

ISSN 1678-0892

Dezembro, 2007

*Empresa Brasileira de Pesquisa Agropecuária
Centro Nacional de Pesquisa de Solos
Ministério da Agricultura, Pecuária e Abastecimento*

Boletim de Pesquisa e Desenvolvimento 119

Mapeamento Digital de Classes de Solos no Estado do Rio de Janeiro

*Maria de Lourdes Mendonça-Santos
Humberto Gonçalves dos Santos
Ricardo de Oliveira Dart
Jerônimo Guedes Pares*

Rio de Janeiro, RJ
2007

Exemplares desta publicação podem ser adquiridos na:

Embrapa Solos

Rua Jardim Botânico, 1.024 Jardim Botânico, Rio de Janeiro, RJ
Fone: (21) 2179-4500
Fax: (21) 2274-5291
Home page: www.cnps.embrapa.br
E-mail (sac): sac@cnps.embrapa.br

Comitê Local de Publicações

Presidente: Aluísio Granato de Andrade

Secretário-Executivo: Antônio Ramalho Filho

Membros: Marcelo Machado de Moraes, Jacqueline S. Rezende Mattos,
Marie Elisabeth C. Claessen, José Coelho de A. Filho, Paulo Emílio
F. da Motta, Vinícius de Melo Benites, Rachel Bardy Prado, Maria
de Lourdes Mendonça S. Brefin, Pedro Luiz de Freitas.

Supervisor editorial: *Jacqueline Silva Rezende Mattos*

Revisor de Português: *André Luiz da Silva Lopes*

Normalização bibliográfica: *Marcelo Machado Moraes*

Editoração eletrônica: *Pedro Coelho Mendes Jardim*

1ª edição

1ª impressão (2007): online

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

631.44
M539

Mendonça-Santos, Maria de Lourdes.

Mapeamento digital de classes de solos no Estado do Rio de Janeiro /
Maria de Lourdes Mendonça Santos ... [et al.]. ã Dados eletrônicos ã Rio
de Janeiro: Embrapa Solos, 2007.

(Boletim de pesquisa e desenvolvimento / Embrapa Solos, ISSN 1678-
0892 ; 119)

Sistema requerido: Adobe Acrobat Reader.

Modo de acesso: <[http://www.cnps.embrapa.br/solosbr/
conhecimentos.html](http://www.cnps.embrapa.br/solosbr/conhecimentos.html)>

Título da página da Web (acesso em 6 dez. 2007).

ISSN 1678-0892

1. Mapa. 2. Pedologia. 3. Rio de Janeiro (Estado). 4. Solo. I. Santos,
Humberto Gonçalves dos. II. Dart, Ricardo de Oliveira. III. Pares, Jerônimo
Guedes. IV. Embrapa. V. Título. V. Série.

© Embrapa 2007

Sumário

Resumo	5
Abstract	7
1. Introdução	9
2. Material e Métodos	11
2.1. Área de estudo	11
2.2. Base de dados auxiliares e de solos	17
2.3. Modelagem e predição de classes de solos	20
3. Resultados e Discussão	22
4. Conclusões	26
5. Agradecimentos	27
6. Referências Bibliográficas	27

Mapeamento Digital de Classes de Solos no Estado do Rio de Janeiro

Maria de Lourdes Mendonça-Santos¹

Humberto Gonçalves dos Santos¹

Ricardo de Oliveira Dart²

Jerônimo Guedes Pares²

Resumo

Com base no banco de dados de solos do Estado de Rio de Janeiro, provenientes de levantamentos feitos pela Embrapa, e atendendo demandas do projeto de pesquisa intitulado *Quantificação da magnitude, distribuição espacial de carbono orgânico dos solos do Estado de Rio de Janeiro utilizando-se tecnologias de modelagem quantitativa, SIGs (Sistemas de Informações Geográficas) e de banco de dados* (Projeto Carbono_RJ foram desenvolvidas atividades de busca, seleção, análise e revisão dos dados para cada perfil de solo descrito na área de estudo, o georeferenciamento de cada perfil (quando as coordenadas não estavam disponíveis) e a entrada de novos dados de perfis de solos descritos durante o projeto. Atualmente, essa base de dados contém 731 perfis de solos, 2.744 horizontes, e 48 atributos e serve a outros projetos, como por exemplo ao mapeamento digital de classes de solos. A base contém poucos dados para a densidade do solo e propriedades hídricas, entre outras. A partir desta base de dados, foram gerados, através de árvores de regressão, 5 modelos quantitativos solo-paisagem. Esses modelos foram comparados quanto à sua performance através de índices estatísticos

¹ Pesquisador da Embrapa Solos, Rua Jardim Botânico 1024, Jardim Botânico, Rio de Janeiro, RJ. CEP: 22460-000. E-mail: loumendonca@cnps.embrapa.br, humberto@cnps.embrapa.br.

² Bolsista da Embrapa Solos, Rua Jardim Botânico 1024, Jardim Botânico, Rio de Janeiro, RJ.

e os modelos foram mapeados numa malha de resolução espacial de 90m. Variáveis ambientais correlatas (imagens de satélite, mapa de litologia, mapa de solos, Modelo Digital de Elevação-MDE e suas derivadas...) foram usadas como preditoras para construir os modelos de acordo com o método *s.c.o.r.p.a.n.* (*soil, climate, organism, relief, parent material, age, space*) (MCBRATNEY et al. 2003). Neste trabalho, a base de dados de solos foi usada para prever classes de solos em nível de Ordem conforme o Sistema Brasileiro de Classificação de Solos - SiBCS (EMBRAPA SOLOS, 2006). O melhor modelo foi o que teve entre suas variáveis preditoras, a informação de um mapa de solos existente, com % de acerto de 97,6% para o conjunto de dados de treinamento e de 96,4% para o conjunto de teste.

Termos para indexação: Mapeamento Digital de Solos, Método de Regressão, Solos, Pedometria.

Digital Soil Classes Mapping in the State of Rio de Janeiro

Abstract

Based in a soil database for the State of Rio de Janeiro, systematizing the data originated from soil surveys made by Embrapa, to meet the demands of the project of research entitled "Quantification of the magnitude, space distribution of organic carbon of the soils of the State of Rio de Janeiro using technologies of quantitative modelling, SIGs (Geographical Information Systems) and of database (Project Carbono_RJ). The main developed activities went to the search, selection, analysis and revision of the data for each soil profile described in the study area, the georeferencing of each profile (when the coordinates were not available) and the input of new data of soil profiles described during the project. Now, the database contains 731 soil profiles, 2.744 horizons, and 48 attributes. Thus, the base contains few data for soil density and hydrological properties, among others. Starting from this database, were generated, through regression trees, 5 quantitative soil-landscape models. These models were compared as to its performance through statistical indexes and mapped in a mesh of 90m space resolution. Correlate environmental variables (satellite images, lithology and of soil map, Digital Elevation Model -DEM and their derivatives...) were used as predictors to build the models in agreement with the method *s.c.o.r.p.a.n.* - soil, climate, organism, relief, parent material, tempo, space - (MCBRATNEY et al. 2003). In this work, the soils database was used to predict classes of soils in Order

level, according to the Brazilian Soil Classification System - SiBCS (EMBRAPA SOLOS, 2006). The best model was the one that had among its predictor variables the information of an existing soil map, with percent of success of 95,4% for the group of training data and of 94,5% for the test group.

Index Terms - Digital Soil Mapping, Regression Tree, Soils, Pedometrics.

1. Introdução

O Mapeamento Digital de Solos foi definido por Lagacherie e McBratney (2006) como *a criação e a população de sistemas de informação espacial de solos por meio de modelos numéricos para inferir as variações espaciais e temporais de classes de solos e suas propriedades, a partir de observações, conhecimento e de dados de variáveis ambientais relacionados*. A principal aplicação desta abordagem é substituir o mapa de solos baseado em polígonos, pelo mapa digital de propriedades e de classes de solos e suas incertezas associadas, em áreas previamente mapeadas assim como em novas áreas. Estes mapas, armazenados e manuseados no formato digital, em ambientes SIG, criam a possibilidade de organizar um amplo conjunto de dados para análise e interpretação em qualquer época.

A predição de classes e propriedades de solos no mapeamento digital fundamenta-se nas relações existentes entre solos e os fatores e processos de formação dos solos como variáveis preditoras. A lógica deste raciocínio tem como base a equação de Jenny (1941) que foi formulada a partir do reconhecimento dos fatores responsáveis pela formação dos solos, em fórmula mais quantitativa,

$$S = f(cl, o, r, p, t)$$

Onde, *S* representa o solo, *cl* clima, *o* organismos inclusive atividades antrópicas, *r* relevo, *p* material originário e *t* tempo.

McBratney et al. (2003) generalizaram e formularam uma equação semelhante, com o objetivo de explicar as variáveis responsáveis pelos processos de formação dos solos, mas para uma descrição quantitativa empírica das relações entre outros fatores espaciais referenciados (ou co-variáveis ambientais), que são usados aqui como funções espaciais de predição. Sete fatores são considerados:

- s*: solo, outras propriedades do solo em um determinado ponto;
- c*: clima, propriedades climáticas do ambiente em determinado ponto;

o: organismos, vegetação ou fauna ou atividades antrópicas;

r: topografia, atributos da paisagem;

p: material originário, litologia;

a: idade, fator de tempo;

n: espaço, georeferenciamento.

O solo (**s**) pode ser considerado como um fator em si, porque pode ser predito a partir de suas propriedades, ou as propriedades podem ser preditas a partir de classes ou de outras propriedades. O modelo *s.c.o.r.p.a.n* pode ser representado como:

$$Sc = f(s.c.o.r.p.a.n) \text{ ou } Sa = f(s.c.o.r.p.a.n)$$

Onde **Sc** é classe de solos e **Sa** atributos de solo. O **s** refere-se a informação de solos proveniente de um mapa existente, de sensores remotos ou através do conhecimento de especialista. Aí implícito estão as coordenadas espaciais **x, y** (localização geográfica) e uma coordenada aproximada de tempo, **ft**. Esta coordenada de tempo pode ser expressa como *t* cerca de um tempo *t_i*. Cada fator será representado por um conjunto de uma ou mais variáveis contínuas ou categóricas, por exemplo, **r** por elevação, declividade ou outro atributo derivado do MDE. As fontes de dados, os métodos para estimar **f**, assim como os passos para executar o *s.c.o.r.p.a.n* são apresentados e amplamente discutidos na revisão feita por McBratney et al. (2003).

Esta base de dados, que representa uma das mais organizadas do Brasil está funcionando como um ensaio para ensino/aprendizagem de DSM, usando uma variedade de métodos baseado no modelo *s.c.o.r.p.a.n* (MCBRATNEY et al. 2003) para prever classes e propriedades de solos. Por questões de exatidão na localização geográfica dos perfis de solos inseridos nesta base de dados, somente 431 perfis de solos foram utilizados nesta aplicação para o ajuste dos modelos construídos para a predição e modelagem de classes de solos do Estado do Rio de Janeiro.

Neste trabalho, os procedimentos de mapeamento digital foram utilizados para prever classes de solos em nível de Ordem de acordo com Sistema Brasileiro de Classificação de Solos (EMBRAPA SOLOS, 2006), utilizando-se o método de árvore de regressão, utilizando o programa See5 (RULEQUEST RESEARCH, 2007), para estimar os modelos. Cinco diferentes modelos foram construídos e comparados quanto à sua performance através de indicadores estatísticos. O melhor modelo encontrado foi utilizado para o mapeamento digital das classes de solos.

2. Material e Métodos

2.1. Área de estudo

A área de estudo é o Estado do Rio de Janeiro, localizado entre 41° e 45° W e 20° 30' e 23° 30' S com cerca de 44.000 km² (Figura 1). Compreende 89 folhas topográficas do IBGE na escala 1: 50.000. A área é caracterizada por oito grandes tipos de paisagem (Figura 2): Faixa Litorânea, Norte-Noroeste Fluminense, Vale do Médio Rio Paraíba do Sul, Região Serrana, Planalto do Alto Itabapoana, Serra dos "rgãos, Serra da Bocaina e Serra da Mantiqueira (LUMBRERAS et al. 2003; RIO..., 2001), que são descritas como se segue:

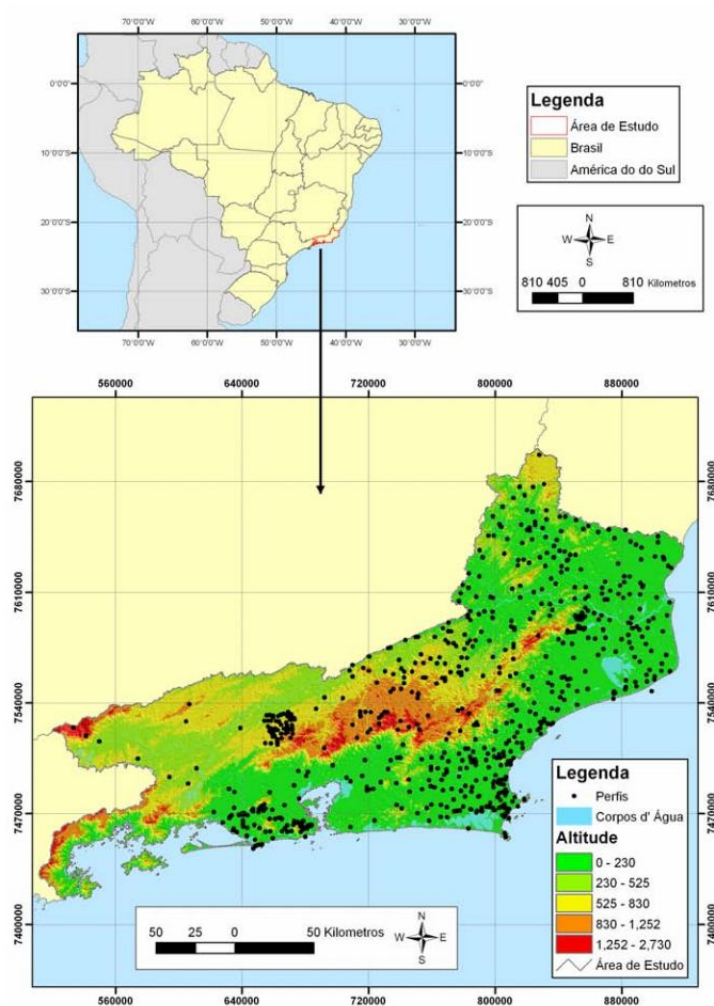


Fig. 1 – Área de estudo e a distribuição de perfis de solos no mapa de elevação, extraída do MDE SRTM (JARVIS et al. 2007) em pixels de resolução a 90 m.

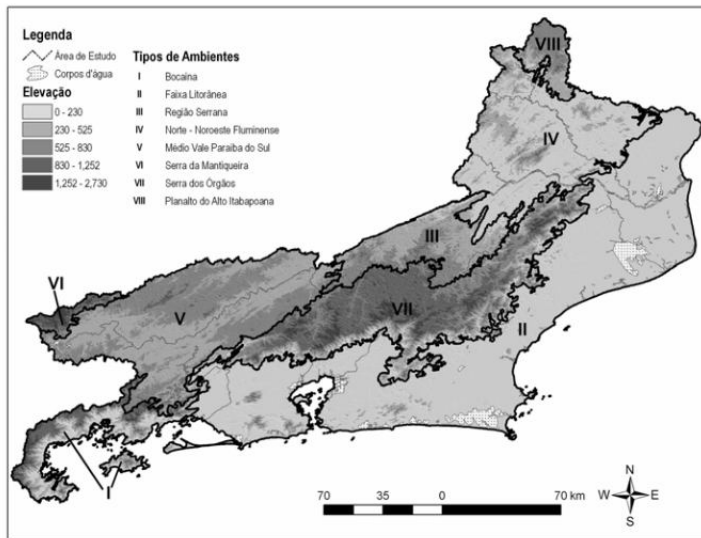


Fig. 2 - Tipos de paisagens do Estado do Rio de Janeiro.

Faixa Litorânea

Este É o domínio fisiográfico mais heterogêneo. Abrange extensas áreas inundáveis tais como baixadas, brejos, além de praias, dunas e até colinas isoladas e alinhamentos serranos com até 1.000 m de altitude.

Na Faixa Litorânea predomina a vegetação natural de floresta e restinga tropical subcaducifólia e próxima ao mar, relacionada ao tipo de clima Aw, tropical quente, com 3 a 6 meses de seca. ...uma área caracterizada pelas altas temperaturas, precipitação pluviométrica moderada e alta taxa de evaporação, onde são encontrados os Gleissolos Melânicos, Háplicos e Tiomúrficos, Planossolos Hidromúrficos, Espodossolos Cárpicos ou Ferrocárpicos. Nesta área, ocorre a mais intensa expansão urbana e industrial na região. Argissolos Vermelhos e Vermelho-Amarelos, Neossolos

Quartzarínicos, Organossolos e Cambissolos Hípicos, também ocorrem. Nas bordas da Faixa Litorânea, é observada uma superfície de relevo tabular, plano, onde solos profundos e bem drenados predominam, identificados como Latossolos e Argissolos Amarelos.

Norte – Noroeste Fluminense

Ambiente tipicamente de relevo ondulado e forte ondulado, interconectado com os alinhamentos serranos ao Norte e ao Sul. Ao Sul faz limite com a Região Serrana e a Serra dos Carregos e ao Norte limita-se com o Planalto do Alto Itabapoana, estendendo-se a Oeste até o Estado de Minas Gerais. Ao leste, limita-se com a Faixa Litorânea.

...uma extensa área de relevo colinoso, ondulado e forte ondulado, onde predominam os Argissolos de fertilidade média a alta (Distróficos e Eutróficos). Neste domínio também encontrados solos mais desenvolvidos e lixiviados (Latosolos e Argissolos Vermelho-Amarelos de baixa fertilidade) caracterizados por horizonte C mais espesso, estendendo-se a grandes profundidades. Nas várzeas fluviais estreitas e descontínuas e no fundo dos vales ocorrem Gleissolos Hípicos, Planossolos Flúvicos e, ocasionalmente, Neossolos Flúvicos, com a presença de sais e/ou sódio em níveis tóxicos em subsuperfície.

Médio Vale do Rio Paraíba do Sul

A região consiste de uma extensa depressão entre a Serra do Mar e as escarpas da Serra da Mantiqueira. Nesta depressão ocorrem os Latossolos Amarelos derivados de depósitos de sedimentos do Terciário caracterizando o relevo local plano, tabular. Entre as colinas de topo plano ocorrem as baixadas fluviais do rio Paraíba do Sul, onde Neossolos Flúvicos e Cambissolos Hípicos de alta fertilidade natural são encontrados. Os solos mais próximos da calha do rio Paraíba do Sul são geralmente Argissolos Vermelhos e Vermelho-Amarelos, usualmente são muito espessos e de boa fertilidade natural (alta saturação por bases). São muito dissecados e erodidos, de tal maneira que em certos locais as voçorocas constituem uma severa limitação ao uso da terra.

Predominam na área os Latossolos e os Argissolos Vermelho-Amarelos de baixa fertilidade. Cambissolos Hípicos e Argissolos Vermelho-Amarelos de alta susceptibilidade à erosão são comuns nesta área.

Região Serrana

A maior parte da área Montanhosa é incluída no relevo íngreme da Serra dos "rgãos, ao norte da Serra do Mar alinhada ao litoral de Estado de Rio de Janeiro. Na área Montanhosa os solos são muito espessos (Cambissolos) ou são muito lixiviados (Latosolos). Nos declives, sob condições de menos umidade e temperaturas mais baixas são encontrados os Argissolos.

Nesta área, o clima úmido e as temperaturas mais amenas de montanha, condicionam a preservação da matéria orgânica dando origem a solos com horizontes A espessos classificados como húmicos.

Na área Montanhosa prevalecem Cambissolos em escarpas e declives íngremes em altitudes mais altas, com solos rasos associados a afloramentos de rochas. Estes solos também ocorrem em áreas de relevos mais suaves constituindo terraços planos (altitudes ao redor 900 metros) onde depósitos de gnaisses pré-temperizados mais espessos dão origem a Latossolos mais profundos. O caráter húmico na superfície destes solos é relacionado ao clima moderado de altitude onde a decomposição da matéria orgânica é mais lenta formando horizontes A bastante espessos nas condições climáticas especificadas.

Serra dos Órgãos

A Serra do Mar atravessa o território inteiro de estado de Rio de Janeiro em uma direção WSW-ENE, acompanhando a direção estrutural do substrato geológico. Em sua porção sul, no limite com o estado de São Paulo, chega muito perto do mar, constituindo um ambiente peculiar, considerado um domínio ambiental diferente, conhecido como Serra da Bocaina - Sul do Rio de Janeiro descrito anteriormente. Nas montanhas e escarpas montanhosas são encontrados Cambissolos Hípicos e, devido aos gradientes de declive muito acentuados e ocorrência de grandes afloramentos de rochas, Neossolos

Litólicos também ocorrem e, em menor extensão, Latossolos Vermelho-Amarelos, em geral não muito espessos como também Argissolos Eutróficos Vermelhos e Vermelho-Amarelos.

Serra da Bocaina

Este domínio inclui um grupo montanhoso representado por parte da Serra do Mar, localmente denominado de Bocaina que se estende do Município de Itaguaí até o limite com o estado de São Paulo. Nesta parte chega muito perto do mar, delineando uma linha de costa recortada por paredes rochosas intercalada por áreas estreitas, ao longo de drenagens, de depósitos flúvio-marinhos. Nestes planos aluviais estreitos ocorrem Neossolos Flúvicos e Gleissolos Háplicos, enquanto que nos bancos de areia, Espodossolos Humilúvicos ou Ferrihumilúvicos são verificados. Solos rasos como Cambissolos Háplicos e Neossolos Litólicos e, em posições de relevo menos acidentado, Latossolos Vermelho-Amarelos. Nestas áreas também são regulares os Cambissolos Húmicos e em menor proporção, o Latossolo Vermelho-Amarelo menos profundo.

Planalto do Alto Itabapoana

Este planalto eleva-se a 700 m de altitude e tem um clima mais úmido e ameno que a extensa depressão adjacente e uma cobertura florestal um pouco mais preservada. O relevo colinoso predominante onde ocorrem Latossolos Vermelho-Amarelos é amplamente utilizado para pastagem e secundariamente, para plantações de café. Áreas de relevos forte ondulados também são comuns e os solos dominantes são Latossolos Vermelho-Amarelos e, em uma menor proporção, Argissolos Vermelho-Amarelos e Cambissolos e Afloramentos de Rocha.

Serra da Mantiqueira

Grupo de montanhas com características semelhantes a Serra Mar, separada daquela pela grande depressão do vale do médio rio Paraíba do Sul, entre os estados de Minas Gerais e São Paulo. Sua característica mais relevante é o relevo de rochas alcalinas de Itatiaia, onde é localizado o cume de elevação

mais alto do estado, a 2.787 metros acima do nível de mar. O clima local e a cobertura vegetal são tipicamente altimontanas, onde Neossolos Litólicos e Cambissolos Húmicos ocorrem com mais frequência. Em áreas circunvizinhas, em relevos relativamente mais baixos, e topografia irregular, ocorrem Cambissolos Húpicos, Afloramentos de Rocha e Latossolos Vermelho-Amarelos menos profundos. Somente em suas vertentes mais baixas ocorre Latossolos Vermelho-Amarelos e eventualmente, Argissolos Vermelho-Amarelos.

2.2. Base de Dados Auxiliares e de Solos

No âmbito deste trabalho um novo banco de dados de solos foi construído em Access e Delphi para facilitar a recuperação, a seleção, a análise e a revisão dos dados de cada perfil de solo descrito na área de estudo, como também o georeferenciamento de cada perfil (quando as coordenadas não estavam disponíveis) e a adição de dados novos e outras medidas descritas de acordo com as novas demandas de projeto.

A partir da base original (731 perfis de solos), somente 431 puderam ser usados nesta aplicação (Figura 1), visto que os outros apresentavam alguns problemas de posicionamento X,Y e, no momento, são objeto de verificação e correção. Como ilustrado na Figura 1, os perfis existentes são desigualmente distribuídos pela área de estudo. Atributos de solos (dados morfológicos, físicos e químicos), informações de vegetação e do terreno registradas por diferentes pesquisadores que executaram pesquisa de solos na área ao longo do tempo foram sintetizados e atualizados por Carvalho Filho et al. (2007). Foram descritos e coletados perfis adicionais durante a execução do Projeto Carbono_RJ (MENDONÇA-SANTOS et al. 2005) que foram somados ao banco de dados.

O Sistema Brasileiro de Classificação de Solos - SiBCS (EMBRAPA SOLOS, 2006) foi usado para classificar e atualizar os perfis de solos em nível de Ordem. Os 431 perfis de solos usados nesta aplicação pertencem a 9 das 13 Ordens de solos do SiBCS (Figura 3). Informação detalhada sobre coleta, procedimentos e métodos de análises físicas e químicas são fornecidos em Métodos de Análises de Solos (EMBRAPA, 1997).

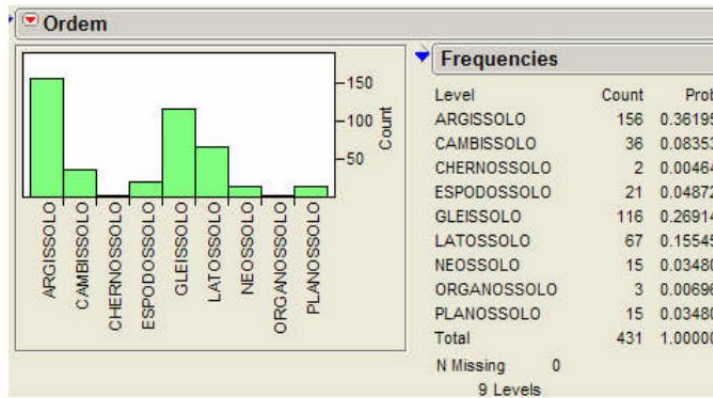


Fig. 3 - Ordens do SiBCS e distribuição de freqüência dos 431 perfis utilizados na aplicação.

Nesta aplicação as seguintes covariáveis foram usadas como variáveis de preditoras para construir os modelos: Mosaico de GeocoverTM (bandas 7, 4 e 2 em RGB), livremente distribuídas pela NASA (NASA, 2007); o index NDVI (usando faixa 2 em vez de 3); Mapa de Uso e Cobertura do Solos (LULC) do Estado de Rio de Janeiro, elaborado por Mendonça-Santos et al. (2003); Mapa de classes litológicas (RIO..., 2001) e SRTM do MDE 90m, obtida do banco de dados de cgiar em <http://srtm.csi.cgiar.org> (JARVIS et al. 2007).

O SRTM do MDE 90m não apresentou nenhum valor de dados nas zonas mais baixas como mostrado na Figura 4. Para preencher esses buracos, foram interpoladas as zonas de MDE que apresentaram este problema usando software Vesper (MINASNY et. al. 2007). O MDE novo de 90 m foi então usado como entrada no software LandMapR (McMillan, 2003), para obter as derivadas MDE a serem utilizadas em modelos preditivos.

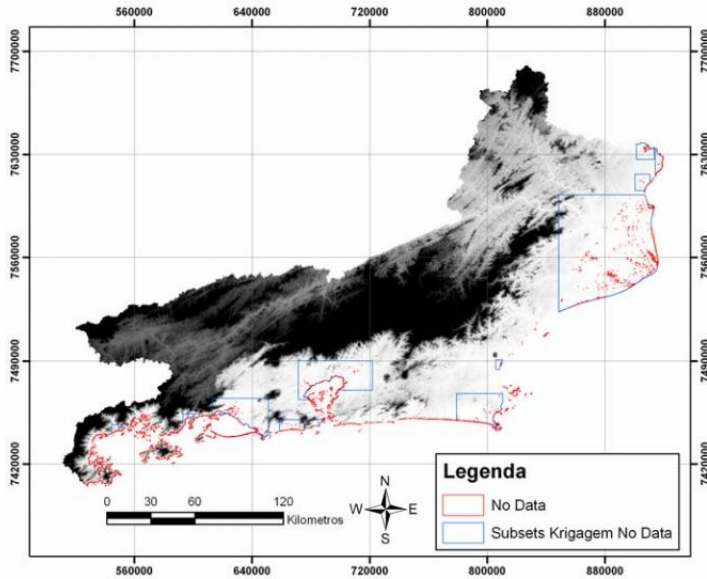


Fig. 4 ã Este mapa mostra a imagem SRTM MDE 90m, destacando em vermelho as áreas sem dados e com dados negativos e em azul as áreas onde foi realizada a krigagem destes valores.

A base de dados de solos foi complementada com as covariáveis de fatores ambientais para cada ponto de dados de solos extraído do programa Erdas Imagine (LEICA GEOSYSTEMS, 2003). Uma base de dados auxiliar que representa toda a área de estudo foi elaborada em um grid de 90m, extraído do SRTM MDE e foi povoado com variáveis ambientais e de solos. Uma análise estatística exploratória foi executada para dados de solos no programa JMP (SALL et al. 2007). A modelagem e predição de classes de solos foram elaboradas através de árvore de regressão, usando o programa See5 (RULEQUEST RESEARCH, 2007). Os dados de produção final foram importados e mapeados no programa de geoprocessamento ArcGIS (ESRI, 2004).

2.3. Modelagem e predição de classes de solos

Para predizer classes de solos por meio de árvore de regressão o programa See5.0 (RULEQUEST RESEARCH, 2007) foi utilizado. A árvore de regressão é um método de estratificação que utiliza os valores de uma ou mais variáveis, para realizar uma classificação de grupos homogêneos, a partir de um conjunto de variáveis pré-selecionados. Conforme cada grupo de dados é dividido a árvore aumenta. A base de dados foi dividida em dados de predição e dados de validação, na proporção de 80 e 20 respectivamente. O programa fornece a avaliação da taxa de erro para ambas bases.

Uma estrutura de árvore foi gerada dividindo repetitivamente os dados em vários grupos, cada divisão é escolhida de maneira a maximizar alguma medida de diferença na variável de resposta nos grupos resultantes. A Tabela 1 ilustra os dados principais e a informação utilizada para construir todos os modelos de predição (M1, M2, M3, M4 e M5) de classes de solos (a nível de Ordem), de acordo com fatores *s.c.o.r.p.a.n.* Todos os modelos de predição foram construídos alterando as variáveis preditoras. Foram obtidas as informações de quase todos os fatores de modelo *s.c.o.r.p.a.n.* como mostrado na Tabela 1, inclusive um polígono preexistente baseado em mapa de classes de solos (CARVALHO FILHO et al. 2003) e um mapa de classes de litologia (RIO..., 2001).

As derivadas de terreno utilizadas no modelo *s.c.o.r.p.a.n.* têm relação direta com os fatores de formação de solo que determinam a natureza e a distribuição das classes de solos na paisagem. Por exemplo, na Tabela 2 constam definições e aplicações das variáveis do terreno, conforme Wilson e Gallant (2000).

Tabela 1. Variáveis predictoras, modelo preditivo, variável preditiva e informação da base de dados de solos para cada modelo construído (M1, M2, M3, M4 e M5).

Modelos	Modelo Preditivo	Variáveis Predictoras para o modelo s.c.o.r.p.a.n	Variáveis Predictoras	Nº de Parâmetros	Perfis de Solo	Data set (3/4)	Test set (1/4)
M1	Árvore de Regressão	S (Mapa de solos - em formato vetorial), O (Landsat ETM - B7, B4, B2 e NDVI), R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE), P (Mapa de litologia - em formato vetorial)	Classe de solos no nível de Ordem de acordo com Embrapa Solos (2006)	30	431	324	107
M2		O (B7, B4, B3, NDVI e mapa de LULC), R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE), P (Mapa de litologia - em formato vetorial)		32	431	324	107
M3		O (Landsat ETM - B7, B4, B2 e NDVI), R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE), P (Mapa de litologia - em formato vetorial)		21	431	324	107
M4		O (Landsat ETM - B7, B4, B2 e NDVI), R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE)		10	431	324	107
M5		R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE)		6	431	324	107

Tabela 2. Derivadas de terreno usadas suas definições e aplicações.

Derivadas de Terreno	Definição	Aplicação
Elevação (ELEV)	Altura acima do nível do mar ou de uma referência local.	Determinação da energia potencial; define variações climáticas - pressão, temperatura, tendência da vegetação e do solo, teores de matéria orgânica, profundidade do solo entre outras.
Declividade (SLOPE)	Taxa de variação da elevação.	As fases do relevo, classificação da potencialidade da terra, tipos de vegetação, resistência no transporte de sedimentos, fluxos de superfície e subsuperfície.
Aspecto (ASPECT)	Orientação do relevo.	Irradiação solar, evapotranspiração, atributos da vegetação, faixas das encostas de acordo a orientação leste-oeste.
Índice de Umidade (QWETI)	Em [uma área específica de captação / tan (declividade)].	Índice de retenção de umidade.
Perfil curvature (PROF)	Taxa de variação da declividade	Aceleração dos fluxos, zonas de erosão acentuada / deposição, vegetação.
Plan curvature (PLAN)	Taxa de variação do aspecto	Convergência / divergência dos fluxos, propriedades de água no solo.

A performance dos modelos foi estimada estatisticamente utilizando-se o AIC (*Akaike's Information Criterion*), que é um índice que considera o número de parâmetros utilizados no modelo e representa um compromisso entre o ajuste e a parcimônia do modelo. O modelo que apresentar o menor AIC será o melhor. O AIC é calculado de acordo com (AKAIKE, 1973):

$$AIC = -2\log \text{like} + 2m$$

Onde *loglike* é o logaritmo da predição, e *m* é o número de parâmetros utilizados no modelo.

O *loglike* para a predição de classes ($k = 1, \dots, K$) é calculado de acordo de acordo com Hastie et al. (2001), como segue:

$$\log \text{like} = \sum_{k=1}^K \log \hat{p}_k$$

onde $\hat{p}_k$ é a proporção corretamente classificada como classe *k*:

$$\hat{p}_k = \frac{1}{N_k} \sum I(y_i = k).$$

O modelo que apresentar menor valor de AIC será utilizado para a predição final e mapeamento do estoque de carbono orgânico do solo, adicionando-se aos valores preditos, os resíduos krigados do ajuste do modelo (método de regressão-krigagem).

3. Resultados e Discussões

A principal diferença entre os modelos é o conjunto de variáveis preditoras usadas para construí-los. No modelo 5 (M5) somente dados do fator R do modelo s.c.o.r.p.a.n foram utilizados para construir o modelo (MDE derivadas: ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE). Em M4, imagens de satélite e índice NDVI (Landsat ETM + 7, bandas 7, 4, 2 e NDVI), foram usadas em adição ao MDE (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE), como fonte de dados em O e R fatores do modelo s.c.o.r.p.a.n. O modelo 3 (M3) foi construído utilizando-se a fonte de dados para O (Landsat ETM + 7, bandas 7, 4, 2 e NDVI); R (ELEV, ASPECT, PLAN, PROF, QWETI, SLOPE) e P

(Mapa de polígonos de classes de litologia), dos fatores do modelo s.c.o.r.p.a.n. Modelo 2 (M2) é similar ao M3, porém possui mais uma fonte de informação para fator O: o mapa de classes LULC. M1 difere de M3 em somente um importante aspecto: inclui informação de solos (um mapa de solos com base em polígonos) para prever classes de solos.

Resultados da predição de classes de solos (no 1º nível do SiBCS – Ordem) e as medidas da performance (AIC) de cada modelo são dados na Tabela 3.

Tabela 3 - Resultados da predição de classes de solos (em nível de Ordem do SiBCS), usando árvore de regressão.

Modelos	Classes de Solos Preditas	training data errors %	test data errors %	AIC
M1	(9 classes): ARGISSOLO, CAMBISSOLO, CHERNOSSOLO, ESPODOSSOLO, GLEISSOLO, LATOSSOLO, NEOSSOLO, ORGANOSSOLO, PLANOSSOLO	2.4	3.6	1616.306
M2	(7 classes): ARGISSOLO, CAMBISSOLO, ESPODOSSOLO, GLEISSOLO, LATOSSOLO, NEOSSOLO, PLANOSSOLO	17.1	45.9	2774.067
M3	(8 classes): ARGISSOLO, CAMBISSOLO, ESPODOSSOLO, GLEISSOLO, LATOSSOLO, ORGANOSSOLO, NEOSSOLO, PLANOSSOLO	11.7	44.1	2681.739
M4	(8 classes): ARGISSOLO, CAMBISSOLO, ESPODOSSOLO, GLEISSOLO, LATOSSOLO, ORGANOSSOLO, NEOSSOLO, PLANOSSOLO	12.3	45.0	2669.068
M5	(7 classes): ARGISSOLO, CAMBISSOLO, ESPODOSSOLO, GLEISSOLO, LATOSSOLO, NEOSSOLO, PLANOSSOLO	21.0	45.0	3249.703

De uma maneira geral todos os modelos fizeram boas predições de classes de solos em nível categórico alto (Ordem) do SiBCS, mostrando valores razoáveis de erros, significando que classes de solos na área de estudo poderiam ser preditas a partir de covariáveis ambientais, que podem ser obtidas facilmente e a baixo custo, como imagens Geocover™ e o SRTM do MDE 90m (Modelos M4 e M5).

Como pode ser observado, o modelo M5 (somente derivadas MDE) teve a pior performance (AIC mais alto) e foi capaz de prever somente 7 classes de um

total de 9 classes representada na base de dados (Fig. 3). Uma performance surpreendente foi mostrada por M4, nas bandas e NDVI do mosaico Geocover™, usadas em adiã, o ao DEM, dando a segunda menor AIC.

O melhor modelo foi o M1 que foi capaz de prever todas as 9 classes incluídas no modelo, com performance muito boa (o menor AIC) assim como sua distribuição espacial sobre toda a área de estudo (classes de solos aparecem onde se esperadas), em um mapa de solos bem generalizado (nível de Ordem do SiBCS). De fato, M1 foi um modelo excepcional devido ao uso da informação de solo já existente (mapa de polígonos de solos) para estimar classes de solos. O uso do fator S do modelo *s.c.o.r.p.a.n*, foi primeiramente proposto por McBratney et al. (2003), mas ainda não havia sido testado para prever classes de solos.

Na Figura 5 são apresentados os mapas resultantes para cada modelo. Mesmo que todos os modelos pudessem ser usados em diferentes situações, e considerando que todos os dados usados neste trabalho não estejam sempre disponíveis, o modelo (M2) mostrou um erro grosseiro ao prever a classe Gleissolo para quase toda a paisagem tipo VII (Serra dos "rgãos) com a maior altitude (Fig. 2). De fato, este tipo de solo poderia ocorrer neste tipo de paisagem somente em pequenas áreas restritas aos fundos de vales, próximas às linhas de drenagem, riachos, canais e rios. Outra falha de classificação foi o fato de que M2 e M3 e M5 não previram (ou não o fizeram bem) a classe dos Espodossolos, diferentemente de M1 (o melhor modelo) ou mesmo M4.

Consistência razoável é observada entre tipos de paisagens e classes em nível de Ordem em todo o mapa, principalmente como fornecido em M1 (Figura 6), conforme previamente descrito no item 2.1, de acordo com Rio (2001).

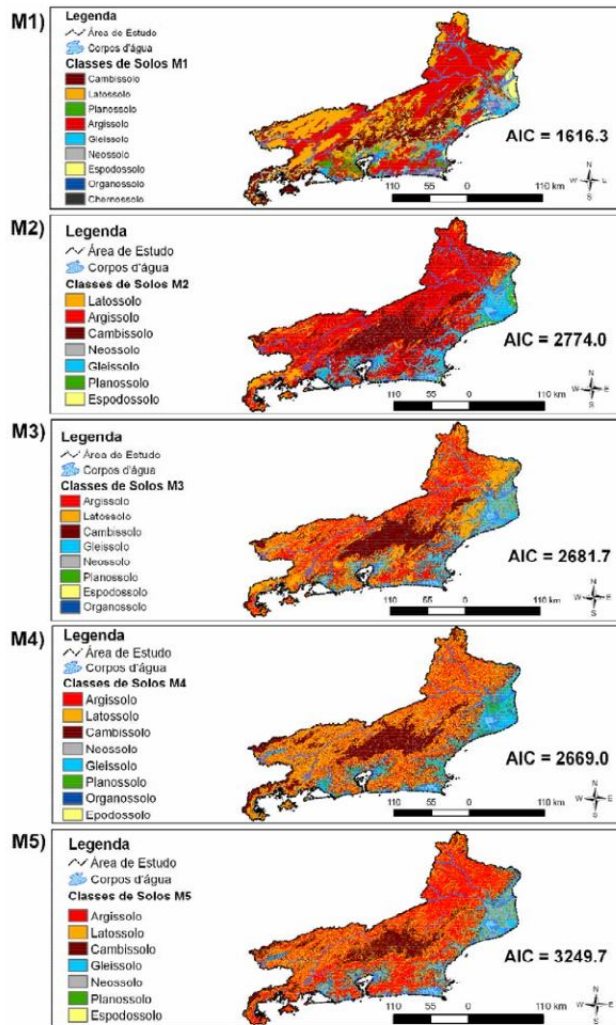


Fig. 5. Mapas resultantes dos cinco modelos de predição de classes de solos.

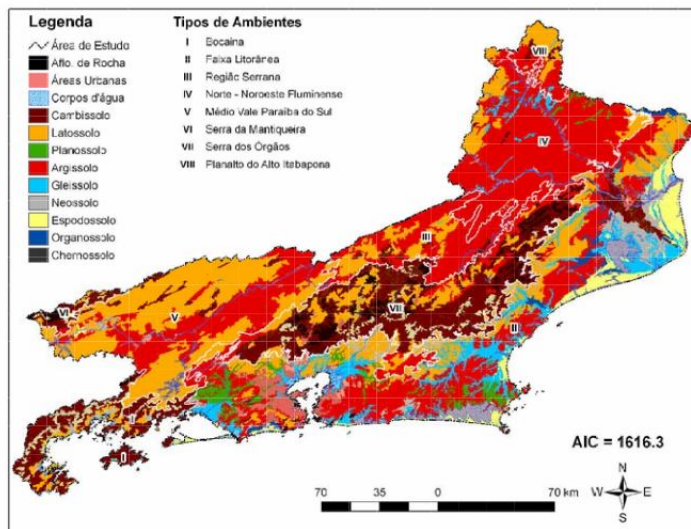


Fig. 6. Mapa Digital de classes de solos (em nível de Ordem do SiBCS) conforme preditas por M1 e tipos de paisagens.

4. Conclusões

Cinco modelos diferentes (M1 a M5) foram construídos e testados usando procedimentos DSM, a fim de prever classes de solos. Mesmo que os perfis de solo não tenham uma distribuição adequada por toda a área de estudo, foi possível prever classes de solos e suas distribuições espaciais por toda a área em alto nível categórico usando 431 perfis e as covariáveis ambientais. Este trabalho demonstra que a correlação solo-paisagem pode ser predita de maneira quantitativa e eficiente, usando a informação existente disponível. Esta abordagem é bastante promissora como procedimento para prever classes e propriedades de solos em áreas carentes de informação de solos, como é o caso do Brasil. Este tipo de informação pode ser diretamente utilizado, como suporte a outros estudos ambientais e de planejamento ou podem auxiliar decisões a respeito da necessidade amostragem adicional para estudos mais detalhados.

5. Agradecimentos

Este trabalho foi realizado no contexto do Projeto de Pesquisa E-26/171.360/2001, "*Modelagem da magnitude e distribuição espacial do carbono orgânico nos solos do Estado do Rio de Janeiro, usando técnicas quantitativas, SIG e Base de Dados*" e E-170.023/2001 "*Aplicação de técnicas quantitativas digitais para otimizar o mapeamento de solos para fins de planejamento e gestão ambiental*", ambos financiados pela FAPERJ ã Fundação Carlos Chagas Filho de Apoio a Pesquisa do Estado do Rio de Janeiro que queremos agradecer e reconhecer o apoio ã pesquisa em DSM desde o início do projeto na Embrapa Solos. Agradecimentos especiais a J.F. Lumbreras pelo mapa de polígonos de solos e apoio nas correções e nas descrições e localização de perfis de solos.

6. Referências Bibliográficas

- AKAIKE, H. **Information theory and an extension of maximum likelihood principle**. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., 1973, Budapest. Budapest: Akadémiai Kiadó, 1973. p. 267-281.
- PETROV, B. N.; CSAKI, F. (Ed.).
- CARVALHO FILHO, A. de; LUMBRERAS, J. F.; WITTERN, K. P.; LEMOS, A. L.; SANTOS, R. D. dos; CALDERANO FILHO, B.; CALDERANO, S. B.; OLIVEIRA, R. P.; AGLIO, M. L. D.; SOUZA, J. S. de; CHAFFIN, C. E. **Mapa de reconhecimento de baixa intensidade dos solos do Estado do Rio de Janeiro**. Rio de Janeiro: Embrapa Solos, 2003. 1 mapa, color. Escala 1:250.000. Disponível em: <<http://www.cnps.embrapa.br/solosbr/sigweb.html>>. Acesso em: 2 maio 2007.
- EMBRAPA. Centro Nacional de Pesquisa de Solos. **Manual de métodos de análises de solo**. 2.ed. Rio de Janeiro, 1997. 212 p.
- EMBRAPA SOLOS. Centro Nacional de Pesquisa de Solos. **Sistema brasileiro de classificação de solos**. 2. ed. Rio de Janeiro, 2006. 306 p.
- ESRI. **Geoprocessing in ArcGIS**. Redlands: Environmental Systems Research Institute, 2004. 368 p.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. New York: Springer, 2001. 552 p.

IPCC. **Mudança do clima 2007**: quarto relatório de avaliação do IPCC. Disponível em: <<http://www.mct.gov.br/index.php/content/view/3881.html>>. Acesso em: 2 maio 2007.

JARVIS, A.; REUTER, H. I.; NELSON, A.; GUEVARA, E. **Hole-filled SRTM for the globe**: version 3: 2006. Disponível em: <<http://srtm.csi.cgiar.org>>. Acesso em: 2 maio 2007.

JENNY, H. **Factors of soil formation**: a system of quantitative pedology. New York: McGraw-Hill, 1941. 281 p.

LAGACHERIE, P.; MCBRATNEY, A. B. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. In: LAGACHERIE, P.; MCBRATNEY, A. B.; VOLTZ, M. (Ed.). **Digital soil mapping**: an introductory perspective. Amsterdam: Elsevier, 2006. 600 p. (Elsevier. Developments in soil science, v. 31).

LEICA GEOSYSTEMS. **ERDAS Tour Guide, version 8.5**: GIS & Mapping: LLC. Atlanta, 2003. 638 p.

LUMBRERAS, J. F.; NAIME, U. J.; CARVALHO FILHO, A. de; WITTERN, K. P.; SHINZATO, E.; DANTAS, M. E.; PALMIERI, F.; FIDALGO, E. C. C.; CALDERANO, S. B.; MEDINA, A. I. M.; PIMENTEL, J.; CHAGAS, C. S. **Zoneamento agroecológico do Estado do Rio de Janeiro**. Rio de Janeiro: Embrapa Solos, 2003. 113 p. (Embrapa Solos. Boletim de pesquisa e desenvolvimento, n. 33).

MCBRATNEY, A. B.; MENDONÇA-SANTOS, M. L.; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, n. 1/2, p. 35-52, Nov. 2003.

MACMILLAN, R. A. **LandMapR® software toolkit - C++ version**: user manual. Edmonton: LandMapper Environmental Solutions, 2003. 110 p.

MENDONÇA A-SANTOS, M. L.; JACQUES, P. D.; COELHO, M. R.; PIMENTEL, J.; SANTOS, H. G.; ALMEIDA, P.; BARBOSA, E. L. A.; COSTA, T. C. C.; VILA, R. M.; SHINZATO, E.; BRANCO, P. C. M. P. A. **Mapeamento do uso atual e cobertura vegetal dos solos do Estado do Rio de Janeiro**. Rio de Janeiro: Embrapa Solos, 2003. 44 p.

MENDONÇA A-SANTOS, M. L.; SANTOS, H. G. dos; COELHO, M. R.; PARES, J. G. **Caracterização de paisagens e solos representativos do Estado do Rio de Janeiro para fins de estimativa de estoques de carbono no solo**. Rio de Janeiro: Embrapa Solos, 2005. 80 p. (Embrapa Solos. Boletim de pesquisa e desenvolvimento, n. 66).

MENASNY, B.; MCBRATNEY, A. B.; WHELAN, B. M. **VESPER 1.5**. Disponível em: <<http://www.usyd.edu.au/su/agric/acpa/pag.htm>>. Acesso em: 2 maio 2007.

NASA. [NASA home page]. Disponível em: <<https://zulu.ssc.nasa.gov/mrsid/>>. Acesso em: 2 maio 2007.

RIO de Janeiro: geologia, geomorfologia, geoquímica, geofísica, recursos minerais, economia mineral, hidrogeologia, estudos de chuvas intensas, aptidão agrícola, uso e cobertura do solo, inventário de escorregamentos, diagnóstico geoambiental. Rio de Janeiro: CPRM: Embrapa Solos: DRM-RJ, 2001. 1 CD-ROM.

RULEQUEST RESEARCH. **Information on See5/C5.0**: [version 1.20]. Disponível em: <<http://www.rulequest.com/see5-info.html>>. Acesso em: 2 maio 2007.

SALL, J.; LEHMAN, A.; CREIGHTON, LEE. **JMP start statistics, version 6**. Cary, North Carolina, 1989-2005. Disponível em: <<http://www.sas.com/>>. Acesso em: 2 maio 2007.

WILSON, J. P.; GALLANT, J. G. (Ed.). **Terrain analysis: principles and applications**. Hoboken: J. Wiley, 2000. 512 p.