

Integração de Meta-heurística e Sistema de Informação Geográfica para a otimização espacial multiobjetivo

Erica Patias Rodrigues¹
Daiane Gracieli Faller¹

¹ COPPE - Universidade Federal do Rio de Janeiro
Caixa Postal 68506 - 21945-970 – Rio de Janeiro - RJ, Brasil
{patias.eric, daianecem}@gmail.com

Abstract. The tool choice for allocation modeling in land use is essential to optimizing processes and assists in decision making. Therefore, the objective of this study is applying genetic algorithms (GA) for land use optimization in a particular geographic region, using data generated by geographic information systems (GIS). The objective functions were considered to minimize the cost of land use reallocation and the minimization of the difference in land use aptitude defined by an expert. Two constraints were applied; the unit contiguity of the same land use type and existence of permanent protection areas. The algorithm was applied to simplified data from Goiânia metropolitan region (an area of 900 km²), considering three types of use: urban, rural and preservation areas. The GA's initial population was 100 individuals, crossover and mutation rate of 80% and 5%, respectively, over 200 generations. The experience acquired from the study case described, we conclude that the GA can be very effective in optimizing spatial allocation problems. Future works may be performed in order to decrease the processing time (e.g., using parallel computing) and in relation to the restriction of minimum contiguous area (considering relevant laws).

Palavras-chave: genetic algorithm, land use optimization, geoprocessing, constraints, algoritmo genético, otimização do uso do solo, geoprocessamento, restrições.

1. Introdução

Devido à crescente industrialização e urbanização, a gestão eficiente do uso do solo é um fator decisivo para o desenvolvimento sustentável. A otimização do uso do solo deve atender diversos objetivos definidos pelas políticas públicas de gestão e são ao mesmo tempo sujeitas a diversas restrições ecológicas e econômicas.

A gestão do uso do solo pode ser definida como o processo de alocação de diferentes tipos competitivos de atividades, como agricultura, indústria, urbanização ou conservação, para diferentes unidades de uma paisagem. Este processo complexo, multi-objetivo e sujeito a uma série de restrições pode ser tratado como um problema de otimização. Os objetivos são definidos em um contexto espacial, portanto, devem incorporar informações de localização e topologia, como relações de vizinhança, aumentando ainda mais a complexidade do problema.

Diversas técnicas de otimização baseadas em meta-heurísticas estão em constante desenvolvimento e aperfeiçoamento, tais como algoritmos genéticos, enxame de partículas, celulares autômatos, regressão logística espacial, entre outros (Janerette e Jianguo, 2001; Han et al., 2005; Luo et al., 2009). Dentre estas, os algoritmos genéticos (AG) são robustos para a identificação de soluções ótimas, particularmente em um espaço de busca grande e complexo (Goldberg, 1989) e na aplicação em problemas multiobjetivos. A utilização dos AGs na otimização do uso do solo tem sido amplamente explorados por diversos estudos (Balling et al., 1999; Xiao et al., 2001; Van Dijk et al., 2002; Vink e Schot, 2002); Sarkar e Midak, 2003; Janssen et al., 2008; Esdrandaly, 2010).

O objetivo deste trabalho é desenvolver uma metodologia de aplicação de algoritmos genéticos para otimização multiobjetivo com restrições do uso do solo e testá-la em um estudo de caso. Neste trabalho, propomos a integração de algoritmos genéticos e sistemas de

informação geográfica (SIG) para encontrar as melhores soluções para o problema de alocação de tipos de uso do solo para uma determinada região geográfica.

2. Metodologia de Trabalho

2.1. Modelagem do problema

O gerenciamento do uso do solo é um problema de otimização multiobjetivo, que pode ser simplificado matematicamente representando-se a paisagem com uma matriz bidimensional, onde cada unidade é representada por uma posição j,i , como mostrado na Figura 1. Para facilitar as operações de cruzamento, a imagem é transformada em um vetor. O cromossomo pode então ser representado por $C = [e_{11}, e_{12}, \dots, e_{1j}, \dots, e_{i1}, e_{i2}, \dots, e_{ij}]$, onde e_{ij} representa o tipo de solo alocado na posição i,j da paisagem, representada por uma matriz de dimensões $R \times C$.

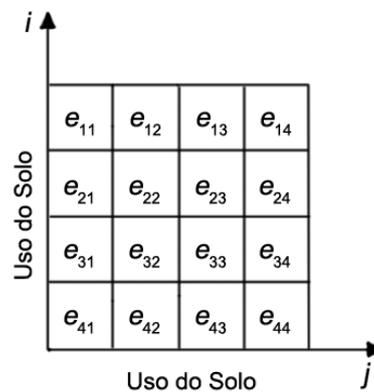


Figura 1. Representação matemática do uso do solo (modificado de Datta e Deb, 2006).

O gerenciamento do uso do solo envolve a alocação de diferentes tipos de uso a diferentes unidades da paisagem baseando-se em objetivos definidos por gestores. É um processo complexo sujeito às restrições físicas e ecológicas que nos leva a pensar na otimização como uma solução para o problema. Neste trabalho consideramos dois objetivos e duas restrições na otimização do uso do solo. As funções objetivo são:

1) *Minimizar o custo de realocação do uso do solo.* Foi definido para cada tipo de solo um custo relativo de realocação para outro tipo de solo baseado no relatório do Zoneamento Ecológico Econômico de Goiânia (2008). Entre os custos estipulados, o mais oneroso é a realocação do uso atual de solo urbana, recebendo valor 10. A realocação de área de agricultura para uso urbano tem custo 1 e para uso de preservação, 5. A realocação de área de preservação para qualquer outro tipo de uso tem o custo relativo mais baixo, 1.

$$f1 = \sum_{i=1}^R \sum_{j=1}^C c_{i,j}$$

Onde c_{ij} é o custo de realocação do uso do solo atual para o uso do solo inferido pela otimização.

2) *Minimizar a diferença de aptidão do uso do solo.* A aptidão do uso do solo é definida por um estudo prévio feito por especialistas do Instituto de Desenvolvimento Tecnológico do Centro Oeste e representada por uma matriz bidimensional de mesma dimensão do uso do solo atual. A aptidão é definida para os seguintes tipos de uso do solo: 2-agricultura, 3-pastagem, 4-preservação.

$$f_2 = \sum_{i=1}^R \sum_{j=1}^C (e_{i,j} - p_{i,j})$$

Onde $p_{i,j}$ é a aptidão da unidade de solo definida previamente por especialistas.

As restrições foram definidas como:

- 1) Um conjunto de unidades contíguas de um tipo de uso do solo deve ter uma área mínima definida para a categoria de uso do solo;
- 2) Áreas de proteção permanente não podem ter seu uso alterado, conforme Lei Complementar nº 31 de 29/12/1994, que “Dispõe sobre o uso e a ocupação do solo nas Zonas Urbana e de Expansão Urbana do Município de Goiânia e estabelece outras providências urbanísticas”.

2.2. Algoritmo Genético (AG)

Neste trabalho o procedimento adotado segue a seguinte estrutura de execução:

1. *Início*: é gerada uma população inicial de n indivíduos aleatoriamente.
2. *Adaptação*: avalia-se o grau de adaptação de cada indivíduo da população através da função objetivo $f(x)$. Como o problema é de minimização, os indivíduos que possuem o menor valor de $f(x)=f_1(x)+f_2(x)$ são considerados os indivíduos mais aptos.
3. *Reprodução*: os indivíduos mais adaptados, isto é, que têm os menores valores de $f(x)$ de acordo com a taxa de cruzamento definida, se reproduzem e criam um novo conjunto de indivíduos. Os novos indivíduos são formados pelo cruzamento clássico, combinando-se os genes dos pais a partir de um ponto de corte definido aleatoriamente.
4. *Mutação*: cada indivíduo pode sofrer mutação em posição aleatória, segundo uma probabilidade de mutação pré-definida. A mutação é um operador genético que garante a diversidade genética, minimizando a convergência prematura, contornando ótimos locais.
5. *Criação da nova população*: a nova geração é composta por 1% de indivíduos selecionados por elitismo, ou seja, por aqueles que possuem os menores valores para a função objetivo $f(x)$ e pelos indivíduos selecionados pela técnica de torneio entre os indivíduos restantes.
6. *Teste da condição de parada e início de nova iteração*: verifica-se se o número total de iterações chegou ao fim. Se sim, a execução do algoritmo é interrompida e o indivíduo mais apto da população atual é a solução do problema. Caso contrário, o algoritmo volta ao passo 2.

Em problemas multiobjetivos, como é o caso do uso do solo, o conjunto de soluções ótimas define um vetor de decisão no qual um objetivo não pode ser melhorado sem que haja a degradação de outro. Este vetor é conhecido como pareto-ótimo e pode ser utilizado para verificar a qualidade da otimização. A qualidade é definida pela convergência e diversidade das soluções na fronteira de pareto, que pode ser visualizada no gráfico onde a abscissa e a ordenada representam os dois objetivos do problema. As soluções que estão na fronteira de pareto são chamadas não dominadas, como as soluções A e B da Figura 2, pois, no caso de problemas de minimização, seus valores para as funções objetivo são menores que para as soluções ditas dominadas, solução C da Figura 2.

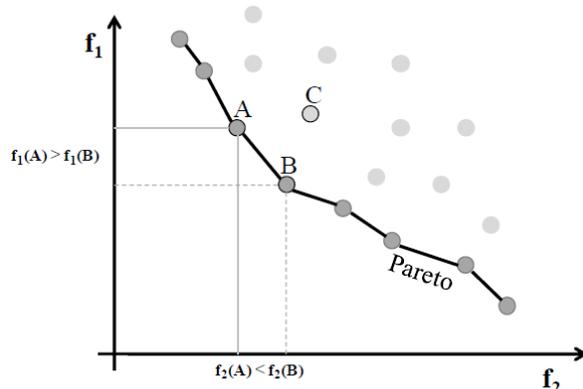


Figura 2. Análise da fronteira de Pareto (fonte: Vieira, 2011).

O uso de mapas gerados em ambiente SIG em conjunto com os AGs possuem grande potencial na otimização de uso do solo. Entretanto, devido à complexidade intrínseca a problemas espaciais, o tempo de cálculo e os recursos computacionais utilizados devem ser considerados. Quanto melhor o processador utilizado para conduzir os experimentos, menor será o custo computacional, como evidenciado por Porta et al. (2012), que utilizaram computação de alto desempenho para realizar a otimização de mapas para a área de Guitiriz (Espanha) e, entre os procedimentos adotados, os autores utilizaram processamento paralelo que auxilia na diminuição do custo computacional.

Em problemas de otimização com AG, devem ser realizados diversos experimentos para testar qual a configuração do algoritmo que melhor se adapta ao problema estudado. Portanto, neste trabalho foram conduzidos diversos testes, modificando os parâmetros do algoritmo. A população inicial foi testada com 12, 60, 100, 200 e 500 indivíduos, cada qual com taxas de cruzamento de 50, 80 e 100%. Para o número de gerações foram testados os valores de 20, 100 e 200 gerações. Para todos os casos a taxa de mutação foi fixada com probabilidade de 5%. O tempo de processamento do algoritmo para uma matriz de dimensões 60x60 executado em um processador Pentium Dual-Core 2.10GHz com 4GB de memória RAM foi em média 24 minutos.

Para o desenvolvimento das etapas do AG e sua aplicação seguindo as funções objetivos e restrições dispostas na Seção 2.1 foi utilizado o *software* MATLAB (The Mathworks, 2011).

2.3. Estudo de caso

Para o nosso estudo, alguns testes foram realizados para identificar se a modelagem proposta neste trabalho se adéqua a problemas espaciais de otimização de uso do solo. Para tanto foram utilizados dados obtidos através do Ministério do Meio Ambiente para a microregião de Meia Ponte, região metropolitana da cidade de Goiânia, representada na Figura 3. A área está compreendida entre as latitudes 17.30°S e 18°S e entre as longitudes 49°W e 49.5°W (datum horizontal WGS-84).

Para aplicar neste trabalho os mapas originais de aptidão e uso do solo obtidos no Ministério do Meio Ambiente foram simplificados. O mapa original de aptidão é composto por três tipos de uso diferentes, sendo: (2) agricultura; (3) pastagem e (4) preservação, e assim foi mantido. Já o mapa de uso atual do solo da região, que originalmente era composto por tipo urbano, diferentes tipos de uso relacionados a agricultura ('Pivo Central', 'Agricultura de Ciclo Curto' e '50 a 80% Agricultura'), diferentes tipos de uso relacionados a pastagem ('> 80% Pastagens' e '> 50% a 80% Pastagens/Vegetação Natural') e diferentes tipos de uso relacionados a vegetação nativa (florestas de diversos tipos, savanas e reflorestamentos), foi classificado em (1) urbano, (2) agricultura; (3) pastagem e (4) preservação.

Para avaliar a eficiência do AG desenvolvido para otimizar o uso do solo foram realizados experimentos utilizando os dados gerados a partir do mapa do estudo de caso, delimitado na Figura 3. Os dados de entrada consistiram de matrizes com 60 linhas e 60 colunas, totalizando 3600 células. Cada célula consiste em uma área aproximada de 500 x 500 m, correspondente a uma área total de 900 km². Através da modificação do tamanho da população inicial, taxa de cruzamento e número de gerações, foram testadas diferentes configurações do algoritmo, visando garantir a eficiência adequada, melhores resultados e uma boa aplicação para o estudo de caso proposto.

Neste trabalho, optou-se por uma simplificação dos mapas do estudo de caso e utilização somente de uma parcela da área total disponível devido à indisponibilidade de recursos computacionais adequados para a execução do algoritmo. Este procedimento, que pode prejudicar os resultados finais, foi adotado devido à complexidade de otimização de problemas espaciais e através da simplificação aplicada, foi possível manter um controle maior sobre a evolução do algoritmo. Um procedimento utilizando os dados reais e uma área maior pode ser útil em processos de tomada de decisão, como já evidenciado para Tongzhou (China, em área de 36x47 km) por Cao et al. (2012), assim como em outros trabalhos conduzidos para problemas de alocação de uso do solo (Esdrandaly, 2010; Janssen et al., 2008).

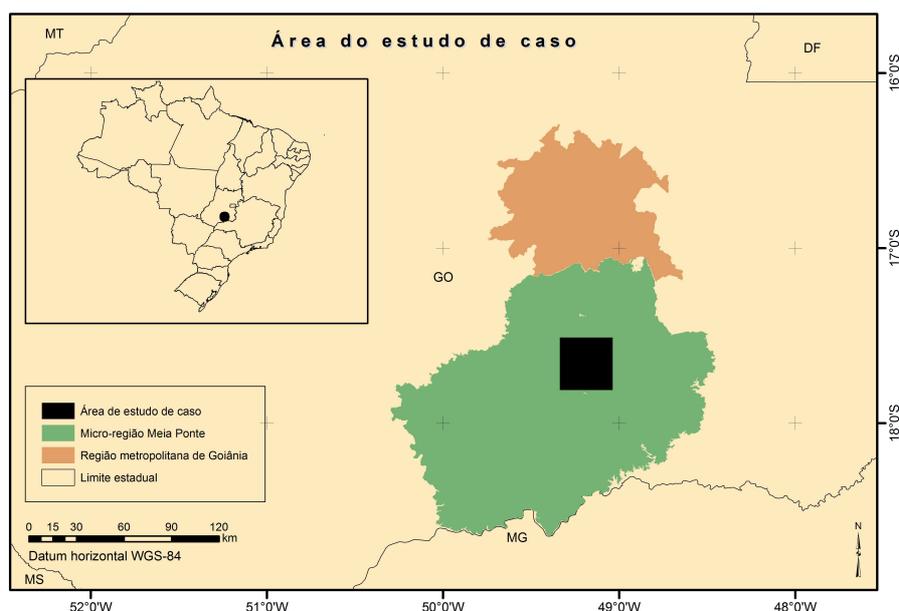


Figura 3. Localização da área de estudo.

Toda a manipulação dos mapas de uso do solo atual e aptidão e dos resultados do algoritmo foi realizada no *software ArcGIS*. O *ArcGIS Desktop* é um Sistema de Informações Geográficas que possui funcionalidades para: visualização de dados, consulta em base de dados, análise espacial, integração de dados, criação e edição de conteúdo geográfico. Ele é desenvolvido pela ESRI (*Environmental Systems Research Institute*) e atualmente encontra-se na versão 10.1.

3. Resultados e Discussão

A matriz de uso do solo atual representada na Figura 4(a) e a aptidão do solo representada na Figura 4(b) foram definidas para quatro categorias, como descritos acima na seção 2 deste trabalho. Baseado nestas duas matrizes, previamente fornecidas ao algoritmo, foi possível realizar diversos testes com o algoritmo de otimização desenvolvido neste trabalho. Os testes

foram realizados para diversos valores de tamanho da população (12, 60, 100, 200, 500), taxa de cruzamento (50%, 80% e 100%) e número de gerações (50, 100, 200). Na Figura 4(c) é apresentado o resultado para a otimização com uma população de 100 indivíduos, taxa de cruzamento de 80% e 200 gerações.

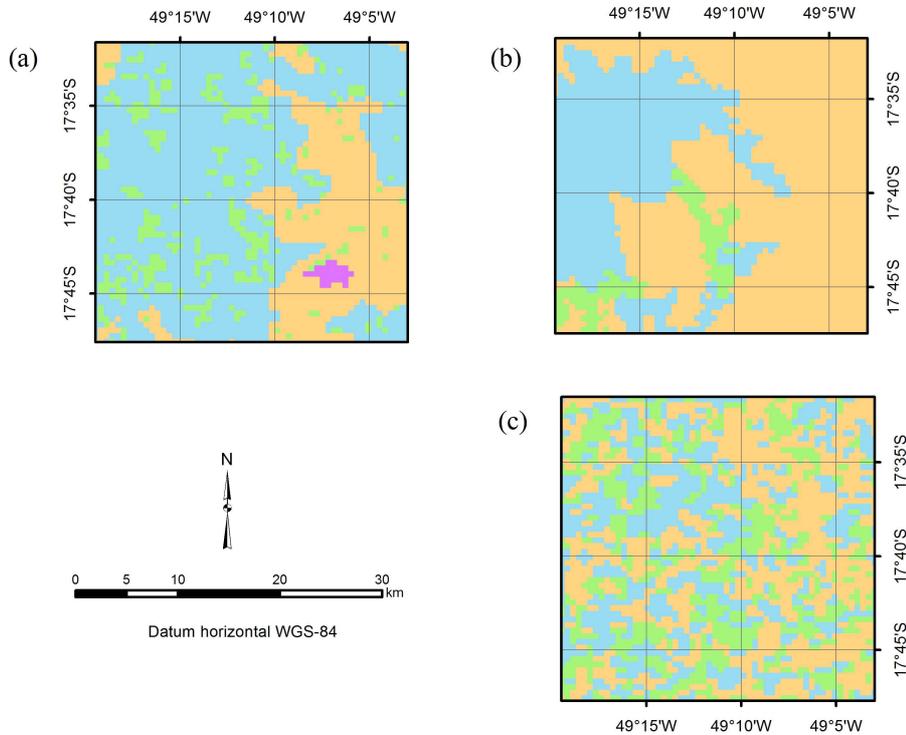


Figura 4. (a) Mapa de aptidão do solo definido por um especialista; (b) Uso do solo atual da região; (c) Resultado obtido para uma população de 100 indivíduos, com taxa de cruzamento de 80% e 200 gerações.

Na Figura 5 podemos analisar os indivíduos da 20ª geração da solução para o algoritmo com os seguintes parâmetros: população com 100 indivíduos e taxa de cruzamento de 80%. Pode-se perceber a tendência de convergência e diversidade das soluções na fronteira de Pareto, o que indica a possibilidade de oferecer ao gestor público uma gama de soluções possíveis para o gerenciamento ótimo do uso do solo.

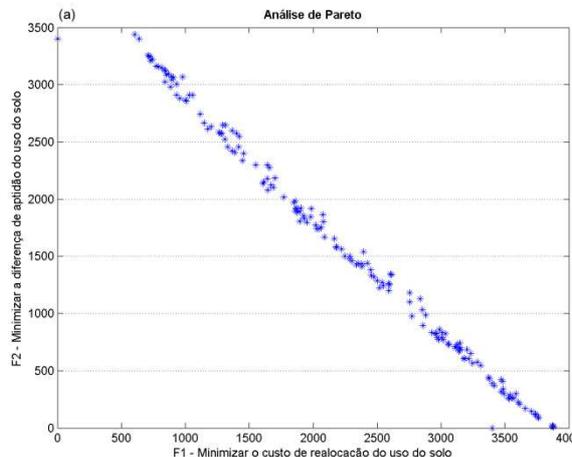


Figura 5. Gráfico representando a análise de Pareto para todas as soluções concorrentes.

Analisando a convergência do algoritmo, verificamos que há vários ótimos locais, mas que a partir da 70ª geração há uma estabilização da solução, como pode ser observado na Figura 4. Particularmente para o problema em questão, as soluções ótimas locais podem ser interessantes, pois a gestão do uso do solo é um problema complexo que envolve custos e está sujeito a diversas restrições e à legislação.

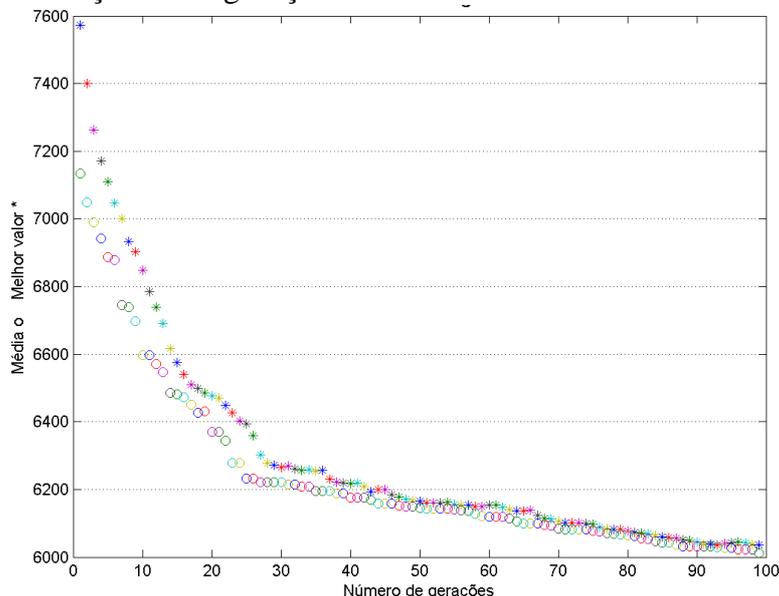


Figura 6. Análise da convergência para o teste com parametrização de 100 indivíduos, taxa de cruzamento de 80% e 100 gerações.

4. Conclusões

Com base na experiência obtida a partir do estudo de caso descrito, podemos concluir que a aplicação de meta-heurísticas pode ser muito eficiente na otimização de problemas de alocação espacial. Entre as vantagens do algoritmo genético, podemos destacar a capacidade de trabalhar com grandes quantidades de dados, como é o caso do dado espacial. Outra vantagem importante do algoritmo genético é que ele produz um conjunto de soluções, que podem ser proveitosamente exploradas para tomada de decisão por gestores em problemas multiobjetivos.

Neste trabalho optou-se pela elaboração de um algoritmo genético clássico para manipular dados espaciais e considerar as relações topológicas de vizinhança intrínsecas ao tipo de dado. O algoritmo foi testado em um estudo de caso com dados simplificados da região metropolitana de Goiânia. Para aplicar o algoritmo desenvolvido para uma problemática real, devem ser realizados ajustes no algoritmo, principalmente no que diz respeito às restrições legais referentes a área de estudo.

Observou-se que a convergência do algoritmo foi prematura para um ótimo local. Neste caso, devem ser implementadas técnicas de diversificação da população, seja através da mutação ou na seleção de soluções. O tempo computacional alto do algoritmo para o estudo de caso nos leva a pensar no processamento paralelo como um trabalho futuro para otimização de áreas maiores. Outro trabalho futuro está relacionado à restrição de área mínima contígua, que deve respeitar as leis vigentes da região estudada e deve ser diferente para cada tipo de uso do solo.

Agradecimentos

Agradecemos às professoras Beatriz de Souza Leite Pires Lima e Solange Guimarães pelo conhecimento transmitido na disciplina de “Algoritmos Computacionais Inspirados na Natureza” do curso de pós-graduação da Engenharia Civil da COPPE/UFRJ, onde surgiu a ideia deste trabalho. E em especial à professora Neida Maria Patias Volpi da Engenharia da Produção da UFPR pela valiosa colaboração durante o desenvolvimento do trabalho.

Referências Bibliográficas

- Balling, R. J.; Taber, J. T.; Brown, M. R.; Day, K.. Multiobjective urban planning using genetic algorithm. **Journal of Urban Planning and Development**. v. 125(2), p. 16–99, 1999.
- Cao, K.; Huang, B.; Wang S.; Lin R. Sustainable land use optimization using Boundary-based Fast Genetic Algorithm, Computers, Environment and Urban Systems. **Computers, Environment and Urban Systems**. v. 36, n. 3, p. 257-269, 2012.
- Datta, D.; Deb, K.. Design of optimum cross-sections for load-carrying members using multi-objective evolutionary algorithms. **International Journal of Systemics, Cybernetics and Informatics (IJSCI)**. p. 57-63, 2006.
- Eldrandaly, K. A GEP-based spatial decision support system for multisite land use allocation. **Applied Soft Computing Journal**. V. 10, p. 694-702, 2010.
- Han, J.; Itoh, R.; Nishiyama, S.; Yamazaki, K.. **Application of Structure Optimization Technique to Aluminum Beverage Bottle Design**. **Structural and Multidisciplinary Optimization**. v. 29(4), p. 304-311, 2005.
- Janerette, G.D.; Wu, J. **Analysis and simulation of land-use change in the central Arizona - Phoenix region, USA**. **Landscape Ecology**. v. 16, p. 611-626, 2001.
- Janssen, R.; Herwijnen, M.V.; Stewart, T.J.; Aerts, J.C. **Multiobjective decision support for land- use planning**. **Environment Planning**. v. 35 (4), p. 740-756, 2008.
- Luo, D.; Xu, Y.Q.; Shao, X.M. Advances and prospects of spatial optimal allocation of land use. **Progress in Geography Journal**. v. 5, p. 791-797, 2009.
- Porta, J.; Parapar, J.; Doallo, R.; Rivera, F.; Santé, I.; Crecente, R. High performance genetic algorithm for land use planning. **Computers, Environment and Urban Systems**. Available online in: <http://www.sciencedirect.com/science/article/pii/S0198971512000506>, 2012.
- Sarkar, D.; Modak, M. Optimization of fed-batch bioreactors using genetic algorithms, **Chemical Engineering Science**. v. 58, p. 2283–2296, 2003.
- Van Dijk, S.; Thierens, D.; Berg, M. Using genetic algorithms for solving hard problems in GIS. **GeoInformatica**. v. 6(4), p. 381–413, 2002.
- Vieira, I. N. **Otimização multi-objetivo de rotas de dutos submarinos através de sistemas imunológicos artificiais**. Tese de doutorado. Programa de Engenharia Civil, COPPE. UFRJ. 2011.
- Vink, K.; Schot, P.P. Multiple-objective optimization of drinking water production strategies using a genetic algorithm, **Water Resources Research**. v. 38(9), n. 1181, 2002.
- Xiao, N.; Bennett, D.A.; Armstrong, M.P. Interactive evolutionary approaches to multiobjective spatial decision making: A synthetic review. **Computers, Environment and Urban Systems**. v. 31, p. 232–252, 2001.