

Álgebra de Geo-Campos e suas Aplicações¹

JOÃO PEDRO CORDEIRO

SILVANA AMARAL

UBIRAJARA MOURA DE FREITAS

GILBERTO CÂMARA

Divisão de Processamento de Imagens - DPI

Instituto Nacional de Pesquisas Espaciais - INPE

P.O.Box 515

12201-010 São José dos Campos, SP, Brasil

{jpedro, silvana, bira, gilberto}@dpi.inpe.br

Abstract. This work presents a proposal for an algebra of geographical fields. Geographical fields are the mathematical expression of continuously-varying spatial data, such as thematic maps, digital terrain models and satellite imagery. The paper describes the initial implementation of the data manipulation language LEGAL, used in SPRING, a geographical information system developed by INPE.

Keywords: Geographical information systems, spatial data bases, spatial data manipulation language.

1 Introdução

A necessidade do entendimento cada vez mais detalhado dos problemas ambientais tem provocado uma demanda crescente na área de Geoprocessamento para que os sistemas de informação geográfica (SIG) passem a contar com ferramentas sofisticadas de modelagem.

Procurando atender a esta demanda, a equipe do INPE engajada no desenvolvimento do SPRING tem procurado caracterizar, de forma ampla, as operações sobre dados geográficos. Deste estudo nasceu o projeto da linguagem LEGAL (Linguagem Espacial para Geoprocessamento Algébrico), descrita em Câmara (1995).

Este trabalho descreve a implementação de operações da álgebra de campos geográficos, como parte da linguagem LEGAL, no sistema SPRING-2.0 (DPI et al., 1995).

Este trabalho está estruturado em cinco partes. Na seção 2, definimos o conceito de *geo-campo* e indicamos suas especializações (dados temáticos, modelos numéricos de terreno e dados de sensoriamento remoto). A seção 3 analisa trabalhos anteriores na literatura. Na seção 4, apresentamos a implementação parcial das operações da álgebra de geo-campos na

versão 2.0 do sistema SPRING. Finalmente, a seção 5 contém exemplos de aplicação.

2 Campos Geográficos

Sob um ponto de vista conceitual, existem duas grandes classes de dados geográficos: *geo-campos* e *geo-objetos* (Goodchild, 1992). Enquanto a primeira constitui o principal assunto deste trabalho, a última representa as entidades individualizáveis do domínio geográfico, que possuem atributos descritivos (usualmente armazenados num banco de dados convencional) e que podem ter várias representações geométricas. A formalização das operações envolvendo geo-objetos e seus atributos constitui extensa área de pesquisa (Egenhofer, 1994; Câmara, 1995) e não será abordada neste artigo.

Definição

Campos geográficos ou *geo-campos* representam a distribuição espacial de uma variável geográfica sobre uma região da superfície terrestre. Um geo-campo é caracterizado por:

- seu *domínio*, a descrição de uma dada região geográfica R ;

¹ Trabalho parcialmente financiado pelo CNPq (programa RHAe e programa Protem/CC- projeto GEOTEC).

- seu *contra-domínio*, que reflete o conjunto V de valores que podem ser assumidos pela variável geográfica;
- um *mapeamento* $\lambda:A \rightarrow V$ entre pontos em A e valores em V .

Dependendo do *contra-domínio* podemos caracterizar as seguintes especializações para os geo-campos:

- TEMÁTICO - um *geo-campo temático* se caracteriza por um mapeamento no qual V é um conjunto finito enumerável de elementos que definem os *temas* do geo-campo (figura 1).
- NUMÉRICO - um *geo-campo numérico* (também chamado de modelo numérico de terreno), tem como característica um mapeamento no qual V é um conjunto dos valores reais.
- DADO DE SENSOR REMOTO - uma instância desta classe (chamado também, por abuso de linguagem, de *imagem*) é caracterizada por um mapeamento cujo contradomínio corresponde à quantização da resposta obtida por um sensor (passivo ou ativo) para uma região geográfica (figura 2).

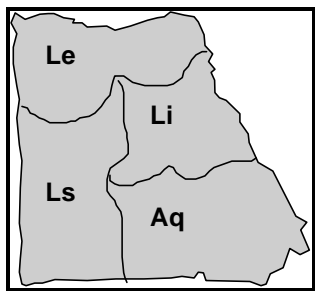


Figura 1 - Exemplo de geo-campo temático (Mapa de Solos)



Figura 2 - Exemplo de dado de sensoriamento remoto (imagem LANDSAT de Manaus)

Representações Geométricas

Geo-campos podem ser representados em um SIG sob vários formatos:

- GEO-CAMPOS NUMÉRICOS (modelos numéricos de terreno) podem ser representados por grades retangulares (regulares) ou triangulares.
- GEO-CAMPOS TEMÁTICOS podem ser representados por um conjunto de vetores estruturados topologicamente ou na forma de matriz.
- DADOS DE SENSOR REMOTO (imagens) são usualmente representadas por uma matriz de valores.

As vantagens e desvantagens de cada opção de armazenamento tem sido discutidas extensivamente na literatura (Worboys, 1995). O consenso é de que as representações matricial e vetorial, bem como por grades retangulares ou triangulares são alternativas úteis, e em geral os SIG devem oferecer todas.

3 Breve Revisão da Literatura

Para realizar operações de modelagem espacial sobre geo-campos, Tomlin (1990) propõe a linguagem MAP (Map Analysis Package), que serve de base para muitas implementações comerciais. A implementação mais difundida das idéias de Tomlin é a linguagem GRID, disponível no sistema ARC/INFO. Esta linguagem permite realizar operações como:

```

OUTGRID = INGRID1 + INGRID2
OUTGRID = INGRID1 XOR 5
OUTGRID = SIN(INGRID1) * 4 / LOG(INGRID2)

```

A linguagem GRID, embora muito flexível, tem o sério inconveniente de não distinguir entre os diferentes tipos de operandos. Isto decorre do fato de estar ligada às estruturas de dados (matrizes) e não levar em conta a semântica das operações. Nos exemplos acima, se INGRID1 for um geo-campo numérico, e INGRID2 um geo-campo temático, o resultado pode não ter sentido. Em GRID, um geo-campo temático no formato matricial é chamado de "grade de inteiros", e um geo-campo numérico de "grade de ponto flutuante", o que estabelece uma confusão entre o dado geográfico e sua representação. Em função deste tipo de problemas, optamos por fazer de LEGAL uma linguagem onde os operadores estão ligados a um contexto semântico, definido pelas diferentes especializações de geo-campos.

4 Álgebra de Geo-Campos no SPRING-2.0

A álgebra de campos proposta em Câmara (1995) foi usada como base para a implementação parcial da linguagem LEGAL no SPRING-2.0, com os seguintes tipos de operações: *transformação*, *booleanas*, *matemáticas* e de *classificação contínua*. Um programa em LEGAL é estruturado em três partes: *declarações*, *instanciações* e *operações*. No que segue, as sentenças em legal serão escritas em fonte Courier New e as palavras reservadas da linguagem começam com Maiúscula.

Declaração

Toda variável em LEGAL deve ser declarada antes de ser utilizada, com a sintaxe:

```
Tematico <var> (<categoria>);
```

```
Numerico <var> (<categoria>);
```

```
Imagem <var> (<categoria>);
```

No SPRING, o termo “*categoria*” identifica uma especialização das classes geográficas básicas, no caso de geo-campos, as classes TEMÁTICO, NUMÉRICO e IMAGEM. Por exemplo, pode-se definir uma categoria “Altimetria” como especialização de GEO-CAMPO NUMÉRICO.

Instanciação

A instanciação é caracterizada pelos operadores *Recupere* e *Novo*. O operador *Recupere* associa uma variável a um geo-campo existente no banco de dados geográfico, a partir de seu nome. O operador *Novo* cria um novo geo-campo no banco de dados e requer parâmetros, dependendo da representação geométrica associada:

- imagem em tons de cinza: resolução horizontal e vertical.
- grade regular: resolução horizontal e vertical, valores máximos e mínimos aceitáveis.
- imagem temática (geo-campo temático em representação matricial): resolução horizontal e vertical e escala.
- representação vetorial: escala.

Operações de Transformação

Operações de transformação são usadas para realizar mapeamentos entre diferentes tipos de campos geográficos:

- PONDERA: transforma uma instancia do campo TEMÁTICO em uma NUMÉRICO;
- FATIA: transforma uma instancia do campo NUMERICO ou IMAGEM em um TEMATICO;
- RECLASSIFICA: transforma uma instancia do campo TEMATICO em outra com diferentes classes.

Como regra geral, os operadores de transformação exigem que o usuário defina um mapeamento entre os domínios de entrada e de saída sob a forma de uma *tabela*. A linguagem permite a definição de tabelas específicas para cada tipo de tais mapeamentos. As especializações de tabelas previstas pela linguagem são: PONDERAÇÃO, FATIAMENTO e RECLASSIFICACAO.

A figura 3 mostra um exemplo de “ponderação” (a conversão de um mapa de solos em um mapa de solos ponderado).

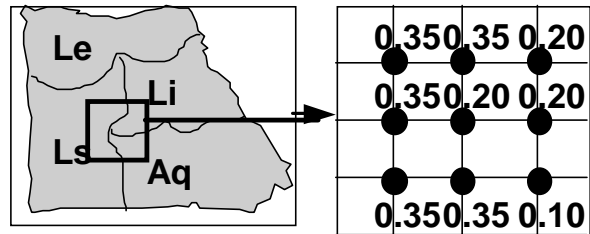


Figura 3 - Exemplo de operação de ponderação.

Operações booleanas

Nas operações booleanas, o valor do geo-campo resultante em cada posição é uma função apenas do valor da posição correspondente nos geo-campos de entrada. A saída de uma operação booleana é um GEO-CAMPO TEMÁTICO, cujos valores do contra-domínio necessitam ser especificados em função dos valores dos campos de entrada. Essas condições são calculadas através do operador *Atribua*. Um exemplo de uso desta operação é dado abaixo, onde um mapa de aptidão agrícola é calculado com base na declividade e no tipo de solo.

```
{
Tematico      mapa_solos ("Solos"),
              mapa_aptidao ("Aptidão");
Numerico      mapa_decl ("Declividade");
mapa_decl = Recupere(Nome="Decl92");
mapa_solos = Recupere(Nome="Solos92");
mapa_aptid = Novo (Nome = "Aptidao92",
                  Representacao = Matricial,
                  ResX = 250, ResY = 250,
                  Escala = 100000);
}
```

```

mapa_aptid = Atribua (CategoriaFim =
"AptidaoAgricola")
{
  "Bom" : mapa_solos.Tema = "Le" E
  mapa_decl >= 0.05;

  "Medio" : mapa_solos.Tema = "Aq" E
  mapa_decl >= 0.1;

  "Ruim" : Outros;
};
}

```

No exemplo acima, o termo "Outros" expressa todos os casos não indicados especificamente nas condições anteriores.

Operações matemáticas

As operações matemáticas sobre geo-campos que são especializações de NUMÉRICO e IMAGEM incluem:

- operações aritméticas: soma (+), subtração (-), multiplicação (*) e divisão (/);
- funções matemáticas: seno (sin), cosseno (cos), tangente (tan), arco tangente (atan), logaritmo (log), exponencial (exp), raiz quadrada (sqrt);
- relações: menor que (<), maior que (>), menor ou igual (<=), maior ou igual (>=), igual (==), diferente (!=).

Classificação Contínua

O uso de técnicas de classificação contínua busca utilizar as noções de conjuntos nebulosos ("fuzzy"), para substituir os processos tradicionais de geração de mapas.

Este aspecto é particularmente evidente no manuseio de mapas temáticos para análises de meio-ambiente. Uma fronteira arbitrária, definida precisamente por uma linha, entre dois tipos de solo, representa erradamente o que é, na realidade, uma variação contínua (Burrough, 1986). Quando realizamos operações de superposição entre mapas temáticos, o erro inerente à divisão arbitrária dos mapas em áreas estanques é propagado.

Em LEGAL, estão disponíveis operador de transformação de um geo-campo numérico num campo nebuloso ("fuzzy"), cujos valores variam entre [0,...1]. Este operador obedece às equações :

$$\mu_A(x) = 1 \quad \text{se } x \geq \beta,$$

$$\mu_A(x) = 1/[1 + \alpha(x - \beta)^2] \quad \text{se } x < \beta.$$

Nas equações acima α e β são parâmetros definidos pelo usuário, de modo a obter uma função nebulosa de forma adequada. A sintaxe da operação será:

```
<mnt>:= FUZZY(<mnt>, <alfa>, <beta>);
```

Para maiores detalhes sobre a operação, o leitor deverá consultar Druck e Braga (1995).

5 Exemplo de Aplicação

Os exemplos de uso apresentados a seguir supõem que o usuário tenha definido suas classes de dados como especializações das classes básicas TEMÁTICO, NUMÉRICO E IMAGEM, através da interface interativa do SPRING.

Apresentamos a seguir dois exemplos de aplicação:

- Avaliação do potencial natural de erosão em bacias hidrográficas, conforme o trabalho de Kuntschik et al. (1995).
- Geração de imagens índice de vegetação a partir de imagens TM.

Potencial Natural de Erosão

Utilizando mapas de solo, altimetria, e informações sobre as condições pluviométricas da área de estudo, a equação de perda de fertilidade do solo (Wischmeier e Smith, 1978) foi aplicada por Kuntschik et al. (1995) num estudo sobre a bacia hidrográfica do Ribeirão das Araras (município de Araras- SP), resultando na identificação do potencial de erosão de cada porção da bacia.

O potencial natural de erosão é dado pela seguinte equação:

$$PNE = R * K * LS$$

onde:

R = erosividade da chuva anual;

K = erodibilidade do solo;

LS = variável calculada a partir do comprimento da encosta L e declividade média S através da fórmula

$$LS = 0,00984 * L^{0,63} * S^{1,18}$$

A erosividade da chuva (R) é dada por:

$$R = 67.335 (p^2/P)^{0,85}$$

onde p é a precipitação média mensal e P é a precipitação média anual.

Neste exemplo para a bacia do Ribeirão das Araras na região de Araras o valor de R corresponde a 6675. A erodibilidade do solo (K), ou seja a resistência deste à ação da chuva, depende diretamente do tipo de solo em questão. Para cada tipo de solo há um valor associado de acordo com a tabela proposta por Lombardi Neto (1995).

A partir do mapa de altimetria, gerou-se um modelo numérico de terreno utilizando-se o interpolador TIN (grade triangular). Desta grade, gerou-se outra grade de declividade e um mapa temático com classes de declividade.

O valor do comprimento da encosta (L), ou percurso da água, foi obtido a partir de um mapa de distância entre o limite da bacia e os níveis mais baixos de altimetria, resultando em um modelo numérico do terreno.

O LEGAL foi utilizado para:

- converter o mapa de solos em uma grade de valores de erodibilidade, utilizando-se a função PONDERE;
- converter o mapa de classes de declividade em uma grade de valores médios de declividade, utilizando o valor central de cada intervalo, também através da função PONDERE;
- aplicar a equação universal de perda de solo considerando todos os parâmetros acima, gerando uma grade onde cada ponto da superfície está associado ao valor de potencial natural de erosão.

Apresenta-se a seguir a sintaxe utilizada:

```
{
//Declaração das variáveis
Tematico solo ("solo"),
    decl ("declive");
Numerico S ("decliv-media"),
    K ("erodibilidade"),
    L ("encosta"),
    LS ("LS"),
    pne ("PNE");
Float R = 6675.;
Tabela tabk (Ponderacao),
    tabs (Ponderacao);
// Transforma Declividade em
// (S )declividade media
decl= Recuperere(Nome = "Declividade");
S = Novo ( Nome = "DeclivMedia",
    Representacao = Matricial,
    ResX = 250, ResY = 250,
    Escala = 100000,
    Min = 0, Max = 50);
tabs = Novo(CategoriaIni="declive",
    "0-3" : 1.5,
    "3-6" : 4.5,
    "6-12" : 9.0,
    "12-20" : 16.0,
    "20-40" : 30.0,
    ">40" : 40.0 );
S = Pondere (decl,tabs);
//Transf. solo em erodibilidade (K)
solo= Recuperere(Nome = "TiposdeSolo");
K = Novo (Nome = "Erodibilidade(K)",
    Representacao = Matricial,
    ResX = 250, ResY = 250,
    Escala = 100000,
    Min = 0, Max = 1);
tabk = Novo (CategoriaIni = "solo",
    "pv2alva" : 0.0462,
    "pv3Olaria" : 0.0280,
    "pv3+pv4" : 0.028,
    "lrd+le1" : 0.0143,
    "lrdbaraogeraldo" : 0.0128,
    "lreribpreto" : 0.0098,
    "le1+lrd" : 0.01514,
    "LV4+LV3" : 0.0132
    "lvsmatodentro" : 0.0246,
    "te" : 0.0181,
    "serrinha" : 0.0462 ,
    "pv2usina" : 0.0462,
    "li3" : 0.0442,
    "pv5" : 0.0462,
```

```

    "li2" : 0.0362 ,
    "urbano" : 0,
    "li2+pv4" : 0.03292,
    "lrd+lre" : 0.0116 ,
    "le1" : 0.0167,
    "agua" : 0,
    "lv4speculas" : 0.0132,
    "LV3LaranjAzeda" : 0.0132);
K = Pondere (solo, tabk);
// Calculo de LS
L= Recuperere(Nome = "L-CompEncosta");
LS = Novo (Nome = "LS",
    Representacao = Grade,
    ResX = 250, ResY = 250,
    Escala = 100000,
    Min = 0, Max = 100 );
LS= 0.00984 * (L^0.63)* (S^1.18);
// Calculo de PNE
PNE = Novo (Nome = "PotNatErosao",
    Representacao = Grade,
    ResX = 250, ResY = 250,
    Escala = 100000,
    Min = 0, Max = 100);
PNE = R*K*LS;
}
```

Esta equação pode ser aplicada a outras áreas desde que os parâmetros de solo, altimetria e precipitação estejam disponíveis para a mesma.

Índice de Vegetação

Outro exemplo de aplicação refere-se ao cálculo de imagens de reflectância aparente de acordo com a equação geral proposta por Markham e Baker (1987) e a geração do Índice de Vegetação de Diferença Normalizada (IVDN) para estas novas imagens.

A partir dos valores de nível digital de imagem TM, são geradas imagens com valores de 0 a 255 e uma grade com os valores reais de reflectância aparente, a partir da fórmula

$$R = (\pi * d^2/E * \cos(t)) * [(ND/ND_{max}) * (L_{max} - L_{min}) + L_{min}]$$

onde R: reflectancia aparente

d: distancia Terra-Sol em unidades astronomicas

t: angulo solar zenital (graus)

E: irradiancia solar média no topo da atmosfera
(mW/cm².Ω.µm)

ND: nivel digital do pixel

ND_{max}: ND maximo possivel

L_{max} : radiância espectral mínima para gerar um valor digital máximo igual a 225 ($mW/cm^2 \cdot \Omega \cdot \mu m$)

L_{min} : radiância espectral correspondente ao número digital zero ($mW/cm^2 \cdot \Omega \cdot \mu m$).

A tabela abaixo indica os valores padronizados de radiância espectral mínima e máxima e de irradiância solar, para o sensor TM do satélite LANDSAT 5.

Banda	Lmin	Lmax	E
TM1	-0.15	15.21	195.7
TM2	-0.28	29.68	182.9
TM3	-0.12	20.43	155.7
TM4	-0.15	20.62	104.7
TM5	-0.037	2.719	21.93
TM6	0.1238	1.56	1.0
TM7	-0.015	1.438	7.452

A imagem utilizada corresponde à passagem 223/67, na data de 15/07/94. Para esta aquisição, temos:

$$\cos(t) = 0.633101$$

$$d = 1.01625$$

O programa abaixo realiza a operação desejada.

```
{
Imagem tm3 ("IMG"), tm4 ("IMG"),
      IV2 ("IMG");
Numerico re3("Reflec"),re4("Reflec"),
      IVDN ("ImgIVDN") ;
Float pi = 3.1415926,
      cost = 0.633101,
      d = 1.01625,
      lmin3 = -0.12,
      lmax3 = 20.43,
      lmin4 = -0.15,
      lmax4 = 20.62,
      e3 = 155.7,
      e4 = 104.7,
      ndmax = 255.;

tm3 = Recuperar (Nome="231_067_B3");
tm4 = Recuperar (Nome="231_067_B4");
re3 = Novo (Nome = "tm3_refl",
      Representacao = Grade,
      ResX = 30, ResY = 30,
      Escala = 100000,
      Min = 0, Max = 50);
```

```
re4 = Novo (Nome = "tm4_refl",
      Representacao = Grade,
      ResX = 30, ResY = 30,
      Escala = 100000,
      Min = 0, Max = 50);
IVDN = Novo (Nome = "IV_Refalg",
      Representacao = Matricial,
      ResX = 30, ResY = 30,
      Escala = 100000,
      Min = 0, Max = 50);
IV2 = Novo (Nome = "IVDN_RefAlg",
      Representacao = Matriz,
      ResX = 30, ResY = 30);
re3 = Numerico ((pi*(d^2)/e3*cost)*
      ((tm3/ndmax)* (lmax3-lmin3)+ lmin3));
re4 = Numerico ((pi*(d^2)/e4*cost)*
      ((tm4/ndmax)* (lmax4-lmin4)+ lmin4));
IVDN = (re4 -re3)/(re4 +re3);
IV2= Imagem ((re4 -re3)/(re4 +re3));
}
```

6 Conclusões e Trabalhos Futuros

Esta proposta representa a primeira versão de uma álgebra para campos geográficos, implementada no SPRING-2.0. A álgebra aqui introduzida é capaz de realizar vários tipos de análise espacial muitas vezes de grande complexidade.

Trabalhos futuros a serem desenvolvidos incluem a definição e implementação de operações de maior complexidade, incluindo: operações de vizinhança (como filtragem e declividade), operações zonais, mapas de distância e funções de processamento de imagens.

7 Agradecimentos

O SPRING é resultado do esforço de uma equipe, cujo arquiteto chefe é Ricardo Cartaxo Modesto de Souza, e cujo gerente do projeto é Juan Carlos Pinto de Garrido e inclui a colaboração dos seguintes profissionais:

No INPE: Ana Paula Dutra de Aguiar, Carlos Felgueiras, Cláudio Clemente Barbosa, Eduardo Camargo, Fernando Mitsuo Ii, Fernando Yutaka Yamaguchi, Flávia Nascimento, Gilberto Camara, Guaraci Erthal, Eugenio Sper de Almeida, João Argemiro de Carvalho Paiva, João Ricardo Freitas Oliveira, João Pedro Cerveira Cordeiro, José Cláudio Mura, Júlio Cesar Lima D'Alge, Laércio Namikawa, Lauro Hara, Leila Garcia, Leonardo Bins, Marina

Ribeiro, Marisa da Motta, Mary Minamoto, Maycira Costa, Missae Yamamoto, Regina Bruno, Silvia Shizue Leonardi, Silvana Amaral, Sergio Rossim, Ubirajara Moura Freitas e Virgínia Correa.

Na IBM Rio: Marco Casanova, Andrea Hemerly, Mauricio Mediano, Marcelo Salim, Claudia Tocantins, Paulo Souza.

Na EMBRAPA: Jaime Tsuruta, Ivan Lucena.

O SPRING conta com o apoio do CNPq, através dos programas RHAÉ e PROTEM/CC (projeto "GEOTEC").

9 Referências

Burrough, P.A (1987). *Principles of geographic information systems for land resources assessment*. Clarendon Press, Oxford.

Câmara, G. *Modelos, Linguagens e Arquiteturas para Bancos de Dados Geográficos*. Tese de Doutorado em Computação Aplicada. São José dos Campos, INPE, 1995.

DPI et al., "SPRING-2.0: Modelagem, Linguagem e Funcionalidade". Submetido ao VII Simpósio Brasileiro de Sensoriamento Remoto, Salvador, 1995.

Druck, S.; Braga, L.P.V. "Análise de Fertilidade de Solos por Geoestatística e Lógica Difusa". Relatório Interno, EMBRAPA, Centro Nacional de Pesquisa em Solos, 1995.

Egenhofer, M. Spatial SQL: A Query and Presentation Language. *IEEE Transactions on Knowledge and Data Engineering*, 6:86-95, 1994.

Goodchild, M. (1992) "Geographical data modeling", *Computers & Geosciences*, 18 (4): 401-408.

Kuntschik, G., Formaggio, A.R., Gleriani, J.M. "Cuantificación de pérdida de suelo por erosión hídrica en la cuenca del Ribeirão das Araras (SP-Brasil) en ambiente SIG, usando imágenes TM. Resultados preliminares". *Simpósio Latinoamericano de Percepción Remota, VII (Memorias- SELPER)*, Puerto Vallarta - México, 796-803, (1995).

Markham, B.L., Baker, J.L. "Landsat MSS and TM Post-Calibration Dynamix Ranges, Exoatmospheric Reflectances and At-Satellite Temperatures", *Landsat Technical Notes*" 1, 3-5, (1987).

Wischmeier, W.H., Smith, D.D. "Predicting rainfall erosion losses: A guide to conservation planning". US Department of Agriculture, *Agriculture handbook No 537*. (1978)

Tomlin, D. *Geographic information systems and Cartographic Modeling*. Prentice Hall, New York, 1990.

Worboys, M. *GIS: A Computing Perspective*. London, Taylor and Francis, 1995.