

## CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ  
Rio de Janeiro - RJ - Brasil  
ISSN 1413-9022 / v. 31, p. 01 - 14, 2011

# ANÁLISE DE SIGNIFICÂNCIA DOS PARÂMETROS DE UM ALGORITMO GENÉTICO ATRAVÉS DO PLANEJAMENTO DE EXPERIMENTOS

Rafael de Carvalho Miranda  
Universidade Federal de Itajubá (UNIFEI)  
rafaelc.miranda@yahoo.com.br

Alexandre Ferreira de Pinho  
Universidade Federal de Itajubá (UNIFEI)  
pinho@unifei.edu.br

José Arnaldo Barra Montevechi  
Universidade Federal de Itajubá (UNIFEI)  
montevechi@unifei.edu.br

### Resumo

*O trabalho aqui desenvolvido utiliza o planejamento de experimentos (DOE) para analisar a significância dos principais parâmetros de um algoritmo genético clássico (AG), na otimização de um modelo de simulação a eventos discretos. Serão analisados os principais parâmetros de um AG: o tamanho da população, o número de gerações, a taxa de crossover e a taxa de mutação. Os parâmetros serão analisados quanto a sua influência na qualidade da solução apresentada pelo algoritmo e no tempo necessário para convergência. Um método de pesquisa voltado à utilização do DOE foi utilizado na condução da experimentação. Pôde-se constatar que os parâmetros tamanho de população e número de gerações são os parâmetros mais significativos dentre os parâmetros investigados, influenciando diretamente na qualidade da solução obtida e no tempo para a convergência do algoritmo. As taxas de crossover e mutação, não foram significativas em nenhum dos experimentos. Com a análise conduzida pode-se auxiliar na definição dos parâmetros de um AG, ou até mesmo, na elaboração de estratégias adaptativas.*

**Palavras-chave:** Algoritmo Genético; Parâmetros do Algoritmo Genético; Planejamento de Experimentos.

## 1. Introdução

Os Algoritmos Genéticos (AG) vem sendo usados com sucesso, para encontrar soluções ótimas, ou quase ótimas, para uma ampla variedade de problemas de otimização (GEN e CHENG, 1997).

Desde sua criação há um interesse crescente na utilização dos AG como uma ferramenta para resolução de problemas complexos de otimização (GEN e CHENG, 2000; HOUPPT e HOUPPT, 1998). Para Asllani e Lari (2007), embora os AG sejam mais gerais e abstratos do que outros métodos de otimização, e, embora, nem sempre ofereçam a solução ideal, eles são considerados flexíveis e aplicáveis a uma grande variedade de problemas.

Segundo Yang *et al.* (2007), apesar dos AG resolverem problemas de otimização eficazmente, muitas vezes estes são sensíveis a seus parâmetros de configuração. Para Neppalli *et al.* (1996), a determinação dos valores dos parâmetros de um AG é um processo complexo, e geralmente, na maioria das aplicações os valores dos parâmetros são ajustados com base em trabalhos previamente conduzidos e relatados.

No entanto, Gosselin *et al.* (2009) destacam que configurações específicas de parâmetros utilizados na configuração de um AG nem sempre estão disponíveis na literatura, podendo complicar a repetibilidade dos resultados ou a extensão de um trabalho para um problema similar. Para estes autores, as escolhas dos parâmetros do AG possui uma grande influência na velocidade de convergência, bem como no sucesso da otimização.

Dessa forma, o trabalho aqui desenvolvido expande o trabalho proposto em Miranda *et al.* (2011), utilizando o planejamento de experimentos (DOE) para análise da significância dos parâmetros de um AG, bem como, suas interações, para a qualidade da solução apresentada pelo algoritmo, e sua influência no tempo de convergência na otimização de um modelo de simulação.

Para cumprir com o objetivo proposto, esse artigo encontra-se dividido em 5 seções. A primeira, já apresentada, trouxe a contextualização do problema de pesquisa e os objetivos. A segunda seção apresenta o referencial teórico, que suporta o trabalho desenvolvido. A terceira traz a metodologia utilizada. A seção 4 apresenta o desenvolvimento do método de pesquisa e, por fim, a quinta seção apresenta as conclusões do trabalho.

## 2. Revisão de Literatura

### 2.1 Parâmetros AG

Yang *et al.* (2007) destacam como principais parâmetros de um AG: o tamanho da população, o critério de parada, a taxa de *crossover* e a taxa de mutação. Azadeh e Tarverdian (2007) acrescentam mais um parâmetro aos já mencionados: o número de gerações.

Segundo Khalik *et al.* (2007), apesar de não existir consenso a respeito das definições dos parâmetros dos AG, sua determinação não deve ser feita ao acaso, pois variações na definição destes parâmetros podem prejudicar seu resultado final. Apesar de não existirem regras, ou consenso, para a definição dos parâmetros dos AG, são apresentados aqui os valores mais comumente utilizados na literatura.

Para Linden (2006), o desempenho do AG é muito sensível ao tamanho da população e alerta que este parâmetro deve ser definido com extremo cuidado. Para este autor, se a população for pequena, não haverá espaço suficiente para uma variedade genética muito grande, o que incapacitará o algoritmo a encontrar boas soluções. Por outro lado, se a população for muito grande, o algoritmo gastará muito tempo até encontrar uma solução.

Para Tanomaru (1995), populações entre 50 a 200 cromossomos resolvem grande parte dos problemas. Michel (1996) indica que o melhor tamanho de uma população está entre 50 e 100. Já no trabalho de Aytug *et al.* (2003), verificaram-se populações variando de 20 a 1024 indivíduos. Aytug e Koehler (2000) trabalharam com um número variável entre 100 e 200. Aryanezhad e Hemati (2008) utilizaram uma população que variou de 20 a 2000. Linden (2006), no entanto, afirma que a maioria dos trabalhos publicados adota o tamanho da população igual a 100.

Com relação ao número de gerações, este varia de acordo com a complexidade do problema que está sendo tratado, e deve ser determinado experimentalmente (YUN e GEN, 2003). Não há um acordo entre os usuários dos AG com relação ao número de gerações, Hwang e He (2006) definem um número de gerações igual a 600, enquanto que Núñez-Letamendia (2007) definem esse valor em 25, Zhang e Ishikawa (2004) utilizaram 10 gerações, Asllani e Lari (2007) e Martin (2009) trabalharam com 100 gerações.

Segundo Linden (2006), o operador de *crossover* deve receber uma

probabilidade maior que o operador de mutação, uma vez que a reprodução é a principal característica de um AG. Para Barboza (2005), normalmente o valor da taxa de *crossover* fica entre 60% e 65%. Já segundo Tanomaru (1995), estudos empíricos mostram que bons resultados são obtidos com valores superiores a 70%. Deb e Tiwari (2008) usaram o valor de 50%. Rees e Koehler (2006) conduziram estudos variando a taxa de *crossover* entre 10% e 90%, enquanto Al-Aomar e Al-Okaily (2006) utilizaram valores entre 60% e 100%.

Para a taxa de mutação, Barboza (2005) afirma que geralmente os valores encontrados na literatura ficam em torno de 0,1% e 5%. Para Tanomaru (1995), essa taxa deve ser menor que 1%. Enquanto que, Azadeh e Tarverdian (2007) recomendam que a taxa de mutação deva estar entre 1% e 5%. Corroborando com essas afirmações, Hwang e He (2006) utilizaram em seus experimentos um valor de 1%. Zhang e Ishikawa (2004) utilizam o valor 0,1%. Já no trabalho de Gosselin *et al.* (2009), várias taxas de mutação e *crossover* foram utilizadas.

## 2.2 Planejamento fatorial completo

Paiva (2008) destaca o planejamento fatorial completo como o arranjo experimental mais comum, para o qual o número de experimentos é igual ao número de níveis experimentais elevado ao número de fatores. No caso típico de fatoriais em dois níveis, o número de experimentos (N) para se avaliar os k fatores é dado por  $N = 2^k$  (PAIVA, 2008). Para Montgomery (2005), esses planejamentos possuem uma análise simplificada e formam a base de muitos outros planejamentos.

De acordo com Montgomery e Runger (2003), no planejamento fatorial completo os fatores são alterados conjuntamente e não um por vez, o que significa que para cada corrida ou réplica completa, todas as combinações dos níveis possíveis são investigadas.

Montgomery (2005) ressalta que os planejamentos fatoriais são mais eficientes que a abordagem de um fator por vez, uma vez que ele permite que os efeitos dos fatores sejam estimados a partir dos níveis dos outros fatores, conduzindo a conclusões válidas dentro do escopo experimental, além de serem a única maneira de descobrir interações entre as variáveis, evitando conclusões erradas quando interações entre fatores estão presentes.

### 3. Metodologia

Montgomery (2005) alerta para a necessidade de que os envolvidos em um experimento tenham uma ideia prévia de seus objetivos, de quais fatores serão estudados, de como o experimento deve ser conduzido e, que tenham pelo menos uma compreensão de como os dados serão analisados. Para um estudo envolvendo o DOE, o autor propõe a utilização das seguintes etapas:

- a) Reconhecimento e definição do problema: desenvolver completamente todas as ideias sobre o problema e sobre os objetivos a serem alcançados com o experimento;
- b) Escolha dos fatores e dos níveis de trabalho: devem-se escolher os fatores, os intervalos sobre os quais esses fatores variarão e os níveis específicos nos quais cada rodada será realizada;
- c) Seleção das variáveis de resposta: o experimentador deve determinar as variáveis de resposta que sejam úteis para o processo em estudo;
- d) Escolha do projeto experimental: essa escolha envolve o tamanho da amostra (número de replicações), seleção da ordem correta das rodadas para as tentativas experimentais, ou se a formação de blocos ou outras restrições de aleatorização estão envolvidas;
- e) Realização dos experimentos: quando da realização dos experimentos é importante monitorar o processo, para garantir que tudo está sendo feito de acordo com o planejado, erros nesta etapa podem destruir a validade do experimento;
- f) Análise estatística dos dados: métodos estatísticos devem ser empregados para analisar os dados, de modo que os resultados e conclusões sejam objetivos e não fruto de opiniões;
- g) Conclusões e recomendações: o experimento deve conduzir a conclusões práticas sobre os resultados e recomendar um plano de ação. Sequências de acompanhamentos e testes de confirmação devem ser conduzidos para validar as conclusões do experimento.

Apesar de genérico, os passos do método de pesquisa podem ser empregados em uma grande variedade de projetos experimentais que usem o DOE. A seção seguinte detalha a aplicação do método neste trabalho, apresentando todas as suas peculiaridades.

#### 4. Análise dos parâmetros do AG no tempo de convergência e na qualidade de resposta

Para a condução dos experimentos será utilizado um AG em sua forma clássica, tal como preconizado por Holland (1975). Um modelo de simulação servirá de objeto de estudo para otimização e análise dos parâmetros. As seções seguintes irão explorar a metodologia proposta por Montgomery (2005) na condução dos experimentos.

##### 4.1 Reconhecimento e definição do problema

Um modelo de simulação a eventos discretos, voltado à área de manufatura, será utilizado para otimização e análise dos parâmetros do AG. O modelo de simulação utilizado refere-se a uma célula de produção de uma empresa multinacional do setor de autopeças. As etapas de construção e validação desse modelo podem ser verificadas em Montevechi *et al.* (2008).

O modelo de simulação, objeto de estudo deste trabalho, será otimizado quanto ao lucro total produzido na célula de produção, ou seja, se buscará a maximização do lucro através da melhor combinação das variáveis de entrada do modelo.

##### 4.2 Escolha dos fatores e dos níveis de trabalho

Foram selecionados quatro parâmetros do AG como fatores para experimentação: tamanho de população, número de gerações, taxa de crossover e taxa de mutação. A escolha desses fatores se deu devido ao fato de a literatura trazê-los como os principais parâmetros dos AG, influenciando diretamente no tempo necessário para se alcançar uma resposta, bem como para a qualidade da resposta alcançada.

Os níveis adotados e as variações para cada fator são apresentados na Tabela 1.

Tabela 1 – Fatores, níveis e variações do projeto de experimentos

Fatores	Níveis	
	-	+
Tamanho da população	20	80
Número de gerações	5	30
Taxa de <i>crossover</i>	50%	90%
Taxa mutação	1%	10%

Para este experimento optou-se em trabalhar com dois níveis para cada fator. Para Montgomery (2005), dois níveis de trabalho são mais do que suficientes para se ter conclusões válidas sobre sua significância. A escolha dos valores para cada um dos níveis foi realizada com base no trabalho de Pinho (2008), o qual desenvolveu um

otimizador de modelos de simulação a eventos discretos, e, como neste trabalho, utilizou-se da experimentação para identificação dos parâmetros críticos de um AG.

#### **4.3 Seleção das variáveis de resposta**

Duas variáveis de resposta serão analisadas nesse projeto experimental. A primeira verificará a qualidade da resposta apresentada pelo AG na otimização do modelo de simulação, já a segunda analisará o tempo necessário para que o algoritmo chegue a uma solução. A qualidade de resposta será avaliada a partir das respostas apresentadas pelo algoritmo quando este converge, já os tempos serão cronometrados desde o momento em que o algoritmo inicia a otimização até o momento de sua convergência. Dessa forma, pretende-se verificar o impacto dos parâmetros escolhidos tanto na qualidade da resposta apresentada pelo AG, como no tempo gasto para tal.

#### **4.4 Escolha do projeto experimental**

Para este trabalho será utilizado o planejamento fatorial completo. Tal escolha se deve ao fato que um fatorial completo permite estimar os efeitos principais dos fatores analisados, bem como quaisquer possíveis interações de qualquer ordem entre os fatores (MONTGOMERY, 2005).

Para o caso em questão com 4 fatores, tem-se um total de 16 experimentos a serem realizados ( $2^4 = 16$ ). As Tabelas 2 e 3 apresentam a matriz experimental, na qual os sinais – e + representam respectivamente os níveis inferior e superior adotados para cada parâmetro, como apresentado na Tabela 1. Foram definidas cinco réplicas para cada experimento, totalizando 80 experimentos para o modelo de simulação.

O número de réplicas foi definido com base no trabalho de Pinho (2008). Em seu trabalho, o autor alcançou bons resultados na identificação dos parâmetros significativos de um AG com a utilização de 5 réplicas, fato este que justifica a escolha aqui adotada.

#### **4.5 Realização dos experimentos**

Para a realização dos experimentos foi necessário desenvolver uma ferramenta computacional que operacionalizasse a realização dos experimentos. Para tal foi implementado um AG em sua forma clássica, como preconizado em sua criação.

A exceção dos parâmetros do AG todos os demais componentes do algoritmo permaneceram constante durante toda experimentação, garantindo assim a mesma

estrutura básica a todos os experimentos. Com relação aos parâmetros, estes foram se alterando conforme a matriz experimental. Os resultados encontrados para cada experimento, realizado no modelo de simulação, são apresentados nas Tabelas 2 e 3.

Tabela 2 – Matriz experimental e resultados da otimização para a variável solução

Exp	Tam. população	No. de gerações	Taxa de <i>Crossover</i>	Taxa de Mutação	Solução 1	Solução 2	Solução 3	Solução 4	Solução 5
1	-	-	-	-	497400	513000	492400	506100	490000
2	+	-	-	-	506600	513000	511400	506100	507300
3	-	+	-	-	493500	513000	511400	498600	503800
4	+	+	-	-	506600	513000	511400	506100	520300
5	-	-	+	-	497400	492400	494800	498000	503800
6	+	-	+	-	506600	506400	511400	506100	507300
7	-	+	+	-	506600	506400	511400	506100	503800
8	+	+	+	-	506600	513000	511400	506100	520000
9	-	-	-	+	492400	492400	511400	506100	520000
10	+	-	-	+	506600	513000	511400	506100	506600
11	-	+	-	+	506600	506400	511400	498600	520300
12	+	+	-	+	506600	513000	511400	506100	520300
13	-	-	+	+	506600	506400	494800	506100	506600
14	+	-	+	+	506600	506400	511400	506100	507300
15	-	+	+	+	506600	513000	511400	506100	520300
16	+	+	+	+	506600	513000	511400	506100	520300

Na Tabela 2, as colunas “Solução” apresentam o valor máximo encontrado para o lucro em cada cenário testado, segundo a matriz experimental em cada réplica conduzida.

Tabela 3 – Matriz experimental e tempo necessário para convergência do algoritmo

Exp	Tam. população	No. de gerações	Taxa de <i>Crossover</i>	Taxa de Mutação	Tempo1 (Seg)	Tempo2 (Seg)	Tempo3 (Seg)	Tempo4 (Seg)	Tempo5 (Seg)
1	-	-	-	-	573	582	595	579	582
2	+	-	-	-	2366	2372	2309	2266	2372
3	-	+	-	-	3588	3487	3381	3609	3546
4	+	+	-	-	13831	13763	13347	13179	13581
5	-	-	+	-	568	602	588	570	584
6	+	-	+	-	2320	2360	2337	2345	2389
7	-	+	+	-	3589	3559	3372	3345	3423
8	+	+	+	-	13369	13903	13388	13147	13751
9	-	-	-	+	568	593	566	573	594
10	+	-	-	+	2321	2379	2326	2386	2428
11	-	+	-	+	3328	3407	3331	3598	3572
12	+	+	-	+	13092	13641	13227	13567	13989
13	-	-	+	+	577	570	572	573	571
14	+	-	+	+	2357	2340	2311	2365	2284
15	-	+	+	+	3345	3766	3306	3378	3419
16	+	+	+	+	13503	13938	13298	13064	13329

Na Tabela 3, as colunas “Tempo” apresentam o valor gasto para o algoritmo convergir para a resposta apresentada na tabela anterior, para cada cenário segundo a matriz experimental em cada réplica conduzida.



Os dados coletados referem-se ao resultado da otimização dos modelos (valor da função objetivo) e o tempo necessário para convergência do algoritmo. Os valores para cada variável de decisão encontrada pelo algoritmo não serão apresentados. Cada uma das tabelas apresenta os resultados das 5 réplicas realizadas na otimização do modelo de simulação.

#### 4.6 Análise estatística dos dados

Antes de se iniciar a análise estatística dos dados obtidos foi necessário realizar a análise dos resíduos. Segundo Montgomery (2005), os resíduos devem ser normais, aleatórios e não correlacionados. Desta forma para cada um dos experimentos realizados foi necessário testar se tais condições foram atendidas.

Para todos os casos foi possível alcançar tais condições, a questão da aleatoriedade e o fato dos resíduos não estarem correlacionados foram facilmente verificados por meio de cartas de controle, onde não foram observados nenhum padrão não aleatório, de agrupamento ou tendencioso por parte dos resíduos. Já a normalidade dos resíduos foi verificada por meio de testes de normalidade, onde após a retirada dos *outliers* das amostras analisadas, pode-se verificar a normalidade dos resíduos.

A análise de variância (ANOVA) para o experimento foi conduzido com um nível de confiança de 95%. Os resultados encontrados para a qualidade da resposta e para o tempo para convergência podem ser visualizados nas Figuras 1 e 2.

Pela análise de variância para Solução (Figura 1), verifica-se que as interações de segunda, terceira e quarta ordem não são significativas ( $p\text{-values} > 0,05$ ) para a qualidade da solução apresentada pelo algoritmo. Apenas os efeitos principais mostraram-se significativos para obtenção de uma solução de elevada qualidade.

Figura 1 – Análise de variância para Solução

	DF	Seq SS	Adj SS	Adj MS	F	P
Efeitos principais	4	1146940500	1146940500	286735125	7,42	0,000
A	1	549152000	549152000	549152000	14,22	0,000
B	1	464648000	464648000	464648000	12,03	0,001
Interações 2a. ordem	6	231833500	231833500	38638917	1,00	0,433
Interações 3a. ordem	4	16035500	16035500	4008875	0,10	0,981
Interações 4a. ordem	1	1300500	1300500	1300500	0,03	0,855
Erro Residual	64	2472228000	2472228000	38628563		
Erro Puro	64	2472228000	2472228000	38628563		
Total	79	3868338000				

Com relação ao Tempo, a análise de variância realizada (Figura 2) mostrou que apenas as interações de segunda ordem mostraram-se significativas, bem como os

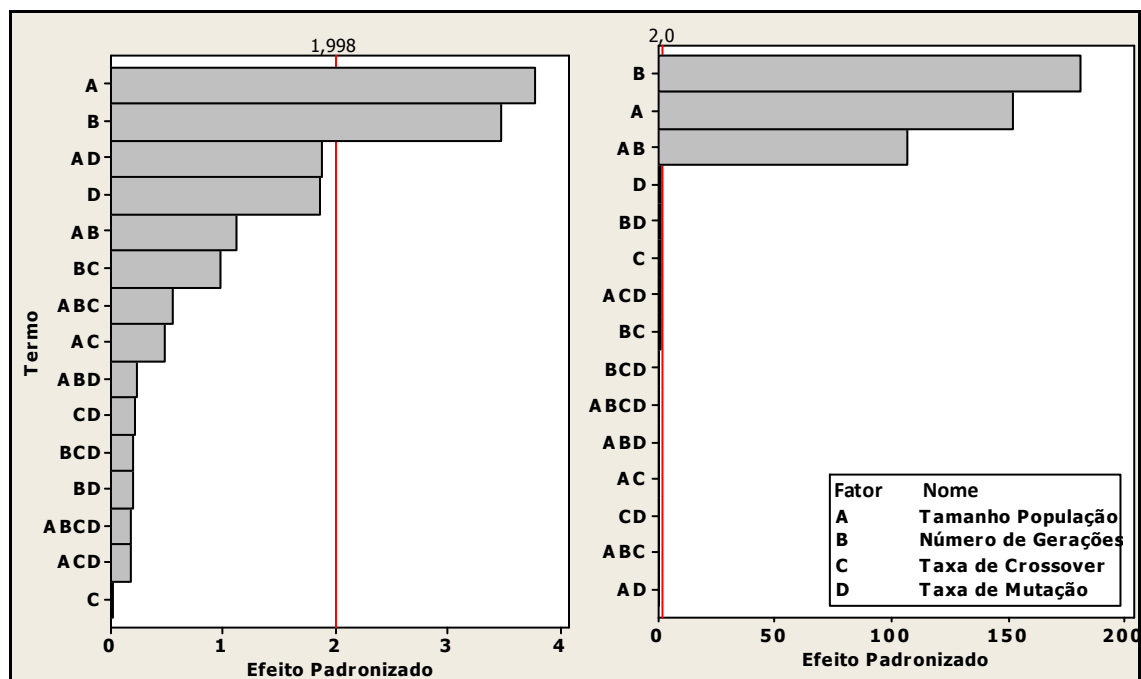
efeitos principais dos fatores analisados.

Figura 2 – Análise de variância para Tempo

	DF	Seq SS	Adj SS	Adj MS	F	P
Efeitos principais	4	1680910057	1680910057	420227514	13965,82	0,000
A	1	695674999	695674999	695674999	23120,03	0,000
B	1	985207901	985207901	985207901	32742,35	0,000
Interações 2a. Ordem	6	341180664	341180664	56863444	1889,80	0,000
A*B	1	341158650	341158650	341158650	11338,05	0,000
Interações 3a. Ordem	4	9811	9811	2453	0,08	0,988
Interações 4a. Ordem	1	1320	1320	1320	0,04	0,835
Erro residual	64	1925742	1925742	30090		
Erro puro	64	1925742	1925742	30090		
Total	79	2024027594				

Os fatores e as interações significativas podem ser visualizados no Gráfico de Pareto para os Efeitos Padronizados (Figura 3).

Figura 3 – Gráfico de Pareto para os efeitos padronizados: Solução e Tempo



Pela análise das Figuras 1 e 3 verifica-se que o tamanho da população (A) e o número de gerações (B) se mostraram significativos para se chegar a soluções de elevada qualidade, ou seja, quanto maior o tamanho da população e o número de gerações, melhor será a solução encontrada pelo AG.

Com relação ao tempo, além dos fatores principais: tamanho da população (A) e número de gerações (B), a interação entre esses fatores (AB) também se mostrou significativa (Figuras 2 e 3). Neste caso, verifica-se que os fatores principais: tamanho da população e número de gerações contribuem significativamente para o aumento do

tempo de convergência do algoritmo, e a interação entre os fatores também possui um efeito no sentido de incrementar o tempo para que o algoritmo alcance uma solução.

Pela análise dos resultados pode-se afirmar que os operadores genéticos de *crossover* (C) e mutação (D) não se mostraram significativos para o tempo de convergência e para a solução encontrada ao nível de confiança de 95% (Figura 3).

## 5. Conclusões

O trabalho aqui desenvolvido expandiu a proposta apresentada em Miranda *et al.* (2011), de utilização do planejamento fatorial completo como meio de análise da significância dos principais parâmetros de um AG, tanto para a qualidade das soluções apresentadas pelo algoritmo, como para o tempo necessário para sua convergência.

Para o objeto de estudo em questão, dois parâmetros se mostraram significativos, tanto para o tempo, quanto para a qualidade de solução: o tamanho de população e o número de gerações. O parâmetro tamanho de população foi o mais significativo para a qualidade de resposta, apresentando um efeito positivo sobre esta variável, ou seja, quanto maior o tamanho da população, melhor a resposta encontrada. Por sua vez, o número de gerações foi o parâmetro que mais influenciou o tempo para convergência do algoritmo.

Com relação ao tempo, verifica-se que a interação entre os fatores, tamanho de população e número de gerações é fortemente significativa para o aumento do tempo de convergência do AG. Com isso, chega-se a um impasse, quanto maior o tamanho da população e o número de gerações, melhor a qualidade da resposta, porém mais tempo o algoritmo demorará a convergir.

Assim como apontado em Miranda *et al.* (2011), os parâmetros taxa de *crossover* e taxa de mutação, não foram significativos em nenhum dos dois experimentos, ao nível de confiança adotado no estudo. Apesar de o parâmetro taxa de mutação e a interação entre taxa de mutação e tamanho de população, estarem próximos de serem significativos para qualidade da solução (Figura 3).

Vale ressaltar que o trabalho aqui desenvolvido, não buscou determinar um conjunto de valores ótimos para o AG, uma vez que o mesmo seria praticamente impossível, devido às inúmeras aplicações deste algoritmo. No entanto, a análise aqui conduzida pode auxiliar na definição dos parâmetros de um AG, ou até mesmo, na elaboração de estratégias adaptativas, que busquem ao mesmo tempo, conciliar bons

resultados, com um tempo admissível para convergência.

## Agradecimentos

Os autores agradecem a CAPES, CNPq e à FAPEMIG.

## Referências

- AL-AOMAR, R.; AL-OKAILY, A. A GA-based parameter design for single machine turning process with high-volume production. **Computers & Industrial Engineering**, v.50, p.317-337, 2006.
- ARYANEZHAD, M. B.; HEMATI, M. A new genetic algorithm for solving nonconvex nonlinear programming problems. **Applied Mathematics and Computation**, v.199, p.186-194, 2008.
- ASLLANI, A.; LARI, A. Using genetic algorithm for dynamic and multiple criteria web-site optimizations. **European Journal of Operational Research**, v.176, p.1767-1777, 2007.
- AYTUG, H.; KOEHLER, G. J. New stopping criterion for genetic algorithms, **European Journal of Operational Research**, n.126, p. 662-674, 2000.
- AYTUG, H.; KHOUJA, M.; VERGARA, F. A review of the use of genetic algorithms to solve production and operations management problems. **International Journal of Production Research**, v.41, n.17, p. 3955-4009, 2003.
- AZADEH, A.; TARVERDIAN, S. Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption. **Energy Policy**, v.35, p.5229-5241, 2007
- BARBOZA, A. O. **Simulação e Técnicas da Computação Evolucionária Aplicadas a Problemas de Programação Linear Inteira Mista**. 2005. 236 f. Tese (Doutorado em Ciências – Informática Industrial). Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, PR, 2005.
- DEB, K.; TIWARI, S. Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization. **European Journal of Operational Research**, v.185, p.1062-1087, 2008.
- GEN, M.; CHENG, R. **Genetic Algorithms and Engineering Design**. New York: John Wiley and Sons, 1997.
- GEN, M.; CHENG, R. **Genetic Algorithms and Engineering Optimization**. New York: John Wiley and Sons. 2000.
- GOSSELIN, L.; TYE-GINGRAS, M.; MATHIEU-POTVIN; F. Review of utilization of genetic algorithms in heat transfer problems. **International Journal of Heat Mass Transfer**, v.52, p.2169-2188, 2009.
- HOLLAND, J. H. **Adaptation in Natural and Artificial Systems**. Ann Arbor: University of Michigan Press, 1975.
- HOUP, R. L.; HOUP, S.E. **Practical Genetic Algorithms**. New York: John Wiley and Sons, Inc., 1998.
- HWANG, S. F.; HE, R. S. A hybrid real-parameter genetic algorithm for function optimization. **Advanced Engineering Informatics**, v.20, p.7-21, 2006.
- KHALIK, M.A.; SHERIF, M.; SARAYA, S.; AREED, F. Parameter identification problem: Real-coded GA approach. **Applied Mathematics and Computation**, v.187, p.1495-1501, 2007.

LINDEN, R. **Algoritmos Genéticos: Uma Importante Ferramenta da Inteligência Computacional**. Rio de Janeiro: Editora Brasport, 2006.

MARTIN, C. H. A hybrid genetic algorithm/mathematical programming approach to the multi-family flowshop scheduling problem with lot streaming. **Omega**, v.37, p.126-137, 2009.

MIRANDA, R. C.; PINHO, A. F.; MONTEVECHI, J. A. B. Utilização do planejamento de experimentos para análise de significância dos parâmetros de um algoritmo genético. In: XVIII SIMPEP, **Anais...** Bauru, SP, 2011.

MONTEVECHI, J.A.B.; PINHO, A.F. DE; LEAL, F.; MARINS, F.A.S.; COSTA, R. F. DA S. Improving a process in a Brazilian automotive plant applying process mapping, design of experiments and discrete events simulation. In: 20 Symposium Europeo de Modelado y Simulacion (SCS), **Proceedings...** Briatico, Italy, 2008.

MONTGOMERY, D.C. **Design and Analysis of Experiments**. 6th ed., Wiley, 2005.

MONTGOMERY, D.C.; RUNGER, G.C. **Estatística Aplicada e Probabilidade para Engenheiros**. 2. ed. Editora LTC, 2003.

NEPPALLI, R. V.; CHEN, C. L.; GUPTA, J. N. Genetic algorithms for the two-stage bicriteria flow shop problem. **European Journal of Operational Research**, v.95, p. 356-373, 1996.

NÚÑEZ-LETAMENDIA, L. Fitting the control parameters of a genetic algorithm: An application to technical trading systems design. **European Journal of Operational Research**, v.179, p.847-868, 2007.

PAIVA, E. J. **Otimização de Processos de Manufatura com Múltiplas Respostas baseada em Índices de Capacidade**. 2008. Dissertação (Mestrado) – Instituto de Engenharia de Produção e Gestão, Universidade Federal de Itajubá, Itajubá, MG, 2008.

PINHO, A. F. **Proposta de um Método de Otimização de Modelos de Simulação a Eventos Discretos**. 2008. 189 f. Tese (Doutorado em Engenharia Mecânica). Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista (UNESP), Guaratinguetá, SP, 2008.

REES, J.; KOEHLER, G. J. Learning genetic algorithm parameters using hidden Markov models. **European Journal of Operational Research**, v.175, p.806-820, 2006.

TANOMARU, J. Motivação, fundamentos e Aplicações de Algoritmos Genéticos. In: II Congresso Brasileiro de Redes Neurais, **Anais...** Curitiba, PR, Brasil, p.373-403, 1995.

YANG, T.; KUO, Y.; CHO, C. A genetic algorithms simulation approach for the multi-attribute combinatorial dispatching decision problem. **European Journal of Operational Research**, v.176, p. 1859-1873, 2007.

YUN, Y.; GEN, M. Performance analysis of adaptive genetic algorithms with fuzzy logic and heuristics. **Fuzzy Optimization and Decision Making**, v.2, n.2, p.161-175, 2003.

ZHANG, H.; ISHIKAWA, M. A solution to combinatorial optimization with time-varying parameters by a hybrid genetic algorithm. **International Congress Series**, v.1269, p.149-152, 2004.

## **SIGNIFICANCE ANALYSIS OF GENETIC ALGORITHM PARAMETERS THROUGH DESIGN OF EXPERIMENTS**

### **Abstract**

*The research presented here uses design of experiments (DOE) to analyze the significance of the main parameters of a classic genetic algorithm (GA) optimization in a discrete event simulation model. The main parameters of a GA are analyzed: population size, the number of generations, crossover rate and mutation rate. The parameters will be analyzed for their influence on the quality of the solution presented by the algorithm and the time required for convergence. A research method focused on the use of the doe was used in conducting the trial. It is evident that the population size and the number of generations are the most significant among those investigated, directly influencing the quality of solution obtained and the time for convergence of the algorithm. The crossover rate and mutation rate were not significant in any of the experiments. With the conducted analysis it's possible to assist in the definition of the parameters of a GA, or even, in the development of adaptive strategies.*

**Key-words:** Genetic Algorithm; Parameters of Genetic Algorithm; Design of Experiments.