumento da quantidade de espécie de acordo com o tamanho da área amostrada. Essa relação é observada desde o século XIX, quando o primeiro gráfico sobre o assunto apareceu no trabalho feito por Watson (WATSON, 1847-59).

Apesar de ser um estudo antigo, ainda não se tem um consenso sobre qual o modelo matemático mais apropriada para descrever a SAR. Em virtude disso, em (MASHAYEKH; MACPHERSON; GRAS, 2014), os pesquisadores analisaram 28 diferentes modelos propostos até o momento, utilizando o simulador de presa-predador EcoSim (GRAS *et al.*, 2009).

A conclusão dos autores foi que os modelos SAR possuem uma alta dependência da escala e estratégias da amostragem, que podem afetar na forma da curva. Além disso, apesar de não existir nenhuma função que descreva melhor todas as relações espécie-área, os pesquisadores afirmam que o modelo mais simples e que melhor se ajusta aos dados é a função exponencial, a qual possui apenas dois coeficientes para serem estimados. A função exponencial foi apresentada por Arrhenius (ARRHENIUS, 1921). Sendo descrita como $S = cA^z$, sendo S a quantidade de espécie, A o tamanho da área amostrada e c e z constantes. No caso, c significa a quantidade de espécies em uma unidade de área. Já a constante z pode variar de acordo com o grupo de espécies estudado. Por exemplo, considerando pássaros, mamíferos e anfíbios, o valor de z varia entre 0.17 – 0.64, segundo o estudo em (STORCH; KEIL; JETZ, 2012). Já em organismos menores, como bactérias, analisado em (HORNER-DEVINE *et al.*, 2004), o valor de z é menor, variando entre 0.008 – 0.04.

A universalidade do modelo exponencial da SAR é um assunto bastante discutido. Estudos teóricos já afirmaram que a inclinação da curva da relação espécie-área pode variar de acordo com a escala espacial (ROSENZWEIG, 1995; CONNOR; MCCOY, 1979). Especificamente, notou-se que, nos sistemas reais, a SAR muda em função da escala espacial de forma trifásica. Isto significa que em escalas intermediárias a SAR se aproxima de uma lei de potência, modelo

Figura 8 – Ilustração da curva trifásica da relação espécie área para diferentes escalas. Figura criada com base em HUBBELL (HUBBELL, 2001).

de ARRHENIUS, porém, o mesmo comportamento não ocorre para pequenas e grandes escalas. A Figura 8 mostra essa curva trifásica para a relação espécie-área. A escala pequena pode ser comparada à 1 acre ($\sim 4.000 \text{ m}^2$), a escala intermediária à 100.000 m^2 e grande escala seria continentes inteiros ou toda a superfície da Terra.

A explicação desse comportamento foi descrito em (HUBBELL, 1997; HUBBELL, 2001). O aspecto para a escala local ocorre em virtude da relação espécie-área ser muito sensível à abundância de espécies, consequentemente, espécies comuns são adicionadas rapidamente, enquanto que espécies raras são coletadas mais lentamente. Com isso, considerando um gráfico log-log, a curva é mais inclinada, decrescente e de forma não-linear. Já para a escala intermediária, a curva é linear nesse mesmo gráfico, que decorre do equilíbrio entre a especiação, dispersão e extinção. Por fim, em grandes escalas, a curva de espécie-área é inclinada e crescente, por conta de processos históricos independentes de evolução e barreiras biogeográficas.

MODELO BASEADO EM ESPÉCIES

Dentre vários modelos baseado em indivíduos, para o desenvolvimento do projeto, iremos analisar a influência da estrutura topológica em dois modelos principais. O primeiro é o apresentado por (CONCEIÇÃO; OLIVEIRA, 2010) e o segundo por (AGUIAR *et al.*, 2009). A escolha desses dois modelos foi motivada por envolverem processos estocásticos e características como especiação, taxa de mutação, competição, interação genética.

Em (CONCEIÇÃO; OLIVEIRA, 2010), os autores consideraram o modelo para a análise da diversidade de espécies, analisando em um espaço e interações regulares, qual a influência de áreas fragmentadas.

4.1 Representação em redes complexas

O modelo que iremos utilizar tradicionalmente são simulados em ambientes regulares, como grades (CONCEIÇÃO; OLIVEIRA, 2010). Entretanto, além da regularidade da área, esse tipo de abordagem considera apenas interações locais, ou seja, cada área só pode ser acessada por quatro vizinhos próximos.

Figura 9 – Ilustração de uma matriz na representação de sítios em uma rede, com nós e arestas. Assim, no caso, uma interação entre as áreas ocorre quando elas compartilham a mesma borda

Na representação dos ambientes em redes complexas cada unidade de área será o vértice da nossa rede, e a existência de uma aresta permite acessar uma área por meio de outra, ver

Figura 9. Ademais, os organismos só podem se locomover em vértices que estão conectados.

4.2 Modelo I

O modelo apresentado por (CONCEIÇÃO; OLIVEIRA, 2010) considera três fatores: taxa de mutação, capacidade da área em ser colonizada e competição entre as espécies. A área a ser colonizada é representada por uma grade sem limite de borda com dimensão $(L \times L)$, sendo assim, tem-se no total $n = L * L$ unidades de área. Cada uma será definido como um sítio x_i , o qual $i = \{1, \dots, n\}$. Assim, as características gerais do modelo são:

- Para o modelo
 - Quantidade de sítios: $n = L * L$
 - Taxa de mutação (μ): informado como entrada.
- Para cada sítio i
 - Capacidade C_i : número aleatório gerado a partir de uma distribuição $U(0, 1)$.
 - Pode existir competição: quando mais de uma espécie é vizinha da mesma área, sendo assim, são indivíduos com potencial de colonização.
 - Possui interação (acesso) à outras k áreas, sendo $k \in \{1, \dots, n\}$
- Para cada espécie w
 - Adaptabilidade da espécie (*fitness*): $f_w = \sum_i C_{i,w}$. Sendo $C_{i,w}$ a capacidade das área já colonizadas pela espécie w .
 - Mutação da espécie: $\mu' = \mu / f$. Observe que essa taxa de mutação diminui com o aumento do *fitness* da espécie, ou seja, a quantidade de áreas colonizadas.

Dadas essas particularidades, descrevemos o modelo fazendo referência as linhas do Algoritmo 1. Para começar o processo (linhas 7-10), escolhe-se, aleatoriamente, algum sítio disponível para a colonização. Em seguida, nesse primeiro sítio, ocupa-se com a 1º espécie, atualizando o valor do seu *fitness*, e inserindo em uma lista, *areas* (linha 11), os possíveis sítios de serem colonizados. Depois, inicia-se um laço de repetição (linhas 12-26) até todos os sítios de *areas* serem ocupados por alguma espécie.

Na lista *areas*, os sítios adicionados são os não ocupados e vizinhos do sítio que acabou de ser colonizado. Para a repetição, tem-se os seguintes passos: seleciona-se aleatoriamente algum sítio da lista *areas* (linha 13). Essa escolha é ponderada pela capacidade C_i desses sítios, sendo assim, é uma escolha aleatória, porém a chance de cada sítio ocorrer é dada como $C_i / \sum_i C_i$.

Algoritmo 1: Dinâmica de Colonização**Data:** Dada uma rede $G(N,E)$

// N: quantidade de sub-regiões. E: interações entre as sub-regiões.

```

1 begin
3   for  $i \leftarrow 1$  to  $N$  do           // Determina a capacidade de cada sub-região
4      $C[i] \leftarrow U(0,1)*$  // Sorteia número aleatório entre (0,1) a partir de
        uma distribuição uniforme
5   end
6    $specie \leftarrow 1$                  // Contador do número de espécies
7    $node \leftarrow U(1,N)$            // Escolhe uma sub-região aleatória para começar
8    $sp \leftarrow specie$              // Identifica qual espécie irá colonizar
9    $habitat[node] \leftarrow sp$            // Ocupa o nó  $node$ 
10   $fitness[sp] \leftarrow fitness[sp] + C[node]$  // Atualiza o  $fitness$  da espécie
11  add to  $areas$  vizinhos do nó  $node$ 
12  while  $areas \neq \emptyset$  do
13     $node \leftarrow$  escolhe um nó dentro de  $areas$  // Escolha aleatória a partir da
        probabilidade com peso das capacidades
14     $competition \leftarrow$  Conjunto de espécies dos vizinhos de  $node$ 
15    if  $competition \neq \emptyset$  then
16       $sp \leftarrow$  escolhe espécie dentro de  $competition$  // Escolha aleatória a
        partir da probabilidade com peso dos  $fitness$ 
17    end
18     $\mu' = \frac{\mu}{fitness[sp]}$            //  $\mu$  é a probabilidade de mutação
19    if ocorre mutação  $\mu'$  then
20       $specie ++$                  // Incrementa o número de espécies
21       $sp \leftarrow specie$ 
22    end
23     $habitat[node] \leftarrow sp$        // Coloniza a sub-área  $node$  com a espécie  $sp$ 
24     $fitness[sp] \leftarrow fitness[sp] + C[node]$ 
25    add to  $areas$  vizinhos de  $node$  que não foram ainda ocupados
26  end
27 end

```

Figura 10 – Fases do Modelo I

Após a escolha do novo sítio s_{new} para a colonização, verifica-se caso os vizinhos de s_{new} já estão ocupados por alguma espécie (linha 14). Se existir mais de uma espécie como vizinho, ocorre o processo de competição. Sendo assim, é necessário escolher qual a espécie que irá colonizar. Para isso, realiza-se uma escolha aleatória (linhas 15-17), a qual cada espécie terá um peso dada pelo seu *fitness*, assim como ocorreu na escolha de s_{new} .

Em seguida, a espécie e , escolhida para colonizar s_{new} , terá uma chance de criar uma nova espécie (linhas 18-22). Essa chance é uma probabilidade $\mu' = \mu / \text{fitness}$. Caso a mutação ocorra, será criada uma nova espécie e_{new} , e ela irá colonizar o sítio s_{new} , se não, será a espécie e .

Por fim, ocupa-se o sítio com a espécie e ou e_{new} , incrementa o *fitness* da espécie, e inclui os sítios vizinhos não ocupados na lista *area* (linhas 23-25).

Nosso primeiro estudo de simulação foi realizado na rede *small-world*, a qual por variação da probabilidade de reconexão, ϕ , é possível obter um modelo totalmente regular (grade) até um aleatório. No caso específico da rede grade, considera-se o modelo com duas dimensões e a quantidade de conexões em $\langle k \rangle = 4$. Assim, utilizamos as número de nós igual à $N = 3600$, e utilizamos a probabilidade de reconexão em $\phi = \{0.0, 0.01, 0.10\}$, ver Figura 11.

Pela figura, concluímos que ao adicionarmos atalhos na rede regular, há um aumento na quantidade de espécies, já que a curva do losango ($\phi = 0.1$) está acima das demais curvas que possuem probabilidade de reconexões mais baixas. Além disso, variando-se a taxa de mutação μ , notamos que a estrutura topológica influencia pouco o modelo para valores muito baixo; menor que ~ 0.1 , ou muito alto; maior que ~ 0.7 . O que é um resultado esperado, pois, independente de como as áreas estão conectadas, dado que uma nova espécie depende de uma probabilidade $\mu' = \frac{\mu}{\text{fitness}}$. Se o valor de μ for pequeno, as novas espécies irão se propagar mais pela rede, aumentando o *fitness* e diminuindo o valor de $\mu' \ll \mu$, assim, a probabilidade de novas espécies será determinado apenas pelo acaso. Já para o caso de α grande, a criação de novas espécies ocorre mais rapidamente que a sua dispersão, consequentemente, o valor de $\mu' \gg \mu$, gerando cada vez mais espécies.

O aumento na quantidade de espécies em valores intermediários de α pode ser causado pelo atalhos criados nas redes, os quais interligam diferentes regiões. Logo, novas espécies podem se dispersar em área com menos competição, tendo mais chances de se propagarem e darem origem a outros indivíduos.

Figura 11 – Resultado das simulações do Modelo I na rede *small-world* com diferentes probabilidades de reconexão (ϕ) e mutação (μ). As figuras representam a relação espécie área em escala Log-Log. Em azul, estão as simulação para taxa de mutação $\mu = 0.1$. Em laranja, $\mu = 0.3$. Em roxo, $\mu = 0.5$. Em verde, $\mu = 0.7$. Em vermelho, $\mu = 0.9$.

Analisou-se também qual propriedade do modelo está influenciando o processo de colonização, simulamos o modelo para valores fixos de quantidade de nós e taxa de mutação, variando apenas a probabilidade de reconexão. Com isso, é possível verificar como o aparecimento de triângulos ou o efeito *small-world* estão correlacionados com o aumento do número de espécies (ver Figura 12).

Comparando a quantidade de espécies final com algumas propriedades de redes complexas, nota-se que, para o modelo Watts-Strogatz, o menor caminho médio é a característica mais acentuada que está influenciando a dinâmica da colonização. Uma vez que verificando os valores numéricos, o menor caminho médio varia entre $5.75 \leq \ell \leq 19.49$, porém o coeficiente de aglomeração $6 \times 10^{-6} \leq \langle cc \rangle \leq 1.7 \times 10^{-3}$.

Assim, pela Figura 12 é possível observar que a diminuição do menor caminho provoca na variação da quantidade de espécies, sendo inversamente proporcional até o ponto crítico em $\phi \approx 0.1$. Esse comportamento ocorre, pois com a diminuição do menor caminho, os nós ficam mais acessíveis e consequentemente a propagação e colonização de uma nova espécie para uma área sem competição é mais provável. Entretanto, se todas as áreas forem facilmente acessadas por diferentes regiões, a probabilidade de competição também aumenta, o que implica na diminuição de espécies a partir de $\phi \approx 0.1$.

Vale lembrar que, para os resultados apresentados, 1000 simulações foram realizadas e utilizou-se a média para representar a tendência central dos dados. Existem outras medidas para caracterizar um conjunto de valores, como a mediana. Entretanto, nas simulações da rede *small-world*, tanto a média quanto a mediana obtiveram resultados similares, como mostra a Figura 13.

Figura 12 – Resultado das simulações do Modelo I para verificar qual propriedade da rede *small-world* está influenciando o processo de colonização. Fixamos $N = 10000$ e $\mu = 0.3$, variamos a probabilidade de reconexão ϕ entre 0.0 e 1.0.

Nota-se que, na relação espécie-área, a curva da média, em azul, está próxima à da mediana, em vermelho. E quando observamos a distribuição de probabilidade da quantidade de espécies, conclui-se que ambas as medidas estão próximos dos valores centrais.

Entretanto, nas análises na rede livre de escala não linear, observou-se o contrário (ver Figura 14), ou seja, os valores de média e mediana são distantes, sendo maior a diferença entre eles conforme o crescimento do parâmetro α da rede. Essa disparidade é consequência da distribuição de probabilidade de espécies não ser mais simétrica e em forma de sino. Assim, para a rede livre de escala, a mediana mostrou-se como sendo o melhor critério para representar os dados obtidos das simulações.

Assim, utilizando a mediana, comparou-se as simulações do Modelo I em redes do tipo pequeno mundo de *Watts-Strogatz* (WS), livre de escala de *Barbási-Albert* (BA) e geográfica livre de escala (SFG), para diferentes taxas de mutação (ver Figura 15).

Nota-se que, como esperado, o processo de geração de espécies é influenciado pela estrutura topológica, porém ele se mostrou invariante para μ próximos de 0.1 e 0.9. Isso ocorre porque com taxas de mutação baixa, a probabilidade de ocorrência de novas espécies já é pequena para todas as redes, independente se alguma espécie consiga propagar para áreas sem competição. Já para taxas de mutação elevadas, a probabilidade de gerar novas espécie é sempre alta, dificultando a propagação de uma única espécie. Sendo isso observado na Figura 15 para

⁰ Para a distribuição de probabilidade com $\alpha = 1.5$ limitamos o eixo x até 2000 para melhor visualização.

¹ Gráficos com outras taxas de mutação podem ser observados no Apêndice.

Figura 13 – Comparação entre a média e mediana de 1000 simulações. Considerando uma rede *small-world* com 1000 vértices, taxa de mutação $\mu = 0.5$. Para cada coluna, tem-se gráficos com uma probabilidades de reconexão ϕ . Em vermelho representa-se a mediana e em azul, a média. No topo, tem-se a relação espécie-área. Embaixo, mostra-se a distribuição de probabilidade da quantidade de espécies.

Figura 14 – Comparação entre a média e mediana de 1000 simulações. Considerando uma rede não linear *scale-free* com 1000 vértices, taxa de mutação $\mu = 0.5$. Para cada coluna, tem-se um α . Em vermelho representa-se a mediana e em azul, a média. No topo, mostra-se a relação espécie-área. Embaixo, tem-se a distribuição de probabilidade da quantidade de espécies¹.

$\mu = 0.9$, a qual a quantidade de espécies está próxima da quantidade de área disponível para ser ocupada.

Figura 15 – Gráfico de comparação do Modelo I da relação espécie-área para diferentes topologia e taxas de mutação μ .

Figura 16 – Gráfico em escala de cor para o valor de z estimado pelo método dos mínimos quadrados da curva $S = cA^z$. Para cada linha é indicado um tipo de rede, e para cada coluna uma variação da taxa de mutação μ . Valores pequenos de z se aproximam da cor azul, e valores alto da cor vermelha.

Na Figura 16 temos as estimativas do valor z da curva espécie-área $S = cA^z$. Nota-se que na primeira e última coluna há pouca variação no valor de z , o que já foi discutido anteriormente quando $\mu = 0.1$ e $\mu = 0.9$. Entretanto para outras taxas de mutação μ , tem-se que, em geral, as redes *small-world* possuem uma quantidade maior de espécies e consequentemente também o seu valor de z , comparado com as outras redes.

Isso pode ser consequência da distribuição heterogênea do grau dos nós, uma vez que os *hubs*, apesar de diminuïrem o menor caminho, eles conectam com quase todas as regiões da rede. Logo, as espécies que ocuparem esse nós centrais, se propagam por várias áreas, diminuindo a probabilidade de geração de novas espécies. Esse caso pode ser observado principalmente para a rede Barabási-Albert quando $\alpha = 1.5$, já que os valores de z são próximos zero para taxas de mutação menores que 0.7.

Além disso, para algumas redes, as curvas espécie-área se intersectam ou sobrepõem, como é o caso da rede BA com $\alpha = 0.5$ e a rede WS com $\phi = 0.0$, ou a rede BA com $\alpha = 1.0$ e a rede SFG de $r_c = 1/20$. Entretanto vale ressaltar que a construção e topologia dessas redes são distintas, resultando em comportamentos diferentes para a dinâmica de propagação das espécies.

Assim, na Figura 17 tem-se em representação matricial o resultado de uma das simulações nos três tipos de redes. Observa-se que, nas redes *small-world*, áreas com as mesmas espécies ocorrem em grupos, formando regiões com determinadas espécies. No caso, quando $\phi = 0.0$, as áreas estão interligadas somente com 4 vizinhos de forma a formar uma topologia de grade. Conforme ϕ aumenta, atalhos são criados na rede, possibilitando que uma espécie colonize áreas mais distantes, como pode ser analisando para $\phi = 0.1$.

Para a rede livre de escala Barabási-Albert a colonização das espécies ocorreu de forma mais dispersa, com uma abundância maior de uma única espécie conforme α aumentava. Como já discutido anteriormente, esse comportamento ocorre devido a existência de *hubs*, os quais possibilitam que uma espécie se propague para várias áreas sem competição, aumentando a probabilidade de colonização. No caso quando $\alpha = 1.5$, isso é mais evidente, visto que pela área escura é possível notar que as ligações se concentram apenas em uma região.

Por fim, as redes geográficas livre de escala obtiveram ambas as características de regiões e dispersão de espécies. Porém, esse comportamento é diretamente relacionado ao tamanho do raio crítico r_c escolhido. Dessa forma, quanto maior o r_c maior a dispersão, por outro lado quanto menor for r_c maior a tendência das espécies se concentrarem em regiões distintas.

Para melhor visualização da colonização nas diferentes topologias tem-se a Figura 20, a qual mostra-se as redes da Figura 17 em outra representação. Com isso, observa-se que uma rede *small-world* limita-se a ligações locais, as quais limitam a dispersão de espécies para outras áreas.

Figura 17 – Representação matricial da colonização de espécies para diferentes topologias, sendo taxa de mutação $\mu = 0.3$ e área total $L \times L = 625$. Sendo (a), (b) e (c) redes *small-world*; (d), (e) e (f) redes livre de escala Barabási-Albert; (g), (h) e (i) redes geográficas livres de escala. Os círculos representam as áreas, as arestas são as ligações e cada cor significa uma espécie.

Entretanto para redes livre de escala Barabási-Albert, não há formação de regiões distintas, favorecendo espécies que ocupam nós com muitas ligações para diferentes áreas. E no caso da rede geográfica livre de escala, encontra-se um modelo o qual possui as características de ligações de curtas e longas distâncias, resultando em uma dinâmica mais realista.

Figura 18 – Representação das redes com a mesma dinâmica e parâmetros da Figura 17.

De fato esse resultado pode ser verificado pela Figura 19, a qual compara-se as curvas espécie-área dos três tipos de redes com dados reais de espécies de peixes localizados em 70 lagos ao redor do mundo (BARBOUR, 1974). Porém, para a análise da topologia da rede é necessário fixar uma taxa de mutação, no caso, utilizou-se $\mu = 0.3$ (ver Figura 20). Verificou-se que as redes que se aproximaram do valor estimado de $z = 0.15$ para o dado real (STEIN; JURITZ, 1988), foram as redes livre de escala geográfica, as quais são as curvas G e H que obtiveram valores de $z = 0.16$ e $z = 0.14$, respectivamente.

Figura 19 – Dados reais da curva espécie-área (BARBOUR, 1974) e curvas estimadas a partir dos resultados obtidos para taxa de mutação $\mu = 0.3$ dos diferentes tipos de redes. Cada curva estimada a partir das redes é indicado por uma letra e linhas tracejadas. A linha azul é a curva original estimada (STEIN; JURITZ, 1988) com $z = 0.15$. E em vermelho tem-se as duas curvas que mais se aproximam da original.

Figura 20 – Curvas obtidas para diferentes valores de μ .

MODELO BASEADO EM AGENTES

5.1 Modelo II

Apesar do modelo anterior apresentar poucas características estocásticas que facilitam o seu entendimento, há uma limitação quando se considera somente populações de espécies, pois ignora-se o comportamento de cada indivíduo.

Dessa forma, para completar as análises, consideramos também o modelo publicado em (AGUIAR *et al.*, 2009), o qual possui como principais particularidades as características individuais e distâncias espaciais. Vale ressaltar que para esse caso não se utilizou a representação por meio de redes complexas, ao contrário disso, variações na dinâmica do modelo foram feitas para o estudo da influência topológica.

Assim, o modelo original possui as definições para cada indivíduo i :

- Raio de acasalamento S .
- Genótipo $g^i = (\sigma_1^i, \sigma_2^i, \dots, \sigma_B^i)$, com B o número de *loci* no genótipo e cada σ_k^i possui o valor binário, ou seja, 1 ou 0.
- Realiza-se uma única reprodução, sendo que nela, o indivíduo é substituído pelo seu descendente.
- Possui uma probabilidade Q de não se reproduzir e com isso, outro indivíduo é escolhido para reproduzir.
- Para reproduzir é necessário também que a distância genética seja menor que o valor definido G .
- Cada espécie pode se dispersar para um dos 20 sítios adjacentes com uma probabilidade D .

Vale ressaltar que dois organismos só geraram novas espécies se as condições espaciais, como o raio de acasalamento S , e a distância genética, limitada pelo o valor crítico G , forem respeitadas. Sendo que para esses requisitos foram utilizados a distância euclidiana, um vez que

cada indivíduo possui uma posição (x, y) no espaço,

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (5.1)$$

e a distância de Hamming, pois a genética de cada um é representada por um vetor de 0s e 1s:

$$dH_{ij} = \sum_{k=1}^B |\sigma_i^k - \sigma_j^k|. \quad (5.2)$$

Assim, a partir de uma área total $(L \times L)$ com n indivíduos geneticamente idênticos e distribuídos de forma aleatória, inicia-se a dinâmica do modelo:

- Processo de reprodução para cada indivíduo i :
 1. Lista-se todos os indivíduos possíveis i_p dentro da sua vizinhança S .
 2. Dentre i_p , seleciona-se apenas aquelas com a distância genética menor ou igual a G , o qual resultada em um nova lista de indivíduo $i_{p,G}$.
 - Caso a quantidade de possíveis parceiros j seja menor que P , aumenta-se a distância genética e espacial até atingir o número mínimo P .
 3. Em $i_{p,G}$, escolhe-se aleatoriamente um indivíduo j para a reprodução e geração do descendente i' .
 - a) O indivíduo i' terá o genoma $g_{i'}$ gerado a partir da recombinação de g_i e g_j . Para isso, escolhe-se algum *loci* aleatório k para permutação.

$$\begin{aligned} g^a &= (\sigma_i^1, \sigma_i^2, \dots, \underline{\sigma_j^k}, \dots, \underline{\sigma_j^B}). \\ g^b &= (\sigma_j^1, \sigma_j^2, \dots, \underline{\sigma_i^k}, \dots, \underline{\sigma_i^B}). \end{aligned} \quad (5.3)$$
 - b) Escolhe-se aleatoriamente entre g^a ou g^b para ser o genoma de i' .
 - c) Calcula-se a probabilidade de mutação em cada loci do genoma de i' .
 4. Com uma probabilidade D , i' pode se dispersar para os 20 sítios mais próximos.
 5. O descendente i' substituirá o indivíduo i na próxima geração.
 6. Início de uma nova geração.

Vale ressaltar que é possível existirem dois indivíduos A e B da mesma espécie, porém com distância genética maior que o valor crítico G . Isso ocorre quando há um terceiro indivíduo C , o qual sua distância genética satisfaz $dH_{A,C}, dH_{B,C} < G$. Com isso, o material genético de A pode ser transmitido à B , de forma indireta, por meio do indivíduo C , aumentando a variabilidade genética dentro da mesma espécie.

O algoritmo que implementa o processo de evolução dos indivíduos é descrito a seguir.

Algoritmo 2: Dinâmica de Evolução**Data:** Dado um espaço bidimensional ($L \times L$) e N indivíduos

```

1 begin
2   origGenes[1..B]  $\leftarrow U(0,1)$            // Define o vetor original de genes
4   for  $i \leftarrow 1$  to  $N$  do                 // Define a posição de cada indivíduo
5     Individuo ind
6     ind.X  $\leftarrow func(0,L)$            // func pode ser aleatório ou outra função
7     ind.Y  $\leftarrow func(0,L)$ 
8     ind.gene  $\leftarrow origGene$            // Inicia todos com os mesmos genes
9     newPop[ $i$ ]  $\leftarrow ind$                // Pop é uma lista com todos os indivíduos
10  end
12  for  $i \leftarrow 1$  to maxCycles do         // para cada ciclo de evolução
13    pop  $\leftarrow NewPop$                    // Define a população atual
14    newPop  $\leftarrow empty$                // Esvazia a nova população
15    while pop is not empty do
16      randInd  $\leftarrow$  seleciona um Indivíduo de pop           // Escolhe de forma
        aleatória
17      pop.remove(randInd)                 // Remove da lista de pop
18      if não se reproduz then             // Probabilidade de reprodução  $Q$ 
19        mVecQ[]  $\leftarrow$  seleciona parceiros (randInd,  $S$ ,  $G$ , pop,  $P$ ) // No mínimo  $P$ 
20        randInd  $\leftarrow$  seleciona um parceiro de mVecQ
21        if randInd is in pop then         // Se randInd ainda não se reproduziu
22          pop.remove(randInd)
23          newPop  $\leftarrow randInd$          // Adiciona randInd na nova população
        // Necessário para manter a quantidade de indivíduos
24        end
        // Parceiro substituirá o Indivíduo escolhido na linha 16
25      end
26      mVec[]  $\leftarrow$  seleciona possíveis parceiros (randInd,  $S$ ,  $G$ , pop,  $P$ )
27      randM  $\leftarrow$  seleciona um parceiro em mVec           // Escolha aleatória
28      newGenes  $\leftarrow$  constrói novo vetor de genes (randInd, randM)
29      newInd  $\leftarrow$  cria novo Indivíduo (newGene, randInd.X, randInd.Y,  $D$ )
        // Probabilidade  $D$  do novo Indivíduo dispersar entre os 20
        sítios mais próximos
30      newPop  $\leftarrow newInd$            // Adiciona novo Indivíduo na nova população
31    end
32  end
33 end

```

Como descrito anteriormente, o modelo original inicializa aleatoriamente os indivíduos dentro do seu espaço $L \times L$. Entretanto, sabe-se que existem outras topologias na natureza que contribuem para o funcionamento e equilíbrio do ecossistema. Dessa forma, considerou-se duas variações no modelo para a simulação de outros ambientes, uma utilizando a lei de potencia para distribuição dos indivíduos e outra a divisão de sub-areas.

Assim, o algoritmo original foi modificado para ter os seguintes passos iniciais:

Algoritmo 3: Estrutura do Indivíduo

```

1 begin
2   struct {
3     genes[1..B]           // Vetor binário de tamanho B
4     X                       // Coordinate X
5     Y                       // Coordinate Y
6   } Individuo
7 end

```

- Define-se:
 - a quantidade total de indivíduos N
 - o expoente da distribuição da lei de potência x
 - o raio da sub-área r
 - o espaço $L \times L$
- Enquanto a quantidade atual de indivíduos for menor que a definida N , popula-se uma nova sub-área de raio r :
 1. Escolhe-se de forma aleatória a localização do centro da sub-área no espaço $L \times L$
 2. Seleciona aleatoriamente a quantidade k de indivíduos a partir da distribuição de lei de potência, limitando-se no máximo à 1000 indivíduos por sub-área.
 3. Distribui-se k indivíduos aleatoriamente dentro da sub-área
 4. Se atingir o número total de indivíduos, a iteração termina
- Após os N indivíduos estiverem distribuídos no espaço $L \times L$, inicia-se o algoritmo de Dinâmica de Evolução

Para variar a estrutura das conexões, consideramos diferentes distribuições de espécies no espaço, ora variando-se o raio das sub-áreas e ora o expoente da distribuição lei de potência, isto é, o tamanho das regiões que as espécies são dispostas são definidas de acordo uma distribuição de Pareto, $P(x) \approx x^\beta$. Essa metodologia gera uma configuração com dimensão fractal β . Entretanto, baseados na literatura (AGUIAR *et al.*, 2009), alguns parâmetros, como a quantidade de indivíduos, o número de *loci* no genótipo, $B = 20$ e raio de acasamento, $S = 5$, foram adotados. Assim, as configurações dos espaços que o modelo foi simulado são ilustradas nas Figuras 21 e 22.

A Figura 21 mostra seis topologias variando-se somente o raio da sub-área. Nota-se que para raios pequenos, essa variação produz ilhas ou regiões distantes uma das outras. Porém, para raios próximos ao tamanho da área, essa alteração influencia pouco na dispersão dos indivíduos. Já na Figura 22, a qual modifica-se o parâmetro da lei de potência, observa-se que quanto menor o expoente, maior a ocorrência de populações menores, resultando em distribuições mais homogêneas, por outro lado, quanto maior o expoente, maior é a chance de aparecerem populações mais densas, consequentemente, mais heterogênea é a disposição dos indivíduos.

Figura 21 – Comparação entre os diferentes tipos de raio das sub-áreas na iteração = 1, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e o expoente da lei de potência $x = 0$

Para os casos descritos anteriormente são apresentados suas configurações finais após 4000 gerações nas Figuras 23 e Figura 24. Observa-se que as populações distantes inicialmente, podem se aproximar após algumas iterações devido ao deslocamento de cada indivíduo, possibilitando trocas de materiais genético que ajudam na evolução da espécie. Porém, algumas vezes, essa dispersão não é suficiente e com isso, ocorre a concentração de determinadas espécies em regiões específicas, por exemplo nas Figuras 23(a) e 24(a).

Por outro lado, quando as distâncias entre as regiões são menores que a distância de acasalamento; ou a densidade de indivíduos é pequena o suficiente para deixar a distribuição mais homogênea, a geração de espécies sofre maior influência das características do modelo, como ocorreu nas Figuras 23(d,e,f) e 24(c,d).

Durante a transição entre esses dois extremos, há a influência tanto das particularidades do modelo quanto da disposição dos indivíduos no espaço $L \times L$. É válido ressaltar que a cada iteração do algoritmo, os indivíduos podem se dispersar entre os 20 sítios mais próximos, com isso, a configuração espacial final é diferente da inicial, como mostra a Figura 23 e 24. Isso possibilita que o material genético de novas espécies possam propagar para novas áreas e indivíduos, diminuindo as suas distâncias genéticas ou permitindo a geração de novas espécies.

Na Figura 25 mostra-se o resultado de comparação das curvas espécie-área para

Figura 22 – Comparação entre os diferentes tipos de expoentes da lei de potência ($P(x) \approx x^\beta$) para a distribuição do raio das regiões, fixando-se a área 256x256, a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$

diferentes valores do expoente da lei de potência, os quais influenciam na distribuição de indivíduos em cada sub-área.

Inicialmente, é possível observar que as curvas se sobrepõem para áreas menores e, em seguida, cada valor do expoente segue seu próprio crescimento. Isso indica que existe uma área mínima para a característica de dispersão influenciar na evolução das espécies de acordo com o raio da sub-área escolhido. Esse resultado era esperado, uma vez que devido ao número fixo de indivíduos, a densidade média para áreas menores é superior ou próximo à 1, com isso, os indivíduos ficam mais próximos fisicamente possibilitando que o mesmo material genético seja compartilhado com maior probabilidade.

Ademais, considerando a variação da lei de potência, nota-se que o número de espécies aumenta quando o valor do expoente diminui, ou seja, quando a distribuição dos indivíduos é mais homogênea. Esse comportamento acontece, porque em distribuições heterogêneas, existem regiões específicas super concentradas, com isso as distâncias espaciais e genéticas, nesses locais, são pequenas e, conseqüentemente, a quantidade de espécies é limitada pela quantidade de sub-áreas geradas. Entretanto, quando os indivíduos estão dispersos, é possível que novos materiais genéticos se propaguem para uma quantidade limitada de vizinhos, criando possibilidades para

Figura 23 – Comparação entre os diferentes tipos de raio das sub-áreas na iteração = 4000, fixando-se a área 256×256 , a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e o expoente da lei de potência $x = 0$

novas espécies que não dependem da sua disposição espacial.

Na Figura 26, o tamanho da área foi fixado em 256×256 e variou-se o raio das sub-áreas para diferentes expoentes. Assim, similar a análise anterior, verificou-se também que expoentes menores tendem a gerar mais espécies e, que o raio da sub-área influencia a evolução das espécies dentro de um intervalo, o qual é limitado pelo tamanho da área total.

Além disso, nota-se que conforme o raio da sub-área aumenta, a quantidade de espécies, inicialmente, diminui e em seguida cresce. Esse comportamento ocorre porque para sub-áreas com tamanho próximo ao raio de acasalamento, as espécies se restringem à essas regiões. Já com sub-áreas maiores, há maior possibilidade de trocas de materiais genéticos, as quais diminuem as distâncias genética entre os indivíduos e consequentemente, o número de espécies. Todavia, é válido ressaltar que o crescimento da curva ocorre devido à própria interação entre o modelo e essa topologia.

Por último, a Figura 27 ilustra a quantidade de espécies para cada iteração do modelo. Nota-se que para sub-áreas próximas ao tamanho da área a topologia possui pouca influência sobre o modelo, uma vez que o comportamento das curvas foi similar nas 4 variações do expoente.

Figura 24 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração = 4000, fixando-se a área 256×256 , a quantidade de indivíduos $N = 2000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$

Figura 25 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração = 4000, variando-se a área e fixando-se a quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$

Notamos ainda que em expoentes menores, as curvas são mais acentuadas, indicando que, para esses casos, as interações foram mais intensas e com isso, houve maior diversidade genética

Figura 26 – Comparação entre os diferentes tipos de expoentes da lei de potência na iteração 4000, variando-se o raio das sub-áreas (r) e fixando-se a área 256×256 , quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$.

e geração de novas espécies. Esse resultado era previsto, já que em distribuições homogêneas, os indivíduos ficam mais próximos, aumentando a chance de transmissão de materiais genéticos e, ocorrência de acasalamento para geração de novas espécies. Esses resultados mostram que a distribuição da espécies no espaço afeta fundamentalmente o número de espécies.

Figura 27 – Comparação entre os diferentes tipos de expoentes da lei de potência, fixando-se a área 256x256, a quantidade de indivíduos $N = 8000$, raio de acasalamento $S = 5$ e raio das sub-áreas $r = 16$

CONCLUSÃO

O estudo da biodiversidade e especiação é uma das áreas mais interdisciplinares na ecologia (SOLE; BASCOMPTE, 2006). Na matemática de modelagem e predição, por exemplo, os pesquisadores utilizam modelos matemáticos para incluir comportamentos observados na natureza, como mutação, dispersão, competição. Apesar dessas características individuais serem importantes, a maioria dos estudos não consideram a influência de fatores externos durante as simulações de propagação das espécies, limitando as discussões sobre o assunto. De fato, como abordado em (HASTINGS *et al.*, 2005; PETERSON *et al.*, 2013), simplificar o habitat ou conexões entre os indivíduos, pode impossibilitar o aparecimento de fenômenos complexos e consequentemente, não representar as interações reais. Dessa forma, o objetivo desse trabalho foi analisar na evolução de espécies, a influência de novos comportamentos que proporcionassem diferentes conexões entre os indivíduos.

Inicialmente, utilizou-se o modelo proposto em (CONCEIÇÃO; OLIVEIRA, 2010), que possui três características principais: mutação, dispersão e competição, entretanto, o artigo considerou apenas áreas regulares, ou seja, redes do tipo grade. Assim, conforme discutido no Capítulo 4, modificou-se o algoritmo para incluir outras formas de conexões, no caso, utilizando como base as redes *Watts-Strogatz*, *Barabási-Albert* e redes geográficas livre de escala. Para esse primeiro modelo, os resultados sugerem que não há diferença na diversidade quando as taxas de mutação (μ) são extremas, ou seja, os valores estão $\mu \gg 0.9$ e $0.1 \ll \mu$. Além disso, observou-se que as redes de *Watts-Strogatz* com probabilidade $\phi = 0.1$ obteve a maior quantidade de espécies. Enquanto que as redes heterogêneas, como a *Barabási-Albert* e as redes geográficas livre de escala, a diversidade de espécies foi menor.

Essa variação ocorreu porque, em redes homogêneas, a distância entre as regiões é grande e, consequentemente, é mais difícil de uma espécie se propagar. Já em redes heterogêneas existem sub-áreas altamente conectados que facilitam a dispersão das espécies. Assim, quanto maior o espaço de colonização, menores são as chances de mutação e o surgimento de novas espécies.

Para completar as análises, utilizou-se também o modelo publicado em (AGUIAR *et al.*, 2009), o qual possui como particularidades, as características de genética, *crossover* e distância máxima para acasalamento. No Capítulo 5, descreveu-se o modelo original e, as alterações realizadas no posicionamento dos indivíduos para verificar a influência da topologia.

O primeiro resultado obtido foi que é necessário uma área mínima para a disposição inicial influenciar na evolução das espécies, uma vez que fixado o raio da sub-área e, utilizando diferentes expoentes na função de distribuição dos indivíduos, as curvas se sobreporam até um valor crítico e em seguida cada uma seguia seu próprio crescimento.

Ademais, observou-se que quanto mais homogênea a distribuição dos indivíduos, maior é a diversidade de espécies, sendo o resultado válido tanto para o caso de expoentes menores quanto para raios de sub-áreas grandes. Por fim, constatou-se que em expoentes menores, o crescimento da diversidade ocorreu de maneira mais acentuada. Porém, não estavam relacionados diretamente com o crescimento na diversidade. O que era esperado, já que sub-área mais densas proporcionam maior interação entre os indivíduos dentro desse raio, porém, devido à heterogeneidade, há pouco relacionamento entre essas regiões.

REFERÊNCIAS

AGUIAR, M. A. M. de; BARANGER, M.; BAPTESTINI, E. M.; KAUFMAN, L.; BARYAM, Y. Global patterns of speciation and diversity. **Nature**, Macmillan Publishers Limited. All rights reserved, v. 460, n. 7253, p. 384–387, Jul 2009. ISSN 0028-0836. Disponível em: <<http://dx.doi.org/10.1038/nature08168>>. Citado 6 vezes nas páginas 16, 17, 29, 43, 46 e 54.

ALBERT, R.; JEONG, H.; BARABÁSI, A. Internet: Diameter of the world-wide web. **Nature**, Nature Publishing Group, v. 401, n. 6749, p. 130–131, 1999. Citado na página 25.

ARRHENIUS, O. Species and area. **a. Journal of Ecology**, v. 9, p. 95–99, 1921. Citado 3 vezes nas páginas 15, 27 e 28.

AZAELE, S.; SUWEIS, S.; GRILLI, J.; VOLKOV, I.; BANAVAR, J. R.; MARITAN, A. Statistical mechanics of ecological systems: Neutral theory and beyond. **Reviews of Modern Physics**, APS, v. 88, n. 3, p. 035003, 2016. Citado na página 16.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p. 509–512, 1999. Disponível em: <<http://www.sciencemag.org/content/286/5439/509.abstract>>. Citado na página 24.

BARABÁSI, A.-L.; ALBERT, R.; JEONG, H. Scale-free characteristics of random networks: the topology of the world-wide web. **Physica A: Statistical Mechanics and its Applications**, v. 281, n. 1-4, p. 69–77, jun. 2000. Citado na página 25.

BARBOUR, J. H. B. C. D. Fish species diversity in lakes. **The American Naturalist**, [University of Chicago Press, American Society of Naturalists], v. 108, n. 962, p. 473–489, 1974. ISSN 00030147, 15375323. Disponível em: <<http://www.jstor.org/stable/2459679>>. Citado 3 vezes nas páginas 11, 39 e 40.

BARTHÉLEMY, M. Crossover from scale-free to spatial networks. **EPL (Europhysics Letters)**, v. 63, n. 6, p. 915, 2003. Disponível em: <<http://stacks.iop.org/0295-5075/63/i=6/a=915>>. Citado na página 26.

COLLEY, E.; FISCHER, M. L. Espéciação e seus mecanismos: histórico conceitual e avanços recentes. **História, Ciências, Saúde-Manguinhos**, scielo, v. 20, p. 1671 – 1694, 12 2013. ISSN 0104-5970. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-59702013000401671&nrm=iso>. Citado na página 15.

CONCEIÇÃO, K. S.; OLIVEIRA, V. M. de. Habitat fragmentation effects on biodiversity patterns. **Physica A: Statistical Mechanics and its Applications**, v. 389, n. 17, p. 3496 – 3502, 2010. ISSN 0378-4371. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0378437110003754>>. Citado 5 vezes nas páginas 16, 17, 29, 30 e 53.

CONCEIÇÃO, K. S.; ULRICH, W.; DINIZ, C. A. R.; RODRIGUES, F. A.; ANDRADE, M. G. d. A generalized approach to the modeling of the species-area relationship. **PLoS One**, v. 9, n. 8, 2014. Citado na página 15.

- CONNOR, E. F.; MCCOY, E. D. The statistics and biology of the species-area relationship. **The American Naturalist**, The University of Chicago Press for The American Society of Naturalists, v. 113, n. 6, p. pp. 791–833, 1979. ISSN 00030147. Disponível em: <<http://www.jstor.org/stable/2460305>>. Citado na página 27.
- COSTA, L. d. F.; JR, O. N. O.; TRAVIESO, G.; RODRIGUES, F. A.; BOAS, P. R. V.; ANTIQUEIRA, L.; VIANA, M. P.; ROCHA, L. E. C. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, Taylor & Francis, v. 60, n. 3, p. 329–412, 2011. Citado na página 16.
- CRAWLEY, M. J. **The Structure of Plant Communities**. Blackwell Publishing Ltd., 1997. 475–531 p. ISBN 9781444313642. Disponível em: <<http://dx.doi.org/10.1002/9781444313642.ch14>>. Citado na página 15.
- ERDŐS, P.; RÉNYI, A. On random graphs. **Publicationes Mathematicae Debrecen**, v. 6, p. 290–297, 1959. Citado na página 22.
- _____. b. **Publ. Math. Inst. Hungar. Acad. Sci.**, v. 5, p. 17–61, 1960. Citado na página 22.
- _____. On the strenght of connectedness of a random graph. **Acta Mathematica Scientia Hungary**, v. 12, p. 261–267, 1961. Citado na página 22.
- FREEMAN, L. C. A set of measures of centrality based on betweenness. **Sociometry**, American Sociological Association, v. 40, n. 1, p. pp. 35–41, 1977. ISSN 00380431. Disponível em: <<http://www.jstor.org/stable/3033543>>. Citado na página 21.
- GRAS, R.; DEVAURS, D.; WOZNIAK, A.; ASPINALL, A. An individual-based evolving predator-prey ecosystem simulation using a fuzzy cognitive map as the behavior model. **Artificial Life**, v. 15, n. 4, p. 423–463, 2009. Citado 3 vezes nas páginas 15, 16 e 27.
- HASTINGS, A.; CUDDINGTON, K.; DAVIES, K. F.; DUGAW, C. J.; ELMENDORF, S.; FREESTONE, A.; HARRISON, S.; HOLLAND, M.; LAMBRINOS, J.; MALVADKAR, U.; MELBOURNE, B. A.; MOORE, K.; TAYLOR, C.; THOMSON, D. The spatial spread of invasions: new developments in theory and evidence. **Ecology Letters**, Blackwell Science Ltd, v. 8, n. 1, p. 91–101, 2005. ISSN 1461-0248. Disponível em: <<http://dx.doi.org/10.1111/j.1461-0248.2004.00687.x>>. Citado na página 53.
- HOLT, R. D.; LAWTON, J. H.; ET. Trophic rank and the species-area relationship. **Ecology**, v. 80, n. 5, p. 1495–1504, 1999. Citado na página 15.
- HORNER-DEVINE, M. C.; LAGE, M.; HUGHES, J. B.; BOHANNAN, B. J. M. A taxa-area relationship for bacteria. **Nature**, v. 432, n. 7018, p. 750–753, Dec 2004. ISSN 0028-0836. Disponível em: <<http://dx.doi.org/10.1038/nature03073>>. Citado na página 27.
- HUBBELL, S. **The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)**. Princeton University Press, 2001. (Monographs in Population Biology). ISBN 9780691021287. Disponível em: <<https://books.google.com.br/books?id=EIQpFBu84NoC>>. Citado 2 vezes nas páginas 10 e 28.
- HUBBELL, S. P. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. **Coral Reefs**, Springer-Verlag, v. 16, n. 1, p. S9–S21, 1997. ISSN 0722-4028. Disponível em: <<http://dx.doi.org/10.1007/s003380050237>>. Citado na página 28.

KITZES, J.; HARTE, J. Beyond the species–area relationship: improving macroecological extinction estimates. **Methods in Ecology and Evolution**, v. 5, n. 1, p. 1–8, 2014. ISSN 2041-210X. Citado na página 15.

KRAPIVSKY, P. L.; REDNER, S.; LEYVRAZ, F. Connectivity of growing random networks. **Phys. Rev. Lett.**, American Physical Society, v. 85, p. 4629–4632, Nov 2000. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.85.4629>>. Citado na página 25.

MASHAYEKH, M.; MACPHERSON, B.; GRAS, R. Species–area relationship and a tentative interpretation of the function coefficients in an ecosystem simulation. **Ecological Complexity**, v. 19, p. 84 – 95, 2014. ISSN 1476-945X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1476945X14000646>>. Citado 2 vezes nas páginas 15 e 27.

MAYR, E. **What evolution is**. New York: Basic Books, 2001. Citado na página 15.

MONTOYA, J. M.; SOLÉ, R. V. Small world patterns in food webs. **Journal of Theoretical Biology**, v. 214, n. 3, p. 405 – 412, 2002. ISSN 0022-5193. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022519301924609>>. Citado na página 16.

NEWMAN, M. E. J. The structure and function of complex networks. **SIAM Review**, v. 45, n. 2, p. 167–256, 2003. Disponível em: <<http://dx.doi.org/10.1137/S003614450342480>>. Citado na página 16.

NEWMAN, M. E. J.; WATTS, D. J. Renormalization group analysis of the small-world network model. **Physics Letters A**, North-Holland, v. 263, p. 341–346, 1999. Citado na página 23.

PETERSON, E. E.; HOEF, J. M. V.; ISAAK, D. J.; FALKE, J. A.; FORTIN, M.-J.; JORDAN, C. E.; MCNYSET, K.; MONESTIEZ, P.; RUESCH, A. S.; SENGUPTA, A.; SOM, N.; STEEL, E. A.; THEOBALD, D. M.; TORGENSEN, C. E.; WENGER, S. J. Modelling dendritic ecological networks in space: an integrated network perspective. **Ecology Letters**, v. 16, n. 5, p. 707–719, 2013. ISSN 1461-0248. Disponível em: <<http://dx.doi.org/10.1111/ele.12084>>. Citado na página 53.

REDNER, S. How popular is your paper? an empirical study of the citation distribution. **The European Physical Journal B - Condensed Matter and Complex Systems**, EDP Sciences, v. 4, n. 2, p. 131–134, 1998. ISSN 1434-6028. Disponível em: <<http://dx.doi.org/10.1007/s100510050359>>. Citado na página 25.

ROSENZWEIG, M. L. Book. **Species diversity in space and time** / Michael L. Rosenzweig. Cambridge University Press Cambridge ; New York, 1995. ISBN 0521496187 0521499526. Disponível em: <<http://www.loc.gov/catdir/toc/cam025/94043750.html>>. Citado 2 vezes nas páginas 15 e 27.

SOLE, R. C.; BASCOMPTE, J. **Self-Organization in Complex Ecosystems**. [S.l.]: Princeton University Press, 2006. Citado na página 53.

STEIN, G. Z.; JURITZ, J. M. Linear models with an inverse gaussian poisson error distribution. **Communications in Statistics - Theory and Methods**, v. 17, n. 2, p. 557–571, 1988. Disponível em: <<http://dx.doi.org/10.1080/03610928808829640>>. Citado 3 vezes nas páginas 11, 39 e 40.

STORCH, D.; KEIL, P.; JETZ, W. Universal species-area and endemics-area relationships at continental scales. **Nature**, v. 488, p. 78 – 81, 2012. Disponível em: <<http://www.nature.com/nature/journal/v488/n7409/full/nature11226.html>>. Citado na página 27.

TJORVE, E. Shapes and functions of species-area curves (ii): A review of new models and parameterizations. **Journal of Biogeography**, v. 36, n. 8, p. 1435–1445, 2009. Citado na página 15.

WATSON, H. C. **Cybele Britannica : or British Plants and their geographical relations**. London: Longman, 1847–59. Citado 2 vezes nas páginas 15 e 27.

WATTS, D.; STROGATZ, S. Collective dynamics of 'small-world' networks. **Nature**, p. 440–442, 1998. Citado 2 vezes nas páginas 21 e 22.