



CRIAÇÃO DE RELAÇÕES FUZZY EM REDES PLANCTÔNICAS POR EVOLUÇÃO DIFERENCIAL

Rogério Pinto Espíndola

rpespindola@uezo.rj.gov.br

UEZO, State University of West Side

Av. Manuel Caldeira de Alvarenga, 1203, bloco II, sala 201, 23070-200, Rio de Janeiro, RJ, Brazil

Lúcio Pereira de Andrade

lucio@vm.uff.br

UFF, Federal Fluminense University

R. Desembargador Ellis Hermydio Figueira, 783, 27213-145, Volta Redonda, RJ, Brazil

Nelson Francisco Favilla Ebecken

nelson@ntt.ufrrj.br

Federal University of Rio de Janeiro (COPPE)

Cidade Universitária, CT Bloco B Sala 100, 21945-970, Rio de Janeiro, RJ, Brazil

Abstract. *Os dados do plâncton marinho, obtidos através da citometria de fluxo, são disponibilizados em larga escala com riqueza de detalhes. Entretanto, as relações existentes entre os organismos detectados não são capturadas pelo citômetro e as variáveis que agem sobre a realidade para compô-las são desconhecidas. Assim, determinar esses relacionamentos requer um grande esforço e é crucial para o estudo das interações de populações planctônicas. Essas interações são inerentemente fuzzy e definir os conjuntos fuzzy que as representem, guardando um sentido ecológico, é o principal desafio da modelagem matemática do comportamento de redes planctônicas reais. Neste trabalho, parâmetros de funções de pertinência fuzzy definidas por partes lineares são otimizados por evolução diferencial de acordo com as características desejadas para elas. Nos experimentos realizados, procurou-se modelar a relação de competição entre organismos e algumas métricas de avaliação de redes estimaram o comportamento da rede final produzida, o que permitiu a comparação com os valores de outras redes encontradas na literatura.*

Keywords: *Complex networks, fuzzy sets, differential evolution, plankton networks*

1 INTRODUÇÃO

As diversas interações que constantemente ocorrem na natureza formam uma vasta e complexa rede, cuja importância tem exigido grandes esforços para estudá-las e compreendê-las. Relações ecológicas entre as diferentes espécies de seres vivos fazem parte do equilíbrio natural que envolve todos no planeta. Os impactos negativos causados por distúrbios naturais e, principalmente, humanos são alvo de preocupação, pesquisa e análise, a fim de modelar, prever e evitar que essas ações tragam consequências piores. A modelagem por redes têm o potencial de transformar um grande sistema em uma poderosa estrutura abstrata reduzida capaz de capturar o básico dentro de um emaranhado de conexões. Na Biologia, as redes são amplamente utilizadas e constituem uma representação conveniente das muitas formas de interação entre os elementos biológicos, tornando-se uma ferramenta importante para a compreensão e quantificação de fenômenos ecológicos (Newman, 2010).

O plâncton é um elemento-chave em todos os ecossistemas marinhos (Odum, 2001). Malkassian (Malkassian *et al.*, 2011) alerta que é necessário obter informações precisas para entender a complexa dinâmica das comunidades planctônicas e seus parâmetros de controle. A abundância e a distribuição do plâncton em um ecossistema dependem intensamente de fatores como a concentração de nutrientes e o estado físico da coluna de água (Agouru & Audu, 2012). Com os recentes desenvolvimentos tecnológicos, uma das principais ferramentas para a captura e a análise de informações sobre o plâncton é o citômetro de fluxo, que é capaz de realizar medições de alta frequência em partículas individuais (Malkassian *et al.*, 2011), produzindo suas assinaturas óticas. Portanto, analisar esse grande volume de dados gerados tem se apresentado como um grande desafio científico, exigindo o desenvolvimento de sistemas computacionais inteligentes.

Neste trabalho, uma metodologia para criar redes planctônicas a partir de dados fornecidos por um citômetro de fluxo é apresentada. Nela, a relação de competição entre os atores da rede é modelada matematicamente pela teoria de conjuntos fuzzy a partir de um processo de otimização bioinspirada conhecido por evolução diferencial (Storn & Price, 1997). A forma da função de pertinência deste conjunto é determinada pela evolução a partir de seu comportamento, especificado pelo pesquisador que analisará a rede modelada.

A evolução diferencial é um método de busca estocástico, baseado em população, o que promove uma busca por melhores soluções em paralelo, ou seja, com diversas áreas do espaço de busca sendo avaliadas simultaneamente, além de serem menos prováveis de ficarem restritas a regiões de mínimos locais. Algumas das características deste algoritmo que o faz ser aplicado são seus poucos parâmetros de controle, sua simplicidade de implementação e sua robustez, que o torna capaz de encontrar o mínimo global em diferentes tipos de problemas de otimização (Das & Suganthan, 2011).

A seção a seguir apresenta a evolução diferencial e como ela foi implantada para encontrar funções de pertinência de conjuntos fuzzy para quaisquer finalidades. Na seção seguinte, a competição entre seres vivos é apresentada e são especificadas as características matemáticas desejadas para uma função de pertinência que a modele. Em seguida, os experimentos e seus resultados são exibidos e discutidos. A última seção trata das conclusões deste trabalho e aborda as questões a serem respondidas em futuras pesquisas.

2 EVOLUÇÃO DIFERENCIAL

Os algoritmos evolucionários (Ma *et al.*, 2013) são métodos de otimização bioinspirada com aprendizado coletivo, em que uma população de indivíduos representa possíveis soluções no espaço de busca de um dado problema. Como os demais algoritmos evolucionários, a evolução diferencial é um método simples, eficiente e robusto, capaz de lidar com funções objetivo complexas (não diferenciáveis, não lineares e multimodais) em espaços de busca multidimensionais. A principal característica deste método é o uso da diferença entre vetores populacionais (soluções candidatas) para criar novos vetores, que perturbam a população, e que, aliado a uma estratégia elitista de preservação de soluções, orienta a busca para regiões onde a função objetivo possa ser melhorada.

Inicialmente, certa quantidade de vetores (indivíduos) é criada aleatoriamente e estes são, em seguida, avaliados pela função de aptidão (função objetivo). Para cada indivíduo ou solução candidata (pai), uma reprodução com um vetor experimental é realizada e o vetor resultante (filho) disputa com o pai a permanência na geração seguinte por meio de uma seleção elitista. Em sua versão mais simples e popular, o vetor experimental é produzido a partir de outros três vetores da população escolhidos aleatoriamente: um dos vetores serve de base e os outros dois produzem o vetor diferencial, resultante da diferença entre eles. A Equação (1) mostra que o vetor experimental é uma combinação linear entre o vetor base e o vetor diferencial:

$$\mathbf{z}_i = \mathbf{v}_{r1} + F \times (\mathbf{v}_{r2} - \mathbf{v}_{r3}) \quad (1)$$

em que $\mathbf{z}_i$ representa o i -ésimo vetor experimental; $\mathbf{v}_{r1}$ é o vetor da população usado como base; F é o fator de escala da diferença entre os vetores $\mathbf{v}_{r2}$ e $\mathbf{v}_{r3}$, e $r1 \neq r2 \neq r3$ são índices selecionados aleatoriamente sem reposição.

A recombinação aqui utilizada para produzir a solução candidata final é a uniforme e a Eq. (2) mostra como tal critério funciona:

$$s_{i,j} = \begin{cases} z_{i,j} & \text{se } (\text{aleat}_j \leq CR) \\ v_{i,j} & \text{caso contrário} \end{cases} \quad (2)$$

cujos índices i e j referem-se aos vetores e aos seus componentes, respectivamente; $s_{i,j}$ é o componente do vetor candidato (filho), $z_{i,j}$ é o componente do vetor experimental e $v_{i,j}$ é o componente do vetor da população (pai); CR é a probabilidade de recombinação; aleat_j representa um valor aleatório de zero a um.

Conforme será explicado na seção a seguir, o objetivo da otimização deste trabalho é a maximização da função objetivo e, entre o vetor da população e o vetor candidato, será preservado na geração seguinte o que possuir maior valor de aptidão. Cada vetor representa uma função de pertinência definida em partes lineares e seus componentes são pares ordenados das extremidades dessas partes. Um exemplo de solução candidata e a função de pertinência que ela representa são apresentados na Fig. 1.

O algoritmo de evolução diferencial possui apenas três parâmetros para controlar o equilíbrio entre a exploração global e local do espaço de busca: o fator de escala diferencial F , a probabilidade de recombinação CR e o tamanho da população. O processo evolutivo segue iterativamente até que algum critério de parada pré-estabelecido seja satisfeito. Geralmente, o critério de parada é definido pela quantidade máxima de gerações.

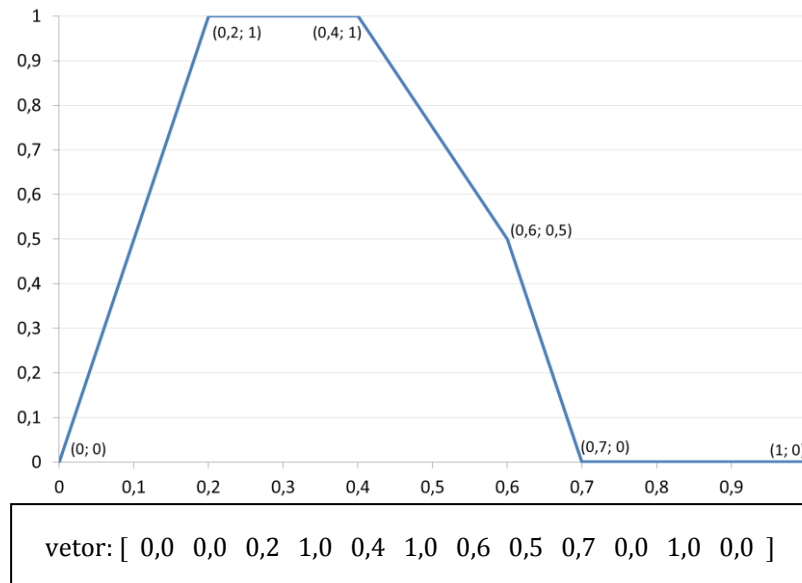


Figura 1. Representação gráfica de uma função de pertinência definida em partes lineares e seu vetor de pares ordenados das extremidades das partes

3 COMPETIÇÃO EM REDES PLANCTÔNICAS

O plâncton marinho é uma comunidade biológica composta por organismos, com dimensões tipicamente muito pequenas, de diversas espécies, entre algas, animais e bactérias, cujo poder de locomoção não é suficiente para deslocá-los contra a corrente das águas (Lalli & Parsons, 1997). Nas águas costeiras e em todo o ecossistema marinho, o plâncton é um elemento chave (Odum, 2001). Os dados do plâncton marinho, obtidos através da citometria de fluxo (Shapiro, 2003), possuem riqueza de detalhes e são disponibilizados em larga escala. Embora os organismos ali representados naturalmente interajam em redes dentro de redes, suas conexões não são detectadas pelo equipamento. Estas são ligações temporárias ou potenciais, sensíveis à composição e ao estado do ambiente e de seus componentes bióticos e abióticos, mutáveis e influenciadas pela entrada e saída de indivíduos no ecossistema. O estabelecimento dessas ligações é um desafio à parte, dada sua natureza fluida e subjetiva. Sua modelagem e os instrumentos de análise exigem recursos que permitam ir além de interpretações determinísticas. Portanto, a dinâmica que rege a vida dos organismos planctônicos impõe uma abordagem que permita contemplar a parcela de imprecisão e incerteza inerente aos sistemas naturais e que guarde um sentido ecológico, formando as arestas necessárias à geração de uma eficiente representação das redes planctônicas.

As relações ecológicas são interações entre organismos que compartilham recursos limitados, como alimento, abrigo e espaço, ocorrendo em uma hierarquia de escalas, individualmente ou coletivamente (Reece, 2011). Elas possuem diferentes características de acordo com o propósito que os organismos assumem ou impõem em suas interações. Embora as relações consumidores-recursos sejam as mais fundamentais entre as espécies, outras relações ecológicas também têm efeitos relevantes no uso dos recursos e na abundância das espécies em um ecossistema. A competição é uma relação de especial interesse, pois ocorre tanto dentro da mesma espécie quanto em espécies diferentes. A primeira é chamada de competição intraespecífica e se manifesta no comportamento territorialista dos organismos, motivada pela escassez das fontes alimentares – quando ocorre desequilíbrio na densidade de consumidores – e, principalmente, de água e luz (Dajoz, 1978). A segunda é chamada de

competição interespecífica e nela há a busca ativa por um mesmo recurso do meio, limitante do crescimento e da sobrevivência das espécies (Begon, Harper & Townsend, 2008).

As redes do mundo real frequentemente se encontram disponíveis em formas que não permitem sua análise imediata: elas podem ser muito complexas (Dempsey, Duraisamy & Ali, 2011) ou extensas, o que torna desejável uma amostra reduzida, mas significativa (Kurant *et al.*, 2011); podem estar incompletas devido à ocultação intencional de informações por parte dos atores (Krebs, 2002); e pode existir significativo ruído a ser filtrado (Dempsey *et al.*, 2012). A presença de obstáculos como estes exigem a adoção de mecanismos que permitam descobrir, inferir e reproduzir seus padrões estruturais originais. Para utilizar a modelagem por redes complexas na representação da competição entre os organismos identificados em dados citométricos, adotou-se a metodologia descrita em (Pereira *et al.*, 2014), observando-se que o objetivo não é a reprodução biológica das interações ocorridas naturalmente, mas a utilização de conceitos reais para fundamentar a modelagem das redes planctônicas. Essa metodologia propôs que os organismos mais semelhantes apresentam relação mais forte de competição e uma função de pertinência triangular modelou a competição linearmente, do indivíduo mais semelhante (com pertinência máxima) para o indivíduo menos semelhante (com pertinência nula). Um corte alfa foi aplicado ao conjunto dos competidores de um indivíduo para promover a eliminação de relações irrelevantes dentro de parâmetros populacionais observados na literatura.

Baseando-se nessa interpretação diretamente proporcional entre semelhança e competição, e procurando oferecer aos pesquisadores maior flexibilidade na definição de uma função que modele a competição de indivíduos segundo seus entendimentos, este trabalho propõe a criação de funções de pertinências por um algoritmo evolucionário a partir da especificação de três de suas características: forma (convexa, côncava, quasiconvexa, quasicôncava, polinomial, exponencial e logarítmica), preservação de relação de ordem (estritamente decrescente, decrescente, crescente, estritamente crescente e qualquer) e abertura do seu suporte. Esta última foi estimada por meio do cálculo da integral definida da função de zero a um. Vale ressaltar que o domínio da função de pertinência é a similaridade normalizada entre um indivíduo e os demais indivíduos da base, isto é, o intervalo real $[0,1]$.

Em especial neste estudo, a relação de competição foi modelada como uma função convexa, decrescente e com a abertura do suporte variável segundo o experimento (maior suporte implica em função mais larga, significando que mais indivíduos são considerados competidores de um indivíduo qualquer, pois terão pertinência positiva ao conjunto de competidores). Ainda, o primeiro par da função foi definido como $(0,1)$ e o último como $(1,0)$, ou seja, o maior competidor hipotético de um indivíduo é outro idêntico a ele e o ser menos semelhante não é considerado competidor efetivamente.

Uma vez produzida uma função de pertinência pela evolução diferencial, ela é aplicada aos dados da amostra para formar uma rede de competição. De uma rede, podem ser obtidos diversos índices e medidas auxiliares em seu entendimento. Neste trabalho foram usados a densidade e o grau médio de separação dos nós presentes no componente gigante da rede.

Em uma rede sem laços, o grau de um nó é a quantidade de arestas que lhe são adjacentes e, se o grau de um nó for igual a zero, ele é um nó isolado. Para um nó com p laços e q arestas, seu grau será igual a $2p+q$ (Balakrishnan, 1997). O grau médio da rede é a média aritmética dos graus de todos os seus nós (Diestel, 2010). No âmbito de teias alimentares, o grau médio de uma rede equivale à quantidade média de relações tróficas não direcionadas e de relações competitivas por espécie (Martinez, 1991). A densidade de uma rede estima a

intensidade de coesão da rede, sendo uma medida global que representa a proporção das arestas existentes (A) em relação ao máximo de arestas possível (Wasserman & Faust, 1994):

$$D = \frac{2A}{n(n-1)} \quad (3)$$

em que n é a quantidade de nós da rede. Em estudos de teias alimentares, uma medida muito similar à densidade chamada conectância é utilizada para estimar esse conceito (Warren, 1994):

$$C = \frac{A}{n^2} \quad (4)$$

em que n e A são, respectivamente, a quantidade de nós e a quantidade de arestas da rede. Pode-se observar que a conectância é praticamente a metade da densidade em redes com muitos nós e é utilizada em inferências sobre a dinâmica de uma rede, tais como: redes com maior conectâncias são menos estáveis (Warren, 1994) e são menos vulneráveis às invasões por parte de outras espécies (Robinson & Valentine, 1979).

A distância entre dois nós conectados de uma rede é o caminho mais curto que os separa no grafo que a representa. A maior distância entre dois nós quaisquer em uma rede define o seu diâmetro (Diestel, 2010). O grau de separação médio, também conhecido como comprimento de caminho característico (Williams *et al.*, 2002), é a média das distâncias entre os nós conectados em uma rede, estimando sua compacidade (Doyle & Graver, 2013) (Erdős, Pach & Spencer, 1988) e pode ser calculada, em uma rede não direcionada, como:

$$\bar{D} = \frac{\sum d(x,y)}{\binom{n}{2}} \quad (5)$$

em que d é a função distância entre dois nós conectados x e y da rede com n nós.

Uma propriedade notável em muitas redes reais é a ocorrência de uma pequena distância entre os nós em relação ao tamanho original da rede (Newman, 2010). Este fenômeno, conhecido como efeito de pequeno mundo, pode ser estimado numericamente através do grau de separação médio da rede e formaliza a ideia de que, em sua maioria, os nós de uma rede podem ser alcançados por caminhos muito curtos. É um efeito bastante disseminado em redes reais, inclusive biológicas (Wagner & Fell, 2001).

4 EXPERIMENTOS

Com os parâmetros populacionais, evolutivos e das funções, as possibilidades de configuração do processo evolutivo são inúmeras. Neste trabalho, uma parte dos experimentos realizados será descrita e discutida para compreender os efeitos da quantidade de partes de uma função e da abertura de seu suporte. Para cada possível configuração estudada, 50 execuções foram realizadas com o objetivo de observar o comportamento médio das funções e das redes formadas a partir delas.

A partir de uma amostra de 90304 assinaturas óticas de partículas, foram utilizados os 20 atributos citométricos mais relevantes derivados da citometria de fluxo (Andrade, 2014), apresentados na Tabela 1. Após essa seleção de atributos, foram identificadas as partículas

com características de fitoplâncton, organismos com capacidade fotossintética, por terem maior representatividade nos dados, resultando em um conjunto com 82635 elementos, cujas estatísticas são apresentadas na tabela. Como pode ser observado, há pouca dispersão nos dados, destacando-se nesse aspecto apenas sete variáveis, marcadas com a cor cinza. Esta pouca dispersão tem como consequência forte similaridade entre as partículas, promovendo, assim, grande competição entre elas. Dessa forma, uma rede de competição formada para esses dados terá grande possibilidade de preservar quase todos os registros (nós) e arestas possíveis, o que impedirá uma das grandes virtudes da modelagem de dados por redes complexas: a análise visual das relações. Quando os elementos de uma rede são em grande número, a visualização é muito prejudicada, pois o que se observa é uma nuvem compacta de arestas e nós.

Tabela 1. Estatísticas da amostra de fitoplânctons

	Média	Mediana	Desvio-padrão	Mínimo	Máximo	Coefficiente de variação
Length FWS	4,740	4,207	1,867	2,055	19,990	0,394
Total FL Yellow	158,492	135,960	245,450	11,444	35620,130	1,549
Total FL Orange	122,830	112,919	81,044	18,769	7959,075	0,660
Total FL Red	25,466	22,431	25,481	7,935	3435,773	1,001
Inertia FL Yellow	0,839	0,840	0,148	0,095	1,612	0,176
Inertia FL Orange	0,840	0,841	0,154	0,211	1,541	0,183
Inertia FL Red	0,823	0,821	0,126	0,178	1,333	0,153
Fill factor FWS	0,580	0,571	0,102	0,089	0,955	0,176
Fill factor SWS	0,578	0,574	0,086	0,189	0,963	0,149
Fill factor FL Yellow	0,849	0,870	0,079	0,236	0,969	0,093
Fill factor FL Orange	0,836	0,851	0,072	0,281	0,964	0,086
Fill factor FL Red	0,851	0,860	0,058	0,302	0,968	0,068
Asymmetry FWS	0,097	0,086	0,068	0,000	0,656	0,699
Asymmetry SWS	0,055	0,047	0,041	0,000	0,540	0,746
Asymmetry FL Yellow	0,086	0,067	0,073	0,000	0,559	0,846
Asymmetry FL Orange	0,084	0,068	0,067	0,000	0,487	0,798
Asymmetry FL Red	0,083	0,072	0,061	0,000	0,427	0,736
# of cells FL Yellow	2,934	2,856	0,886	1,028	15,919	0,302
# of cells FL Orange	2,872	2,789	0,860	1,034	20,525	0,300
# of cells FL Red	2,873	2,770	0,821	1,084	19,048	0,286

Para a realização dos experimentos, com o objetivo de usufruir dos recursos de visualização das relações que uma rede proporciona, decidiu-se escolher uma amostra com 1000 partículas que tivessem uma dispersão maior do que a amostra original, de modo que fosse possível notar mais rapidamente os efeitos das mudanças das aberturas das funções. Ao proceder com amostragens aleatórias sem reposição de 1000 registros, as estatísticas não diferiram significativamente dos valores apresentados acima. Dessa maneira, utilizou-se o agrupamento de dados pelo tradicional algoritmo *k-means* com distância euclidiana para selecionar os dados aleatoriamente dentro dos grupos. Diversas experimentações foram realizadas com diferentes quantidades de grupos e formas de seleção de dados. A estratégia que produziu as 1000 partículas com maiores dispersões foi obtida a partir de 80 grupos: os

1000 registros mais distantes de seus respectivos centros de grupos foram selecionados, cujas estatísticas são apresentadas na Tabela 2. As informações apresentadas nesta tabela e ilustradas na Fig. 2, permitem observar que o coeficiente de variação de todas as variáveis aumentou, sendo expressivos esses aumentos em 15 delas, destacadas em cinza na tabela.

Com a amostra definida, definiu-se a configuração dos parâmetros evolutivos que serão objetos de análise. Os componentes lineares de uma função de pertinência definida em partes são identificados pelas suas extremidades. Portanto, uma função com maior quantidade de extremidades/partes tem mais capacidade de ajuste aos critérios de busca (forma, preservação de ordem e abertura do suporte) do que uma função com poucos elementos. Para os experimentos desse estudo, especificou-se que a competição é modelada por uma função convexa decrescente. Além disso, a função deve iniciar no ponto (0,1) e terminar no ponto (1,0), garantindo pertinência máxima a indivíduos idênticos e pertinência nula a indivíduos menos semelhantes.

Tabela 2. Estatísticas da amostra de 1000 registros de fitoplânctons

	Média	Mediana	Desvio-padrão	Mínimo	Máximo	Coeficiente de variação
Length FWS	7,857	5,946	4,581	2,220	19,948	0,583
Total FL Yellow	482,076	140,156	2072,687	17,701	35620,130	4,300
Total FL Orange	258,551	119,785	554,579	27,161	7959,075	2,145
Total FL Red	67,421	26,170	190,709	9,613	3435,773	2,829
Inertia FL Yellow	0,824	0,827	0,265	0,095	1,612	0,322
Inertia FL Orange	0,820	0,829	0,269	0,211	1,541	0,328
Inertia FL Red	0,786	0,793	0,214	0,178	1,310	0,273
Fill factor FWS	0,618	0,595	0,169	0,089	0,955	0,274
Fill factor SWS	0,593	0,580	0,141	0,225	0,963	0,238
Fill factor FL Yellow	0,744	0,759	0,138	0,236	0,969	0,186
Fill factor FL Orange	0,746	0,754	0,119	0,281	0,953	0,160
Fill factor FL Red	0,776	0,802	0,117	0,302	0,952	0,151
Asymmetry FWS	0,136	0,096	0,122	0,000	0,646	0,897
Asymmetry SWS	0,088	0,062	0,089	0,000	0,529	1,017
Asymmetry FL Yellow	0,134	0,103	0,116	0,000	0,559	0,861
Asymmetry FL Orange	0,112	0,082	0,097	0,000	0,487	0,864
Asymmetry FL Red	0,112	0,098	0,087	0,000	0,427	0,774
# of cells FL Yellow	3,099	2,672	1,680	1,028	13,178	0,542
# of cells FL Orange	3,109	2,700	1,720	1,034	14,667	0,553
# of cells FL Red	3,018	2,667	1,524	1,099	14,716	0,505

Outro parâmetro variado neste estudo foi o peso da estimativa da abertura por meio da integral. A expectativa é que funções com maiores integrais tenham suportes mais amplos, o que promove a retenção de mais nós quando aplicadas aos dados das partículas. Para uma rede de competição de espécies, adotou-se peso negativo para essa característica, provocando sua minimização – pois essa relação é forte entre seres semelhantes e fraca entre os distintos – e induzindo funções potencialmente mais estreitas, de menor suporte. Após alguns experimentos iniciais com a parametrização apresentada na Tabela 3, observou-se que as funções produzidas apresentavam suporte muito extenso, mas com baixíssimos valores de

pertinência, o que provocava a indesejada preservação de muitos nós com graus de competitividade irrelevantes. Assim, as pertinências abaixo de 0,15 foram tornadas nulas.

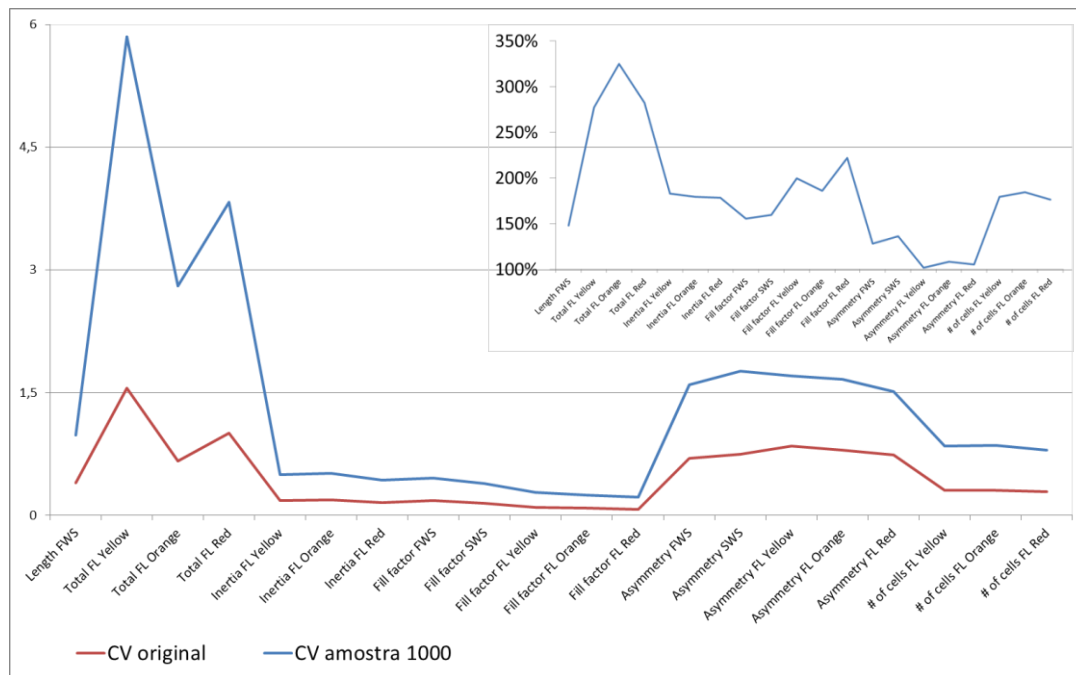


Figura 2. Comparação da dispersão por variável entre a amostra original e a reduzida. No detalhe, o aumento do coeficiente de variação, em porcentagem.

Tabela 3. Parametrização do processo evolutivo

Parâmetro	Valor	Parâmetro	Valor
Tamanho da população	30	Recombinação	Uniforme
Quantidade de gerações	1000	Probabilidade de recombinação	0,5
Quantidade de variáveis da função objetivo	3	Pesos das variáveis de forma e de preservação da ordem da função	1
Quantidade de pontos a otimizar	7, 10 ou 15	Peso da abertura da função	-0,5; -1,5; -3
Comprimento dos vetores	14, 20 ou 30	Peso do vetor diferencial	0,3

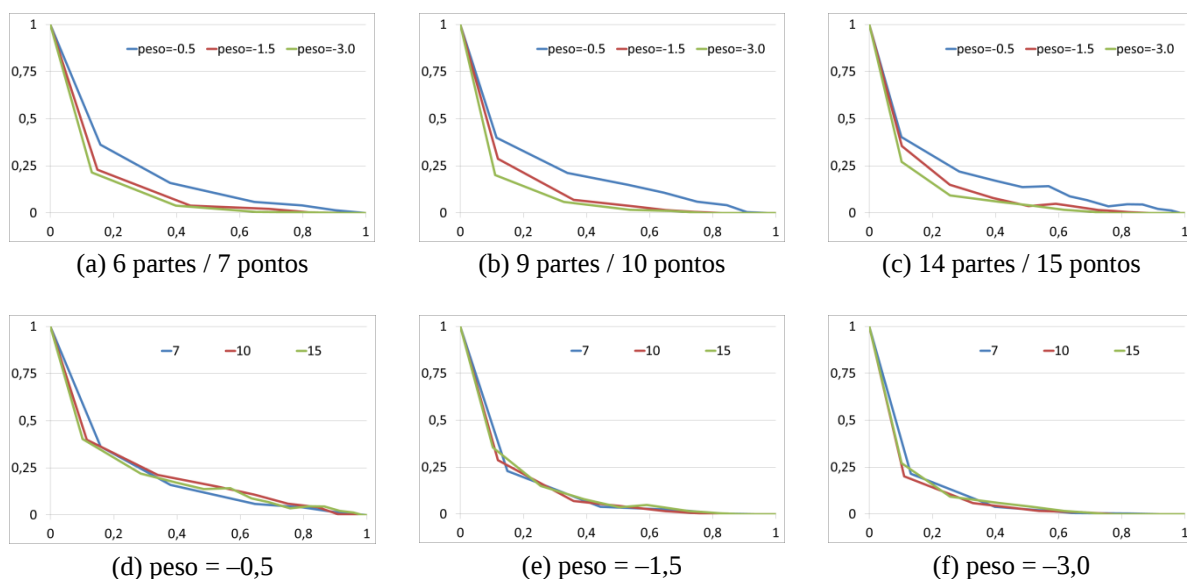
5 ANÁLISE DE RESULTADOS

Antes das considerações sobre as funções produzidas, vale informar que o volume de resultados produzidos dos experimentos é muito grande e apenas uma pequena parte será tratada neste trabalho. Inicialmente, procurou-se observar o comportamento do parâmetro peso da abertura da função de pertinência: poderia um aumento da importância dessa característica na função objetivo promover a criação de uma função mais estreita? Se sim, uma rede formada deveria possuir relativamente poucos nós, já que nem todos terão pertinência positiva. A Tabela 4 apresenta os valores médios das integrais das funções de pertinências encontradas ao final de 50 execuções de cada configuração da evolução diferencial, bem como seus desvios-padrão.

Tabela 4. Médias e desvios-padrão das integrais das funções geradas pela evolução diferencial

Integral da função	Peso da abertura da função		
Pontos	-0,5	-1,5	-3,0
7	0,182±0,084	0,131±0,059	0,103±0,044
10	0,175±0,111	0,117±0,059	0,091±0,039
15	0,162±0,111	0,115±0,066	0,094±0,051

Analisando os resultados por linha dessa tabela, a hipótese anterior é confirmada para cada categoria da função de pertinência (entende-se por categoria a quantidade de partes lineares da função a serem otimizadas). Nos itens (a)-(c) da Fig. 3 são exibidas as médias das funções de pertinências encontradas e pode-se visualizar o estreitamento da função à medida que mais importância é dada ao peso da abertura da função. Analisando os resultados por coluna dessa tabela, os valores das integrais médias são muito próximos e suas diferenças são inferiores aos desvios-padrão observados, independente da quantidade de partes otimizadas. Isso sugere uma região de convergência para as funções de pertinência, em menor escala quando o peso da abertura é pequeno em valores absolutos, onde se nota uma maior dispersão de resultados. Nos itens (d)-(f) da Fig. 3 observa-se a forte similaridade das médias das funções encontradas, por peso de abertura.

**Figura 3. Médias das funções de pertinência em 50 execuções da evolução diferencial**

Ao observar a Tabela 5 nota-se que os percentuais médios de nós remanescentes nas componentes gigantes das redes planctônicas produzidas exibem o mesmo padrão observado nas médias das integrais, ou seja, quanto maior a importância do peso (negativo) da abertura, menor o seu valor. Este comportamento está coerente com as informações apresentadas na Tabela 4, pois funções (supostamente) mais estreitas tendem a filtrar mais nós das redes originais, deixando mais nós fora de suas componentes gigantes. Observa-se que quanto maior o aumento nas quantidades de partes das funções a serem otimizadas, menores foram as retenções de nós, mesmo com os discretos aumentos das integrais das funções conforme Tabela 4. Isto pode indicar uma insuficiência da integral da função de pertinência como

indicador da abertura do seu suporte. Há de se investigar, ainda, o discreto comportamento contrário manifestado quando o peso da abertura foi -1,5 em funções com 14 partes a serem otimizadas.

Observando a Tabela 6, os graus médios dos nós diminuíram consideravelmente com um peso maior (em valores absolutos) da abertura da função, informando que redes menores (Tabela 5) e menos conectadas foram obtidas com funções mais estreitas. Portanto, redes menos complexas foram obtidas. Tal comportamento também foi observado quando se aumentou a quantidade de partes a serem otimizadas: quanto mais partes a função possui, mais relações são excluídas, o que pode indicar que mais partes permitem que a função consiga detectar melhor relações irrelevantes.

Tabela 5. Percentual médio dos nós originais que permaneceram nas redes planctônicas geradas a partir das funções de pertinência e seus desvios-padrão

% de nós originais	Peso da abertura da função		
Pontos	-0,5	-1,5	-3,0
7	99,4±02,7	87,8±32,5	86,6±30,3
10	81,7±37,9	80,4±37,7	75,0±39,9
15	77,3±41,2	80,7±35,7	73,6±41,1

Tabela 6. Graus médios dos nós das redes planctônicas geradas a partir das funções de pertinência

Graus médios	Peso da abertura da função		
Pontos	-0,5	-1,5	-3,0
7	647,9±389,9	494,5±394,8	373,0±386,5
10	620,1±402,1	319,6±386,2	200,8±289,3
15	495,7±414,7	239,6±332,9	117,2±222,0

Aliando as informações das Tabelas 5 e 6 às da Tabela 7, que apresenta as densidades encontradas nas redes obtidas, conclui-se que se o peso da abertura for aumentado em termos absolutos, redes menos complexas e densas serão produzidas já que houve redução da densidade com o aumento dos parâmetros observados. Além disso, a adequação das redes obtidas por este trabalho é sugerida pela compatibilidade dessas densidades com as das redes biológicas reais, como as 16 teias alimentares analisadas em (Dunne, Williams & Martinez, 2002) com conectâncias entre 0,026 e 0,315, o que equivaleriam a densidades entre 0,052 e 0,630 aproximadamente. Nota-se, também, a grande dispersão de resultados, sobretudo com os maiores valores de peso de abertura.

Tabela 7. Densidades médias das redes planctônicas geradas a partir das funções de pertinência e seus desvios-padrão

Densidades	Peso da abertura da função		
Pontos	-0,5	-1,5	-3,0
7	0,659±0,380	0,521±0,375	0,336±0,361
10	0,620±0,369	0,397±0,356	0,282±0,313
15	0,473±0,364	0,300±0,334	0,215±0,291

A Tabela 8 apresenta as médias dos graus de separação média dos nós das redes obtidas. Observa-se uma grande proximidade topológica entre os nós, confirmando os resultados de alguns estudos, que apontam a presença dessa topologia em redes ecológicas, como as redes de polinização citadas em (Olesen *et al.*, 2006), em que os graus de separação média são muito reduzidos, entre 1,00 e 2,31, para redes de sete a 840 nós. Além disso, nota-se uma inversão da tendência apresentada nas medidas de rede anteriores: quantos maiores os parâmetros, maiores os graus de separação média. Embora seja uma inversão da tendência apresentada nas medidas de rede anteriores, esse fato não é estranho e pode ser creditado à densidade de cada rede: densidades menores apresentam menos arestas, o que causa redução na quantidade de caminhos, aumentando o distanciamento entre os nós. Novamente, é de se destacar a grande dispersão observada nessa métrica de rede, assunto que será abordado em trabalhos futuros.

Tabela 8. Médias dos graus de separação médio das redes planctônicas geradas a partir das funções de pertinência e seus desvios-padrão

Graus de separação médio	Peso da abertura da função			
	Pontos	-0,5	-1,5	-3,0
7		1,610±1,133	1,691±0,862	2,699±2,534
10		1,711±1,640	2,465±1,937	2,971±2,495
15		1,900±1,494	2,967±2,519	3,123±2,166

Com o objetivo de visualizar a estrutura de uma rede de modo compreensível, foi selecionada uma rede com uma quantidade razoável de nós e baixa densidade. A Fig. 4 mostra uma rede de competição obtida a partir de uma função de pertinência definida em 14 partes com 544 nós, grau médio de 3,62, densidade de 0,007 e grau de separação médio de 11,78. Este é bem maior que as médias apresentadas na Tabela 8 devido à baixa densidade da rede, que leva a uma ocorrência menor de caminhos alternativos.

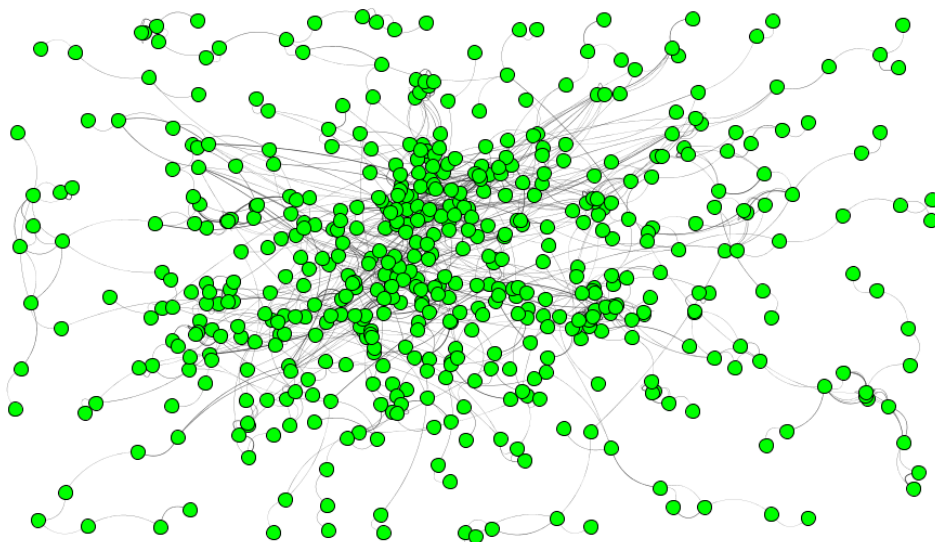


Figura 4. Rede planctônica de competição com 544 nós.

6 CONCLUSÃO

A evolução diferencial mostrou-se promissora na obtenção de uma função de pertinência *fuzzy* definida em partes lineares de acordo com os interesses de um pesquisador, em especial uma função para modelar a competição em plâncton marinho. Antes as funções eram triangulares e adotava-se um corte-alfa para promover a eliminação de relações irrelevantes dentro de parâmetros populacionais observados na literatura. As vantagens da abordagem apresentada neste trabalho são a eliminação do corte-alfa e a possibilidade de se especificar uma função a partir de sua forma, preservação de ordem e abertura, o que oferece aos pesquisadores maior flexibilidade na definição de uma função que modele uma relação de indivíduos segundo seus entendimentos. Para esse trabalho, procuraram-se funções decrescentes, convexas e com suporte pequeno para modelar as relações de competição entre plânctons. Quando aplicadas aos dados citométricos de uma amostra de 1000 partículas, as funções encontradas pelo processo evolutivo produziram redes coerentes com as observadas na literatura. De acordo com as parametrizações utilizadas, as características de forma e de preservação da ordem foram atendidas com louvor, entretanto a estimativa da abertura de uma função pela sua integral se mostrou insuficiente para modelar tal característica. Em continuidade a este trabalho, maior acompanhamento do processo evolutivo será realizado, com a adoção de parâmetros auto adaptativos; uma nova modelagem da abertura da função será feita ao se combinar a amplitude do suporte com a integral da função; outras características da função objetivo serão implantadas para encontrar funções de pertinência mais complexas; serão exploradas outras métricas de redes, como o coeficiente de agrupamento, de forma a avaliar melhor as redes produzidas; e será realizada uma análise estatística mais aprofundada das redes produzidas para compreender as dispersões observadas em suas métricas.

REFERENCES

- Agouru, C. U., Audu, G., “Studies on the Range of Plankton in River Benue, North Central, Nigeria, Western Africa”, *Greener Journal of Biological Sciences*, V.2, n. 2, pp. 028-034, 2012.
- Andrade, Lúcio P., 2014. *Análise de Redes Planctônicas Utilizando Multigrafos Fuzzy*. Tese de Doutorado, Universidade Federal do Rio de Janeiro.
- Balakrishnan, V. K., 1997. *Schaum's Outline of Graph Theory: Including Hundreds of Solved Problems*. McGraw Hill Companies, Inc.
- Begon, M., Harper, J. L., & Townsend, C. R, 2008. *Essentials of Ecology*. 3 ed. Blackwell Publishing.
- Dajoz, Róger, 1978. *Ecologia geral*. 3 ed. Editora Vozes Ltda.
- Das, S., Suganthan, P. N., “Differential Evolution: A Survey of the State-of-the-Art,” *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4-31, 2011
- Dempsey, K., Duraisamy, K., Ali, H., et al, 2011. A parallel graph sampling algorithm for analyzing gene correlation networks. In Sato, M., Matsuoka, S., Sloot, P. M., et al, eds, *International Conference on Computational Science*, pp. 136-145.
- Dempsey, K.; Bhowmick, S., & Ali, H., 2012. Function-preserving Filters for Sampling in Biological Networks. In Ali, H., Shi, Y., Khazanchi, D., et al, eds, *International Conference on Computational Science*, pp 587-595.
- Diestel, Reinhard, 2010. *Graph Theory*, 4. electronic edition. Springer-Verlag Heidelberg.

- Doyle, J. K. & Graver, J. E., 2013. Mean distance in a graph. *Discrete Mathematics*, v. 17, n. 2, pp. 147-154, 1977. Available: <<http://www.sciencedirect.com/science/article/pii/0012365X77901443>> . Access: 25th July 2015.
- Dunne, J. A., Williams, R. J. & Martinez, N. D., 2002. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, vol. 99, n. 20, pp.12917-12922.
- Erdős, P., Pach, J. & Spencer, J. H., 1988. On the mean distance between points of a graph. 250th Anniversary Conference on Graph Theory, 64, pp. 121-124.
- Kurant, M., Gjoka, M., Butts, C., et al, 2011. Walking on a Graph with a Magnifying Glass: Stratified Sampling via Weighted Random Walks. In *Proceedings of ACM SIGMETRICS '11*, pp. 281-292.
- Lalli, Carol M., & Parsons, Timothy R., 1997. *Biological oceanography: an introduction*. 2 ed. Elsevier Butterworth-Heinemann.
- Ma, H., Simon, D., Fei, M. et al. “On the equivalences and differences of evolutionary algorithms”, *Engineering Applications of Artificial Intelligence*, vol 26, no. 10, pp. 2397–2407, 2013.
- Malkassian, A., Nerini, D., van Dijk, M. A., et al, 2011. Functional analysis and classification of phytoplankton based on data from an automated flow cytometer. *Cytometry A*, vol. 79, n. 4, pp. 263-275.
- Martinez, Neo D., 1991. Artifacts or Attributes? Effects of Resolution on the Little Rock Lake Food Web. *Ecological Monographs*, vol. 61, n. 4, pp. 367-392.
- Newman, Mark E. J., 2010. *Networks: An Introduction*. Oxford University Press.
- Odum, Eugène P., 2001. *Fundamentos de Ecologia*. 6 ed. Fundação Calouste Gulbenkian.
- Olesen, J. M., Bascompte, J., Dupont, Y. L. et al, 2006. The smallest of all worlds: pollination networks. *Journal of Theoretical Biology*, vol. 240, n. 2, pp. 270-276.
- Pereira, G. C., Andrade, L. P., Espíndola, R. P., et al, 2014. Structural Analysis and Static Simulation of Coastal Planktonic Networks. *Journal of Intelligent Learning Systems and Applications*, vol. 6, n. 2, pp. 113-124.
- Reece, Jane B., et al., 2011. *Campbell Biology*. 9 ed. Pearson Education, Inc.
- Robinson, James V. & Valentine, Walter D., 1979. The concepts of elasticity, invulnerability and invadability. *Journal of Theoretical Biology*, vol. 8, n. 1, pp. 91-104.
- Shapiro, Howard M., *Practical Flow Cytometry*. 4 ed. New Jersey, USA, John Wiley & Sons, Inc., 2003.
- Storn, R., Price, K., “Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces,” *Journal of Global Optimization*, vol. 11, no. 4, pp. 341-359, 1997.
- Wagner, A. & Fell, D. A., 2001. The small world inside large metabolic networks, *Proceedings of the Royal Society B Biological Sciences*, vol. 268, n. 1478, pp. 1803-1810.
- Warren, Philip H., 1994. Making connections in food webs, *Trends in Ecology & Evolution*, vol. 9, n. 4, pp. 136-141.
- Wasserman, Stanley & Faust, Katherine, 1994. *Social Network Analysis: methods and applications*. Cambridge University Press.
- Williams, R. J., Berlow, E. L., Dunne, J. A., et al, 2002. Two degrees of separation in complex food webs, *Proceedings of the National Academy of Sciences*, vol. 99, n. 20, pp. 12913-12916.