

Um algoritmo heurístico híbrido para criação de modelos com representação explícita de relacionamentos espaciais

Adair Santa Catarina¹
Antônio Miguel Vieira Monteiro²
João Ricardo de Freitas Oliveira²

¹ Colegiado de Informática - Universidade Estadual do Oeste do Paraná - UNIOESTE
Rua Universitária, 2069 – Jardim Universitário – 85819-110 – Cascavel – PR, Brasil
asc@unioeste.br

² Divisão de Processamento de Imagens - Instituto Nacional de Pesquisas Espaciais – INPE
São José dos Campos – SP, Brasil
{miguel, joao}@dpi.inpe.br

Abstract. A fundamental idea to analyze and interpret spatial data is spatial dependence. The spatial dependence results from that “everything is related to everything else, but near things are more related than distant things”. Spatial data analysis uses heuristic algorithms as Genetic Algorithms; however, these algorithms ignore the spatial dependence. In this work we present an architecture for a Hybrid Heuristic Algorithm (HHA) with explicit representation of spatial relationships. The architecture proposed includes a Generalized Proximity Matrix that represents the spatial association between objects in space, in other words the spatial relationships. The HHA embody two heuristics: Genetic Algorithms and Simulated Annealing. It creates two types of models: Simple Linear Regression and Weighted Linear Regression. The first type is equal to Statistic Linear Regression and the second type is a Linear Regression that computes the spatial relationships gives by the Generalized Proximity Matrix. We tested the HHA producing Linear and Weighted Linear Regression. The two types of models adjusted rightly show the relations between dependent and independent variables. These results allow us to conclude that HHA can be used as a tool for spatial analysis.

Palavras-chave: model breeder, spatial analysis, genetic algorithm, simulated annealing, construtor de modelos, análise espacial, algoritmo genético.

1. Introdução

A capacidade para produzir, armazenar e recuperar dados espaço-temporais cresceu significativamente nos últimos anos. Contribuíram para isto o aumento na oferta de imagens de satélites em diferentes resoluções espaciais, espectrais e temporais, o uso de GPS para coleta direta de dados e o acesso a bases de dados demográficas mais detalhadas, como o censo brasileiro de 2000, que dividiu os municípios em diversas unidades censitárias.

Apesar dos SIG atuais incrementarem nossa capacidade de analisar dados espaciais, esses ainda estão no começo da efetiva integração entre métodos de exploração e análise deste tipo de dado. Os métodos para explorar e analisar dados espaciais advém de áreas tradicionais como a Estatística e de áreas emergentes como a Inteligência Computacional, Sistemas Complexos e outras (Openshaw e Openshaw, 1997; Couclelis, 1998; Openshaw e Abrahart, 2000; Câmara e Monteiro, 2001).

A busca por informações contidas em bancos de dados geográficos, usando computação intensiva, originou uma nova área de conhecimento: a geocomputação [Openshaw e Abrahart 1996].

Os Algoritmos Genéticos (AG) são uma classe de algoritmos computacionais intensivos usados em geocomputação. AG são usados em software como os *Model Breeders* (Openshaw e Openshaw, 1997; Santa Catarina *et al.*, 2005) e GARP (*Genetic Algorithm to Rule-set Prediction*) (Stockwell e Peters, 1999). A combinação entre um AG e outras técnicas heurísticas como o *simulated annealing*, *hill climbing* e busca tabu, por exemplo, define um algoritmo heurístico híbrido (AHH). A combinação de diferentes algoritmos visa utilizar o

que de melhor há neles. Espera-se, então, que o novo algoritmo seja capaz de produzir melhores resultados do que aqueles encontrados pelos algoritmos aplicados isoladamente.

Um princípio essencial para compreender e analisar problemas espaciais é originado da Primeira Lei da Geografia¹ (Tobler, 1970). A expressão quantitativa deste princípio é o efeito da dependência espacial (Câmara e Monteiro, 2001).

Este princípio mostra a importância em considerar o espaço próximo ao problema em estudo. Entretanto o espaço próximo, e portanto a dependência espacial, é negligenciada nos *Model Breeders* e no GARP. Esta negligência ocorre não apenas nestes softwares; ela deve-se a ausência de um AG que incorpore, em seu mecanismos evolutivos, os relacionamentos espaciais.

O objetivo principal deste trabalho é apresentar a arquitetura de um Algoritmo Heurístico Híbrido (AHH) com representação explícita de relacionamentos espaciais. Os relacionamentos espaciais serão representados numa matriz de proximidade generalizada (MPG) que permitirá o cômputo dos efeitos da dependência espacial. O AHH será empregado na análise de dados espaciais, construindo modelos que capturem os relacionamentos entre variáveis.

Além disso, através de representação explícita dos relacionamentos espaciais, é possível inserir o conhecimento prévio sobre os elementos naturais e artificiais presentes no espaço, cujos efeitos são significativos para o problema em estudo.

1.1 Algoritmos Genéticos

AG é um algoritmo de busca que simula o mecanismo evolucionário dos sistemas biológicos naturais, onde os indivíduos mais aptos têm maior probabilidade de se reproduzir gerando descendentes. Em problemas de otimização os AG tem apresentado boas soluções naqueles problemas onde as técnicas tradicionais de busca falham (Holland, 1975; Goldberg, 1989).

Inicialmente um AG gera um conjunto de soluções aleatórias codificadas em cadeias de dígitos binários, chamado população. Cada indivíduo da população é chamado de cromossomo, representando uma solução candidata para o problema. A avaliação de cada cromossomo determina seu índice de aptidão; indivíduos mais aptos têm maior probabilidade de gerarem filhos. Os cromossomos evoluem através de iterações sucessivas chamadas de gerações.

O processo de seleção e os operadores de cruzamento e mutação são os responsáveis por criar as novas gerações. O processo de seleção visa selecionar indivíduos, de acordo com seu índice de aptidão, para combiná-los através do operador de cruzamento, gerando novos indivíduos que mantêm características de seus pais. Posteriormente estes novos indivíduos podem ser modificados pelo operador de mutação.

Após várias gerações o AG pode produzir soluções aceitáveis para um problema. A Figura 1 mostra o fluxograma básico de um AG.

1.2 Simulated Annealing

Esta heurística é uma metáfora de um processo térmico utilizado para obtenção de estados de baixa energia num sólido. O processo consiste de duas etapas: na primeira a temperatura do sólido é aumentada para um valor máximo no qual ele se funde; na segunda o resfriamento deve ser realizado lentamente até que o material se solidifique.

Quando a temperatura baixa, os átomos gradualmente movem-se em direção a uma estrutura regular; somente com pequena probabilidade incrementarão suas energias. Esse processo foi simulado em computador, com sucesso, por Metropolis *et al.* (1953). Kirkpatrick

¹ “Everything is related to everything else, but near things are more related than distant things” [Tobler, 1970, p.236].

et al. (1983) desenvolveram um algoritmo de utilização genérica análogo ao de Metropolis, denominado Algoritmo *Simulated Annealing*, apresentado na Figura 2.

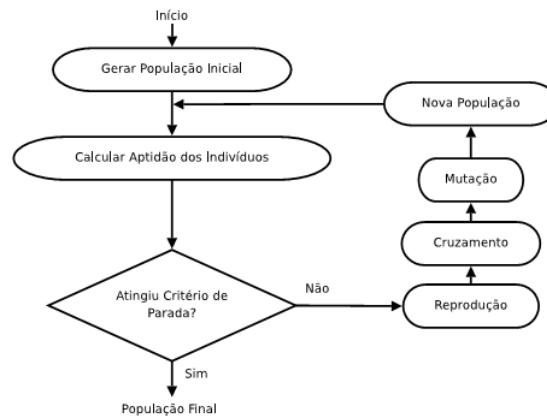


Figura 1. Fluxograma básico de um Algoritmo Genético

```

Ler  $\alpha$  e NR; //Constante de resfriamento ( $\alpha < 1$ ) e número de repetições
S = S(); //Conjunto aleatório de soluções iniciais
T = LS; //Limite superior
TMIN = LI; //Limite inferior

Enquanto (T > TMIN) faça
  Para i = 1 até NR faça
    Gerar uma solução S' a partir de S; //perturbação de S
    Avaliar a variação de energia; // $\Delta E = g(S') - g(S)$ ;
    Se (variação de energia  $\leq 0$ ) então S = S'
    Senão
      Gerar aleatoriamente Rnd; //no intervalo [0, 1]
      Se (Rnd < exp(-variação de energia / T) então S = S';
    Fim se;
  Fim Para;
  T = T *  $\alpha$ ;
Fim enquanto;
  
```

Figura 2. O algoritmo *Simulated Annealing*

Segundo o algoritmo apresentado na Figura 2, se uma solução candidata S' é melhor que a solução corrente S , ou seja $g(S') \leq g(S)$, esta é aceita com probabilidade 1. Caso contrário a solução candidata é aceita com uma dada probabilidade. Essa probabilidade é maior à medida em que for menor a variação de energia, definida por ΔE . Ao mesmo tempo, à medida que há um decréscimo da temperatura T , o algoritmo torna-se mais seletivo, passando a aceitar, com menor frequência, soluções que apresentem grande aumento na variação de energia, isto é, soluções que sejam muito piores que a solução corrente. Essa probabilidade tende a zero à medida que a temperatura se aproxima do ponto de congelamento definido por TMIN.

O algoritmo *Simulated Annealing* pode ser considerado como uma extensão do método original de busca local. Porém, dispõe de um mecanismo que permite sua escapada de um ótimo local. O aceite de uma nova solução não depende única e exclusivamente do seu valor; mesmo apresentando um valor pior que o da solução corrente, uma nova solução pode ser aceita de forma probabilística.

1.3 Matriz de Proximidade Generalizada

A MPG é uma variação da matriz de proximidade. Estas matrizes expressam quanto dois objetos geográficos são vizinhos, considerando suas posições no espaço. Geralmente esta medida de vizinhança é estimada em termos de adjacência topológica ou distância Euclidiana.

Representamos a relação de vizinhança entre dois objetos O_i e O_j por W_{ij} . As opções mais comuns para definir W_{ij} são:

- $W_{ij} = 1$, se O_i é topologicamente adjacente a O_j ; $W_{ij} = 0$ nos outros casos;
- $W_{ij} = 1$, se a distância $(O_i, O_j) < d$; $W_{ij} = 0$ nos outros casos.

A MPG é composta por um conjunto de objetos geográficos O , um grafo G e uma matriz de proximidade V . Células regulares ou polígonos, de acordo com a representação espacial empregada, representam o conjunto de objetos geográficos O . Nós e arcos compõem o grafo G ; cada nó representa um objeto e os arcos representam as relações espaciais entre dois nós (Pedrosa, 2003).

2. Arquitetura do Algoritmo Heurístico Híbrido com Representação Explícita de Relacionamentos Espaciais

Na arquitetura proposta a representação explícita dos relacionamentos espaciais dá-se pela incorporação de uma MPG e um peso associado a cada variável independente. A Figura 3 mostra a arquitetura proposta.

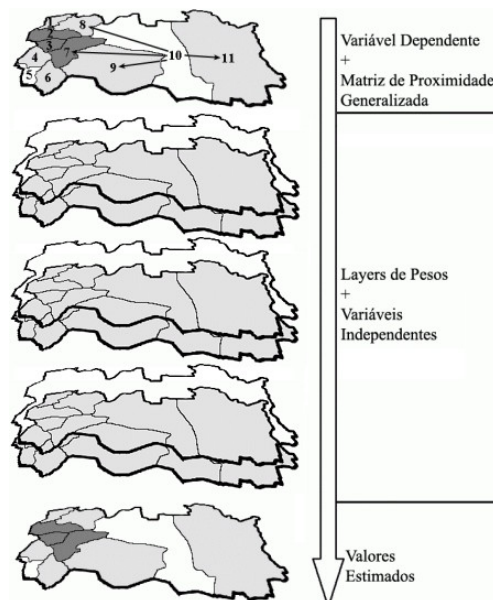


Figura 3. Arquitetura do AHH com representação explícita de relacionamentos espaciais

A MPG atende a dois objetivos: incorporar os relacionamentos espaciais existentes entre os objetos geográficos e a representação do conhecimento sobre os elementos naturais e artificiais presentes no espaço. Os pesos W_{ij} na MPG são responsáveis por representar o conhecimento pré-existente sobre os elementos no espaço. Os pesos, representados pelas *layers* de pesos, quantificam o efeito das variáveis independentes sobre a variável dependente.

2.1 Dados de Entrada e sua Representação

Esta subseção mostra a estrutura dos dados de entrada e exemplifica como os relacionamentos espaciais são representados. Os dados de entrada são: objetos geográficos e seus relacionamentos espaciais (MPG), as variáveis dependente e independentes. Para cada objeto geográfico haverá um conjunto de dados de entrada conforme apresentado na Figura 3.

IdPol	NRE	RelEsp	W_{ij}	X_0	X_1	X_2	...	X_n
-------	-----	--------	----------	-------	-------	-------	-----	-------

Figura 3. Estrutura dos dados de entrada

onde:

- *IdPol*: número que identifica o objeto geográfico;
- *NRE*: Número de objetos aos quais o objeto *IdPol* está relacionado
- *RelEsp*: conjunto com os *NRE* objetos espacialmente relacionados com o objeto *IdPol* ;
- *W_{ij}*: conjunto com *NRE* pesos, correspondendo à relação de vizinhança entre dois objetos *O_i* e *O_j*;
- *X₀*: variável dependente;
- *X_k*: variáveis independentes. *k* = 1..*n*;
- *n*: número de variáveis independentes;

A Figura 4 mostra um exemplo de entrada de dados para o objeto geográfico da Figura 2, cujo *IdPol* = 10.

10	5	7, 8, 9, 10, 11	1, 2, 1, 1, 2	1,318	0,508	0,705	1,1
----	---	-----------------	---------------	-------	-------	-------	-----

Figura 4. Exemplo de uso da estrutura de dados de entrada

Esta figura mostra que o polígono com *Id* = 10 tem *NRE* = 5 relacionamentos de vizinhança com os polígono 7, 8, 9, 10 e 11 e $W_{ij} = \{1; 2; 1; 1; 2\}$. O valor da variável dependente (*X₀*) para o polígono 10 é 1,318 e os valores das variáveis independentes $X_k = \{0,508; 0,705; 1,1\}$.

2.2 Codificação, Avaliação da Aptidão e Operadores Genéticos

Num AG o processo evolutivo ocorre sobre os cromossomos. Os cromossomos codificam as soluções possíveis para o problema. Na arquitetura proposta o processo evolutivo ocorre sobre os pesos associados às variáveis independentes; portanto são estes pesos que devem ser codificados nos cromossomos. A MPG é um dado de entrada estático; ela representa o conhecimento sobre os relacionamentos espaciais que não mudam durante o processo evolutivo.

A codificação utilizada é decimal; cada gene do cromossomo corresponde a um peso associado à uma variável independente. O último gene do cromossomo corresponde à constante do modelo. A Figura 5 mostra a estrutura geral de um cromossomo.

Peso X_1	Peso X_2	...	Peso X_n	Const
1,453	2,317	...	-2,112	0,015

Figura 5. Estrutura geral de um cromossomo

Os cromossomos são avaliados por duas equações diferentes, de acordo com o tipo de modelo que se está ajustando.

No caso de regressão linear simples o objeto geográfico *O_i* está relacionado apenas consigo mesmo e com relação de vizinhança $W_{ij} = 1$. Os cromossomos para este tipo de modelo são avaliados pela Equação 1.

$$\hat{X}_{0i} = \sum_{k=1}^n (PesosX_k \cdot X_{ki}) + Const \quad (1)$$

No caso de regressão linear ponderada os cromossomos são avaliados pela Equação 2.

$$\hat{X}_{0i} = \sum_{k=1}^n \left(PesosX_k \cdot \left(\frac{\sum_{j=1}^{NRE_i} (W_{ij} \cdot X_{kj})}{\sum_{j=1}^{NRE_i} W_{ij}} \right) \right) + Const \quad (2)$$

De acordo com a Equação 2, para cada objeto geográfico as variáveis independente X_k são estimadas através de uma média ponderada, considerando os pesos advindos das relações de vizinhança descritas na MPG.

A aptidão final de um cromossomo é calculada pela soma dos erros ao quadrado, de acordo com a Equação 3, onde m é o número total de objetos geográficos. O objetivo final do AHH é minimizar o valor de Aptidão.

$$Aptidão = \sum_{i=1}^m (X_{0i} - \hat{X}_{0i})^2 \quad (3)$$

Os operadores genéticos usados são a seleção, o cruzamento aritmético de Michalewicz (1994) e a mutação. A probabilidade de seleção é proporcional a aptidão de cada cromossomo. Os parâmetros tamanho da população, taxas de cruzamento e mutação, temperatura mínima, temperatura máxima e a constante de resfriamento são definidos empiricamente. O algoritmo pára após a realização de um número pré-definido de ciclos.

A Figura 6 apresenta o núcleo de otimização do AHH, mostrando a hibridização de um AG com SA.

```
void AGSA::Run(){
    CreatePop(); //Criando a população inicial
    EvalPopAG(); //Avaliando a população inicial

    for (i = 0; i < NumCiclos; ++i){
        EvolPopAG(); //Gerando uma nova população através de uma iteração do AG
        EvalPopAG(); //Avaliando a nova população gerada

        while (TAtual > TMinima){ //Núcleo do SA
            EvolPopSA(); //Executando uma iteração do SA
            EvalPopSA(); //Avaliando o resultado obtido na iteração do SA
        }
        Populacao[0] = GBest; //Restaura o melhor indivíduo obtido até o momento
        ResetTAtual(); //Reinicializa temperatura no SA --> TAtual = TMax
    }
}
```

Figura 6. O núcleo de otimização do AHH

3 Testes Realizados com o AHH

Para testar o AHH utilizamos um conjunto de dados obtidos de <http://www.sidra.ibge.gov.br/cd/default.asp>. Os dados provêm do Censo Demográfico 2000 e são apresentados na Tabela 1.

Os testes aqui realizados visam encontrar um modelo que relaciona a variável dependente X_0 = “número de filhos nascidos vivos”, com as variáveis independentes X_1 = “número de domicílios com banheiro”, X_2 = “número de domicílios cujo responsável tem 8 ou mais anos de estudo” e X_3 = “número de domicílios onde a renda é maior que 3 salários mínimos”.

Para utilização com o AHH os dados da Tabela 1, referentes às variáveis dependente e independentes, foram padronizados através da Equação 4.

$$X_p = \frac{X_i - \bar{x}}{s} \quad (4)$$

Os dados apresentados na Tabela 1 incluem informações sobre a MPG. Para o teste com o modelo de regressão linear simples basta desconsiderar as informações da MPG, assumindo que cada objeto relaciona-se apenas consigo mesmo, com $W_{ij} = 1$.

Os parâmetros do AHH foram ajustados com os seguintes valores: tamanho da população = 50, número de ciclos = 10, temperatura mínima = 0,001, temperatura máxima = 3, constante de resfriamento = 0,9, número de repetições para cada ciclo do SA = 3, tamanho da elite da população = 1, taxa de cruzamento = 80% e taxa de mutação = 1%.

Tabela 1. Dados utilizados nos testes do AHH

Estado	Id	NRE	RelEsp			W _{ij}				X ₀	X ₁	X ₂	X ₃
Rondônia	1	2	1	2		1	1			1.202.863	209.451	274875	174857
Acre	2	2	2	1		1	1			519.328	49.142	106099	53380
Amazonas	3	1	3			1				2.423.692	300.476	611136	250182
Roraima	4	1	4			1				265.621	47.634	83363	39898
Pará	5	1	5			1				5.538.281	580.556	1211265	537327
Amapá	6	1	6			1				405.330	52.668	126069	51216
Tocantins	7	1	7			1				1.063.302	159.651	232773	101387
Maranhão	8	1	8			1				5.577.060	403.502	912504	307044
Piauí	9	1	9			1				2.926.783	316.964	457217	175372
Ceará	10	1	10			1				7.693.171	1.013.537	1433765	552265
Rio Grande do Norte	11	1	11			1				2.935.527	481.658	612604	244271
Paraíba	12	2	12	13		1	1			3.847.081	595.675	619956	260108
Pernambuco	13	2	13	12		1	1			8.015.103	1.446.109	1753714	703468
Alagoas	14	1	14			1				2.982.700	427.277	443524	188763
Sergipe	15	1	15			1				1.770.399	335.513	344433	144054
Bahia	16	2	16	17		1	1			12.761.579	2.095.830	2499704	1049684
Minas Gerais	17	4	17	16	19	20	1	1	1	15.763.185	4.329.948	4840801	2574913
Espírito Santo	18	1	18			1				2.677.347	786.522	921201	468228
Rio de Janeiro	19	3	19	17	20		1	1	1	11.131.946	4.114.484	5496427	2797601
São Paulo	20	4	20	17	19	21	1	1	1	28.499.542	10.176.557	13662174	7543605
Paraná	21	3	21	20	22		1	1	1	8.313.533	2.419.203	2992480	1569988
Santa Catarina	22	3	22	21	23		1	1	1	4.538.950	1.397.872	1651501	1025313
Rio Grande do Sul	23	2	23	21			1	1		8.266.032	2.780.508	3252382	1909129
Mato Grosso do Sul	24	2	24	25			1	1		1.790.394	502.762	546199	296185
Mato Grosso	25	2	25	24			1	1		2.075.588	508.834	622448	356853
Goiás	26	1	26				1			4.218.333	1.243.066	1329209	729428
Distrito Federal	27	1	27				1			1.499.905	518.354	865115	401411

3.1 Modelo de Regressão Linear Simples

O modelo de regressão linear simples ajustado pelo AHH para os dados da Tabela 1, padronizados pela Equação 4, é apresentado na Equação 5. Este modelo apresentou aptidão mínima igual a 0,9298. A aptidão medida corresponde à soma dos desvios quadráticos.

$$\hat{X}_i = 1,5605 \cdot X_1 + 2,7114 \cdot X_2 - 3,3185 \cdot X_3 + 0,0025 \quad (5)$$

Comparando este resultado com o modelo ajustado através do método dos mínimos quadrados (Equação 6) percebemos que são similares. A soma dos desvios quadráticos deste modelo foi igual a 0,9282.

$$\hat{X}_i = 1,5931 \cdot X_1 + 2,6806 \cdot X_2 - 3,3269 \cdot X_3 \quad (6)$$

3.2 Modelo de Regressão Linear Ponderado

O modelo de regressão linear ponderado ajustado pelo AHH para os dados da Tabela 1, padronizados pela Equação 4, considerando os relacionamentos de vizinhança representados pela MPG, é apresentado na Equação 7. A soma dos desvios quadráticos deste modelo foi igual a 1,0946.

$$\hat{X}_i = 1,4262 \cdot X_1 + 3,4663 \cdot X_2 - 4,0000 \cdot X_3 - 0,0312 \quad (7)$$

Os relacionamentos e pesos apresentados na Tabela 1, correspondentes à MPG, foram definidos arbitrariamente. Imaginamos que alguns Estados da Federação exercem algum tipo de influência sobre seus vizinhos mais próximos. Por exemplo, São Paulo exerce influência

sobre os Estados de Minas Gerais, Rio de Janeiro e Paraná, bem como é influenciado por estes.

Apesar deste modelo não ser comparável ao modelo linear simples, observamos que a aptidão do mesmo não é muito diferente àquela apresentada pelo modelo linear simples, mostrando que pode também ser utilizado para estimar adequadamente os valores de X_i a partir do conjunto de variáveis independentes.

4. Conclusões

O Algoritmo Heurístico Híbrido (AHH) proposto apresentou bons resultados, tanto ajustando modelos lineares simples quanto modelos lineares ponderados. A comparação do modelo linear simples com o modelo ajustado pelo método dos mínimos quadrados confirma que o AHH apresenta resultados confiáveis.

Entretanto, a principal vantagem apresentada pelo AHH é o fato dele trabalhar com diferentes tipos de modelos utilizando a mesma estrutura de codificação. Ou seja, mostrou-se uma solução genérica para o problema de analisar dados espaciais procurando relações existentes entre variáveis dependentes e independentes, com e sem relacionamentos espaciais.

O AHH apresenta outra característica favorável; ele é extensível. Os testes realizados ajustaram apenas modelos lineares, mas ele pode ser utilizado para ajustar modelos não lineares. Para tanto basta mudar a estrutura do cromossomo (Figura 5) incluindo um número maior de coeficientes e a escrever uma função de avaliação do cromossomo, considerando termos não lineares.

Como trabalho futuro aplicaremos o AHH no processo de modelagem de distribuição de espécies. A incorporação da MPG representa uma inovação que permitirá computar os efeitos dos relacionamentos espaciais sobre a distribuição potencial das espécies estudadas.

5. Referências Bibliográficas

- Câmara, G.; Monteiro, A. M. V. Geocomputation techniques for spatial analysis: are they relevant to health data? *Cadernos Saúde Pública*, v. 17, n. 5, p. 1059-1081, 2001.
- Couclelis, H. Geocomputation in context. In: Longley, P. A.; Brooks, S. M.; McDonnell, R.; MacMillan, B. **Geocomputation: a primer**. New York : John Wiley & Sons, 1998.
- Goldberg, D. E. **Genetic algorithms in search, optimization & machine learning**. Reading : Addison-Wesley, 1989.
- Holland, J. H. **Adaptation in natural and artificial systems**. Ann Arbor : University of Michigan Press, 1975.
- Kirkpatrick, S.; Gelatt, D. C.; Vecchi, M. P. Optimization by simulated annealing. *Science*. v. 220, n. 4598. p. 671-680, 1983.
- Metropolis, W.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, E. Equation of state calculations by fast computing machines. *Journal of chemical physics*. v. 21, n. 6, p. 1087-1092, Jun. 1953.
- Michalewicz, Z. **Genetic algorithms + data structures = evolution programs**. 3.ed. Springer-Verlag, 1994.
- Openshaw, S.; Abrahart, R. J. Geocomputation. In: 1st. International Conference on Geocomputation, 1996, Leeds. **Proceedings...** Leeds : University of Leeds, 1996.
- Openshaw, S.; Openshaw, C. **Artificial intelligence in geography**. West Sussex : John Wiley & Sons, 1997.
- Openshaw, S.; Abrahart, R. J. Eds. **GeoComputation**. London : Taylor & Francis, 2000.
- Pedrosa, B. M. **Ambiente computacional para modelagem dinâmica**. 2003, 71 p. Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2003.
- Santa Catarina, A.; Oliveira, J. R. F.; Monteiro, A. M. V. Model Breeder: um algoritmo genético para criação de modelos. Disponível em: <http://hermes2.dpi.inpe.br:1905/col/dpi.inpe.br/hermes2@1905/2005/10.03.20.04/doc/Model%20Breeder%20-%20Worcap2005.pdf>. Acesso em: 01.set.2007.
- Stockwell, D. e D. Peters. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, v. 13, n. 2, p. 143-158, 1999.
- Tobler, W. R. A computer model simulation of urban growth in the Detroit region. *Economic Geography*, v. 46, n. 2, p. 234-240, 1970.