

Influência dos parâmetros da metaheurística algoritmo genético em um problema de planejamento florestal

Lisandra Maria Alves Matos^{1*}, Carlos Alberto Araújo Júnior¹, Adriana Leandra de Assis¹, Christian Dias Cabacinha¹, Paulo Henrique Batista Ferreira¹, Emanuely Canabrava Magalhães¹

¹ Universidade Federal de Minas Gerais, Avenida Presidente Antônio Carlos, 6627, Pampulha, Belo Horizonte - MG, 31270-901

*Autor para correspondência: lisandraamatos@gmail.com

Recebido: Fevereiro 2018 / Aceito: Outubro 2019 / Publicado: Dezembro 2019

Resumo

Este trabalho objetivou avaliar o comportamento de diferentes parâmetros da metaheurística Algoritmo Genético para solução de problemas do planejamento florestal. O estudo considerou um horizonte de planejamento de 16 anos, analisando-se como parâmetros da metaheurística a utilização de elitismo, o tipo de cruzamento (1 ponto de corte e uniforme), o tipo de seleção dos pais para cruzamento (roleta e torneio) e tipo de mutação (escolha aleatória do gene e gene a gene), totalizando 16 combinações. Adicionalmente foram avaliados o tamanho da população inicial (20, 50 e 80 indivíduos) e o critério de parada (100, 300 e 500 gerações), totalizando 9 combinações. Cada uma dessas combinações foi considerada como um tratamento e processada com 30 repetições. O problema objetivou encontrar um cenário de planejamento da produção florestal que retornasse o máximo valor presente líquido sujeito às restrições de idade de corte (entre 5 e 7 anos), demandas mínima e máxima anual (140.000 m³ e 160.000 m³, respectivamente) e integridade. O processamento foi efetuado utilizando-se o software MeP. Aplicou-se o teste não-paramétrico de Kruskal-Wallis. Os resultados mostram que há diferença estatisticamente significativa, a 5% de probabilidade, entre as combinações dos parâmetros. Conclui-se que a qualidade das soluções geradas pela metaheurística algoritmo genético é dependente da seleção de uma boa configuração de seus parâmetros. A melhor solução foi encontrada utilizando a seguinte configuração: elitismo, cruzamento de 1 ponto, seleção do tipo torneio, mutação do tipo gene a gene, população inicial com 50 indivíduos e critério de parada com 500 gerações.

Palavras-chave: Inteligência artificial, Planejamento florestal, Manejo e gerenciamento de florestas, Programação matemática.

Abstract

This work aimed to evaluate the behavior of different parameters of the Genetic Algorithm metaheuristic to solve problems of forest planning. The study was developed considering a planning horizon of 16 years, analyzing as parameters of the metaheuristic the use of elitism, the type of crossing (1 cut-off point and uniform), the type of parent selection for crossing (roulette and tournament) and mutation type (random gene choice and gene to gene), totaling 16 combinations. In addition, the initial population size (20, 50 and 80 individuals) and the stopping criterion (100, 300 and 500 generations) were evaluated, totaling 9 combinations. Each of these combinations of genetic algorithm parameters was considered as a different treatment and processed with 30 replicates. The problem was to find a scenario of forest production planning that would return the maximum Net Present Value subject to the restrictions of age of cut (between 5 and 7 years), minimum and maximum annual demands (140,000 m³ and 160,000 m³, respectively) and integrity. Processing was done using MeP software. To compare the results, the non-parametric Kruskal-Wallis test

was used. The results show that there is a statistically significant difference, at 5% probability, between the parameter combinations. We conclude that the quality of the solutions generated by the metaheuristic genetic algorithm is dependent on the selection of a good configuration of its parameters. The best solution was found using the following configuration: elitism, 1-point crossing, tournament-type selection, gene-to-gene mutation, initial population with 50 individuals, and 500-generation stop criterion.

Keywords: Artificial intelligence, Forest planning, Forest management, Mathematical programming.

Introdução

O manejo florestal está associado à utilização de forma sustentável e sensata dos recursos florestais, de maneira que as gerações futuras possam usufruir ao menos dos mesmos benefícios da geração presente (Scolforo 1998). Nesse contexto, é necessário que os gestores florestais conheçam as técnicas de planejamento e reconheçam sua importância para o sucesso do empreendimento. É por meio do planejamento que é possível organizar, racionalizar e otimizar as operações, tornando possível abordar todos os fatores que podem atingir positivamente ou negativamente as atividades, identificando e antecipando alguns problemas que viriam a afetá-las. Em geral, o planejamento é feito seguindo uma estrutura hierárquica, sendo dividido em estratégico (longo prazo), tático (médio prazo) e operacional (curto prazo).

O processo de produção florestal é caracterizado pelo longo prazo, principalmente pelo período de maturação da matéria-prima trabalhada, e isso é uma das dificuldades que tornam ainda mais complexos os problemas de planejamento florestal, introduzindo um maior grau de incerteza na previsão dos fatores envolvidos e consequente instabilidade (Arce 1997). Esse fato gera insatisfação nas empresas florestais quanto à correlação entre as ações planejadas e as executadas, principalmente quando a questão é sequenciamento de atividades.

A longo prazo, um dos quesitos mais importantes no planejamento florestal é a definição do modelo de regulação considerado e a indicação das ferramentas de programação matemática mais apropriada para solucioná-lo. O conceito de floresta regulada está ligado à sustentabilidade, com a viabilização da continuidade do processo e dos usos associados aos produtos madeireiros ou não-madeireiros. A conversão da floresta não regulada para uma situação que o seja é a chamada regulação florestal. Pode-se dizer, então, que o principal objetivo da regulação florestal seja a definição e a aplicação de prescrições de manejo, de modo que se tenha o suprimento constante e contínuo da demanda de bens e serviços dentro dos limites da capacidade produtiva de cada área (Davis 1966; Carvalho 2015).

A complexidade, em conjunto à quantidade de variáveis envolvidas nos modelos de regulação, incentivou a introdução de diversas ferramentas da pesquisa operacional na área florestal, possibilitando a redução de riscos e incerteza

na tomada de decisão (Rodrigues 2004). Dentre as técnicas aplicadas no manejo florestal estão a Programação Linear (PL) e a Programação Linear Inteira (PLI) (Bettinger et al. 2009; Troncoso et al. 2016). Apesar da grande aplicação da PL em casos de ordenamento e otimização da produção florestal, a necessidade de incorporação da restrição de integridade nos modelos de regulação florestal tem levado à limitação do uso da técnica (Silva et al. 2003). Dessa forma, modelo de programação inteira (PLI) utilizando algoritmos exatos (branch and bound) e metaheurísticas (algoritmo genético, busca tabu, simulated annealing) tem sido aplicados na resolução de problemas de gerenciamento de recursos florestais (Bettinger et al. 2009; Yoshimoto et al. 2016; Jin et al. 2016; Dong et al. 2016; Ezquerro et al. 2016; Kaya et al. 2016).

A metaheurística algoritmo genético (AG) é uma ferramenta de busca e otimização que se baseia nos preceitos Darwinianos de seleção natural e reprodução genética, utilizando princípios de sobrevivência e reprodução dos indivíduos mais aptos para produzir melhores soluções a cada geração (Hillier e Lieberman 2006). O AG se enquadra nos métodos heurísticos por se basear em um procedimento que provavelmente vai encontrar uma excelente solução viável, mas não necessariamente uma solução ótima para o problema em questão. Porém, a geração de uma solução para um problema complexo mais rapidamente, por estar ligada a um bom desempenho e performance computacional, é uma importante razão pela qual a metaheurística é selecionada como uma ferramenta de planejamento, mesmo não garantindo a otimalidade da resposta (Bettinger et al. 2009).

A maior dificuldade em se utilizar esse algoritmo gira em torno do ajuste de seus parâmetros (Gomide et al. 2009). Por isso, é necessário que haja o desenvolvimento de pesquisas que forneçam subsídios para escolha das melhores configurações para os algoritmos de otimização que vêm sendo utilizados no planejamento florestal. Nesse contexto, o principal objetivo do presente trabalho foi identificar a melhor parametrização para a metaheurística algoritmo genético por meio de combinações de seus parâmetros para solução de um problema de ordenamento da produção florestal.

Material e Métodos

A área utilizada para as simulações foi de 4.269,29 ha, distribuída em 120 talhões, com uma distribuição irregular para as classes de idades existentes no povoamento: 339 ha com 1 ano; 768 ha com 2 anos; 1031 ha com 3 anos; 601 ha com 4 anos; 958 ha com 5 anos; e 513 ha com 6 anos.

Os valores de custos silviculturais foram obtidos de Araújo Júnior (2012): R\$ 4.059,05 ha⁻¹ no primeiro ano; R\$ 1.627,81 ha⁻¹ no segundo ano; R\$ 757,95 ha⁻¹ no terceiro ano; e R\$ 88,12 ha⁻¹ a partir do quarto ano do desenvolvimento dos plantios. Os valores de receita de venda da madeira foram: R\$ 20,00 m⁻³ para florestas com idades até 3 anos; R\$ 30,00 m⁻³ para florestas com 4 anos de idade; R\$ 40,00 m⁻³ para florestas com 5 anos de idade; e R\$ 80,00 m⁻³ para florestas com idade igual ou superior a 6 anos.

O problema proposto tem como objetivo encontrar um cenário de planejamento da produção florestal que retorne o máximo Valor Presente Líquido (VPL). O modelo matemático de otimização considerou a estrutura proposta por Johnson e Scheurman (1977):

$$\text{Max VPL} = \sum_{i=1}^M \sum_{j=1}^P \text{VPL}_{ij} \cdot X_{ij} \quad (1)$$

Sujeito a

$$\sum_{i=1}^M \sum_{j=1}^P X_{ij} = 1 \quad (2)$$

$$\sum_{i=1}^M \sum_{j=1}^P V_{ij} X_{ij} \geq \text{Dmin}_k \quad (3)$$

$$\sum_{i=1}^M \sum_{j=1}^P V_{ij} X_{ij} \leq \text{Dmax}_k \quad (4)$$

$$X_{ij} \in \{0,1\} \quad (5)$$

em que: VPL_{ij} é o VPL da unidade de manejo *i* segundo o regime de manejo *j*; X_{ij} é a variável de decisão do tipo binária, recebendo valor 1 quando a alternativa de manejo *j* é assinalada à unidade de manejo *i* ou 0, caso contrário; M é o número total de unidades de manejo; P é o número total de alternativas de manejo para a unidade de manejo *i*; *k* é o período do horizonte de planejamento; Dmin_k e Dmax_k são, respectivamente, a demanda mínima e máxima de madeira (m³) no período *k* do horizonte de planejamento.

A função objetivo (1) busca maximizar o retorno financeiro e está sujeita à restrição binária (5), restrição de integridade (2), onde cada uma das unidades de manejo só devem receber uma única prescrição e restrição de demanda volumétrica, onde o volume colhido durante o ano deveria respeitar um limite mínimo (3) e máximo (4). Para o presente trabalho foi estabelecida uma demanda anual máxima de 160.000 m³ e mínima de 140.000 m³, permitido o corte das unidades de manejo com idades entre 5 e 7 anos e definido um horizonte de planejamento com 16 anos. A taxa de juros considerada foi de 8%.

A função de aptidão do AG utilizado levou em consideração o método de penalidades, conforme Gomide et al. (2009), Rodrigues (2004) e Binoti et al. (2014). O valor adotado nesse trabalho foi de R\$ 500,00/m³ por volume excedido ou faltante em relação ao intervalo pré-estabelecido pelas restrições impostas ao modelo. Assim, a equação que representa a função objetivo é dada por:

$$\text{Max VPL} = \sum_{i=1}^M \sum_{j=1}^P \text{VPL}_{ij} \cdot X_{ij} - P \cdot \sum_{k=1}^T d_k \quad (6)$$

em que: P é a penalidade em R\$/m³ de madeira; T é o tempo total do planejamento; e d_k é o valor absoluto do desvio volumétrico de madeira em m³ em um período *k* do horizonte de planejamento.

Foram avaliadas 16 diferentes combinações de parâmetros do AG, com o objetivo de chegar à melhor delas. Em relação à seleção dos pais para cruzamento, avaliou-se o método da roleta (Ro) e o método do torneio (To). No processo de reprodução considerou-se o cruzamento com um ponto de ruptura (1p) e o cruzamento uniforme (Un). A taxa de cruzamento considerada foi de 80%. Já para o processo de mutação avaliou-se a mutação aleatória (Al) e a mutação gene a gene (Ge), sempre com taxa de mutação igual a 1%. Ainda, a seleção dos indivíduos para a próxima geração foi feita com (Ce) e sem elitismo (Se). Para análise dos resultados, considerou-se o melhor indivíduo de cada repetição, independente da geração na qual o mesmo foi obtido.

Inicialmente foram testadas todas as configurações com 50 indivíduos na população inicial e critério de parada em 100 gerações. Posteriormente, estabelecida a melhor combinação dos parâmetros, foram testadas populações com 20, 50 e 80 indivíduos e critério de parada com 100, 300 e 500 gerações. O principal objetivo de testar tamanhos variados foi analisar melhor de que forma esse parâmetro afeta a qualidade da resposta final e sua eficiência quanto ao tempo de processamento.

O processamento de todas as configurações foi efetuado através do aplicativo MeP - *Metaheuristics for Forest Planning*, em desenvolvimento pelo Laboratório de Pesquisa Operacional e Modelagem Florestal (LPM) da Universidade Federal de Minas Gerais.

Os dados de *fitness* e tempo de processamento foram analisados a partir do valor mínimo, médio e máximo de cada configuração avaliada, proveniente de 30 repetições. Na análise foi aplicado o teste não paramétrico desenvolvido por

Kruskal e Wallis (1952), o qual utiliza *ranks* dos dados ao invés de seus valores brutos para o cálculo da estatística de teste. Para a comparação múltipla, foi utilizado o teste de Tukey, conforme indicado em Zar (1999).

Assim como Rodrigues et al. (2004), para analisar a acurácia do resultado da melhor combinação de parâmetros, o mesmo foi comparado à solução gerada pelo modelo de programação inteira *branch and bound* processado no *software* Lingo, interrompido após 96 horas de processamento. Dessa forma, a eficácia (*Ef*) foi calculada segundo a fórmula:

$$Ef = \frac{f_{AG}}{f_0} \times 100 \quad (7)$$

Tabela 1 - Valores máximo, médio, mínimo, desvio padrão e eficácia dos *fitness* gerados pelas combinações entre os parâmetros do algoritmo genético.

Combinação	Máximo (R\$)	Médio (R\$)	Mínimo (R\$)	Desvio (R\$)	Eficácia (%)
04 - To_1p_Ge_Ce	30.582.271	30.336.526	29.950.852	127.846	95,7
02 - Ro_1p_Ge_Ce	30.514.441	30.268.322	29.981.728	130.474	95,5
06 - Ro_Un_Ge_Ce	30.514.023	30.267.666	29.991.120	117.126	95,5
08 - To_Un_Ge_Ce	30.528.960	30.280.900	29.922.636	122.637	95,5
16 - To_Un_Ge_Se	30.527.415	30.065.664	29.665.328	177.517	95,5
12 - To_1p_Ge_Se	30.385.421	29.983.441	29.639.960	169.485	95,1
01 - Ro_1p_Al_Ce	30.290.620	28.666.138	24.667.924	1.502.946	94,8
11 - To_1p_Al_Se	30.228.176	26.868.620	10.451.915	3.848.382	94,6
03 - To_1p_Al_Ce	30.187.268	26.824.866	19.368.386	3.123.281	94,4
05 - Ro_Un_Al_Ce	27.280.659	9.252.245	-14.504.788	9.732.754	85,3
09 - Ro_1p_Al_Se	25.753.525	4.835.537	-9.627.649	7.916.626	80,6
07 - To_Un_Al_Ce	25.704.270	9.310.407	-20.566.389	9.912.923	80,4
15 - To_Un_Al_Se	23.168.314	1.061.475	-21.316.736	11.776.938	72,5
10 - Ro_1p_Ge_Se	19.402.347	2.911.672	-14.093.762	7.511.151	60,7
14 - Ro_Un_Ge_Se	14.542.186	-4.510.935	-19.343.447	9.561.605	45,5
13 - Ro_Un_Al_Se	12.185.194	-31.352.348	-57.713.235	15.575.244	38,1

A melhor solução obtida pela combinação 4 atendeu à restrição de regulação de acordo com a demanda anual definida, enquanto a melhor solução pela combinação 13 apresentou irregularidades em termos de produção florestal ao longo dos anos. A diferença mostrada confirma que a configuração correta dos parâmetros da metaheurística é relevante para eficiência e funcionamento dos algoritmos genéticos.

Nesse contexto de comparação entre as combinações é importante avaliar o comportamento dos resultados ao longo das gerações (Figura 2). Pode-se notar que ocorreram diferenças significativas em relação à evolução das soluções

em que: f_{AG} é o valor da melhor solução obtida pelo algoritmo genético (R\$); e f_0 é o valor da melhor solução gerada pelo algoritmo de *branch and bound* (R\$).

Resultados e discussão

A eficácia das combinações variou de 38,1% (combinação 13) a 95,7% (combinação 4), o que ressalta a importância da correta escolha da configuração dos parâmetros na performance do algoritmo (Tabela 1). A solução encontrada pelo algoritmo de *branch and bound* apresentou um valor igual a R\$ 31.964.100.

Ainda sobre a expressiva variação entre os *fitness* gerados pelas combinações testadas, a Figura 1 mostra a produção em volume ao longo dos 16 anos de horizonte de planejamento considerando a configuração com a pior eficácia, combinação 13 (Figura 1a), e a de melhor eficácia, combinação 4 (Figura 1b). A melhor solução obtida pela combinação 4 atendeu à (Figura 1b).

obtidas em cada configuração. As figuras (i), (j), (m) e (n) mostram que não houve uma melhoria contínua das soluções encontradas pelo algoritmo ao longo das gerações, variando em torno de um valor médio abaixo de zero. Isso pode ser associado principalmente à combinação do método de seleção por roleta e a não utilização do elitismo. Em observação aos resultados obtidos nesse trabalho, é possível notar que o mecanismo de seleção tanto por roleta quanto por torneio apresenta diferentes comportamentos quando há ou não a presença do elitismo. Apesar disso, os demais gráficos mostraram tendência à obtenção de indivíduos mais aptos com o passar das gerações.

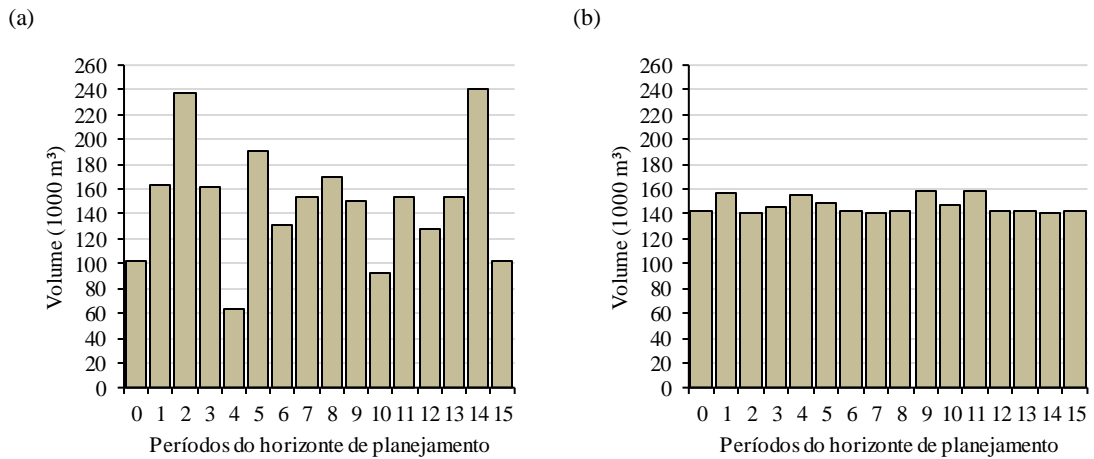
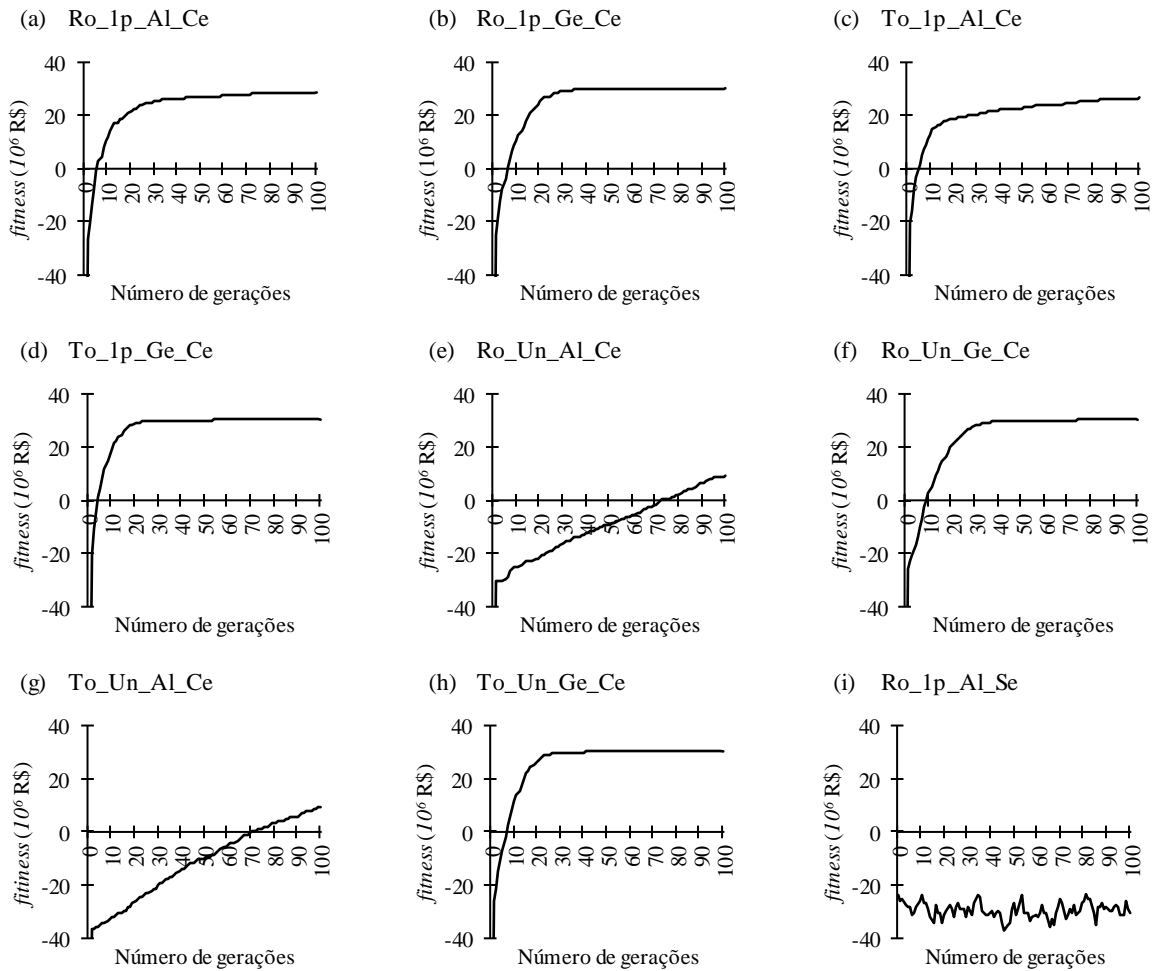


Figura 1 – Volumes obtidos pela combinação 13 (a) e pela combinação 4 (b) distribuídos ao longo dos 16 anos do horizonte de planejamento.



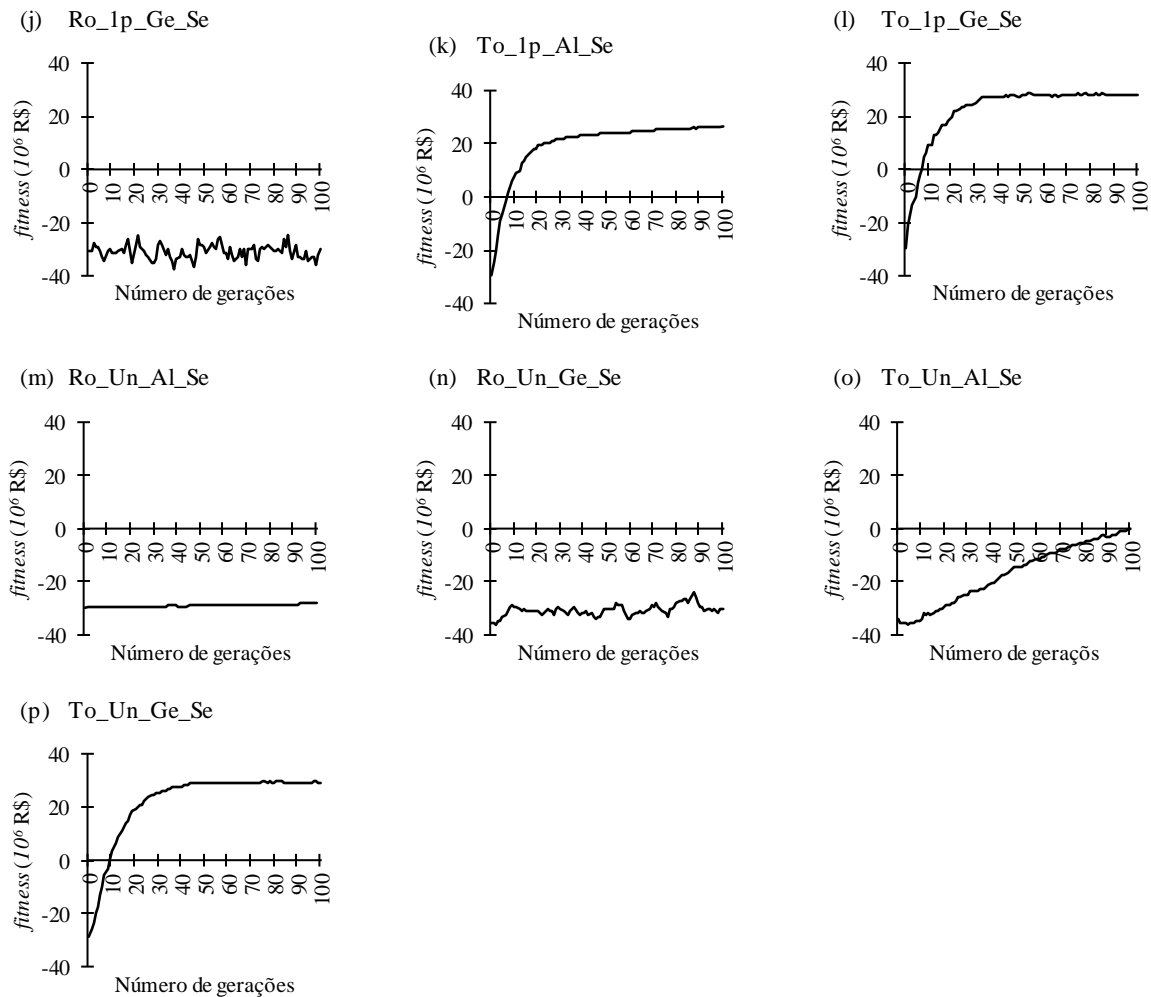


Figura 2 - Estrutura da evolução das buscas pelos valores médios de *fitness* obtidos pelo AG nas 16 combinações testadas ao longo de 100 gerações.

Os resultados do teste Kruskal Wallis, juntamente com o teste de Tukey para comparar as combinações entre os métodos de seleção, cruzamento, mutação e elitismo, apresentaram diferenças estatísticas significativas entre si para o problema analisado a um nível de significância igual a 5% (Tabela 2). A combinação 4 não foi estatisticamente diferente da combinação 8, que por sua vez não apresentou diferenças

estatisticamente significativas em relação às combinações 2 e 6. O tempo de processamento não apresentou grandes variações, sendo 2,47, 2,78, 2,52 e 2,43 minutos para as combinações 2, 4, 6 e 8, respectivamente. Apesar da semelhança, a combinação 4 tende a ser mais onerosa quanto ao tempo.

Tabela 2 - Resumo da análise estatística não paramétrica Kruskal Wallis aplicado nas variáveis *fitness* para as 16 combinações dos parâmetros.

Combinação	rank	Fitness médio (R\$)		Combinação	rank	Fitness médio (R\$)	
04 - To_1p_Ge_Ce	428,53	30.336.526	a	03 - To_1p_Al_Ce	255,00	26.824.866	d
08 - To_Un_Ge_Ce	409,70	30.280.900	ab	07 - To_Un_Al_Ce	150,17	9.310.407	e
02 - Ro_1p_Ge_Ce	405,60	30.268.322	b	05 - Ro_Un_Al_Ce	148,13	9.252.245	e
06 - Ro_Un_Ge_Ce	405,53	30.267.666	b	09 - Ro_1p_Al_Se	122,30	4.835.537	f
16 - To_Un_Ge_Se	344,20	30.065.664	c	10 - Ro_1p_Ge_Se	114,70	2.911.672	f
12 - To_1p_Ge_Se	325,57	29.983.441	c	15 - To_Un_Al_Se	107,70	1.061.475	f
01 - Ro_1p_Al_Ce	274,83	28.666.138	d	14 - Ro_Un_Ge_Se	78,50	-4.510.935	g
11 - To_1p_Al_Se	255,13	26.868.620	d	13 - Ro_Un_Al_Se	22,40	-26.917.572	h

Médias seguidas pelas mesmas letras não apresentam diferenças estatisticamente significativas ($p > 0,05$).

Os métodos de seleção por torneio (*fitness* médio igual a R\$ 23.091.487) e por roleta (*fitness* médio igual a R\$ 9.346.634) apresentaram diferenças estatisticamente

significativas entre si a um nível de 5% de probabilidade. Lopes e Takahashi (2011) corroboram com o resultado encontrado nesse trabalho e ainda intitulam a seleção por

roleta como o “pior método existente”. A menos que haja uma boa razão para utilização de outro método, o torneio é certamente a melhor das escolhas (Rennard 2007; Sivanandam e Deepa, 2008), por ser o operador de seleção mais simples de ser implementado e ainda assim ser capaz de gerar resultados muito bons (Bento e Kagan 2008). No entanto, os resultados apresentados mostram que há mudança de comportamento dos métodos de seleção quando há presença ou não do mecanismo de substituição dos indivíduos para a geração seguinte, o elitismo. As combinações sem elitismo apresentaram melhores resultados quando utilizado o método de seleção por torneio, porém, quando analisados somente os dados com elitismo o método por roleta mostrou melhor desempenho.

Utilizando-se do método de elitismo em um problema de planejamento florestal, Rodrigues et al. (2004) testaram cinco métodos de seleção, Boltzman, escalonamento, ranqueamento, roleta e torneio, e o último foi o que gerou os piores resultados. Segundo os autores, a principal causa foi a aleatoriedade do método, não utilizando nenhuma informação do problema para seleção dos indivíduos ao longo da busca. Gomide et al. (2009) observaram o mesmo comportamento dos resultados, onde o método por roleta gerou melhores resultados que o método por torneio. Os autores alegam que o operador de seleção por roleta obteve uma maior diversidade de indivíduos ao longo das iterações e, por isso o fato de se obter as melhores respostas em *fitness*.

Os tipos de cruzamento também apresentaram diferenças estatísticas significativas entre si pelo teste de Kruskal Wallis a um nível de 5% de probabilidade. O teste apontou o mecanismo de 1 ponto (*fitness* médio igual a R\$ 22.586.960) como superior ao uniforme (*fitness* médio igual a R\$ 9.851.231), principalmente por apresentar apenas uma posição de ruptura, diferente do método uniforme que segue uma probabilidade para determinar de qual dos pais serão copiadas as características para geração do filho, possibilitando um maior número de cortes para cruzamento. No geral, quanto maior for o número de rupturas, menor a semelhança entre os pais e filhos, o que dificulta a convergência do algoritmo (Costa Filho e Poppi 1999).

Quando o problema apresenta genes com comportamento epistático, ou seja, exerce ação inibitória, o cruzamento uniforme deve ser evitado. Na florestal, isso pode ser dado como verdade, já que para respeitar-se as restrições, a seleção de uma unidade para corte depende da prescrição adotada pelas outras unidades de manejo. Por outro lado, tem sido mostrado que um único ponto de cruzamento pode ser problemático por talvez não ser suficiente para a criação de boas combinações. Nos casos em que esse problema pode acontecer, um *crossover* de 2 pontos é geralmente mais recomendado (Rennard 2007). Este tipo de cruzamento segue a mesma estrutura de funcionamento do tipo de 1 ponto, testado nesse trabalho.

Os tipos de mutação também apresentaram diferenças estatísticas significativas entre si pelo teste de Kruskal Wallis a um nível de 5% de probabilidade, apontando o método gene a gene (*fitness* médio igual a R\$ 22.450.407) como melhor que o aleatório (*fitness* médio igual a R\$ 9.987.715). O mecanismo de mutação gene a gene percorre cada uma das características associando um número aleatório que indica se o gene será ou não mutado seguindo uma probabilidade pré-definida, direcionando melhor a população a uma maior diversidade, e quanto maior essa diversidade, maiores são as chances de se obter indivíduos mais adaptados ao problema. A diferença entre os tipos de mutação também pode ser observada na Figura 2 pela comparação entre os gráficos (e) e (g) e os gráficos (f) e (h). Os dois primeiros tendem a

encontrar melhores resultados a cada nova geração por estar sendo utilizado o elitismo. Porém, seria necessário um número de gerações significativamente maior que o testado para se alcançar bons resultados. Os gráficos (f) e (h) por sua vez, mostram uma convergência em uma boa solução viável em um número muito menor de gerações.

As combinações que tiveram o elitismo introduzido (*fitness* médio igual a R\$ 24.400.884) se diferiram estatisticamente daquelas que não possuíam essa característica (*fitness* médio igual a R\$ 8.037.238), sendo a utilização da técnica mais recomendada. Muitos indivíduos podem ser perdidos se eles não forem selecionados para reprodução ou se eles forem destruídos pelo cruzamento ou mutação (Mitchell 1999; Mendes 2013). O elitismo é aplicado a fim de garantir a permanência do melhor indivíduo encontrado a cada nova geração.

Apesar da garantia de não se perder o melhor indivíduo durante o processo evolutivo, onde o resultado não piora, o elitismo possui a desvantagem de possibilidade de forçar a busca, pela presença de mais uma cópia do melhor indivíduo, na direção de algum ponto ótimo local que tenha sido descoberto antes do global. Nesses casos, considerar não se utilizar o parâmetro de elitismo pode ser interessante, uma vez que a melhor solução encontrada durante a evolução é guardada, para no final designá-lo como indivíduo ótimo encontrado, mesmo que ele não esteja presente na última geração da execução.

Identificada a melhor combinação entre os parâmetros elitismo, método de seleção, tipo de cruzamento e tipo de mutação, constatou-se variações entre as combinações entre tamanho da população inicial e número de gerações de 95,3% (combinação 1) a 97,1% (combinação 6), resultado obtido em menos de 12 minutos (Tabela 3). Silva et al. (2009) obtiveram menor e maior eficácia de 96,6% e 98,41%, respectivamente, enquanto Rodrigues et al. (2004) encontraram variação de eficácia entre 90,0% e 98,48% quando comparados diferentes tamanhos de população inicial.

Os resultados do teste Kruskal Wallis para comparar as combinações entre o tamanho da população inicial e o número de gerações apresentaram diferenças estatísticas significativas entre si para o problema analisado a um nível de significância igual a 5% (Tabela 4). Os resultados do teste mostraram que a combinação de número 9, que contava com 80 indivíduos na população inicial e 500 gerações como critério de parada, não se diferiu estatisticamente da combinação de número 6 (50 indivíduos na população inicial e 500 gerações), que por sua vez não se diferiu estatisticamente da combinação de número 8 (80 indivíduos e 300 gerações). Apesar da semelhança estatística, a combinação de número 6 demandou um menor tempo de processamento, pois gerou um bom resultado em 11,74 minutos enquanto o tempo gasto pelas combinações 8 e 9 foram de 16,91 e 27,57 minutos, respectivamente.

Pelo teste de Kruskal Wallis a um nível de significância de 5%, a população com apenas 20 indivíduos (*fitness* médio igual a R\$ 30.399.728) se diferiu das demais combinações, 50 (*fitness* médio igual a R\$ 30.520.497) e 80 (*fitness* médio igual a R\$ 30.568.458) indivíduos, que por sua vez não diferiram entre si. Esse parâmetro deve ser escolhido cautelosamente de modo a permitir que o algoritmo seja capaz de explorar todo o campo de busca. Quando a população é muito pequena o espaço de busca é limitado, provocando uma queda no desempenho. Porém, quanto maior a população, maior o tempo de processamento e os recursos computacionais necessários (Rodrigues et al. 2004).

Tabela 3 - Valores máximo, médio, mínimo, desvio padrão e eficácia dos *fitness* gerados e tempo de processamento das combinações entre tamanho da população inicial e número de gerações.

Combinação	Máximo (R\$)	Médio (R\$)	Mínimo (R\$)	Desvio (R\$)	Eficácia (%)	Tempo (min)
1 - 20_100	30.460.748	30.236.810	29.987.893	134.400	95,30	0,67
2 - 20_300	30.637.644	30.421.669	30.165.908	98.515	95,85	1,85
3 - 20_500	30.766.261	30.540.704	30.356.725	117.983	96,25	2,91
4 - 50_100	30.517.447	30.332.460	30.122.940	98.200	95,47	2,62
5 - 50_300	30.780.028	30.538.297	30.305.287	111.714	96,30	7,32
6 - 50_500	31.024.769	30.690.734	30.482.052	140.229	97,06	11,74
7 - 80_100	30.586.855	30.352.364	30.139.824	109.371	95,69	6,08
8 - 80_300	30.765.532	30.614.299	30.294.983	97.884	96,25	16,91
9 - 80_500	30.957.746	30.738.711	30.515.949	110.609	96,85	27,57

Tabela 4 - Resumo da análise estatística não paramétrica Kruskal Wallis aplicado às combinações entre tamanho de população inicial e número de gerações.

Combinação	rank	Fitness médio (R\$)	Combinação	rank	Fitness médio (R\$)		
9 - 80_500	230,37	30.738.711	a	2 - 20_300	101,40	30.421.670	d
6 - 50_500	211,63	30.690.734	ab	7 - 80_100	73,03	30.352.364	e
8 - 80_300	189,37	30.614.299	b	4 - 50_100	63,70	30.332.460	e
3 - 20_500	155,57	30.540.704	c	1 - 20_100	39,67	30.236.810	f
5 - 50_100	154,77	30.538.297	c				

Médias seguidas pelas mesmas letras não apresentam diferenças estatisticamente significativas ($p > 0,05$).

Nesse trabalho, como as duas primeiras combinações, 50 e 80 indivíduos, não se diferiram entre si estatisticamente, a população com 50 indivíduos pode ser mais vantajosa por exigir menores esforços computacionais e tempo de processamento.

Rodrigues et al. (2004) avaliaram populações com 25 a 100 indivíduos e concluíram que tamanhos de 75 a 100 proporcionam melhores eficácias para o AG, reforçando a ideia de que o fator população inicial foi o que exerceu maior influência na aproximação da resposta obtida pelo algoritmo genético da resposta encontrada pelo algoritmo exato. Da mesma forma, Silva et al. (2009) encontraram maior eficácia para os maiores tamanhos de população inicial, observando tendência de piora do desempenho em termos de eficácia com a redução do tamanho da população. Assim, é possível inferir que uma população de maior extensão pode preservar melhor a diversidade, evitando a convergência prematura (Rennard 2007). Por fim, para todo e qualquer problema, o tamanho da população irá depender de sua complexidade (Sivanandam e Deepa 2008).

Houve diferenças estatísticas significativas pelo teste de Kruskal Wallis a um nível de significância de 5% entre os números de gerações testados. O critério de parada considerando 500 gerações (*fitness* médio igual a R\$ 30.656.717) apresentou melhores resultados que os processamentos considerando 300 (*fitness* médio igual a R\$ 30.524.755) e 100 (*fitness* médio igual a R\$ 30.307.212) gerações. Isso já era esperado, uma vez que valores muito baixos podem diminuir a eficácia do algoritmo pela redução do número de soluções viáveis encontradas nesse menor intervalo, diminuindo, assim, as chances de encontrar solução de melhor qualidade entre as soluções candidatas. Por outro lado, valores muito altos, apesar de aumentar as chances de encontrar uma solução de melhor qualidade, podem prejudicar a eficiência do AG, causando depreciação no tempo de processamento da solução (Rodrigues et al. 2004).

Conclusões

A qualidade das soluções obtidas com a metaheurística algoritmo genético é dependente da seleção de uma boa configuração de seus parâmetros.

A melhor combinação de parâmetros apontada pelo ranqueamento foi a que tinha como método de seleção o torneio, mutação do tipo gene a gene, utilização de elitismo, tipo de cruzamento de 1 ponto, população inicial com 50 indivíduos e critério de parada com 500 gerações.

Referências

- Almeida LPA, Alvarenga AA, Castro EM, Araújo Júnior CA (2012) *Simulação multiagentes aplicada ao planejamento da produção florestal sustentável*. Dissertação, Departamento de Engenharia Florestal, Universidade Federal de Viçosa. 75 p.
- Arce JE (1997) *Um sistema de programação do transporte principal de multiprodutos florestais visando a minimização de custos*. Dissertação, Departamento de Engenharia Florestal, Universidade Federal do Paraná. 98p.
- Bento EP, Hagan N (2008) Algoritmos genéticos e variantes na solução de problemas de configuração de redes de distribuição. *Revista Controle & Automação*, 19(3):302-315. doi: dx.doi.org/10.1590/S0103-17592008000300006
- Bettinger P, Boston K, Siry LP, Grebner RL (2009) *Forest management and planning*. 1ª Edição. Londres: Academic. 331 p.
- Binoti BHB, Binoti MLM, Leite HG, Gleriani JM, Ribeiro CAAS (2014) Inclusão e influência de características espaciais em modelos de regulação florestal. *Cerne*, 20(1):157-164. doi: dx.doi.org/10.1590/S0104-77602014000100019

- Carvalho KHA, Silva ML, Leite HG, Binoti DHB, Silva LF (2015) Influência da taxa de juros e do preço da madeira em modelos de regulação florestal. *Pesquisa Florestal Brasileira*, 35(82):143-151. doi: dx.doi.org/10.4336/2015.pfb.35.82.554
- Costa Filho PA, Poppi RJ (1999) Algoritmo genético em química. *Química Nova*, 22(3):405-411. doi: dx.doi.org/10.1590/S0100-40421999000300019
- Davis KP (1966) *Forest management: regulation and valuation*. 2ª Edição. New York: McGraw-Hill. 519 p.
- Dong L, Bettinger P, Lui Z, Qin H, Zhao Y (2016) Evaluating the neighborhood, hybrid and reversion search techniques of a simulated annealing algorithm in solving forest spatial harvest scheduling problems. *Silva Fennica*, 50(4):1-20. doi: doi.org/10.14214/sf.1622
- Ezquerro M, Pardos M, Diaz-Balteiro L (2016) Operational research techniques used for addressing biodiversity objectives into forest management: an overview. *Forests*, 7(10):229-247. doi: doi.org/10.3390/f7100229
- Gomide LR, Arce JE, Silva ACL (2009) Uso do algoritmo genético no planejamento florestal considerando seus operadores de seleção. *Cerne*, 15(4): 460-467.
- Hillier FS, Lieberman GJ (2006) *Introdução à pesquisa operacional*. 8ª Edição. São Paulo: MacGraw-Hill. 850 p.
- Jin X, Pukkala T, Li F (2016) Fine-tuning heuristic methods for combinatorial optimization in forest planning. *European Journal of Forest Research*, 135(1):765-779. doi: doi.org/10.1007/s10342-016-0971-x
- Johnson KN, Scheurman HL (1977) Techniques for prescribing optimal timber harvest and investment under different objectives - discussion and synthesis. *Forest Science*, 18(1):1-31.
- Kaya A, Bettinger P, Boston K, Akbulut R, Ucar Z, Siry J, Merry K, Cieszewski C (2016) Optimisation in Forest Management. *Current Forestry Reports*, 2(1): 1-17, 2016. doi: 10.1007/s40725-016-0027-y
- Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260): 583-621.
- Lopes HS, Takahashi RHC (2011) *Computação evolucionária em problemas de engenharia*. 1ª Edição. Curitiba: Omnipax. 385 p.
- Mendes JB (2013) *Uma abordagem multiobjetivo para o problema de despacho de caminhões em minas a céu aberto*. Tese, Departamento de Engenharia Elétrica, Universidade Federal de Minas Gerais. 133p.
- Mitchell M (1999) *An introduction to genetic algorithms*. 1ª Edição. Londres: Massachusetts Institute of Technology. 162 p.
- Rennard JP (2007) *Handbook of research on nature-inspired computing for economic and management*. 1ª Edição. Londres: Idea Group Reference. 998 p.
- Rodrigues FL, Leite HG, Santos HN, Souza AL, Silva GF (2004) Metaheurística algoritmo genético para solução de problemas de planejamento florestal com restrições de integridade. *Revista Árvore*, 28(2):233-245. doi: dx.doi.org/10.1590/S0100-67622004000200010
- Scolforo JSR (1998) *Modelagem do crescimento e da produção de florestas plantadas e nativas*. 1ª Edição. Lavras: UFLA/FAEPE. 451 p.
- Silva GF, Leite HG, Silva ML, Rodrigues FL, Santos HN (2003) Problemas com o uso de programação linear com posterior arredondamento da solução ótima, em regulação florestal. *Revista Árvore*, 27(5):677-688. doi: dx.doi.org/10.1590/S0100-67622003000500010
- Silva GF, Piassi LC, Mora R, Martins L, Teixeira AF (2009) Metaheurística algoritmo genético na solução de modelos de planejamento florestal. *Revista Brasileira de Ciências Agrárias*, 4(2):120-166. doi: doi.org/10.5039/agraria.v4i2a7
- Sivanandam SN, Deepa SN (2008) *Introduction to genetic algorithms*. 1ª Edição. New York: Spring. 453 p.
- Takahashi RHC, Vasconcelos JA, Ramirez JA, Krahenbuhl L A (2003) Multi-objective Methodology for Evaluating Genetic Operators. *IEEE Transactions on Magnetics*, 39(3):1321-1324. doi: doi.org/10.1109/TMAG.2003.810371
- Troncoso JJ, Weintraub A, Martell DL (2016) Development of a threat index to manage timber production on flammable forest landscapes subject to spatial harvest constraints. *Information Systems and Operational Research*, 54(3):262-281. doi: doi.org/10.1080/03155986.2016.1197543
- Yoshimoto A, Asante P, Konoshima M (2016) Stand-level forest management planning approaches. *Current Forestry Reports*, 2(3):163-176. doi: doi.org/10.1007/s40725-016-0041-0
- Zar JH (2010) *Biostatistical analysis*. 5ª Edição. New Jersey: Pearson Prentice Hall. 944 p.