





ARTIGO ORIGINAL

Classificação da capacidade produtiva de florestas inequidêneas por meio de mapas auto-organizáveis

Classification of the productive capacity of uneven-aged forests through self-organizing maps

Diego dos Santos Vieira¹ , Marcio Leles Romarco de Oliveira¹ , João Ricardo Vasconcellos Gama² , Bruno Oliveira Lafetá³ 

¹Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, Diamantina, MG, Brasil

²Universidade Federal do Oeste do Pará – UFOPA, Santarém, PA, Brasil

³Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais – IFMG, São João Evangelista, MG, Brasil

Como citar: Vieira, D. S., Oliveira, M. L. R., Gama, J. R. V., & Lafetá, B. O. (2022). Classificação da capacidade produtiva de florestas inequidêneas por meio de mapas auto-organizáveis. *Scientia Forestalis*, 50, e3794. <https://doi.org/10.18671/scifor.v50.28>

Resumo

O objetivo desse estudo foi avaliar e comparar o uso do mapa auto-organizável de Kohonen (método Kohonen) e análise de agrupamento (método Souza) para classificação da capacidade produtiva de florestas inequidêneas na Amazônia. Os dados foram coletados a partir de um inventário sistemático com 110 parcelas de 0,75 ha, alocadas sistematicamente em 42.159,34 ha. A análise discriminante foi usada para confirmação da distinção e classificação das classes de produtividade obtida pelos dois métodos. Ambos os métodos registraram três classes de produtividade, a saber: baixa, média e alta. O volume das classes baixa e alta, entre métodos, apresentaram diferenças significativas. A análise discriminante mostrou que 98,2% das parcelas foram corretamente classificadas nas classes de produtividade do método Souza. Esse percentual foi de 99,1% para o método Kohonen. A similaridade de alocação de parcelas nas classes pelos dois métodos foi de 74,5%. Ambos os métodos podem ser utilizados para a classificação da capacidade produtiva de florestas inequidêneas. Mas, o método Kohonen apresenta maiores vantagens, a saber: maiores acertos na classificação das parcelas dentro das classes de produtividade; menores chances de obter-se erros de classificação discrepantes; possibilidade de classificação de parcelas que não estavam presentes durante o treinamento do mapa auto-organizável; produz estruturas de visualização mais informativas; e possibilidade de inserção de variáveis categóricas. O método Souza deve ser calculado preferencialmente com poucas variáveis, para que não ocorram os erros de classificação discrepantes.

Palavras-chave: Mapas de Kohonen; Análise de agrupamento; Análise discriminante; Estratificação volumétrica; Amazônia.

Abstract

The aim of this study was to evaluate and compare the use of the self-organizing map Kohonen (Kohonen method) and cluster analysis (Souza method) to classify the productive capacity of uneven forests in the Amazon area. Data were collected from a systematic inventory with 110 plots of 0.75 ha, systematically allocated in 42,159.34 ha. Discriminant analysis was used to confirm the distinction and classification of productivity classes obtained by the two methods. Both methods registered three classes of productivity, namely: low, medium and high. The volume of low and high classes between methods displayed significant differences. Discriminant analysis showed that 98.2% of the plots were correctly classified in the productivity classes by the Souza method. This percentage was 99.1% for the Kohonen method. The similarity of allocation of plots in classes by the two methods was 74.5%. Both methods can be used to classify the productive capacity of uneven forests. However, the Kohonen method has greater

Fonte de financiamento: Nenhuma.

Conflito de interesse: Nada a declarar.

Autor correspondente: diegovieir4@gmail.com

Recebido: 11 novembro 2021.

Aceito: 6 julho 2022.

Editor: Mauro Valdir Schumacher.



Este é um artigo publicado em acesso aberto (*Open Access*) sob a licença *Creative Commons Attribution*, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o artigo científico seja corretamente citado.

advantages, namely: greater success in classifying the parcels within the productivity classes; less chances of getting outlying misclassifications; possibility classification of plots that were not present during the training of the self-organizing map; produces more informative visualization structures; and possibility of inserting categorical variables. The Souza method should preferably be calculated with few variables, so that discrepant classification errors do not occur.

Keywords: Kohonen map; Cluster analysis; Discriminant analysis; Volumetric stratification; Amazon area.

INTRODUÇÃO

A Amazônia possui diversas fitofisionomias que apresentam alta variabilidade, sobretudo de espécies, idades, tamanhos e condições gerais de crescimento e produção (Corrêa et al., 2018). Essa variabilidade, principalmente no que se refere à idade das árvores, impede que métodos de classificação da capacidade produtiva usados para florestas equiâneas (e.g., relação altura dominante *versus* idade) sejam aplicados naquelas inequiâneas. Mas, a capacidade produtiva de florestas inequiâneas pode ser determinada de outras maneiras, a saber: por meio de fatores edafoclimáticos; pela vegetação indicadora; e pela produção volumétrica (Campos & Leite, 2017).

A produção volumétrica é a mais usada em florestas inequiâneas e normalmente associa-se a métodos multivariados, a saber: cluster e análise discriminante. Esses métodos dividem a floresta em classes de sítio com capacidade produtiva baixa, média e alta (Souza & Souza, 2006), o que permite melhores planejamentos e controles da produção (Souza et al., 2014; Souza & Souza, 2006). Além disso, são os mais usados porque relacionam-se com a separação de indivíduos (e.g., parcelas e talhões) e com alocação desses em grupos predeterminados. A análise de cluster constrói os agrupamentos a partir da matriz multivariada, enquanto a discriminante aloca indivíduos em classes predeterminadas (Souza & Souza, 2006).

Muitos autores comprovaram a viabilidade desses métodos (Corrêa et al., 2018; Corrêa et al., 2020; Lima et al., 2014; Souza et al., 2014; Souza & Souza, 2006), porém nenhum avaliou algoritmos que incorporassem variáveis categorias e numéricas, juntas. A maioria usa apenas o volume, enquanto variáveis como classes diamétricas e qualidade de fuste das árvores são descartadas. Essas variáveis forneceriam ao administrador uma classificação da produtividade que mostra, além da produção volumétrica, os percentuais de volume formado por árvores com qualidade de fuste superior e que possuem diâmetros maiores ou menores que 50 cm. Visualmente, traria uma avaliação e interpretação mais rápida, precisa e eficiente da produtividade de florestas na Amazônia.

O método multivariado fornece essas informações, mas somente na forma de tabelas ou parágrafos de texto curto, o que os torna inadequados para o reconhecimento rápido de padrões e desvios interessantes para muitas variáveis (Klobucar & Subasic, 2012). Além disso, fornecem resultados que são difíceis de interpretar e muitas vezes não representam as variáveis de uma grande base de dados de maneira fácil e compreensível (Annas et al., 2007). O mapa auto-organizável de Kohonen (MAK) é um método que soluciona esses problemas, pois combina plausibilidade biológica à aplicabilidade, solucionando muitos problemas de processamento e otimização de informações (Stümer et al., 2010).

O MAK é usado para detecção e visualização de agrupamentos em um conjunto de dados, além de representar o mesmo conjunto em um mapa bidimensional, de uma maneira que preserva as relações não lineares das variáveis (Klobucar & Subasic, 2012). O objetivo desse estudo foi avaliar o uso do MAK para classificação da capacidade produtiva de florestas inequiâneas na Amazônia. Além disso, realizou-se a comparação do MAK com o método multivariado. A hipótese nula (H_0) que norteou essa pesquisa foi que o MAK é mais eficiente que o método multivariado para classificação da capacidade produtiva de florestas inequiâneas, enquanto a alternativa (H_1) considerou a rejeição da hipótese H_0 .

MATERIAL E MÉTODOS

Área de estudo

O estudo foi realizado na Zona de Manejo Florestal da Floresta Nacional do Tapajós, localizada às margens da BR-163, estado do Pará. Trata-se de uma área com 42.159,3 ha, localizada em regiões de baixo platô com árvores emergentes. O clima da região, conforme a classificação de Köppen, é Am, ou seja, tropical úmido com variação térmica anual inferior a 5,0°C (Alvares et al., 2013). A temperatura, umidade relativa do ar e precipitação pluviométrica média anual são de 25,5°C, 88% de 1.820 mm, respectivamente. O solo predominante na área é o Latossolo Amarelo Distrófico (Santos et al., 2017). A vegetação predominante é a Floresta Ombrófila Densa de Terra Firme, caracterizando-se pela dominância de árvores de grande porte e pela abundância de lianas lenhosas, palmeiras e epífitas (Santos et al., 2017).

Coleta de dados

O processo de amostragem usado foi o sistemático (Loetsch et al., 1964). A alocação das parcelas considerou um modelo digital de elevação (MDE), construído a partir da técnica de interferometria de radar e uma imagem com características altimétricas do tipo *Shuttle Radar Topography Mission* (SRTM), correspondente à carta topográfica SB-21-X-D na escala de 1:250.000 (Andrade et al., 2015). O MDE serviu para identificar áreas inacessíveis para a colheita, considerando a tecnologia utilizada pela Cooperativa Mista da Floresta Nacional do Tapajós, e para instalação de cinco faixas de inventário, dispostas longitudinalmente, equidistantes 4.000 m e referenciadas ao Meridiano Central correspondente à Zona 21M, no sistema de quadrículas *Universal Transversa de Mercator* (UTM) e *datum* horizontal WGS 84. A partir desse critério áreas com declividades maiores do que 20° não foram amostradas (Andrade et al., 2015).

As classes de declividade ($C_1 = 0^\circ - 10^\circ$, $C_2 = 10^\circ - 20^\circ$, $C_3 = 20^\circ - 45^\circ$, $C_4 = 45^\circ - 90^\circ$) foram definidas por meio do *software* ArcMap™ 10.5, onde foram obtidas as curvas de nível de 20 em 20 metros, e, posteriormente, por meio da ferramenta *Create TIN*, o MDE com as classes C_1 , C_2 , C_3 e C_4 . Após a disposição das faixas sobre o MDE, elas foram exportadas para um GPS Garmin 60 Cxs, para posterior identificação na área. As faixas foram a base para instalação de 110 parcelas de 30 m x 250 m, equidistantes 500 m. Em cada parcela, foram inventariadas as seguintes classes de tamanho (CT) de diâmetro a 1,30 m do solo (DAP): CT₁: árvores com 10 cm \leq DAP < 25 cm em subparcelas de 30 m x 50 m; CT₂: árvores com 25 cm \leq DAP < 50 cm em subparcelas de 30 m x 100 m; e CT₃: árvores com DAP \geq 50 cm na parcela de 30 m x 250 m.

As informações coletadas durante o inventário amostral foram o nome científico, circunferência do tronco à altura de 1,30 m do solo (CAP), altura comercial (Hc), longitude (X) e latitude (Y) das parcelas e qualidade de fuste (QF). Essa última foi dividida em três categorias, a saber: QF₁ = fustes com aproveitamento de 80 - 100%; QF₂ = fustes com aproveitamento de 50 - 79%; e QF₃ = fustes com aproveitamento abaixo de 50%. As árvores foram identificadas em campo, em nível de espécie, e aquelas que suscitaram dúvidas tiveram sua determinação taxonômica feita por especialistas, por meio de comparações, no herbário da Universidade Federal do Oeste do Pará (UFOPA). O sistema de classificação adotado foi o APG III (Chase et al., 2016).

Análise de dados

As variáveis de entrada usadas para classificação da produtividade da Zona de Manejo Florestal foram a média aritmética do volume (V), área basal (G) e a densidade de árvores (DA) por hectare de cada parcela. Além dessas, optou-se pela adição da variável categórica qualidade de fuste (QF). A conversão de categórica para numérica ocorreu se obtendo a média ponderada arredondada do número de árvores nas classes de QF. Os pesos para QF₁, QF₂ e QF₃ foram 3, 2 e 1, respectivamente. Logo, se uma parcela apresenta 25 árvores na classe QF₁,

30 árvores na classe QF_2 e 10 árvores na classe QF_1 , então sua média ponderada é aproximadamente de 24,0.

A partir das variáveis de entrada, elaborou-se uma matriz X de dados, em que cada variável x_{ij} representa o j -ésimo volume, área basal, densidade de árvores e qualidade de fuste na i -ésima parcela, para $i = 1, 2, 3, \dots, 110$. O primeiro método, aqui denominado de Souza, usou a análise de *cluster* para classificação da capacidade produtiva (Souza et al., 2014; Souza & Souza, 2004). A partir do dendrograma traçou-se uma linha de corte a um nível de homogeneidade de 25%, de modo que fossem formadas três classes de produtividade. O segundo método, denominado de Kohonen, fez uso do mapa auto-organizável de Kohonen - MAK (Kohonen, 2001). O treinamento iniciou com a determinação dos parâmetros topológicos, a saber: número de neurônios, topologia e dimensão do MAK.

O número de neurônios foi calculado para aproximadamente $5 \times \sqrt{\text{número de classes iniciais de DAP}}$ (Vesanto et al., 2000). A topologia selecionada foi a hexagonal regular, baseado no conselho de boas práticas de Kohonen (2001), que afirma que mapas hexagonais devem ser preferidos porque as posições finais de cada neurônio dependem de mais vizinhos do que topologias regulares, retangulares ou triangulares. A dimensão do MAK foi calculada em função do número de neurônios, de forma que a largura \times comprimento fosse, aproximadamente, igual ao número de neurônios calculados (Kohonen, 2001).

Outros parâmetros usados durante o treinamento foram a inicialização aleatória do vetor de pesos inicial, medida de dissimilaridade distância euclidiana, função de vizinhança bolha, 1000 iterações e taxa de aprendizagem inicial e final de 0,05 e 0,01, respectivamente. Esses últimos parâmetros são *default* do *software* utilizado, o R. Após a determinação dos parâmetros, foram executadas oito etapas para o treinamento do MAK, a saber:

- Etapa 1: Escolha dos parâmetros topológicos;
- Etapa 2: Inicialização do vetor de pesos $[w^{(j)}]$ de cada neurônio aleatoriamente;
- Etapa 3: Especificação da função de vizinhança $[\alpha^{(t)}]$ e taxa de aprendizagem $[\lambda]$;
- Etapa 4: Obtenção do conjunto de amostras $[x^{(k)}]$ para o treinamento;
- Etapa 5: Inicialização do contador de números de iteração (iteração $\leftarrow 0$);
- Etapa 6: Repetição das instruções a seguir até que não haja mudanças significativas no vetor de pesos sináptico final:
 - Em todas as amostras de entrada, fazer:
 - Calcule as distâncias euclidianas entre amostra k e os vetores de pesos, conforme a expressão a seguir:

$$\text{dist}_j^{(k)} = \sqrt{\sum_{i=1}^n (x_i^{(k)} - w_i^{(j)})^2} \tag{1}$$

em que: $\text{dist}_j^{(k)}$ = distância euclidiana entre o vetor de entrada, representando a k -ésima amostras $[x^{(k)}]$, em relação ao vetor de pesos do i -ésimo neurônio $[w^{(j)}]$.

- Declare como neurônio vencedor o neurônio j que obter a menor distância euclidiana:

$$\text{vencedor} = \arg \min_i \{ \|x^{(k)} - w^{(i)}\| \} \tag{2}$$

- Ajuste do vetor de pesos do neurônio vencedor conforme a expressão a seguir:

$$w_{(t+1)}^v = w_{(t)}^v + \eta \cdot (x^{(k)} - w_{(t)}^v) \tag{3}$$

em que: $w_{(t+1)}^v$ = vetor de pesos atualizado do neurônio vencedor na iteração $t+1$, $w_{(t)}^v$ = vetor de pesos do neurônio vencedor na iteração t ; η = parâmetro de taxa de aprendizagem; e $x^{(k)}$ = vetor de entrada da k -ésima amostra.

- Ajuste o vetor de pesos dos neurônios vizinhos ao vencedor, determinados pela função de vizinhança bolha [$\alpha^{(\Omega)}$], conforme a expressão a seguir:

$$w_{(t+1)}^{\Omega} = w_{(t)}^{\Omega} + \eta \cdot \alpha^{(\Omega)} \cdot (x^{(k)} - w_{(t)}^{\Omega}) \quad (4)$$

em que: $w_{(t+1)}^{\Omega}$ = vetor de pesos atualizado do neurônio vizinho ao vencedor na iteração t+1, $w_{(t)}^{\Omega}$ = vetor de pesos do neurônio vizinho ao vencedor na iteração t; η = parâmetro de taxa de aprendizagem; $\alpha^{(\Omega)}$ = função de vizinhança centrada em torno do neurônio vencedor; e $x^{(k)}$ = vetor de entrada da k-ésima amostra.

- Iteração \leftarrow Iteração +1;

Etapa 7: Análise do MAK para extração de características;

Etapa 8: Identificação de agrupamentos que possibilitem a definição de classes de produtividade.

Outros detalhes sobre o algoritmo são descritos por Kohonen (2001) e Asan & Ercan (2012). O treinamento teve como resposta um vetor de pesos sinápticos, que foi agrupado em três classes de produtividade. A medida de dissimilaridade e o método de agrupamento adotados para formação dessas classes foram a distância euclidiana e o método do vizinho mais distante, respectivamente. O MAK foi avaliado considerando-se o progresso de treinamento, histograma de ativação, mapa de qualidade e planos componentes. Análises discriminantes foram realizadas, posteriormente, para confirmação e distinção da classificação das classes de produtividade obtidas pelos métodos Souza e Kohonen.

Após a determinação das classes de produtividade, aplicou-se o teste de Welch, a 95% de probabilidade, para comparar as médias de volume, área basal e densidade de árvores das classes entre os métodos de Kohonen e Souza. A área das classes de produtividade dos métodos Souza e Kohonen foi calculada por meio da proporção do número de parcelas agrupadas em cada classe, em relação à área total (Barbosa et al., 2017). Isso somente foi possível devido ao esquema sistemático de distribuição das parcelas, que possuem a mesma probabilidade de seleção. Todas as análises foram realizadas no software R versão 4.0.3 (R Development Core Team, 2022). O treinamento do MAK, as análises de agrupamento e discriminantes foram realizadas por meio dos pacotes "kohonen" (Wehrens & Kruisselbrink, 2018), "vegan" (Oksanen et al., 2019), e "MASS" (Venables & Ripley, 2002), respectivamente.

RESULTADOS E DISCUSSÃO

Foram registradas 4.678 árvores, distribuídas em 205 espécies e 45 famílias botânicas. A densidade de árvores, área basal e volume por hectare foram de 176,8 árv ha⁻¹, 13,7 m² ha⁻¹ e 121,0 m³ ha⁻¹, respectivamente. O dendrograma do método de Souza formou três grupos homogêneos e distintos denominados de classes de produtividade baixa, média e alta, as quais apresentaram 62, 34 e 14 parcelas. A análise discriminante foi significativa e mostrou que 95,5% das parcelas foram corretamente distribuídas nas classes de produtividade. Após a reclassificação, por meio das probabilidades posteriores, as classes baixa, média e alta apresentaram 60, 37 e 13 parcelas. Essas classes apresentaram área de 22.996,0, 14.180,8 e 4.982,3 hectares, respectivamente.

A média volumétrica das classes baixa, média e alta foi de 106, 130,6 e 164,4 m³ ha⁻¹, respectivamente. A análise discriminante manteve-se significativa e apresentou Wilk's Lambda λ de 0,160, indicando elevado poder discriminatório das variáveis de entrada na formação das classes de produtividade. A taxa de acerto aumentou 2,7%, alcançando 98,2%. A análise discriminante avaliou, auxiliou e melhorou a alocação das parcelas nas classes cuja probabilidade de pertencer é maior (Souza et al., 2014; Souza & Souza, 2006). Apenas duas parcelas não foram alocadas corretamente, isso porque, de acordo com a análise discriminante, apresentaram maiores probabilidades posteriores de pertencerem a outra

classe de produtividade. Isso diminuiu a taxa de acerto das classes baixa e média para 98,2 e 97,3%, respectivamente (Tabela 1).

Tabela 1. Estatística Wilk's Lambda, volume médio e análise discriminante após a reclassificação das classes de produtividade calculadas por meio dos métodos Souza e Kohonen, Floresta Nacional do Tapajós, estado do Pará.

Método	Classes	Wilk's Lambda	Volume ¹	Parcelas	Matriz de confusão			Acertos (%)
					Baixa	Média	Alta	
Souza	Baixa	0,160	106,0	60	59	1	0	98,2
	Média		130,6	37	1	36	0	97,3
	Alta		164,4	13	0	0	13	100,0
Total	-	-	121,0	110	60	37	13	98,2
Kohonen	Baixa	0,150	89,60	56	56	0	0	100,0
	Média		141,4	44	0	43	1	97,7
	Alta		207,6	10	0	0	10	100,0
Total	-	-	121,0	110	56	43	11	99,1

¹volume, em m³ ha⁻¹.

O progresso de treinamento do MAK mostrou que conforme as iterações progredirem, a distância do vetor de pesos de cada neurônio para as amostras representadas por esse neurônio (i.e., o erro de quantização) é reduzida (Figura 1A). O ideal é que a distância atinja um patamar mínimo. Assim, se a curva diminui continuamente, sem patamar, mais iterações serão necessárias (Wehrens & Kruisselbrink, 2018). Observou-se o patamar a partir de, aproximadamente, 7.100 iterações, sugerindo que o número de iterações foi suficiente e que a distância média para o neurônio mais próximo, isto é, o vencedor, estabilizou a partir da 8.100ª iteração, quando atingiu o valor de 0,119. Quanto mais próximo de zero, melhor. O mapa de frequência de ativação indicou que todos os neurônios foram ativados, ou seja, que pelo menos uma amostra foi mapeada para cada neurônio (Figura 1B).

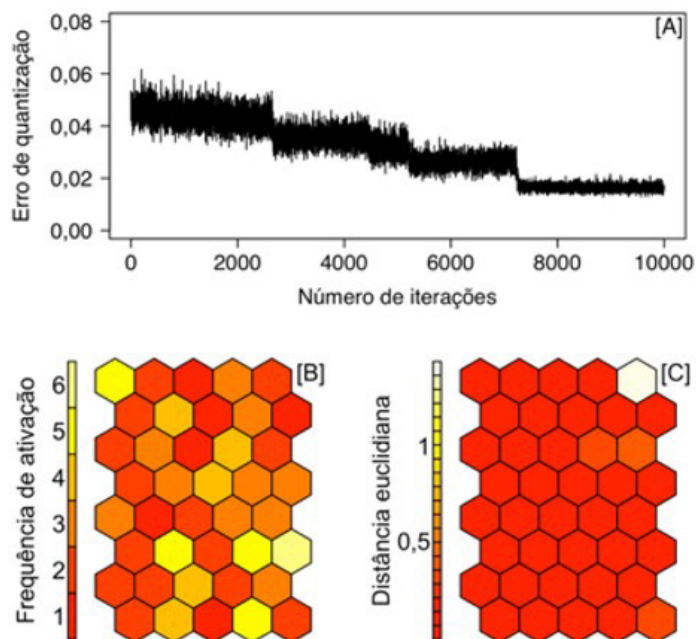


Figura 1. Progresso de treinamento [A], histograma de ativação [B] e mapa de qualidade [C] de treinamento do mapa auto-organizável de Kohonen, Floresta Nacional do Tapajós, estado do Pará.

Além disso, observou-se uma frequência de ativação relativamente uniforme, com a maioria dos neurônios ativados 2 e 3 vezes. Espera-se que a frequência de ativação seja relativamente uniforme, pois neurônios que são ativados muitas vezes sugerem que mapas de dimensões maiores devem ser considerados, enquanto neurônios com ativação zero indicam que um mapa menor deve ser considerado (Wehrens & Kruisselbrink, 2018). O mapa de qualidade mostra que a maioria das amostras apresentam distâncias euclidianas médias pequenas em relação ao vetor de pesos dos neurônios (Figura 1C). Teoricamente, quanto menores as distâncias, melhores as amostras são representadas pelos vetores de peso. Logo, como a finalidade é obter maior proximidade entre amostras e vetores de peso, de forma que esse vetor represente a distribuição das amostras, o treinamento foi considerado ótimo.

Assim, o método de Kohonen agrupou as parcelas em três grupos de neurônios, denominados de classes de produtividade baixa, média e alta, que apresentaram 56, 44 e 10 parcelas, respectivamente (Figura 2A). A taxa de acerto foi de 98,2%. Isso porque duas amostras foram alocadas nas classes erradas, uma na média e outra na alta. Após a reclassificação, as mesmas classes apresentaram 56, 44 e 10 parcelas, o que originou áreas de 21.462,9, 16.863,7 e 3.832,7 hectares, respectivamente. A média volumétrica das classes baixa, média e alta foi de 89,6, 141,4 e 207,6 m³ ha⁻¹, respectivamente. A análise discriminante manteve-se altamente significativa e apresentou Wilk's Lambda λ de 0,150, indicando elevado poder discriminatório de DA, V, G e QF. A taxa de acerto aumentou 0,9%, alcançando 99,1%. Apenas uma amostra foi alocada de forma errada, pois apresentava maior probabilidade de pertencer à classe de produtividade alta (Tabela 1).

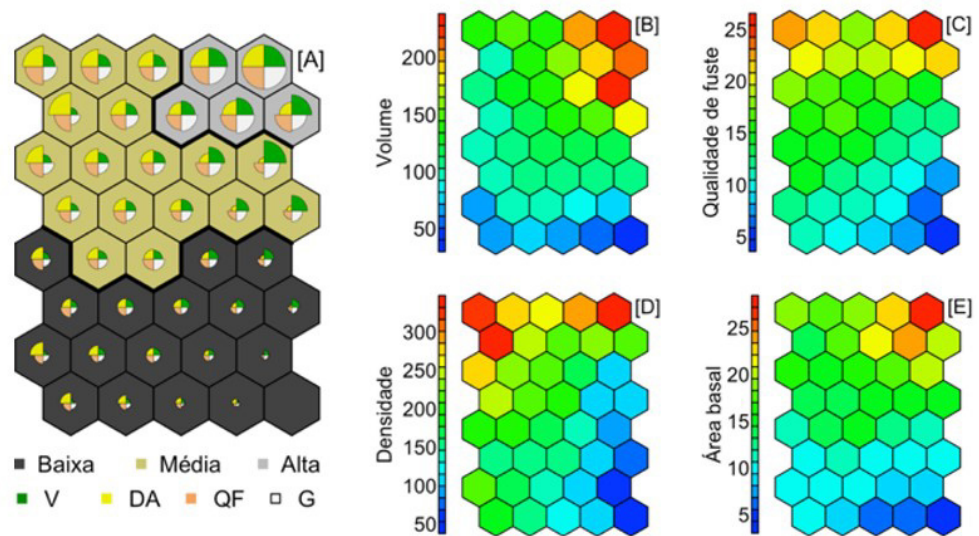


Figura 2. Mapa auto-organizável de Kohonen [A] e planos componentes para o volume [B], qualidade de fuste [C], densidade de árvores [D] e área basal [E], Floresta Nacional do Tapajós, estado do Pará.

Além de maiores acertos, antes e após a reclassificação, o método Kohonen apresenta a vantagem de uma avaliação rápida das características estruturais da área, pois neurônios pertencentes a mesma classe são constituídos de amostras com médias próximas. A semelhança pode ser verificada por meio do gráfico de pizza no centro dos neurônios, onde o tamanho de cada “pedaço da pizza” representa a magnitude da variável no vetor de pesos (Figura 2A). O “tamanho da pizza” indica a influência na delimitação das classes de produtividade. Os neurônios do MAK mostram que a classe de produtividade alta, localizada na região superior direita, é formada por parcelas que foram distribuídas em áreas com médias de V, DA, QF e G altas.

Além disso, considerando que a variável categórica QF foi calculada pela média ponderada arredondada do número de árvores nas classes de QF, acredita-se que nessas parcelas o V, G e DA são formados, principalmente por árvores com percentual de

aproveitamento de 80 - 100% e 50 - 79%. A presença de parcelas com médias de V, QF, DA e G altas são mais facilmente visualizadas nos planos componentes (Figura 2B-2E). Os planos componentes mostram também que as variáveis mudam ao longo do MAK no mesmo sentido das classes formadas, mostrando sua alta contribuição para formação das classes de produtividade.

As variáveis V, QF e G foram as que mais contribuíram para a formação da classe alta, pois as médias mais altas dessas variáveis estão localizadas nos mesmos neurônios que constituem a classe de produtividade alta. Eles permitem também avaliação da correlação entre variáveis. A mudança da variável ao longo do plano componente permite ainda avaliar a correlação entre variáveis. O volume é mais correlacionado com área basal e qualidade de fuste do que com a densidade de árvores. A densidade de árvores é mais correlacionada com a qualidade de fuste que com o volume e área basal.

A avaliação das classes média e baixa seguem o mesmo raciocínio, ou seja, formadas por parcelas com valores de V, DA, QF e G médios e baixos, respectivamente. Todavia, é possível que em alguns neurônios a magnitude da variável seja maior que o esperado para a classe, mas isso ocorre por causa da magnitude das outras variáveis na formação da classe de produtividade. A densidade de árvores, por exemplo, apresenta altos valores nos neurônios do canto superior esquerdo da classe média, indicando, provavelmente, alta abundância de árvores com menor diâmetro (Figura 2D).

A similaridade de alocação de parcelas nas classes de produtividade calculadas pelos métodos Kohonen e Souza foi de 74,5% (Tabela 2). A classe alta apresentou a menor percentagem de similaridade, com 50%, pois 4 de 10 amostras foram alocadas na classe média por meio do método Souza. Essas mesmas amostras foram alocadas na classe alta pelo método Kohonen. Além disso, uma amostra, com volume de $260,6 \text{ m}^3 \text{ ha}^{-1}$, alocada na classe alta pelo método Kohonen, foi alocada na baixa pelo método Souza. Acredita-se que isso seja um "erro de classificação discrepante", pois a média volumétrica da classe baixa do método de Souza foi de $89,6 \text{ m}^3 \text{ ha}^{-1}$.

Tabela 2. Similaridade de alocação das amostras nas classes de produtividade calculadas por meio dos métodos Kohonen e Souza, Floresta Nacional do Tapajós, estado do Pará.

Comparação	Classes	Classes			Similaridade (%)
		Baixa	Média	Alta	
Kohonen vs. Souza	Baixa	50	6	0	89,3
	Média	9	27	8	61,4
	Alta	1	4	5	50,0
Total	-	60	37	13	74,5

O esperado é que as amostras alocadas na classe baixa apresentassem volumes próximos à média volumétrica de sua classe, ou seja, $89,6 \text{ m}^3 \text{ ha}^{-1}$. Isso também ocorreu na classe média do método Souza, onde parcelas que supostamente estão mais próximas aos limites da classe média e superior foram alocadas nas classes de produtividade inferior (Figura 3A). O método de Kohonen não apresentou erros de classificação excessivos, ou seja, de amostras com altos volumes alocados na classe baixa, ou de baixo volume na classe alta. Foi observado apenas que algumas amostras da classe média apresentaram volume maiores que algumas da classe alta, mas isso era esperado devido ao uso de múltiplas variáveis para a formação das classes de produtividade (Figura 3B).

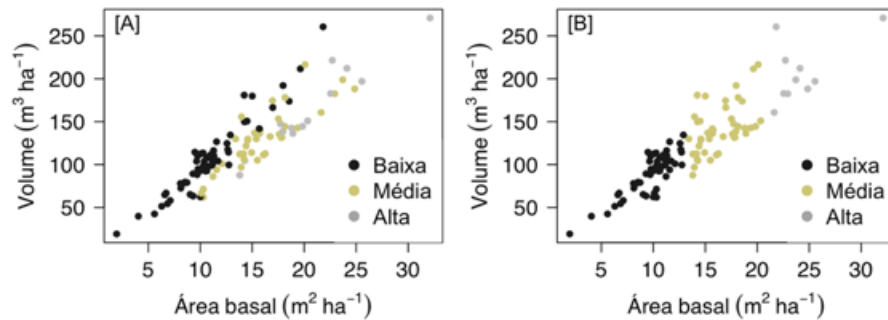


Figura 3. Distribuição de amostras nas classes de produtividade obtidas pelos métodos Souza [A] e Kohonen [B], Floresta Nacional do Tapajós, estado do Pará.

Ademais, embora a similaridade entre os métodos tenha sido de 74,5%, as médias de volume das classes de produtividade baixa e alta calculadas pelos métodos Souza e Kohonen foram significativamente diferentes, a 95% de probabilidade (Tabela 3). Em relação à área basal e densidade de árvores, apenas a classe de produtividade alta apresentou diferenças. Esses resultados mostram que as classes de produtividade não são iguais, principalmente naquela de maior interesse, ou seja, a de maior volume comercial.

Tabela 3. Volume, área basal, densidade de árvore e teste t de Welch para as classes de produtividade obtidas pelos métodos Souza e Kohonen, Floresta Nacional do Tapajós, estado do Pará.

Classes	Volume (m³ ha⁻¹)		Área basal (m² ha⁻¹)		Densidade (árv ha⁻¹)	
	Kohonen	Souza	Kohonen	Souza	Kohonen	Souza
Baixa	89,60 a ¹	106,0 b	9,8 a	10,9 a	126,8 a	117,6 a
Média	141,4 a	130,6 a	16,4 a	16,0 a	222,1 a	221,5 a
Alta	207,6 a	164,4 b	24,2 a	20,7 b	260,1 a	321,4 b

¹Médias com letras diferentes nas linhas representam as diferenças significativas entre classes de produtividade.

Ambos os métodos são eficientes para a classificação da capacidade produtiva de florestas inequidêneas. Mas, algumas considerações devem ser realizadas. O método Souza apresentou os chamados erros de classificação discrepantes, causados provavelmente pelo uso de três ou mais variáveis na análise de *cluster*. Esse método é comumente calculado apenas com o volume (Lima et al., 2014; Souza & Souza, 2006) ou volume e área basal (Corrêa et al., 2018; Corrêa et al., 2020; Souza et al., 2014), mostrando alta eficiência e ausência de erros de classificação discrepantes.

Assim, se a opção for o método Souza, recomenda-se que a matriz X seja formada apenas com a variável volume, conforme Souza & Souza (2006), o que permite a determinação da amplitude volumétrica das classes de produtividade. Ao se utilizar outras variáveis além do volume, a amplitude volumétrica tornar-se-á indeterminável, fazendo com que as classes sejam representadas pela média volumétrica das amostras que as constituem. Além disso, o método Souza não prevê a classe de produtividade de uma nova amostra, caso necessário (Rhys, 2020). Isso é importante quando se adiciona novas unidades de gerenciamento florestal (e.g., UTs) e deseja-se saber sua classe de produtividade.

O método Kohonen minimizou os erros de classificação discrepantes e o problema de multiplicidade de variáveis constatados para o método Souza. Além disso, prevê a classe de produtividade de novas amostras (Rhys, 2020) e produz estruturas de visualização muito mais informativas (e.g., mapa auto-organizável e planos componentes) que as partições simples (e.g., dendrograma) retornadas no método Souza. Isso permite uma ligeira e precisa avaliação dos parâmetros estruturais da vegetação durante a tomada de decisão no gerenciamento de florestas inequidêneas.

A desvantagem do método Kohonen é sua repetitividade (Giraudel & Lek, 2001; Klobucar & Subasic, 2012), pois durante a inicialização do algoritmo de Kohonen as amostras são conectadas a um vetor de pesos originado aleatoriamente, fazendo com que o mesmo conjunto de amostras produza mapas diferentes. Mas, se os parâmetros topológicos foram determinados corretamente, os vários mapas auto-organizáveis serão diferentes, mas as posições dos neurônios será aproximadamente a mesma. Ademais, o método Kohonen não é recomendado para classificação da capacidade produtiva de áreas com poucas amostras, pois MAK pequenos são difíceis de serem interpretados (Rhys, 2020). Isso não parece ser um problema para o método de Souza.

As classes de produtividade calculadas por meio dos métodos Souza e Kohonen são úteis para diversos propósitos, a saber: determinação de corte; planejamento e realização de atividades de colheita, pois se colhe mais em área de alta a média produtividade (Corrêa et al., 2018; Corrêa et al., 2020); aplicação de tratamentos silviculturais (e.g., desbastes e enriquecimento de espécies comerciais), uma vez que sua aplicação é concentrada em área de baixa produtividade; e para redução de custos de inventários florestais, pois obtêm-se classes mais homogêneas que toda a população. Isso promove a redução da variância total, aumenta a precisão das estimativas de um inventário florestal e reduz o tamanho da amostra, o que consequentemente reduz o tempo e os recursos humanos e financeiros na execução de inventários por amostragem em até 50%.

CONCLUSÕES

Ambos os métodos podem ser utilizados para a classificação da capacidade produtiva de florestas inequidistantes. Mas, o método Kohonen apresenta maiores vantagens, a saber: maiores acertos na classificação das unidades de amostrais dentro das classes de produtividade; menores chances de obter-se erros de classificação discrepantes; possibilidade de classificação de unidades amostrais que não estavam presentes durante o treinamento do mapa auto-organizável; e produz estruturas de visualização mais informativas. O método Souza pode ser utilizado, mas preferencialmente deve ser calculado com poucas variáveis, para que não ocorra os aqui chamados erros de classificação discrepantes.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6), 711-728. <http://dx.doi.org/10.1127/0941-2948/2013/0507>.
- Andrade, D. F., Gama, J. R. V., Melo, L. O., & Ruschel, A. R. (2015). Inventário florestal de grandes áreas na Floresta Nacional do Tapajós, Pará, Amazônia, Brasil. *Biota Amazônia*, 5(1), 109-115. <http://dx.doi.org/10.18561/2179-5746/biotaamazonia.v5n1p109-115>.
- Annas, S., Kanai, T., & Koyama, S. (2007). Principal component analysis and self-organizing map for visualizing and classifying fire risks in forest regions. *Agricultural Information Research*, 16(2), 44-51. <http://dx.doi.org/10.3173/air.16.44>.
- Asan, U., & Ercan, S. (2012). An introduction to self-organizing maps. In: C. Kahraman (Ed.), *Computational intelligence systems in industrial engineering* (pp. 295-315). Paris: Atlantis Press. http://dx.doi.org/10.2991/978-94-91216-77-0_14.
- Barbosa, G. P., Nogueira, G. S., Oliveira, M. L. R., Machado, E. L. M., Castro, R. V. O., & Dutra, G. C. (2017). Pós-estratificação em inventário florestal da vegetação arbórea-arbustiva. *Scientia Forestalis*, 45(115), 445-453. <http://dx.doi.org/10.18671/scifor.v45n115.03>.
- Campos, J. C. C., & Leite, H. G. (2017). *Mensuração florestal: perguntas e respostas*. (363 p.). Viçosa: Editora UFV.
- Chase, M. W., Christenhusz, M. J. M., Fay, M. F., Byng, J. W., Judd, W. S., Soltis, D. E., Mabberley, D. J., Sennikov, A. N., Soltis, P. S., & Stevens, P. F. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*, 181(1), 1-20. <http://dx.doi.org/10.1111/boj.12385>.

- Corrêa, K. K. S., Gama, J. R. V., Silva-Ribeiro, R. B. S., & Ximenes, L. C. (2018). Ajuste e classificação do potencial volumétrico de *Lecythis lurida* (Miers) Mori, Flona do Tapajós. *Nativa*, 6(4), 395-401. <http://dx.doi.org/10.31413/nativa.v6i4.5450>.
- Corrêa, K. K. S., Ximenes, L. C., Ribeiro, R. B. S., & Gama, J. R. V. (2020). Classificação do potencial volumétrico de uma área submetida a manejo florestal na Floresta Nacional do Tapajós. *Revista Ibero-Americana de Ciências Ambientais*, 11(2), 43-52. <http://dx.doi.org/10.6008/CBPC2179-6858.2020.002.0005>.
- Giraudel, J. L., & Lek, S. (2001). A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling*, 146(1-3), 329-339. [http://dx.doi.org/10.1016/S0304-3800\(01\)00324-6](http://dx.doi.org/10.1016/S0304-3800(01)00324-6).
- Klobucar, D., & Subasic, M. (2012). Using self-organizing maps in the visualization and analysis of forest inventory. *iForest - Biogeosciences and Forestry*, 5(5), 216-223. <http://dx.doi.org/10.3832/ifer0629-005>.
- Kohonen, T. (2001). *Self-organizing maps*. (502 p.). Berlin: Springer. <http://dx.doi.org/10.1007/978-3-642-56927-2>.
- Lima, R. B., Aparício, P. S., Ferreira, R. L. C., Silva, W. C., Guedes, M. C., Oliveira, C. P., Silva, D. A. S., & Batista, A. P. B. (2014). Volumetria e classificação da capacidade produtiva para *Mora paraensis* (Ducke) no estuário amapaense. *Scientia Forestalis*, 42(101), 141-154.
- Loetsch, F., Haller, K. E., Zöhrer, F., Brünig, E. F., & Panzer, K. F. (1964). *Forest inventory*. (436 p.). Munich: BLV Verlagsgesellschaft.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., & Wagner, H. (2019). *Vegan: community ecology package. Version 2.5-6*. Vienna: R Foundation for Statistical Computing.
- R Development Core Team. (2022). *R: a language and environment for statistical computing. Version 4.1.2*. Vienna: R Foundation for Statistical Computing.
- Rhys, H. I. (2020). *Machine learning with R, the tidyverse, and mlr*. (536 p.). Shelter Island: Manning Publications.
- Santos, L. S., Martorano, L. G., Gutierrez, C. B. B., Pontes, A. N., Silva, O. M., & Gutierrez, D. M. G. (2017). Aspecto fisiográficos da Floresta Nacional do Tapajós e seu entorno - oeste do Pará, Brasil. *Revista Espacios*, 38(2), 26.
- Souza, A. L., & Souza, D. R. (2006). Análise multivariada para estratificação volumétrica de uma floresta ombrófila densa de terra firme, Amazônia Oriental. *Revista Árvore*, 30(1), 49-54. <http://dx.doi.org/10.1590/S0100-67622006000100007>.
- Souza, A. L., Medeiros, R. M., Matos, L. M. S., Silva, K. R., Corrêa, P. A., & Faria, F. N. (2014). Estratificação volumétrica por classes de estoque em uma floresta ombrófila densa, no município de Almeirim, estado do Pará, Brasil. *Revista Árvore*, 38(3), 533-541. <http://dx.doi.org/10.1590/S0100-67622014000300016>.
- Souza, D. R., & Souza, A. L. (2004). Estratificação vertical em floresta ombrófila densa de terra firme não explorada, Amazônia Oriental. *Revista Árvore*, 28(5), 691-698. <http://dx.doi.org/10.1590/S0100-67622004000500008>.
- Stümer, W., Kenter, B., & Köhl, M. (2010). Spatial interpolation of in situ data by self-organizing map algorithms (neural networks) for the assessment of carbon stocks in European forests. *Forest Ecology and Management*, 260(3), 287-293. <http://dx.doi.org/10.1016/j.foreco.2010.04.008>.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. (498 p.). New York: Springer. <http://dx.doi.org/10.1007/978-0-387-21706-2>.
- Vesanto, J., Himberg, J., Alhoniemi, I. E., & Parhankangas, J. (2000). SOM toolbox for Matlab 5 documentation. In *Proceedings of the Matlab DSP Conference* (pp. 35-40). Espoo: Helsinki University of Technology.
- Wehrens, R., & Kruisselbrink, J. (2018). Flexible self-organizing maps in kohonen 3.0. *Journal of Statistical Software*, 87(7), 1-18. <http://dx.doi.org/10.18637/jss.v087.i07>.

Contribuição dos Autores: DSV: conceptualization, funding acquisition, methodology, writing, data curation, formal analysis; MLRO, JRVG and BOL: conceptualization, methodology, supervision.