

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO**

**MAPEAMENTO SUPERVISIONADO DE SOLOS ATRAVÉS
DO USO DE REGRESSÕES LOGÍSTICAS MÚLTIPLAS E SISTEMA DE
INFORMAÇÕES GEOGRÁFICAS**

**Samuel Ribeiro Figueiredo
(Dissertação de Mestrado)**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DO SOLO**

**MAPEAMENTO SUPERVISIONADO DE SOLOS ATRAVÉS
DO USO DE REGRESSÕES LOGÍSTICAS MÚLTIPLAS E SISTEMA DE
INFORMAÇÕES GEOGRÁFICAS**

**SAMUEL RIBEIRO FIGUEIREDO
Engenheiro Agrônomo (UDESC)**

**Dissertação apresentada como um dos requisitos à obtenção do
Grau de Mestre em Ciência do Solo**

**Porto Alegre (RS), Brasil
Abril, 2006**

*Dedicado à Janaína,
aos meus pais Luiz e Carmen e
à minha irmã Nanci.*

AGRADECIMENTOS

A Deus pelo dom da vida e pela presença e auxílio constante em todos os momentos.

Ao professor Élvio Giasson, pela paciência, amizade, ensinamentos e orientação desta dissertação de mestrado.

Aos meus pais, Luiz (*in memorian*) e Carmen, pelo amor, carinho, ensinamentos, sustento, apoio incondicional e por não terem medido esforços para verem meus sonhos realizados.

A Janaína pelo amor, amizade, companheirismo, compreensão e apoio incondicional em todos os momentos.

Aos meus irmãos Vânia, Tânia, Nanci (*in memorian*), Luis Carlos, Marcos e Mateus bem como aos demais familiares pelo apoio, amor, amizade e auxílios.

Ao meu “irmão gêmeo” Luis (Lula).

Aos amigos Jeane, Fabíola, Tatiana, Debiasi, Osmar, José Baptista, Verônica (*in memorian*), Telmo e Edir pela amizade e companheirismo durante estes dois anos.

Ao colega e amigo Carlos Gustavo Tornquist pelo auxílio neste trabalho.

Aos demais professores do PPG em Ciência do Solo pelo auxílio, ensino e amizade.

Aos funcionários do PPG em Ciência do Solo, em especial a Jader Amaro pela pronta disponibilidade em auxiliar.

Aos professores participantes da banca de defesa, professor Egon Klamt, professor Ricardo Simão Diniz Dalmolin, professor Gustavo Merten e professor Paulo César Nascimento pelas valiosas contribuições e pela pronta disponibilidade em contribuir com esta dissertação.

Ao povo brasileiro que através de seu pagamento de impostos proporcionou ao CNPq condições de me conceder uma bolsa de estudos, bem como patrocina a excelência e a qualidade do ensino ministrado no PPG em Ciência do Solo da UFRGS.

MAPEAMENTO SUPERVISIONADO DE SOLOS ATRAVÉS DO USO DE REGRESSÕES LOGÍSTICAS MÚLTIPLAS E SISTEMA DE INFORMAÇÕES GEOGRÁFICAS^{1/}

Autor: Samuel Ribeiro Figueiredo

Orientador: Professor Elvio Giasson

RESUMO

Regressões logísticas múltiplas (RLM), que estabelecem relações matemáticas entre variáveis resposta discretas e variáveis explanatórias contínuas, foram avaliadas quanto ao seu potencial em prever a ocorrência e distribuição de classes de solos nas regiões da Bacia Hidrográfica de Arroio Portão, de Ibirubá e Quinze de Novembro e de Viamão. A partir de modelo numérico de terreno digital (MNT) com 90 m de resolução, foram calculadas variáveis de terreno topográficas (elevação, declividade e curvaturas) e hidrográficas (distância dos rios, índice de umidade topográfica, índice de poder de córrego, comprimento de fluxo de escoamento) de cada área de estudo. Foram então estabelecidas RLM entre as classes de solos das regiões com base em levantamento tradicional e as variáveis de terreno. As regressões serviram para calcular a probabilidade de ocorrência de cada classe de solo, sendo que o mapa final de solos estimado foi produzido com base na seleção da classe de solo com maior probabilidade de ocorrência. A acurácia geral (AG), que avaliou a capacidade dos mapas estimados reproduzirem os respectivos mapas originais, foi de 40,3% (Viamão), 58,3% (Ibirubá) e 58,7% (Arroio Portão), e o coeficiente Kappa de Cohen foi de 32,6% (Viamão), 36,6% (Ibirubá) e 41,7% (Arroio Portão). A simplificação da legenda dos mapas elevou a acurácia geral dos mapas para 61,4% (Iburubá) e 67,7% (Arroio Portão) e o coeficiente de Kappa para 38,3% (Ibirubá) e 50,3% (Arroio Portão) de Kappa. Os melhores resultados foram verificados para a bacia do Arroio Portão, em escala maior que os demais (1:50.000) e que possui unidades de mapeamento de solos bem diferenciadas e, ocupando extensas áreas do terreno. Conclui-se que as RLM têm potencial como método preditivo para predição de ocorrência de classes de solos, produzindo resultados de qualidade dependente da qualidade do MNT, da escala do mapa original e do grau de relação entre as classes de solos e as variáveis de terreno.

^{1/} Dissertação de Mestrado em Ciência do Solo. Programa de Pós-Graduação em Ciência do Solo, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul. Porto Alegre, RS. Brasil. (93p.) – Abril, 2006.

SUPERVISIONED SOIL MAPPING USING MULTIPLE LOGISTIC REGRESSION AND GEOGRAPHIC INFORMATION SYSTEM^{2/}

Author: Samuel Ribeiro Figueiredo

Advisor: Professor Elvio Giasson

ABSTRACT

Multiple logistic regressions (MLR), which establish mathematical relationships between discrete response variables and continuous explanatory variables, were appraised regarding their potential for predicting the occurrence and distribution of classes of soils in the areas of Arroio Port~ao Basin, of Ibirubá and Quinze de Novembro, and of Viamão. From a digital elevation model (DEM) with resolution of 90m, topographical (elevation, steepness, and curvatures) and hydrologic (distance to rivers, Topographic Wetness Index, Stream Power Index, and flow length) were calculated for each study area. MLR were established between classes of soils mapped by traditional soil survey techniques and the terrain derived variables. The regressions calculated the probability of occurrence of each soil class, and the final map predicted the occurrence of soils classes based on the selection of the soil class with larger probability of occurrence. General accuracy evaluated the capacity of the estimated maps to reproduce the respective original maps, and had values of 40,3% (Viamão), 58,3% (Ibirubá), and 58,7% (Arroio Portão), and Kappa Index was of 32,6% (Viamão), 36,6% (Ibirubá), and 41,7% (Arroio Portão). The simplification of the original map legends elevated the general accuracy to 61,4% (Iburubá) and 67,7% (Arroio Portão), and Kappa Index to 38,3% (Ibirubá) and 50,3% (Arroio Portão). The best results were verified for the Arroio Portão Basin, with larger scale (1:50.000) than the others areas and having soil mapping units more differentiated and occupying large area extensions. MLR have potential as method for prediction of occurrence of soil classes, reproducing results with quality differentiated in function of quality of the DEM, scale of the original soil map, and degree of dependence between occurrence of soil classes and terrain variables.

^{2/}Soil Science Master Dissertation. Soil Science Graduate Program, Agronomy, do Rio Grande do Sul Federal University. Porto Alegre, RS. Brasil. (93p.) – Abril, 2006.

SUMÁRIO

1. INTRODUÇÃO	1
2. REVISÃO BIBLIOGRÁFICA.....	4
2.1. Mapeamento tradicional de solos.....	4
2.1.1. Princípios gerais	4
2.1.2. Uso dos fatores de formação para o mapeamento tradicional de solos	5
2.1.3. Métodos empregados	5
2.1.4. Avaliação do processo e produto final	6
2.2. Mapeamento supervisionado de solos	7
2.2.1. Princípios gerais	7
2.2.2. Pedometria	7
2.2.3. Fontes de dados	9
2.2.4. Principais variáveis de terreno utilizadas	11
a) direção de fluxo de escoamento:	13
b) acumulação de fluxo:	14
c) distância dos rios:	14
d) comprimento de fluxo:.....	14
2.2.5. Mapeamento de classes de solos.....	14
2.2.6. Métodos matemáticos de predição de ocorrência e distribuição de classes de solos	16
a) Árvores de decisão:	16
b) Redes neurais:.....	17
c) Lógica difusa:	18
d) Regressões logísticas múltiplas:.....	18
2.2.7. Regressões logísticas múltiplas.....	19
2.2.8. Avaliação do erro	22
3. HIPÓTESE E OBJETIVOS	25
3.1. Hipóteses:	25
3.2. Objetivos	25
3.1.1. Objetivo geral:.....	25
3.1.2. Objetivos específicos:.....	26

4. MATERIAL E MÉTODOS.....	27
4.1. Áreas de estudo	28
4.1.1. Área de estudo 1: Município de Ibirubá e parte do município de Quinze de Novembro	30
a) situação, limites e extensão:.....	30
b) geologia:	30
c) geomorfologia:	30
d) clima:.....	31
e) vegetação:	31
f) solos:.....	31
4.1.2. Área de Estudo 2: Bacia Hidrográfica de Arroio Portão	34
a) Situação, limites e extensão:	34
b) geologia:	35
c) geomorfologia:	35
d) clima:.....	35
e) vegetação:	35
f) solos:.....	36
4.1.3. Área de estudo 3: Área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão, RS.....	38
a) Situação, limites e extensão:	38
b) geologia:	38
c) geomorfologia:	39
d) clima:.....	39
e) vegetação:	40
f) solos:.....	40
4.2. Procedimentos	43
4.3. Amostragem	48
4.4. Regressões Logísticas Múltiplas	49
4.5. Estatística da acurácia	51
5. RESULTADOS E DISCUSSÃO	52
5.1. Resultados e discussão por Caso de Estudo	52
5.1.1. Caso de Estudo 1: Região de Ibirubá e Quinze de Novembro	52
Modelo de probabilidade	54
Total.....	61
5.1.2. Caso de Estudo 2: Bacia Hidrográfica do Arroio Portão	61
5.1.3. Caso de Estudo 3: Área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão, RS	72
5.2. Resultados e discussão geral.....	79
5.2.1. Atributos do terreno empregados no estudo.....	79
5.2.2. Acurácia dos mapas	80
5.2.3. Vantagens da metodologia	81
5.2.5. Perspectivas da metodologia.....	83
6. CONCLUSÕES.....	84
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	85

RELAÇÃO DE TABELAS

1. Matriz de erros exemplificando a comparação entre a verdade de campo e o mapa estimado para cálculo de estatística de acurácia de mapas.	23
2. Unidades de mapeamento de solos do “Levantamento e Utilização Agrícola dos Solos do Município de Ibirubá” (adaptado de Santos et al., 1970)	34
3. Unidades de mapeamento de solos do levantamento da Bacia do Arroio Portão (Klamt et al., 1992)	38
4. Unidades de mapeamento do levantamento de solos da área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão,RS (Schneider et al., 1992).	42
5. Parâmetros e coeficientes de modelo (<i>função logit</i>) empregado na predição da ocorrência de classes de solos na região do Município de Ibirubá e parte do Município de Quinze de Novembro, RS.....	54
6. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na região de Ibirubá e Quinze de Novembro, RS ...	58
Tabela 7. Matriz de erro entre o mapa original com legenda simplificada e o respectivo mapa estimado por regressões logísticas múltiplas na região de Ibirubá e Quinze de Novembro, RS.....	61
Tabela 8. Parâmetros e coeficientes de modelo (<i>função logit</i>) empregado na predição da ocorrência de classes de solos na região da Bacia Hidrográfica de Arroio Portão, RS	64
Tabela 9. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS	68
Tabela 10. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na Bacia Hidrográfica do Arroio Portão, RS, utilizando legenda simplificada	72
Tabela 11. Parâmetros e coeficientes de modelo (<i>função logit</i>) empregado na predição da ocorrência de classes de solos na região de Viamão, RS ..	74
Tabela 12. Matriz de erro entre a legenda do mapa original e a legenda do mapa estimado por regressões logísticas múltiplas na região de Viamão, RS	78

RELAÇÃO DE FIGURAS

Figura 1. Esquema representando a interação entre SIG, estatística e a Ciência do Solo na construção da Pedometria (Hengl, 2003).....	8
Figura 2. Submatriz de elevação, empregada nos cálculos de variáveis topográficas do terreno (Zevenbergen & Tornen, 1987).	11
Figura 3. Esquema representando a utilização do mapeamento supervisionado de solos para a extrapolação de uma informação de classes de solos pré-existente com base no estabelecimento de relações matemáticas entre estas classes e variáveis topográficas e hidrográficas obtidas através do modelo numérico do terreno.....	15
Figura 5. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo município de Ibirubá e parte do município de Quinze de Novembro e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.	33
Figura 6. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo bacia hidrográfica do Arroio Portão e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.	37
Figura 7. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo Viamão e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.	41
Figura 8. Fluxograma dos procedimentos gerais desenvolvidos para o mapeamento supervisionado de solos.	45
Figura 9. Quadro esquemático representando o produto da extensão DEMAT a partir do MNT.	46
Figura 10. Quadro esquemático representando o produto da extensão Terrain Analysis: o IUT (<i>Topographic Wetness Index</i>) e o IPC (<i>Streampower Index</i>).	47
Figura 11. Quadro esquemático mostrando os procedimentos de obtenção de variáveis hidrológicas com base em modelo numérico de terreno através de extensões do SIG ArcView 3.2.....	48

Figura 12. Algoritmo do script empregado para executar as regressões logísticas múltiplas em ambiente ArcView 3.2.	50
Figura 13. Variáveis oriundas da modelagem topográfica e hidrográfica a partir MNT na área correspondente ao Município de Ibirubá e parte do Município de Quinze de Novembro, RS.	53
Figura 14. Probabilidades de ocorrência de unidades de mapeamento na região do Município de Ibirubá e parte do Município de Quinze de Novembro, RS. A escala apresenta-se em tons de cinza, variando entre probabilidade de 0% (branco) a probabilidade de 100% (preto).	55
Figura 15. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas na região do município de Ibirubá e parte do município de Quinze de Novembro.	57
Figura 16. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas com legenda simplificada na região do município de Ibirubá e parte do município de Quinze de Novembro, RS.	60
Figura 17. Variáveis oriundas da modelagem topográfica e hidrográfica a partir de MNT na Bacia Hidrográfica do Arroio Portão, RS.	63
Figura 18. Probabilidade de ocorrência das unidades de mapeamento de solos na região da Bacia Hidrográfica do Arroio Portão, RS. A escala apresenta-se em tons de cinza, variando de 0% (branco) a probabilidade de 100% (preto).	66
Figura 19. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS.	67
Figura 20. Mapa de solos, com legenda simplificada, original (a) e estimado (b) por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS.	71
Figura 21. Variação espacial dos atributos de terreno do Caso de Estudo 3, região de Viamão, RS.	73
Figura 22. Probabilidade de ocorrência das unidades de mapeamento de solos na região de Viamão, RS. A escala apresenta-se em tons de cinza, variando de 0% (branco) a probabilidade de 100% (preto).	76
Figura 23. Mapas de solos com legenda simplificada, original (a) e estimado (b) por regressões logísticas múltiplas na região de Viamão, RS.	77

1. INTRODUÇÃO

As informações advindas de um levantamento de solos são importantes para o planejamento e otimização da expansão urbana e rural e não menos importante com questões concernentes a gestão de recursos naturais e preservação ambiental. A aplicabilidade de um mapa de solos está diretamente relacionada com a escala com a qual este solo está mapeado, que é uma medida de seu grau de detalhamento (Dalmolin et al., 2004).

Atualmente, o Brasil tem todo o seu território coberto pelo levantamento de solos produzido pelo Radam Brasil (1986) em escala 1:1.000.000, sendo que no Rio Grande do Sul, também estão disponíveis dados do Levantamento de Reconhecimento dos Solos do Rio Grande do Sul de 1973 em escala de 1:750.000 (Brasil, 1973). Mapas de solo em escalas pequenas, tais como 1:750.000 ou 1:1.000.000, não são adequadas para o planejamento de uso e manejo de solos em municípios, bacias hidrográficas ou propriedades, sendo para esse fim necessários mapas em escalas mais detalhadas (Dalmolin et al., 2004).

No Brasil, mapas em escalas iguais ou maiores que 1:100.000 existem apenas em áreas restritas, como por exemplo, no Estado do Rio Grande do Sul temos os municípios de Ibirubá, Portão, Dois Irmãos, Sentinela do Sul e Alegrete, entre outros. Portanto, deve-se fazer um esforço no sentido de se produzir mapas em escala reduzida para todo o Brasil, em virtude da aplicabilidade destes mapas

em conformidade com a demanda crescente por informações que auxiliem no processo de tomada de decisão relacionadas com a expansão de municípios e com questões relacionadas com a preservação ambiental.

Entretanto, os custos envolvidos e o volume e tempo de trabalho associados à necessidade de recursos humanos com elevado potencial técnico, tornam a execução destes levantamentos mais detalhados pelos métodos tradicionais praticamente inexecutáveis para grandes extensões.

Dessa forma, métodos computacionais baseados na descrição de parâmetros derivados de modelo numérico de terreno e a associação destes com a ocorrência e distribuição de solos, como é o caso do mapeamento supervisionado de solos, revestem-se de grande importância para o país, pela relevância das informações resultantes da produção de mapas em maior escala associados à economicidade e rapidez de aquisição, processamento e fornecimento de resultados através de um sistema de informações geográficas (Hengl, 2003).

Vários métodos têm sido usados para estabelecer relações matemáticas e estatísticas entre solos e variáveis do terreno, entre eles temos a lógica difusa (Zhu et al., 1997a), as redes neurais (Zhu et al., 1997b), a análise discriminante (Hastie et al., 2001), as árvores de decisão (Moran & Bui, 2002), a geoestatística (MacBratney et al., 2003), a engenharia do conhecimento (Bui, 2004), a análise por componentes principais (McBratney et al., 2003) e, mais recentemente, as regressões logísticas (Bailey et al., 2003). Estes métodos têm sua eficiência relativa condicionada à qualidade dos dados obtidos, a acurácia do mapa original de solos empregado e à resolução do modelo numérico do terreno (MNT) (Zhu et al., 2001), tendo sido extensamente discutidos e expostos por McBratney et al. (2003) e Zhu et al. (1997a).

Recentemente, tem crescido o interesse no emprego de regressões logísticas múltiplas como método preditivo em estudos de mapeamento supervisionado de solos. Estas regressões estabelecem relações matemáticas entre variáveis independentes discretas ou contínuas com variáveis dependentes discretas. Campling et al. (2002), por exemplo, empregaram regressões logísticas

na predição da probabilidade da ocorrência espacial de classes de drenagem de solos usando atributos derivados do modelo numérico de terreno e índices de vegetação através de imagens LANDSAT-5 em uma região tropical úmida da Nigéria, com concordância entre as classes de drenagem originais e estimadas variando entre 95% e 65%, reduzindo na medida que se aumentava o número de classes de drenagem do solo.

No Brasil, também se tem tentado definir métodos alternativos em relação aos métodos convencionais que auxiliem no mapeamento de solos. Tais métodos se concentram no uso de sensoriamento remoto espectral (Lima et al., 2001; Demattê et al., 2004) e na correlação entre geoformas e pedoformas (Reis et al., 2004; Ippoliti R. et al., 2005). Entretanto estudos que avaliem o potencial de métodos matemáticos preditivos e mapeamento supervisionado são ainda incipientes.

2. REVISÃO BIBLIOGRÁFICA

2.1. Mapeamento tradicional de solos

2.1.1. Princípios gerais

Dada a grande diversidade de tipos de solos e da importância de diferentes características do solo para a exploração sustentável de áreas agrícolas ou preservação ambiental, os solos precisam ser mapeados, a fim de mapear-se a sua ocorrência e distribuição e usar-se essa informação para o planejamento de uso das terras. As informações sobre os tipos de solos e sua distribuição geográfica, provenientes de levantamentos de solos, são destinadas a atender a uma série de finalidades, tais como: (i) equacionamento de problemas agrícolas regionais ou locais; (ii) planejamento de propriedades rurais; (iii) planejamento de programas de reflorestamento, colonização e desenvolvimento, irrigação e drenagem, taxaço de impostos, avaliação ambiental e outros; (iv) elaboração de mapas interpretativos para objetivos específicos: aptidão para culturas, programas conservacionistas, descartes de resíduos, urbanização, localização de estradas, reservas ambientais e outros; e (v) uso em atividades de ensino, pesquisa e extensão agrícola.

Portanto é fundamental conhecer os solos, e sabendo que estes variam espacialmente em função da interação de diversos fatores ambientais, estes devem ser mapeados, ou seja espacializados dentro do terreno.

2.1.2. Uso dos fatores de formação para o mapeamento tradicional de solos

De acordo com Bui (2004), os mapas de solos são a representação gráfica do conhecimento dos pedólogos acerca das pedoformas no terreno, baseados na equação estocástica determinística de Jenny (1941), na qual representa-se que o solo formado e suas características são função da interação de cinco fatores de formação do solo: $S = f$ (material de origem, clima, relevo, organismos vivos, tempo), significando que mudanças em um ou mais dos fatores de formação do solo causam alteração no tipo de solo formado. Da mesma forma, permite inferir que a possibilidade de quantificação destas mudanças nos fatores de formação permitiria fazer-se inferências sobre a distribuição espacial dos solos.

2.1.3. Métodos empregados

O mapeamento de solos exige o conhecimento da gênese do solo e de sua classificação taxonômica, bem como de aerofotogrametria para delimitação dos limites de cada classe de solo. Um requisito essencial para o mapeamento é a disponibilidade de materiais cartográficos que sirvam como mapa base, por exemplo aerofotografias, cartas topográficas, mapas geológicos, mapas da rede viária, entre outros.

Na execução de levantamentos de solos, os limites entre as unidades de mapeamento de solos são localizados através de prospecções a campo tanto mais intensas quanto maior for a escala de mapeamento, sempre baseadas na interpretação de fotografias aéreas através de estereoscopia. A identificação dos solos a campo é correlacionada com padrões de cor, textura e localização nas fotografias aéreas e sua correlação com estes padrões e com a geologia, relevo,

drenagem e vegetação, permite a extrapolação destas informações para áreas com padrões similares.

2.1.4. Avaliação do processo e produto final

O mapeamento tradicional de solos tem recebido muitas críticas em virtude de ser um processo essencialmente qualitativo e de certa forma vedado aos não iniciados, além de ser de caráter subjetivo, sendo necessário a introdução de maior proporção de componentes quantitativos para que haja sua popularização (Bonish et al., 2004).

Além disso, a qualidade de um mapa de solos obtido tradicionalmente está correlacionada com sua escala. Todo o mapeamento de solos possui certo nível de erro inerente, por exemplo, mapas de reconhecimento em escala 1:50.000 podem ter até 25% de erro, ou seja, em 75% da área mapeada como sendo pertencente a uma determinada UM pode-se ser verificada esta UM, sendo os outros 25% compostos por inclusões e variações.

Como visto anteriormente, o mapeamento tradicional de solos é altamente dependente da experiência pedológica do mapeador, além disso a informação concernente à ocorrência e distribuição de solos tem grande demanda por recursos para a obtenção de informações, fotos aéreas, análises de laboratório, inspeções a campo, bem como o uso de pessoal especializado e treinado para executar estes tipos de levantamentos, tipificados na pessoa do pedólogo.

O grande volume de fontes de informação e mão-de-obra especializada torna o processo oneroso. Em áreas extensas, o processo tradicional de levantamento de solos é lento. Segundo Zhu et al. (1997a) os EUA possuem 0,9 bilhões de hectares e uma taxa de mapeamento de solos de 0,4 milhões de hectares por ano, o que demandaria 220 anos para que todo o seu território fosse coberto por levantamentos de solos em escala adequada. No Brasil, com 0,8 milhões de hectares e uma menor taxa de mapeamento anual de solos do que os EUA, seria necessário muito mais tempo para ter todo o seu

território mapeado em escala detalhada, o que vem em sentido contrário a vocação agrícola do país.

2.2. Mapeamento supervisionado de solos

2.2.1. Princípios gerais

O avanço da computação e da tecnologia da informação gerou uma grande quantidade de dados, informações e ferramentas computacionais em todos os campos do conhecimento (McBratney et al., 2003). Grandes avanços tem sido alcançados dentro da Ciência do Solo graças ao advento e disponibilização da constelação de satélites do Sistema de Posicionamento Global (GPS), desenvolvimento de Sistemas de Informações Geográficas (SIG) através do geoprocessamento e imageamento da superfície terrestre através do Sensoriamento Remoto.

Com o advento e popularização destas novas tecnologias, o campo da cartografia desenvolveu-se grandemente. Através do emprego das técnicas de geoprocessamento, SIGs complexos podem manusear com grande rapidez e facilidade um grande número de mapas através da execução de operações algébricas. Esta rapidez e facilidade de processamento de informações geográficas permitiu, então, o advento e crescimento da pedometria como ramo da Ciência do Solo (Cook et al., 1996).

2.2.2. Pedometria

A pedometria é um dos mais novos ramos da Ciência do Solo, e que estabelece relações matemáticas e (geo)estatísticas entre classes e propriedades do solo com variáveis quantitativas preditoras destas classes e propriedades. Este neologismo foi conceitualizado por McBratney (1996) e, segundo o próprio autor, é a aplicação de métodos matemáticos e estatísticos para o estudo da distribuição e gênese de solos.

A pedometria começou a popularizar-se e a desenvolver-se como ciência no início da década de 1990, sendo que em 1994 foi realizada sua primeira conferência cujos trabalhos foram publicados no volume 62 da revista *Geoderma* no mesmo ano. Por estes anos, surgiu um grupo de trabalho intitulado de *Pedometrics Society* vinculada a *Intenational Union of Soil Science* (IUSS). Após uma década de existência, o grupo de trabalho foi elevado à condição de Comissão da Sociedade Internacional, no XVII Congresso Mundial de Ciência do Solo. Apesar dos ramos da Pedometria estenderem-se em vários campos de atuação variando da simulação da gênese de solos até aplicações dentro da agricultura de precisão, em literatura temos visto predominantemente a bifurcação em dois campos de atuação, uma relacionada à predição de propriedades do solo e outra relacionada a predição de classes taxonômicas ou unidades de mapeamento de solos (Figura 1).

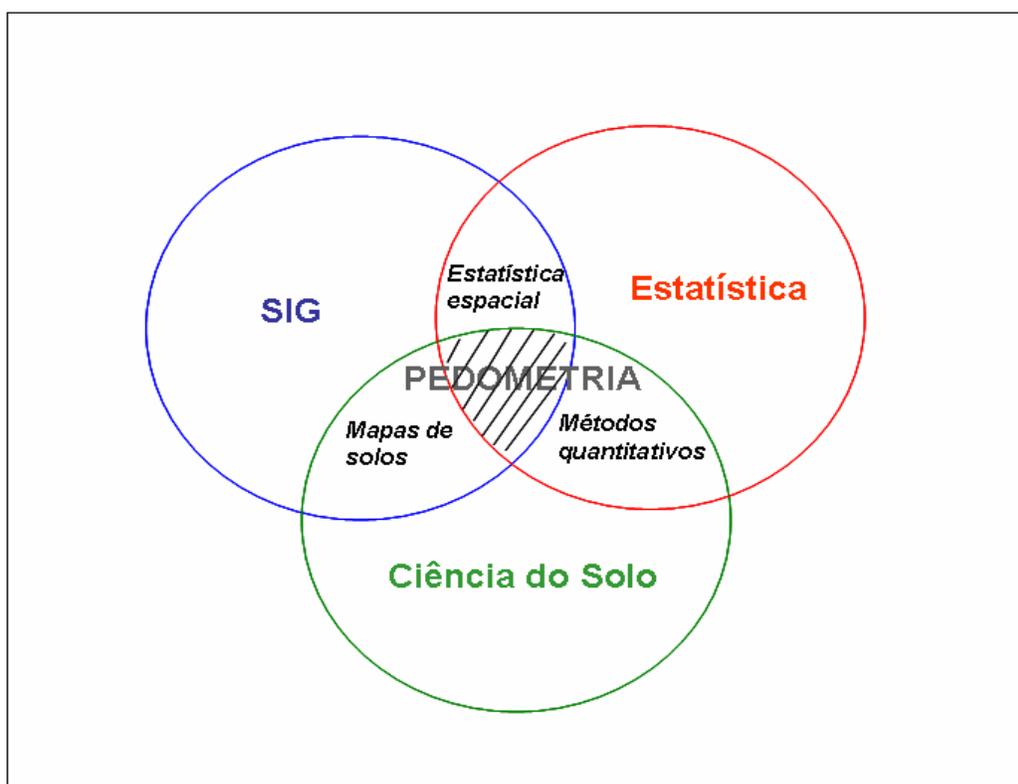


Figura 1. Esquema representando a interação entre SIG, estatística e a Ciência do Solo na construção da Pedometria (Hengl, 2003).

Em relação a atributos do solo, por exemplo, variáveis de terreno topográficas e hidrográficas, além de mapas geológicos e de vegetação, tem sido utilizados para predizer a concentração de carbono no solo (Somaratne et al., 2005), valores de pH (Hengl et al., 2002) e a profundidade do horizonte A.

Em relação a estimativas de carbono, Somaratne et al. (2005) estabeleceram relações matemáticas, através de redes neurais, entre variáveis climáticas e topográficas do terreno com o teor de carbono em amostras de solos provenientes de vários tipos de uso do solo no Sri Lanka. O melhor modelo de predição por redes neurais correlacionou-se com a verdade de campo em uma magnitude de R^2 de 0,92. Os autores concluíram ser as técnicas pedométricas importantes para se estimar o teor de carbono orgânico com reduzido custo, trabalho e tempo.

Hengl et al. (2002), utilizaram variáveis climáticas como precipitação anual e temperatura anual, índices de vegetação e sua variação sazonal, mapa geológico e atributos do terreno (elevação, declividade, índice de umidade topográfica e curvaturas), além do mapa de solos tradicional na região da Croácia. Os autores utilizaram estes dados para predição de vários atributos do solo, entre estas as concentrações de pH em água. Um modelo preditivo por regressões múltiplas estimou o teor pH em água do solo com uma concordância de 88% em relação à verdade de campo.

2.2.3. Fontes de dados

Segundo McBratney et al. (2003), os dados empregados para o mapeamento supervisionado de solos seguem, ao todo ou em partes, um modelo estocástico determinístico expandido a partir da equação de Jenny (1941) e representados pela expressão: $Sc = f(s, c, o, r, p, a, n)$. Onde Sc é a classe de solo pretendida que é uma função da interação entre s (mapa de solos pré-existente), c (clima), o (organismos vivos), r (relevo), p (material de origem), a (tempo) e n (componente espacial, localização geográfica).

Dessa forma, como fonte de dados são empregadas: a) mapas de solos pré-existentes da área, geralmente com cobertura de 10% da área total a ser mapeada (Hengl & Rossiter, 2003); b) mapas geológicos em escala igual ou maior do que o mapa de solos (Zhu et al., 1997a), quando disponíveis; c) modelos numéricos de terreno em escala maior ou igual ao do mapa de solos (Hastie et al., 2001) e d) bandas de reflectância espectral através de satélites imageadores (Bishop & McBratney, 2001).

O MNT, que está disponível praticamente em todo o mundo com resolução espacial de 90m x 90m, equivalente a uma área de pixel de 0,81ha através do Shuttle Radar Topographic Mission – SRTM, é o mapa de elevação da área do qual podem ser calculadas cerca de 72 variáveis de terreno topográficas e hidrográficas (Park et al., 2001) das quais as mais comumente empregadas em mapeamento supervisionado são a própria elevação, as curvaturas, a declividade, o aspecto, o índice de umidade topográfica, a direção de escoamento, o acúmulo de escoamento e o comprimento de fluxo do escoamento (Florinsky et al., 2002).

A resolução espacial do MNT SRTM na Europa e EUA é maior do que os disponíveis na América Latina, sendo de 30m x 30m (0,09ha), chegando a algumas áreas a ser de 10m x 10m (0,01ha). Provavelmente, nos próximos anos uma nova varredura dos satélites sensores do SRTM irão produzir mapas de elevação em escalas mais detalhadas para a América Latina. Além do SRTM, existem outras formas de obtenção do MNT tais como o uso de imagens ASTER, digitalização de cartas topográficas de levantamento planialtimétrico, percorrer a área de interesse com GPS específico para levantamento topográfico ou mapeamento digital da elevação através de estações totais geodésicas. No Brasil, apesar de ter uma resolução média, o uso de SRTM torna-se vantajoso em virtude da amplitude do levantamento, atingindo todo o território brasileiro, a facilidade de obtenção pois é de disponibilização gratuita e a operacionalidade da informação pois está em formato suportado pelos mais diversos SIGs, além de possuir diversos tutoriais facilitando o seu manuseio.

2.2.4. Principais variáveis de terreno utilizadas

Estas variáveis controlam importantes aspectos relacionados à pedogênese do solo e à distribuição de sua ocorrência na paisagem, tais como acúmulo de água, erosão, escarpamento de morros, ondulações do terreno, concentração de matéria orgânica, pH, profundidade do horizonte A, entre outros (Florinsky et al., 2002).

Para a modelagem geomorfológica, geralmente distinguem-se os atributos calculados entre atributos primários e secundários. Sendo os atributos primários computados diretamente a partir do MNT, já os atributos secundários envolvem a combinação de atributos primários e constituem geralmente índices empíricos adimensionais que podem caracterizar a variabilidade espacial de processos específicos ocorrendo na paisagem (Moore et.al., 1993).

As variáveis curvatura, perfil de curvatura, curvatura planar, declividade e aspecto são obtidas com base em modelo de submatriz de elevação conforme exposto em Zevenbergen & Toren (1987) (Figura 2).

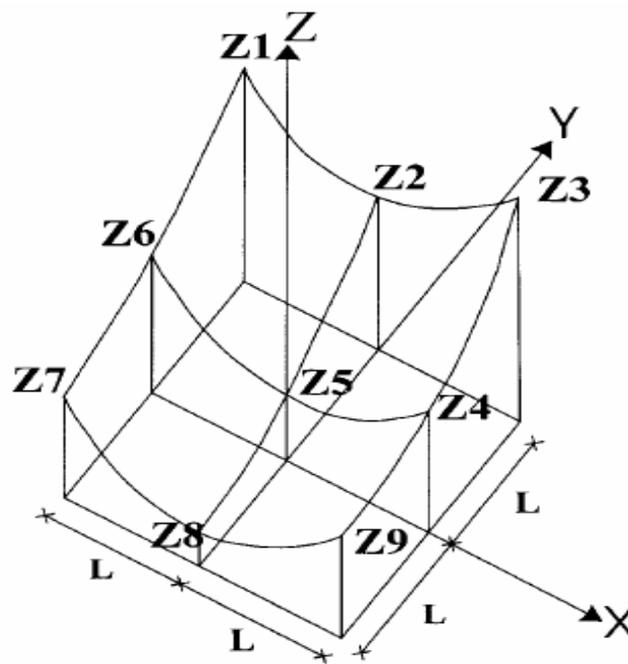


Figura 2. Submatriz de elevação, empregada nos cálculos de variáveis topográficas do terreno (Zevenbergen & Toren, 1987).

Nesta figura, representa-se uma pendente côncava com uma unidade básica formada pela junção de quatro *pixels*, tendo cada *pixel* a resolução mínima do MNT. Os componentes X e Y da figura representam as coordenadas retangulares de latitude e longitude, os componentes Z são os pontos de elevação da pendente. Os componentes L representam a distância entre duas cotas de elevação, na prática é a distância do lado do *pixel*.

Para o cálculo das variáveis de terreno topográficas, vários parâmetros intermediários (A a H) são calculados com base nos componentes da submatriz de elevação, a saber:

$$A = [\{(Z_1+Z_3+Z_7+Z_9)/4 - (Z_2+Z_4+Z_8+Z_{10})/2\} + Z_5]/L^4$$

$$B = [\{(Z_1+Z_3-Z_7-Z_9)/4 - (Z_2-Z_8)/2\} + Z_5]/L^3$$

$$C = [(-Z_1+Z_3-Z_7+Z_9)/4 - (Z_4-Z_6)/2]/L^3$$

$$D = [(-Z_4+Z_6)/2 - Z_5]/L^2$$

$$E = [(Z_2+Z_8)/2 - Z_5]/L^2$$

$$F = [(-Z_1+Z_3+Z_7-Z_9)]/4L^2$$

$$G = [-Z_4+Z_6]/2L$$

$$H = [Z_2+Z_8]/2L$$

$$H = Z_5$$

Os parâmetros primários do terreno são calculados com base nas variáveis acima, da seguinte forma:

$$d = \arctangente [(G^2 + H^2)^{0,5}]$$

$$cp = -2[(DG^2 + EH^2 - FGH)/(G^2 + H^2)]$$

$$pc = 2[(DH^2 + EG^2 - FGH)/(G^2 + H^2)]$$

$$a = 180 - \arctangente(H/G) + 90(G/G)$$

onde: d = declividade; cp = curvatura planar; pc = perfil de curvatura e a = aspecto.

Os parâmetros secundários do terreno índice de umidade topográfica (IUT) e índice de poder de escoamento (IPC) foram calculados da seguinte forma:

$IUM = \ln(A/tand)$; de acordo com Kirkby (1975) e O`Laughlin (1981)

$IPC = \ln(A*tand)$, de acordo com Moore (1991)

IUM e IPC são atributos que quantificam o papel desempenhado pela topografia em redistribuir a água da chuva sobre o relevo bem como a modificação da quantidade de radiação solar recebida na superfície. Estes atributos podem afetar as características do solo, por que a pedogênese é afetada pelo caminho percorrido pelo movimento da água em uma topossequência resultando em vários fatores tais como distribuição e abundância da água no solo, suscetibilidade á erosão hídrica e, não menos importante, a distribuição e abundância da flora e da fauna.

Um modelo hidrológico pode ser definido como uma representação matemática do fluxo de água e seus constituintes sobre alguma parte da superfície e/ou subsuperfície terrestre. A hidrografia de uma paisagem tem grande influência na pedogênese, em função da identificação de áreas de acúmulo ou de excessiva drenagem bem como áreas de suscetibilidade a erosão. As seguintes variáveis são comumente empregadas na modelagem hidrológica:

a) direção de fluxo de escoamento:

A direção de fluxo de uma célula é a direção na qual a água irá fluir para fora da célula, sendo que esta definição está atrelada a um modelo no qual a célula em questão está circundada por outras oito células vizinhas, sendo que a direção do fluxo vai depender basicamente do valor de elevação das células vizinhas.

A direção de fluxo do escoamento superficial é determinada pelo resultado da direção da água das cotas mais elevadas para as cotas mais baixas em cada célula. Esta é calculada através das mudanças nos valores de $z/distância * 100$; sendo z a elevação. A distância é determinada entre os centros das células.

Dessa forma, se o tamanho da célula é 1, a distância entre duas células ortogonais é 1 e a distância entre duas células diagonais é 1,414214 (Jenson & Domingue, 1988).

b) acumulação de fluxo:

A acumulação de fluxo é por definição a quantidade de água acumulada em uma superfície sobre a qual ocorreu um fluxo de água. Ao se requisitar a acumulação de fluxo, o SIG gera um mapa de acumulação de fluxo de cada célula, através do acúmulo do peso de contribuição das células em cotas acima, responsável pelo escoamento declive abaixo. Este método de derivação do acúmulo de fluxo através do MNT foi apresentado em Jenson & Domingue (1988).

c) distância dos rios:

A distância dos rios é a menor distância horizontal entre uma célula qualquer e uma célula de rio, sem considerar a tortuosidade do terreno.

d) comprimento de fluxo:

O comprimento de fluxo conceitualmente é o caminho que água percorre desde a sua célula de origem até um ponto de estagnação de água onde o fluxo se encerra, acompanhando a tortuosidade do terreno. A função comprimento de fluxo calcula distância ao longo de um caminho que percorre o fluxo para cada célula, com base no mapa de elevação. O mapa resultante define a impedância de movimento da água através de cada célula.

2.2.5. Mapeamento de classes de solos

O mapeamento supervisionado de solos é um processo que se utiliza de ferramentas computacionais para a predição de ocorrência de classes de solo na produção de mapas de solos a partir de relações quantitativas entre classes de solo com variáveis ambientais (McBratney et al., 2003).

Para exemplificar o fluxograma de procedimentos envolvidos no mapeamento de classes de solo por métodos pedométricos, pode ser verificado o quadro esquemático compreendido na Figura 3. Neste, temos tipificado o mapeamento supervisionado de solos, onde, a partir do MNT correspondente a área do mapa de solos são extraídos atributos topográficos e hidrográficos e posteriormente são estabelecidas relações matemáticas entre o mapa de solos e estes atributos. Com base nestas relações é estimado novo mapa de solos com a possibilidade de extrapolação da informação para as áreas circunvizinhas.

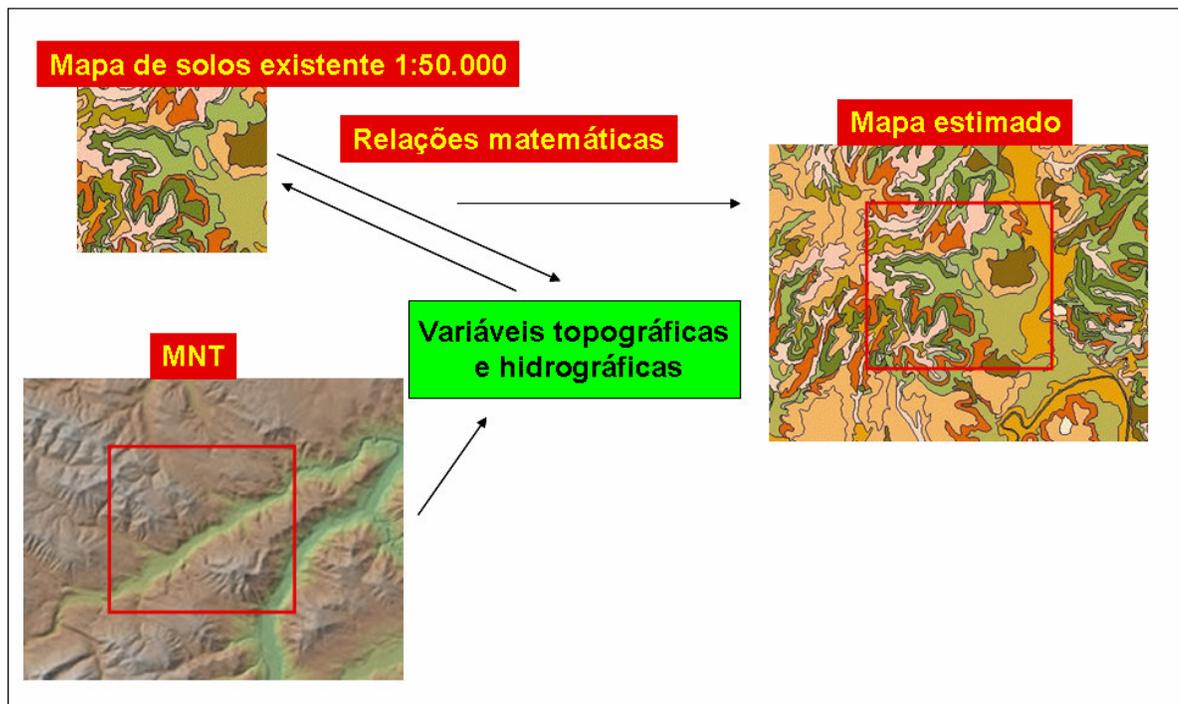


Figura 3. Esquema representando a utilização do mapeamento supervisionado de solos para a extrapolação de uma informação de classes de solos pré-existente com base no estabelecimento de relações matemáticas entre estas classes e variáveis topográficas e hidrográficas obtidas através do modelo numérico do terreno.

2.2.6. Métodos matemáticos de predição de ocorrência e distribuição de classes de solos

Segundo Rossiter (2004), em virtude da complexidade das relações matemáticas entre classes de solo e elementos do ambiente ao qual estas estão inseridas, as ferramentas matemáticas e estatísticas necessitam, portanto, apresentar o mesmo grau de complexidade.

Através do advento e crescimento da pedometria, popularizaram-se entre os pedométricos diversas ferramentas matemáticas complexas, destacando-se as análises multivariadas que consideram o peso concomitante de múltiplos fatores preditores em torno de uma única variável resposta. São exemplos destas análises multivariadas: árvores de decisão, redes neurais, lógica difusa, algoritmos genéticos e regressões logísticas (McBratney et al., 2003).

a) Árvores de decisão:

A árvore de decisão é um modelo preditivo que comporta-se como um mapeamento de observações acerca de um item; cada nó interno da árvore corresponde a uma variável, um arco em direção aos ramos representa um possível valor para a variável; as folhas representam o valor preditivo para a variável resposta dado pelos valores das variáveis preditoras ao longo do caminho partindo desde a raiz da árvore (Tan & Dowe, 2004). As vantagens dessa ferramenta é que árvores de decisão são geralmente simples de entender e fáceis de interpretar, processa grande número de dados em curto espaço de tempo, a preparação ou transformação de dados é básica ou desnecessária, relaciona variáveis categóricas e nominais e sua validação é obtida através de métodos estatísticos. (Comley & Dowe, 2005).

Lagasherie & Holmes (1997) discutiram a aplicação de árvores de decisão em mapeamento de solos bem como sua sensibilidade ao erro. Da mesma forma Bui & Moran (2003), utilizaram esta ferramenta para mapear classes de solos na Austrália. Mais recentemente Scull et al. (2005) aplicaram árvores de decisão na predição de classes de solos no deserto de Mojave na Califórnia, os autores utilizaram como fontes de dados mapas de solos pré-existent na escala

1:205.000 com 12 classes de solo, mapa geológico, variáveis climáticas, índices de vegetação e variáveis topográficas como índice de umidade topográfica, declividade, elevação e aspecto, obtendo uma acurácia geral e acurácia pelo coeficiente Kappa de Cohen de 65,2 % e 52%, respectivamente, para a região plana da área e de 88% e 44%, respectivamente, para a região montanhosa.

b) Redes neurais:

As redes neurais são conceitualmente um paradigma de processamento de informações que é inspirado na maneira como sistemas nervosos biológicos, como os do cérebro humano, processam informações (Gershenfeld, 1999). O elemento chave das redes neurais é sua estrutura do sistema de processamento de informações, que é composto de um grande número de elementos processantes altamente interconectados chamados de neurônios, os neurônios trabalham em conjunto para solução de um problema específico (Chang & Islam, 2000). Assim como ocorre com os seres humanos, as redes neurais “aprendem” através de exemplos. Cada rede neural é configurada para uma aplicação específica, tais como reconhecimento de padrões ou classificação de dados, através de um processo de aprendizagem. A aprendizagem, dentro de sistemas biológicos, envolve o ajustamento das conexões sinápticas existentes entre os neurônios (Minasny & McBRatney, 2002). Estas unidades de ligações extraem informações úteis das variáveis de entrada do sistema para predição da variável de saída, ou variável resposta, neste caso as classes de solos (Kohonen, 1982).

Zhu (2000), utilizou-se de redes neurais para prever a probabilidade de ocorrência de classes de solos através de variáveis ambientais nas quais estes estavam inseridos. Em uma área da floresta de Lubrecht, nas regiões montanhosas do estado de Montana, o autor utilizou como fonte de dados mapa de solos pré-existente, geologia, vegetação e variáveis do terreno (elevação, declividade, aspecto e perfil de curvatura). Como resultados, ao realizar amostragem de campo de 64 pontos na área, o autor verificou que destes 64 pontos, 49 pontos (77%) eram concordantes com mapa estimado por redes

neurais e apenas 39 pontos (61%) eram concordantes com mapa de solos obtido por métodos tradicionais.

c) Lógica difusa:

A lógica difusa se ocupa em representar a incerteza nos atributos ou classes predictoras e/ou preditas. Através da lógica difusa, variáveis de entrada são relacionadas a uma variável de saída esperada. Os passos do processo que envolvem uma análise por lógica difusa usualmente compreendem “difusãoamento” das variáveis predictoras de entrada, definição do operador difuso, aplicação de um método de implicação, agregação dos dados de saída e “desdifusãoamento” da variável resposta de saída (Dobermann & Oberthur, 1997).

Em trabalho de Zhu et al. (2001) a lógica difusa foi utilizada para inferir classes de solo a partir de variáveis ambientais, como material de origem, elevação, aspecto, declividade, perfil de curvatura, curvatura planar e índice de umidade topográfica, entre outros. Os autores usaram duas áreas de estudo, uma na floresta de Librecht na região montanhosa de Montana e a outra área na região sudeste de Wisconsin. Na primeira área o mapeamento estimado por lógica difusa concordou com 52 pontos (81%) dos 64 pontos de amostragem de campo da área, sendo que o mapa de solos tradicional concordou em 49 pontos. Já na segunda área, dos 37 pontos de amostragem, 89% da amostra concordou com o mapa estimado e 73% da amostra concordou com o mapa tradicional. Os autores concluíram que a lógica difusa tem a habilidade de identificar maior grau de detalhamento em um mapa de solos, entretanto, a concordância entre os dois mapas, o estimado e o observado, foi baixa.

d) Regressões logísticas múltiplas:

As regressões logísticas múltiplas são regressões nas quais a variável resposta é descontínua ou categórica, ou seja, qualitativa, como por exemplo classes de solo. As variáveis predictoras do modelo regressional podem ser tanto qualitativas quanto quantitativas, sendo a segunda forma a mais usual. O item

2.2.7 versará especificamente sobre este método, detalhando-o conforme necessário.

2.2.7. Regressões logísticas múltiplas

Pesquisadores estão frequentemente interessados em buscar ferramentas que possam estabelecer regressões nas quais a variável resposta é categórica ou descontínua.

As regressões logísticas podem estabelecer tais relações matemáticas entre uma, duas ou mais variáveis preditoras contínuas e até mesmo descontínuas com uma variável resposta descontínua apresentando várias classes, sendo o caso de várias classes uma adaptação ao modelo original conhecido por regressão logística binária.

Estas regressões se assemelham às regressões lineares comuns, entretanto diferem quanto ao tipo de variável resposta discreta e por se caracterizarem por um comportamento sigmóide de sua curva. As regressões logísticas apresentam duas funções principais: (i) descritiva, descrevendo a natureza do relacionamento entre a resposta e uma (ou mais) variáveis regressoras e (ii) preditiva, sendo capaz de identificar a reincidência de determinada resposta se as variáveis regressoras se repetirem (Hosmer & Lameshow, 1989).

As funções resposta são, portanto, denominadas de regressões logísticas múltiplas (RLM), cuja expressão para prever a ocorrência de uma das classes de uma variável resposta a partir de uma variável preditora, num exemplo com três classes da variável resposta, são:

$$\log \frac{p_{i1}}{p_{i3}} = \beta_1 x_i$$

$$\log \frac{p_{i1}}{p_{i2}} = \beta_2 x_i$$

$$\log \frac{p_{i2}}{p_{i3}} = \beta_3 x_i$$

Onde: p_{i1}/p_{i3} é a probabilidade de ocorrência da classe 1 sobre a probabilidade de ocorrência da classe 3, p_{i1}/p_{i2} é a probabilidade de ocorrência da classe 1 sobre a probabilidade de ocorrência da classe 2 e p_{i2}/p_{i3} é a probabilidade de ocorrência da classe 2 sobre a probabilidade de ocorrência da classe 3; β é a constante e x_i a variável do terreno para explicar a variável discreta.

Sendo a probabilidade de ocorrência de cada classe calculada conforme o modelo a seguir:

$$p_{i1} = \frac{e^{\beta_1 x_i}}{1 + e^{\beta_1 x_i} + e^{\beta_2 x_i} + \dots + e^{\beta_{(n-1)} x_i}}$$

$$p_{i2} = \frac{e^{\beta_2 x_i}}{1 + e^{\beta_1 x_i} + e^{\beta_2 x_i} + \dots + e^{\beta_{(n-1)} x_i}}$$

$$p_{i3} = \frac{1}{1 + e^{\beta_1 x_i} + e^{\beta_2 x_i} + \dots + e^{\beta_{(n-1)} x_i}}$$

Onde: $p_{i1} \dots p_{i3}$ é probabilidade de ocorrência da classe 1... classe 3, β é o peso multiplicador da variável e x_i é a variável preditora da classe.

As regressões logísticas têm a desvantagem de não possuírem um valor de r^2 como as regressões comuns, necessitando-se assim de estratégias mais complexas para excluir variáveis preditoras desnecessárias do modelo, bem como para verificar a adequação do modelo quanto à estimativa da resposta (Wong & Mason, 1985).

O método para seleção de variáveis preditoras foi o método iterativo “*stepwise*” ou passo a passo, interagindo-se com a resposta das análises de regressão, para adicionar ou remover variáveis explanatórias do modelo, assim como efeitos de curvatura e de interação. O método é similar ao modelo de regressão linear, onde se verifica se um subconjunto das variáveis X podem ser retiradas do modelo de regressão logística múltiplo, isto é, testar se os coeficientes de regressão b_k são iguais à zero, para este fim usa-se o *Teste da Razão de Verossimilhança*, que é baseado na estatística chamada de *Deviance do modelo*.

A “*deviance*” (desvio) de um modelo compara o logaritmo da verossimilhança deste modelo com o logaritmo da verossimilhança do modelo completo. Um modelo completo é um modelo que se ajusta completamente aos dados, isto é, para cada observação tem-se um parâmetro.

Para validação do modelo deve-se considerar que novos dados (uma nova amostra) ou, então, uma amostra reservada dos dados, deveria ser usada para verificar se o mesmo modelo pode ser usado com estes dados novos, se os coeficientes de regressão e os erros padrões são similares e se as mesmas conclusões ou inferências seriam obtidas.

Entretanto, ao se trabalhar com álgebras de mapas, utilizando-se RLM para isso, outra forma de se verificar a adequação seria através do uso de métodos de verificação de acurácias dos solos, como o método de acurácia geral, acurácia do mapeador, acurácia do usuário de mapas e o coeficiente Kappa de Cohen (Rossiter, 2004).

Dentro da Ciência do Solo, o emprego de regressões logísticas múltiplas é ainda incipiente e se concentra mais a regressões binárias e a predição de atributos de solos. Gessler et al. (1995) usaram regressões logísticas binárias para prever a probabilidade da presença ou ausência do horizonte E utilizando o índice de umidade topográfica e a curvatura planar como variáveis ambientais no sul da Austrália. Ainda em relação a regressões binárias, King et al. (1999) as utilizaram para prever a presença e ausência de argila não calcária em horizontes franco

argilosos na França central usando atributos de terreno calculados em MNT com 20 m de resolução.

Em relação a regressões logísticas múltiplas e classes de solo, Bailey et al. (2003) as utilizaram para estimativa de classes de solos na Inglaterra. Os autores usaram um mapa de solos pré-existente na escala 1:250.000, com base no MNT calcularam as variáveis elevação ponderada, declividade e aspecto, os autores verificaram correspondências entre os valores estimados e a verdade de campo variando de 47% a 65% em três formas de relevo: planalto, vale e escarpado. Os autores concluíram serem as regressões logísticas múltiplas um método válido para estimativa de classes de solo e ainda afirmam ser o método válido no sentido de aumentar o detalhamento de mapas de solos pré-existent.

2.2.8. Avaliação do erro

Medidas de acurácia de mapas destinam-se a verificar a magnitude de concordância de um mapa com a realidade a qual este pretende representar.

Existem diversos métodos para avaliação da acurácia de mapas, podendo ser classificados em métodos binomiais ou multinomiais (Rossiter, 2004). Conforme Stehman & Czaplewski (1998), são quatro as principais considerações em relação ao estudo da acurácia de um mapa estimado: (i) a freqüência de erro, ou seja, qual a proporção de área no mapa que não concorda com a realidade; (ii) a natureza do erro, ou seja, quais classes não são mapeadas corretamente e com quais classes são estas confundidas; (iii) qual a magnitude do erro, ou seja, o quão importante é o erro para o usuário e (iv) qual a fonte do erro, que é o porquê deste ocorrer.

Quatro métodos estatísticos multinomiais de se medir a acurácia de mapas podem contemplar as considerações acima e são: a acurácia geral (AG), acurácia do usuário (AU), acurácia do mapeador (AM) e a acurácia pelo coeficiente Kappa de Cohen (Kappa) (Rossiter, 2004).

Para verificarmos a aplicação prática destes métodos, tomemos o modelo representado na matriz da Tabela 1, que mostra a concordância entre um mapa e a realidade que o mesmo representa.

A partir de amostragens comparativas do mapa com a verdade de campo, constrói-se a matriz colocando-se nas linhas as i classes do mapa ($i = 1,2,3,4,\dots,x$) e na coluna as j classes de verdade de campo ($j = 1,2,3,\dots,y$) e na intersecção da linha i com a coluna j conta-se o total de pontos de amostragem obtidos para cada combinação de x e y .

Tabela 1. Matriz de erros exemplificando a comparação entre a verdade de campo e o mapa estimado para cálculo de estatística de acurácia de mapas.

		Verdade de campo			
		A	B	C	Total
Mapa Estimado	A	X_{11}	X_{12}	X_{13}	$\sum X_{i1}(i=1,2,3)$
	B	X_{21}	X_{22}	X_{23}	$\sum X_{i2}(i=1,2,3)$
	C	X_{31}	X_{32}	X_{33}	$\sum X_{i3}(i=1,2,3)$
	Total	$\sum X_{1j}(j=1,2,3)$	$\sum X_{2j}(j=1,2,3)$	$\sum X_{3j}(j=1,2,3)$	$\sum X_{ij}$

A acurácia geral (AG) é expressa pela somatória da diagonal dos valores concordantes entre o encontrado no campo e o estimado, ou seja, $AG = [(X_{11} + X_{22} + X_{33}) / \sum X_{ij}]$. Esta estatística, é uma medida geral de adequação do modelo preditivo.

A acurácia do usuário (AU) em cada classe é o valor de concordância entre o real e estimado, dividido pelo número total de observações na linha, ou seja, $AU_A = X_{11} / \sum X_{i1}(i=1,2,3)$. Esta estatística mede o quanto um usuário ao observar o mapa irá encontrar campo a classe mapeada..

A acurácia do mapeador (AM) de cada classe, é o valor de concordância entre o campo e o estimado, dividido pelo total de observações da coluna, por exemplo, $AM_A = X_{11} / \sum X_{1j}(j=1,2,3)$.

A acurácia pelo coeficiente Kappa de Cohen é uma estatística que leva em consideração as informações empregadas para se obter a AG, bem como a AU e a AM.

Inicialmente ocorre a somatória das concordâncias na diagonal, $\sum cd = (X_{11} + X_{22} + X_{33})$. Posteriormente, a somatória dos produtos entre o total das linhas e total das colunas de cada classe dividido pelo número total de observações, na forma de frequência de concordância, ou seja, $\sum fc = \sum [(\sum X_{1(j=1,2,3)} * \sum X_{1(i=1,2,3)}) / \sum + \dots + \sum X_{3(j=1,2,3)} * \sum X_{3(i=1,2,3)}]$.

O coeficiente Kappa de Cohen é calculado por:

$$\text{Kappa} = (\sum cd - \sum fc) / (\sum X_{ij} - \sum fc)$$

Em literatura tem-se observado que modelos com Kappa ao redor de 0,7 são considerados como satisfatórios.

3. HIPÓTESE E OBJETIVOS

3.1. Hipóteses:

Com base nas relações solo-paisagem, regressões logísticas múltiplas podem estabelecer relações matemáticas entre o tipo de solo e variáveis de terreno, sendo que estas relações podem ser usadas para elaboração de mapas de solos com elevado grau de semelhança a mapas produzidos por métodos tradicionais.

3.2. Objetivos

3.1.1. Objetivo geral:

Produzir mapas de solos através do estabelecimento de relações matemáticas entre variáveis dependentes discretas (tipo de solos) e variáveis independentes contínuas (atributos de terreno) utilizando-se de mapa de solos de referência e regressões logísticas múltiplas em sistemas de informações geográficas.

3.1.2. Objetivos específicos:

a) Estabelecer uma base de dados através da obtenção do modelo numérico do terreno (MNT), bem como mapas de solos na forma digital de regiões mapeadas em escalas adequadas ao planejamento municipal;

b) A partir do MNT, calcular variáveis topográficas (elevação, declividade, curvatura, perfil de curvatura, plano de curvatura) e hidrográficas (direção de escoamento, acúmulo de escoamento, índice de umidade topográfica e índice de poder de córrego);

c) Estabelecer relações matemáticas entre a distribuição e ocorrência das classes de solos com variáveis de terreno topográficas e hidrográficas através de regressões logísticas, selecionando as variáveis significativas ao modelo;

d) Produzir mapa preditivo da ocorrência de solos por estimativa através das regressões estabelecidas com base em mapa em escala original;

e) Avaliar a qualidade do mapa produzido por comparação com o mapa original e com a distribuição dos solos a campo.

4. MATERIAL E MÉTODOS

Considerando que as relações solo-paisagem são importantes para identificar a ocorrência e distribuição de unidades de mapeamento dentro da paisagem e que a paisagem pode ser matematicamente representada, os procedimentos deste estudo buscaram estabelecer relações matemáticas entre as formas do terreno (geoformas e hidroformas) e a ocorrência de unidades de mapeamento de solos (UM) na paisagem (pedoformas). Foram selecionadas três áreas representativas de diferentes regiões edafocológicas do Estado do Rio Grande do Sul. Na seleção das áreas de estudo, além da disponibilidade de mapas de solos procuraram-se áreas com feições distintas quanto ao relevo, geologia, clima e, portanto, com fatores e processos de formação de solos diferenciados.

Utilizaram-se como variáveis mapas de solos pré-existentes mapeados na forma tradicional e variáveis de terreno derivadas da modelagem topográfica e hidrográfica da área com base na informação de elevação constante no modelo numérico do terreno (MNT) de cada área. Entre a variável resposta solo e as variáveis preditoras topográficas e hidrográficas foram estabelecidas regressões logísticas múltiplas, estas regressões calcularam espacialmente a probabilidade de ocorrência de cada UM individualmente em relação às demais. Para cada

localização espacial foi atribuída a ocorrência da UM de solos que apresentou a maior probabilidade de ocorrência. A justaposição das predições de ocorrência de todas as UM estimadas deu origem ao mapa de solos estimado. A avaliação da acurácia dos mapas foi executada comparando as correspondências entre o mapa original e o mapa estimado através do uso de matrizes de erro.

4.1. Áreas de estudo

As três regiões estudadas constituem-se de áreas diferenciadas ecologicamente, sendo representativas do Planalto Médio, da Depressão Central e da Encosta da Serra do Nordeste e da Planície Litorânea do Estado do Rio Grande do Sul (Figura 4, mapa do estado com localização das áreas).

A região do município de Ibirubá e parte do município de Quinze de Novembro, representando a região edafoecológica do Planalto Médio, caracterizada por apresentar topossequência variando entre solos profundos e bem drenados como os Latossolos nas áreas planas em maior altitude, regiões escarpadas com associação Neossolos-Chernossolos e áreas planas com solos aluviais próximos aos rios.

A região da Bacia Hidrográfica de Arroio Portão, que se localiza em uma transição entra a Depressão Central e as Encostas da Serra do Nordeste, apresentando solos Gleissolos e Planossolos em áreas de baixadas e planas, Neossolos e Chernossolos em áreas com terreno escarpado e Argissolos em áreas com declive moderado.

A área localizada no município de Viamão está compreendida dentro da Planície Litorânea do Estado, tem o predomínio da ocorrência de solos que se desenvolvem em áreas de baixada e áreas planas influenciados pela grande quantidade de fração areia por deposições lacustres. Ocorrem nas áreas planas solos como os Neossolos Quartzarênicos, Organossolos, Gleissolos, nas áreas declivosas os Argissolos, além de tipos de terreno como as dunas.

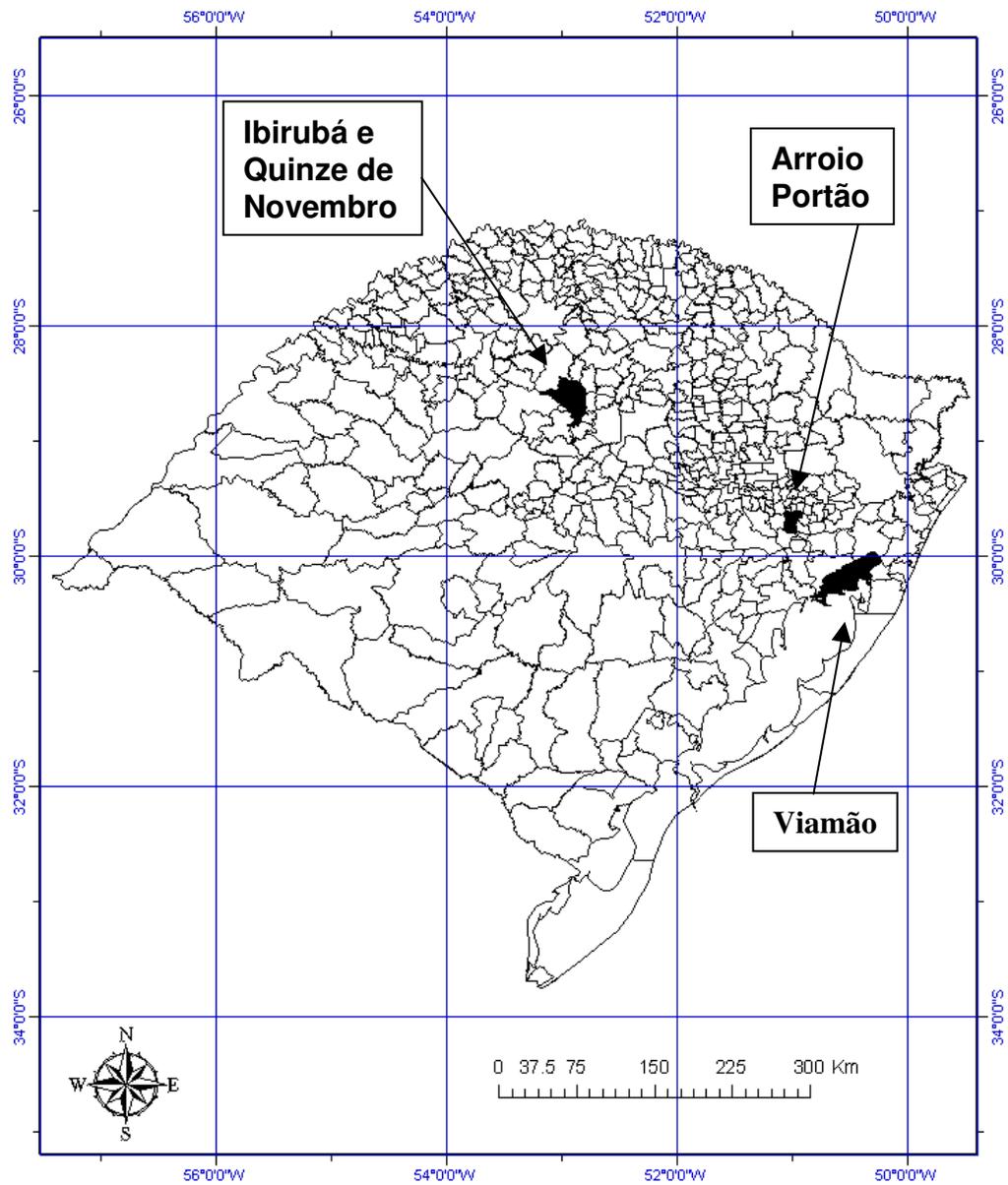


Figura 4. Localização geográfica das três Áreas de Estudo no Estado do Rio Grande do Sul.

4.1.1. Área de estudo 1: Município de Ibirubá e parte do município de Quinze de Novembro

a) situação, limites e extensão:

A área de estudo 1, atualmente compreende o município de Ibirubá e parte do município de Quinze de Novembro. Tendo uma extensão de aproximadamente 72.000 ha, está situada entre as latitudes 28° 28' e 28° 56' Sul e longitudes 53° 00' e 53° 34' Oeste, na região fisiográfica do Planalto Médio do RS (Figura 5). Limita-se ao Norte com Santa Bárbara do Sul e Colorado, ao Sul com Espumoso e Cruz Alta, e Este com Colorado, Selbrach e Espumoso e, a Oeste, com Santa Bárbara do Sul e Cruz Alta.

b) geologia:

Derrames basálticos da formação Serra Geral compõem a geologia ao Norte da área. Ao Sul e Oeste, encontram-se a formação Tupanciretã que é uma seqüência sedimentar constituída por conglomerados, arenitos e argilas, que cobrem os derrames basálticos da Serra Geral. Nas regiões de encontro entre as duas formações existe uma mistura de basalto com sedimentos arenosos.

c) geomorfologia:

A região apresenta uma geomorfologia caracterizada por coxilhas tabulares, que se desenvolveram sobre os derrames basálticos da Formação Serra Geral. Estas elevações caracterizaram-se por apresentarem topos achatados quanto ao perfil e arredondados em planta, relevo ondulado a suave ondulado e padrão de drenagem retangular. Entretanto, em áreas de maior altitude ocorre relevo forte ondulado.

d) clima:

O clima é do tipo subtropical, com temperatura média anual de 17,5°C A precipitação anual situa-se na faixa de 1750 mm e acompanha o regime de chuvas de inverno. Ocorrem geadas entre abril e novembro. Não existe períodos de deficiência hídrica ao longo do ano.

e) vegetação:

A vegetação é constituída por floresta latifolia tropical entremeada por pinheiros nas áreas de maior altitude, junto aos cursos de água encontram-se matas de galeria, constituída por vegetação densa de composição semelhante à mata de latifolia tropical. A vegetação de campo, localizada a Oeste do município, apresenta-se com espécies de compostas pequenas, verbenáceas, ciperáceas, leguminosas e gramíneas

f) solos:

As informações sobre os solos desta área de estudo, estão contidas no “Levantamento e Utilização Agrícola dos Solos do Município de Ibirubá” (Santos et al., 1970). Neste levantamento, foram definidas oito unidades de mapeamento para a região em escala 1:80.000. Na Tabela 2 e na Figura 6, é apresentada uma legenda adaptada, composta de seis das oito classes, em virtude da junção de unidades de mapeamento para a realização deste estudo, sendo a UM Batu incorporada à UM LVd2 e a UM Intergrade incorporada à UM LVd1. Esta incorporação fez-se necessário em virtude da baixa ocorrência da Batu e da dificuldade de dissociação das unidades que compõem a Intergrade, gerando equações de regressão com modelos de baixa acurácia, ao passo que adotando esta simplificação inicial na legenda destas unidades, passou-se a obter regressões com maior grau de acurácia.

Desenvolvendo-se a partir dos derrames de basalto encontramos predominantemente as unidades taxonômicas Latossolo Vermelho aluminoférrico

típico sobre os topos achatados de morros em relevo ondulado e forte ondulado e associação Chernossolo argilúvico-Neossolos litólicos nas zonas escarpadas, nestas áreas a vegetação natural é a de floresta latifolia entremeada por pinheiros. Ao Sul e Oeste da área, sobre a formação Tupanciretã origina-se a UM Latossolo Vermelho distrófico típico (LVd1), com relevo suave ondulado e vegetação semelhante às unidades anteriores. Nas regiões de encontro entre as duas formações geológicas, ocorrem a UM LVd2, com vegetação de campos. Nas regiões de baixada e próximas aos rios desenvolvem-se solos aluviais com horizontes gleizados.

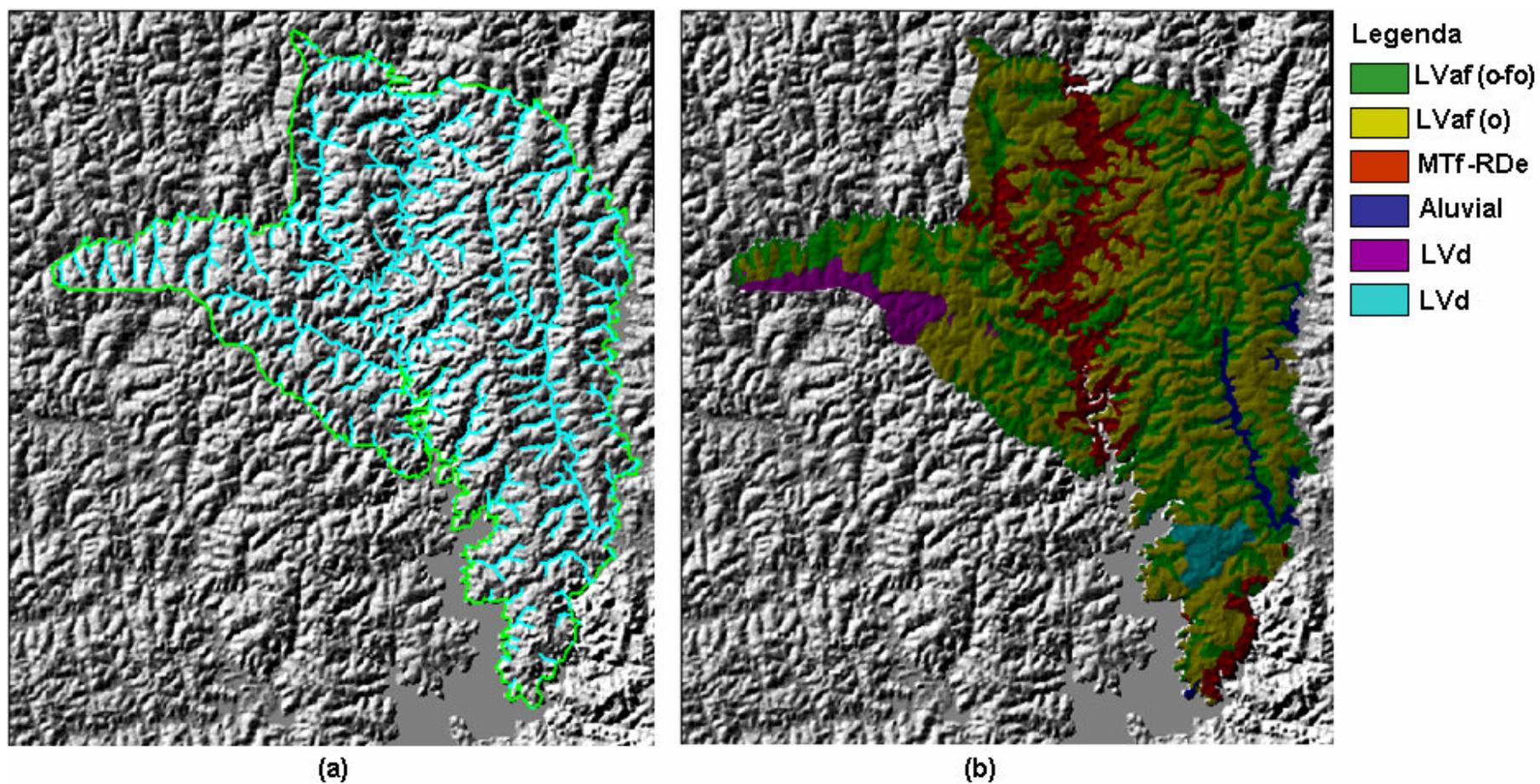


Figura 5. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo município de Ibirubá e parte do município de Quinze de Novembro e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.

Tabela 2. Unidades de mapeamento de solos do “Levantamento e Utilização Agrícola dos Solos do Município de Ibirubá” (adaptado de Santos et al., 1970)

Unidade de mapeamento	Classificação taxonômica segundo EMBRAPA (1999)	Extensão (ha)	Distribuição (%)
LVaf (fo)	Latossolo Vermelho aluminoférrico (ondulado)	26.732	35,5
LVaf (o)	Latossolo Vermelho aluminoférrico (forte-ondulado)	29.879	39,6
LVd1	Latossolo Vermelho distrófico, textura média.	2.366	3,2
LVd2	Latossolo Vermelho distrófico, textura arenosa	1.492	2,0
MTf-RDe	Chernossolo argilúvico férrico típico e Neossolo litólico eutrófico chernossólico	10.287	13,7
Aluvial	Solos hidromórficos	1.234	1,6
Total		72.000	100

4.1.2. Área de Estudo 2: Bacia Hidrográfica de Arroio Portão

a) Situação, limites e extensão:

A área em estudo, compreende aproximadamente 22.400 ha., abrangendo parte dos municípios de Estância Velha, Portão, Ivoti, São Leopoldo e Canoas. A área situa-se entre as coordenadas 29°39' e 29°48' de latitude sul e 51°18' e 51°10' de longitude oeste, abrangendo toda a área da bacia hidrográfica do Arroio Estância Velha/Portão, e também toda a área das bacias hidrográficas dos pequenos Arroio Boa Vista e Arroio da Estância, sendo os três, afluentes da margem direita do Rio dos Sinos.

b) geologia:

Na área ocorrem as formações Botucatu, Rosário do Sul e Serra Geral e os aluviões. A formação Botucatu é constituída por arenitos, apresenta estratificações eólicas e cores variando do róseo ao vermelho. A formação Rosário do Sul compreende faces fluvial com arenitos vermelhos com estratificação cruzada e intercalações de siltitos avermelhados. A formação Serra Geral compreende a zona de escarpas e platôs, sendo constituída por rochas vulcânicas, diques e sills de diabásio associado. Os aluviões são sedimentos atuais e sub-atuais depositados em planícies de inundação e barras, sendo constituídos basicamente por cascalhos, areia, silte e argila.

c) geomorfologia:

A área compreende a região fisiográfica da Depressão Central e a de transição para a Serra Geral. Verificando-se a existência de três formas de relevo predominante: áreas baixas e planas, ao longo dos arroios e rios, formando extensas várzeas ao sul e ao sudeste da área; áreas constituídas por relevo suavemente ondulado a ondulado, compreendendo a maior parte da área onde predominam os arenitos (região mais ao centro da área) e zona escarpada, com relevo acidentado, compreendendo a zona de transição entre os arenitos e a formação Serra Geral, ao norte e nordeste da área.

d) clima:

O clima da região é “Cfa” segundo Koeppen. A temperatura do mês mais quente pode ser superior a 22º C e a do mês mais frio oscila entre -3ºC e 18ºC, sendo a média anual de 19,3º C. A precipitação média anual é de 1.400 mm. Podem ocorrer períodos de déficit hídrico nos meses de janeiro a abril, havendo reposição de abril a julho, excedente de julho a outubro, com retirada e possibilidade de déficit de novembro a dezembro.

e) vegetação:

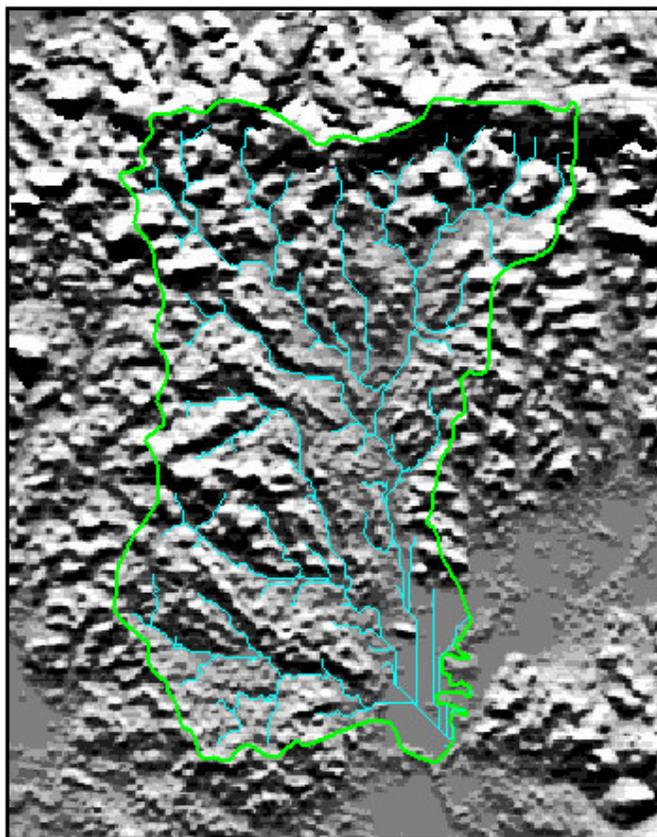
A vegetação nativa compreendia campos nativos (estepe) na Depressão

Central. A vegetação atual, antrópica, compreende a policultura com predomínio do milho, melancia, abóbora, mandioca e de reflorestamentos por eucalipto, acácia, pinus e pomares de citrus. As áreas baixas compreendem campos nativos inundáveis e banhados com porções de reserva biológica associada à criação de gado. Subindo em direção à área de transição entre os sedimentos e basaltos, originariamente florestas nativas eram encontradas (umbrófila densa e umbrófila temperada). Atualmente encontra-se manchas de floresta nativa em terrenos íngremes e nas zonas de policulturas.

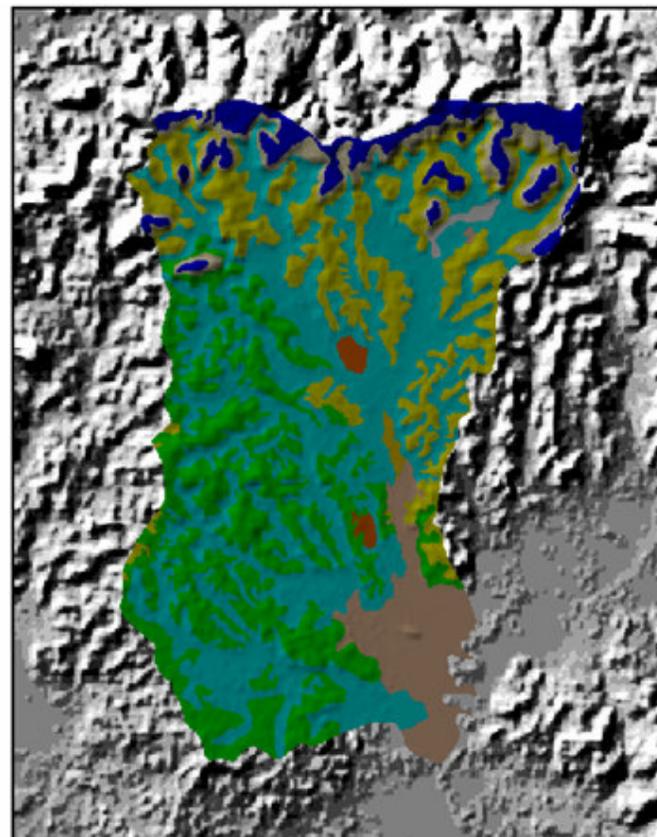
f) solos:

As informações referentes aos solos desta bacia hidrográfica são provenientes do relatório de levantamento de solos elaborado por Klamt et al. (1992). Em 1992, os autores realizaram um levantamento de solos semi-detalhado na bacia de Arroio Portão na escala de 1:50.000. Estes verificaram a ocorrência originalmente de 18 classes de solo distribuídas em 13 unidades de mapeamento, sendo cinco unidades compostas por associações. Estas 13 unidades originais foram agrupadas neste estudo em nove unidades de mapeamento apresentadas na Tabela 3 e Figura 8.

Nas áreas planas e baixas ocorrem Planossolos de textura média e Gleissolos; nas áreas de relevo suave ondulado na região central da área, encontram-se Argissolos Vermelho-Amarelos, enquanto na zona escarpada ocorrem associações como as de Neossolos Litólicos com Chernossolos e Cambissolos.



(a)



(b)

Legenda

-  MXo/Rle
-  Pva
-  PVd
-  GMd
-  Cha
-  CXf/Rle
-  PVa/Gme
-  Aluvial
-  SXd

Figura 6. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo bacia hidrográfica do Arroio Portão e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.

Tabela 3. Unidades de mapeamento de solos do levantamento da Bacia do Arroio Portão (Klamt et al., 1992)

Unidade de mapeamento	Classificação taxonômica segundo Embrapa (1999)	Extensão (ha)	Distribuição (%)
MXo/Rle	Chernossolo Háplico órtico/Neossolo litólico eutrófico	1.372,6	6,0
PVa	Argissolo Vermelho amarelo	3.511,4	15,7
PVd	Argissolo Vermelho distrófico	4.756,7	21,2
GMd	Gleissolo melânico distrófico	1.768,2	8,4
CHa	Cambissolo Húmico álico	88,1	0,4
CXf/RLe	Cambissolo Háplico	1225,3	5,8
PVa/GMe	amarelo/Gleissolo melânico	9.255,9	41,3
Aluvial	Solos hidromórficos	60,9	0,3
SXd	Planossolo Háplico distrófico	121,8	0,6
Total		22.300	100

4.1.3. Área de estudo 3: Área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão, RS

a) Situação, limites e extensão:

A área objeto deste levantamento, com cerca de 100.000 ha, localiza-se no município de Viamão, na região compreendida entre o divisor de águas da Coxilha das Lombas e as margens norte da Lagoa dos Patos e do Casamento. A linha do divisor de águas inicia no Morro da Fortaleza (Itapoã) e segue pela Coxilha das Lombas até o limite municipal entre Viamão e Santo Antônio da Patrulha. As margens consideradas da porção norte da Lagoa dos Patos e do Casamento são aquelas que iniciam na margem do Morro da Fortaleza, passam pelo Pontal das Desertas e vão até a Lagoa Capivari.

b) geologia:

Praticamente restrita à área de Itapuã, encontramos unidades graníticas isoladas ou associados a migmatitos ou gnaisses, cortados por diversos diques de

riolito, microgranitos e aplitos além de diabásios.

A formação Itapoã ocorre na Coxilha das Lombas, sendo formada por areias quartzosas. A formação Chuí ocorre nas bordas da Lagoa dos Patos e Casamento, é composta por areias siltico-argilosas com laminação plano-paralela incipiente.

Também na borda das margens da Lagoa dos Patos e Casamento ocorrem depósitos eólicos ativos de dunas lagunares, depósitos de cristas e praias lagunares (areias quartzosas finas a muito finas) e depósitos lagunares. Junto da Coxilha das Lombas, na porção sudoeste, se encontram leques aluviais, depósitos lagunares e paludais e depósitos de planície e canal aluvial.

c) geomorfologia:

Na área em estudo ocorrem três barreiras e três terraços lagunares, a primeira barreira, corresponde à própria Coxilha das Lombas e é formada por colinas baixas, suaves, com vertentes convexas e cotas médias entre 40 e 50 metros, sendo que seu terraço lagunar se encontra ausente na área. A segunda barreira, consiste em uma região baixa, de superfícies aplainadas e cotas entre 10 e 15 m. A terceira barreira, possui ocorrência bastante restrita, formando uma única faixa estreita e isolada a sudoeste da área.

O planalto rebaixado marginal forma morros e coxilhas com altitudes entre 50 e 260 m, constituindo um prolongamento da região geomorfológica Planalto Sul-Riograndense, onde se desenvolve ao redor dos relevos mais elevados denominados de planalto residuais Canguçu - Caçapava do Sul, que nesse prolongamento seriam representados pelos morros mais altos ao redor de Porto alegre e a sudoeste da cidade de Viamão.

d) clima:

O tipo fundamental é o Cfa₂ de Koeppen (subtropical úmido) com temperatura média anual entre 18 e 19 graus centígrados, precipitação média anual de 1200-1300mm e insolação média de 2300 a 2400 horas/ano. Predominam ventos do quadrante leste e pode ocorrer deficit hídrico de 100mm em 70% dos anos.

e) vegetação:

A vegetação da área é constituída de campos nativos com predomínio de gramíneas, mata subtropical arbustiva em capões e mata de galeria, e, também ocorrendo vegetação higrófila de aguapés e ciperáceas em áreas mal drenadas.

f) solos:

Foram identificadas originalmente 25 UM, estas foram agrupadas por semelhança de características para 10 UM, estas estão listadas na Tabela 4 e espacializadas na Figura 7. Nas áreas mal drenadas temos a predominância de Planossolos Hidromórficos e nas áreas bem drenadas de Neossolos Flúvicos. Em pequenas elevações ocorrem Argissolos Vermelho-Amarelos com textura média sob relevo ondulado e substrato de areias eólicas, tendo como inclusões os Gleissolos.

Sobre substrato granítico ocorrem Neossolos litólicos de textura média com ocorrência de afloramentos rochosos em relevo ondulado a forte ondulado. Encontram-se ainda entre as classes predominantes na região os Neossolos Quartzarênicos, Organossolos e Gleissolos Melânicos existindo várias associações entre estes, bem como alguns tipos de terreno do tipo dunas.

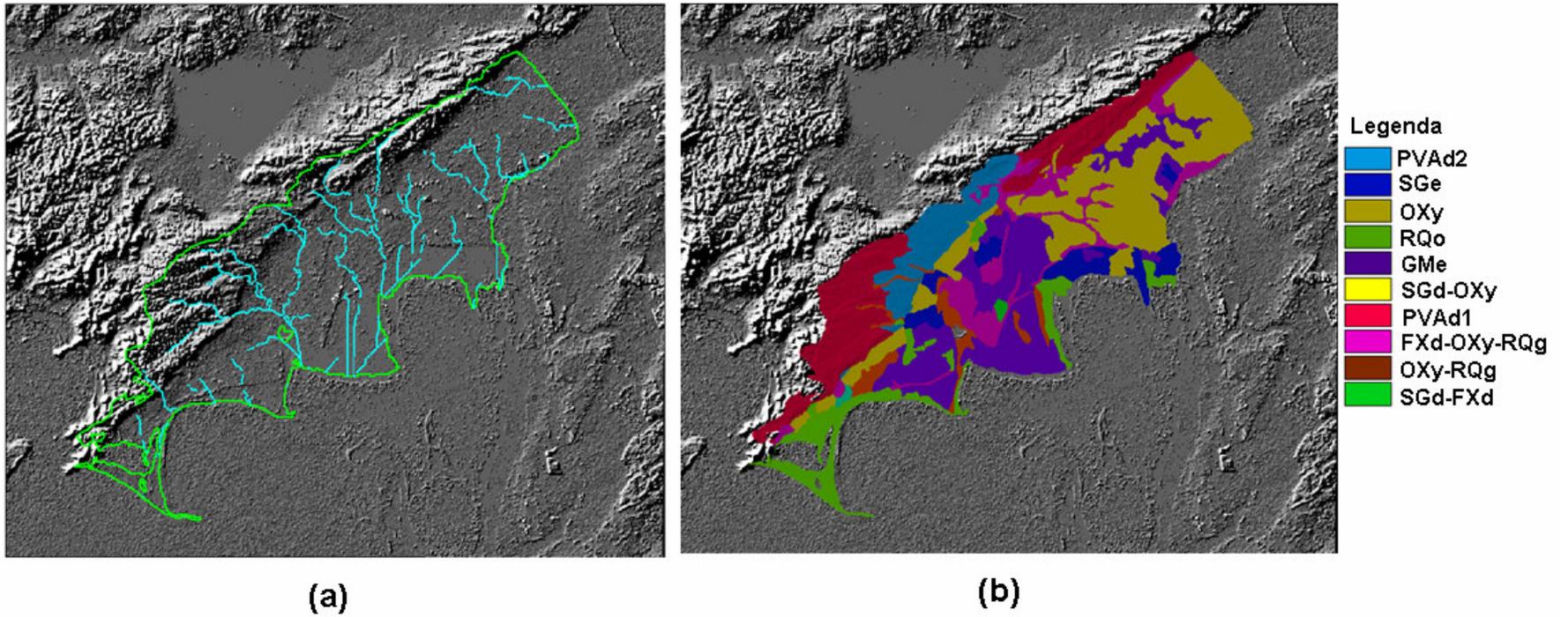


Figura 7. (a) Aspecto do relevo, hidrografia e delimitações da área de estudo Viamão e (b) mapa de solos inserido sobre o relevo da paisagem com respectiva legenda.

Tabela 4. Unidades de mapeamento do levantamento de solos da área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão,RS (Schneider et al., 1992).

Unidade de mapeamento	Classificação Taxonômica segundo Embrapa (1999)	Extensão (ha)	Distribuição (%)
PVAd2	Argissolo Vermelho amarelo distrófico arênico	6.850	6,3
SGd	Planossolo Hidromórfico distrófico arênico	5.840	5,4
OXY	Organossolo Háptico hêmico típico	5.840	5,4
RQo	Neossolo Quartzarênico órtico típico	100	0,1
GMd	Gleissolo Melânico distrófico típico	22.630	21,2
SGd-OXY	Planossolo Hidromórfico distrófico arênico- Organossolo Háptico hêmico típico	20.380	18,7
PVAd1	Argissolo Vermelho-Amarelo distrófico típico	18.740	17,2
FXd-OXY-RQg	Plintossolo Háptico distrófico típico- Organossolo Háptico hêmico típico - Neossolo Quartzarênico hidromórfico típico.	5.660	5,3
OXY-RQg	Organossolo Háptico hêmico típico - Neossolo Quartzarênico hidromórfico típico	9.140	8,4
SGd -FXd	Planossolo Hidromórfico distrófico típico - Plintossolo Háptico distrófico típico	310	0,3
Total		96.000	100%

4.2. Procedimentos

Uma parte fundamental no mapeamento supervisionado de solos é a seleção das fontes de informação que irão alimentar o SIG, bem como as ferramentas computacionais que serão empregadas.

O SIG empregado neste estudo foi o ArcView 3.2 (ESRI, 1995), que possui um grande número de usuários em todo o mundo, vários tutoriais e tem sido amplamente empregado em estudos de geoprocessamento, pertencente à família de softwares SIG mais utilizada no mundo atualmente. Além do programa básico de funções comuns a qualquer SIG, o ArcView 3.2 possui recursos de extensões e “scripts”. As extensões são arquivos executáveis de programas que rodam dentro da interface do ArcView 3.2 como parte integrante do programa, são escritos em diversas linguagens de programação, já os scripts são programações escritas e editadas dentro do próprio ArcView 3.2 usando linguagem de programação Avenue e destinadas a executar funções específicas.

Dentre as extensões disponíveis para o ArcView 3.2, a extensão mais comumente empregada na modelagem geomorfológica e hidrológica do terreno quando utilizando dados matriciais é a extensão *Spatial Analyst*, que propicia ao SIG ferramentas para criar, manipular, realizar cálculos algébricos, analisar e mapear células agrupadas em qualquer sistema de arquivos do tipo *raster*.

A outra ferramenta computacional empregada foi o pacote estatístico Minitab 11 (Minitab Inc., 1996). É um programa que roda em ambiente Windows e agrupa dados em forma de tabela e possui uma família de funções de regressões que inclui as regressões logísticas múltiplas.

Como o objetivo deste estudo envolve a utilização de material facilmente disponível e acessível para a maioria dos pedólogos, com custos relativamente reduzidos, utilizaram-se mapas de solos de levantamentos tradicionais pré-existent e disponíveis em escalas variadas encontrados em mapotecas pedológicas. Os solos das áreas de estudo 1, 2 e 3 foram originalmente mapeados nas escalas 1:80.000, 1:50.000 e 1:100.000, respectivamente.

Os mapas originais em papel foram copiados para a forma de imagem digital através de *scanner de rolo* e importados para o SIG ArcView 3.2. As imagens foram georreferenciadas utilizando-se a extensão Image Analyst do ArcView 3.2. Com o auxílio do sistema de edição do SIG, as linhas de contorno das UM de solos na imagem foram digitalizadas na tela criando-se arquivo do tipo vetorial (*shape*), sendo este processo conhecido como vetorização. Possuindo o arquivo vetorial uma tabela de atributos, nesta classificaram-se cada um dos polígonos em função do tipo de solo correspondente, gerando o mapa digital de solos. Posteriormente, o mapa vetorial foi convertido para o formato *raster (grid)*, formato em que o mapa é composto por milhares de pequenas unidades celulares (pixels), sendo que este é o formato necessário para executar análises algébricas de mapas.

A outra fonte de informação utilizada, foi o mapa de elevação da área, também conhecido como modelo digital de elevação ou modelo numérico do terreno (MNT). O MNT empregado neste estudo tem uma resolução mínima de 3 arc sec ou de 90m de tamanho de pixel (tendo cada *pixel* uma área de 0,81ha), obtido do Shuttle Radar Topograph Mission (SRTM/USGS).

As variáveis para uma modelagem digital geomorfológica e hidrológica do terreno, aspectos fundamentais no mapeamento supervisionado de solos, foram obtidas através do MNT, constituindo-se de variáveis geomorfológicas primárias (declividade, elevação, aspecto e curvaturas), geomorfológicas secundárias na forma de índices calculados (índice de umidade topográfica e índice de poder de córrego) e hidrológicas (direção do fluxo, acúmulo do fluxo, comprimento do fluxo e distância dos rios) (Wilson & Gallant, 2000).

As variáveis topográficas primárias e secundárias foram calculadas a partir do MNT, usando-se as extensões *Digital Elevation Model Analysis Tools* (DEMAT) (para cálculo de variáveis primárias) e *Terrain Analysis* (para cálculo de variáveis secundárias) do *Spatial Analyst* em ambiente ArcView 3.2 (ESRI, 1995).

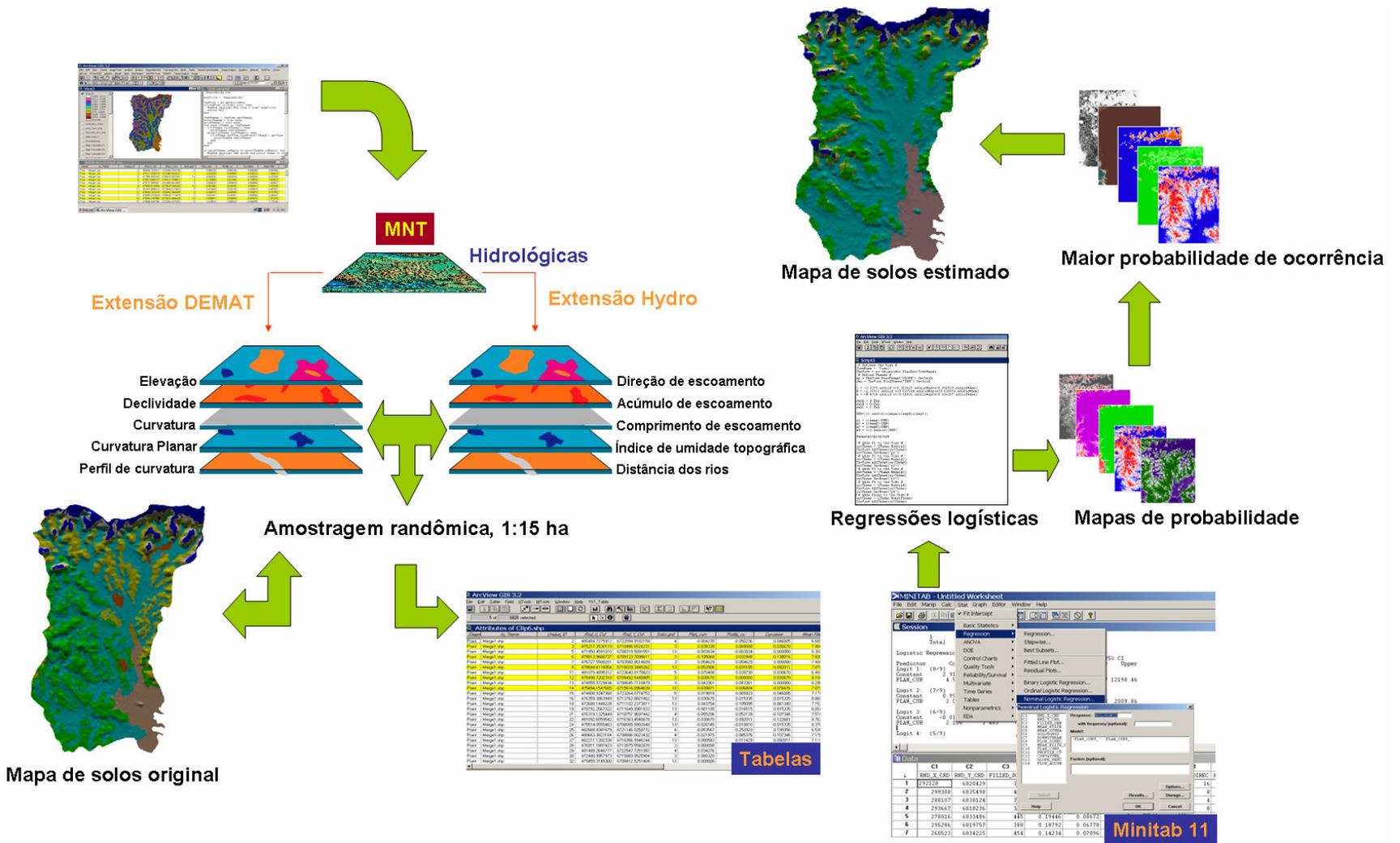


Figura 8. Fluxograma dos procedimentos gerais desenvolvidos para o mapeamento supervisionado de solos.

A extensão DEMAT é uma interface entre o *Spatial Analyst* e o usuário para modelagem geomorfológica do terreno, permitindo calcular a declividade, a curvatura, a curvatura planar, a curvatura horizontal e o aspecto a partir do MNT (Figura 9).

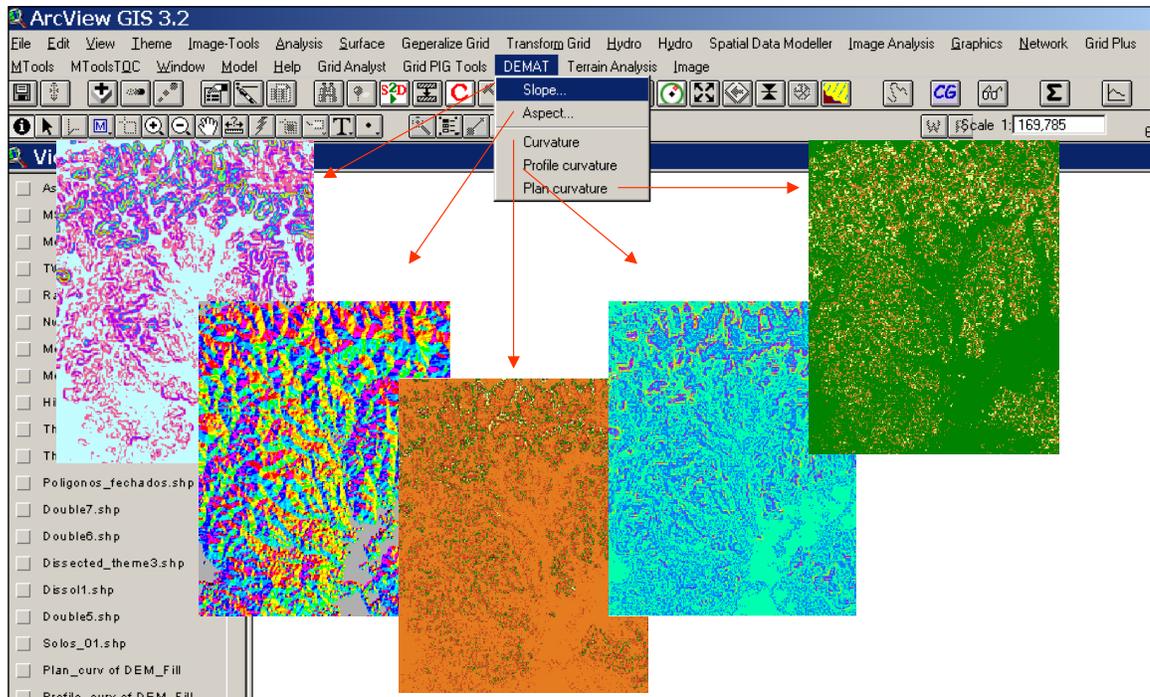


Figura 9. Quadro esquemático representado o produto da extensão DEMAT ao partir do MNT.

Já a extensão *Terrain Analysis* é a interface entre o usuário e o *Spatial Analyst* para cálculo das variáveis topográficas secundárias IUT e IPC (Figura 10).

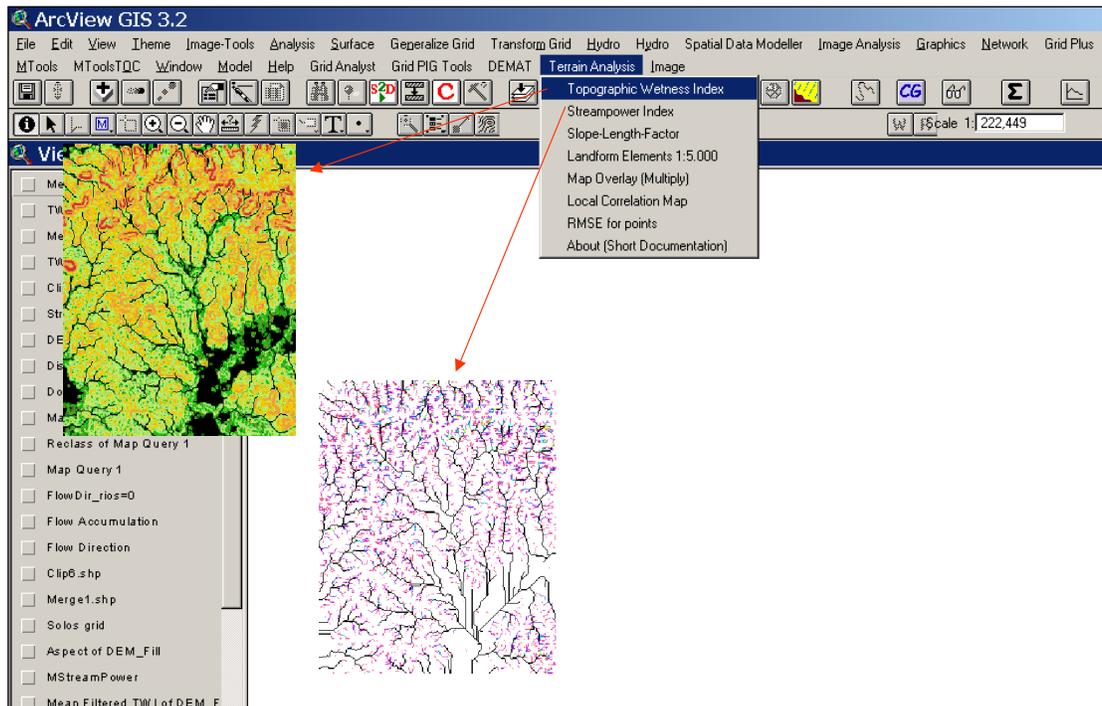


Figura 10. Quadro esquemático representando o produto da extensão Terrain Analysis: o IUT (*Topographic Wetness Index*) e o IPC (*Streampower Index*).

As modelagens hidrológicas das áreas de estudo foram executadas através das seguintes variáveis: direção de fluxo de escoamento; acumulação de fluxo; distância dos rios e comprimento de fluxo.

As variáveis para modelagem hidrológica foram obtidas com base no MNT empregando-se a extensão *Hydrologic Modeling Sample* (Hydro) do *Spatial Analyst* do ArcView 3.2). A extensão Hydro é a interface entre o *Spatial Analyst* e o usuário para modelagem hidrológica.

Através da extensão Hydro, a direção de fluxo foi gerada automaticamente a partir da elevação (MNT) pela função *Flow direction*. Sobre a direção de fluxo foi calculada a acumulação de fluxo através da função *Flow Accumulation*. Para a obtenção da variável distância dos rios, os rios foram gerados a partir do mapa de acumulação de fluxo, por seleção de células com

acúmulo de fluxo maior que 100 células para as regiões de estudo 1 e 2 e maior que 500 células para região de estudo 3. A partir do mapa de rios, através da função *Find Distance* do Spatial Analyst foram calculadas as distâncias de cada célula em relação às células determinadas como rios.

Multiplicando-se o mapa dos rios pela direção de fluxo, obteve-se o mapa de direção de fluxo onde os rios têm valor igual a zero. Sobre este mapa foi calculado o comprimento de fluxo através da função *Flow Length*.

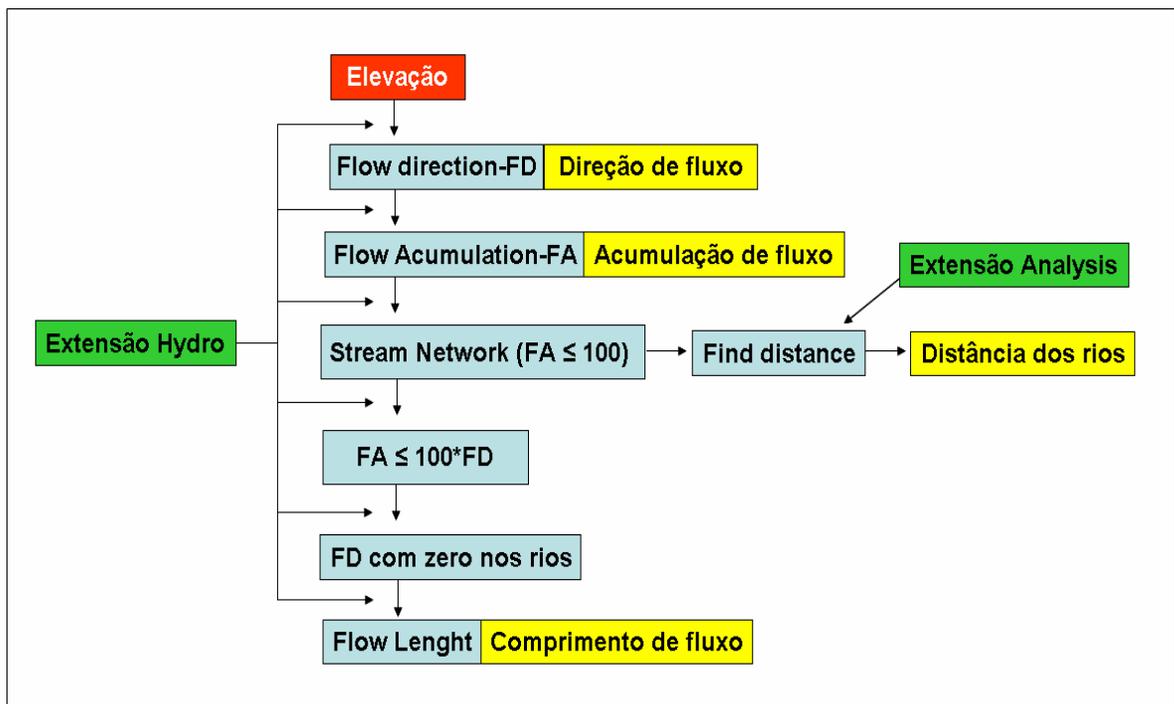


Figura 11. Quadro esquemático mostrando os procedimentos de obtenção de variáveis hidrológicas com base em modelo numérico de terreno através de extensões do SIG ArcView 3.2.

4.3. Amostragem

Com o objetivo de tabular as informações obtidas no cálculo das variáveis empregadas neste estudo a fim de se proceder ao cálculo das regressões logísticas múltiplas, foi criado um mapa de pontos aleatórios de

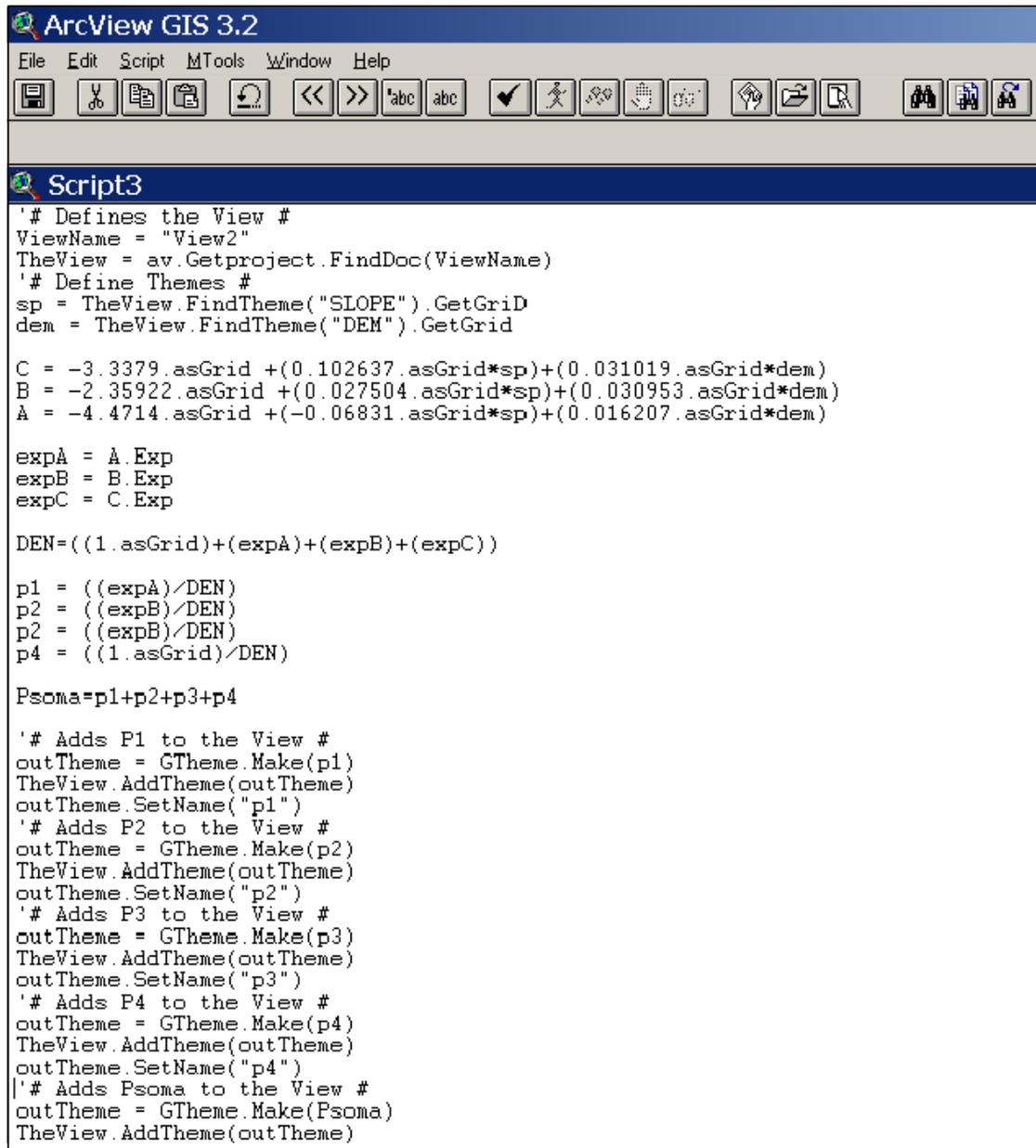
amostragem através da extensão *Randon Points Generator*, correspondente a um ponto de amostragem a cada 15 ha. A este mapa de pontos, esteve vinculada uma tabela com a posição geográfica de cada ponto. Através do uso do *script Grid Sampler* foram sendo adicionadas colunas a esta tabela com valores obtidos sobre os mapas de cada uma das variáveis topográficas e hidrográficas sendo referente a cada ponto de amostragem, bem como foi adicionada uma coluna para a informação unidade de mapeamento obtida sobre o mapa de solos. Esta tabela foi exportada em formato Data Base File (DBF) para uso no programa estatístico Minitab 11 (Minitab Inc.).

4.4. Regressões Logísticas Múltiplas

A tabela exportada do ambiente ArcView 3.2. na forma de DBF, foi importada para o ambiente Minitab 11. No programa, em ambiente Windows, foi selecionado o menu das famílias de regressões e então selecionada para a opção regressão logística múltipla. Neste pacote estatístico, a variável UM foi selecionada como variável resposta e os atributos de terreno como variáveis preditoras em diferentes combinações.

A seleção das variáveis preditoras mais significativas foi feita com base nos critérios de seleção através de um processo interativo, procedendo-se a análise com diversos conjuntos de variáveis, eliminando-se aquelas que não influenciavam significativamente o comportamento da variável resposta. Este procedimento foi repetido até obter-se um conjunto de variáveis preditoras com significativa influência sobre a variável resposta, que não eram funcionalmente redundantes nem geravam conflitos entre si.

Os dados correspondentes às equações geradas foram transcritos do ambiente Minitab novamente para o ambiente ArcView, onde as equações foram escritas através de linguagem de programação Avenue específica para a produção de arquivos de programação e execução do tipo *scripts* (Figura 12).



```

ArcView GIS 3.2
File Edit Script MTools Window Help
[Icons]

Script3
'# Defines the View #
ViewName = "View2"
TheView = av.Getproject.FindDoc(ViewName)
'# Define Themes #
sp = TheView.FindTheme("SLOPE").GetGrid
dem = TheView.FindTheme("DEM").GetGrid

C = -3.3379.asGrid +(0.102637.asGrid*sp)+(0.031019.asGrid*dem)
B = -2.35922.asGrid +(0.027504.asGrid*sp)+(0.030953.asGrid*dem)
A = -4.4714.asGrid +(-0.06831.asGrid*sp)+(0.016207.asGrid*dem)

expA = A.Exp
expB = B.Exp
expC = C.Exp

DEN=((1.asGrid)+(expA)+(expB)+(expC))

p1 = ((expA)/DEN)
p2 = ((expB)/DEN)
p2 = ((expB)/DEN)
p4 = ((1.asGrid)/DEN)

Psoma=p1+p2+p3+p4

'# Adds P1 to the View #
outTheme = GTheme.Make(p1)
TheView.AddTheme(outTheme)
outTheme.SetName("p1")
'# Adds P2 to the View #
outTheme = GTheme.Make(p2)
TheView.AddTheme(outTheme)
outTheme.SetName("p2")
'# Adds P3 to the View #
outTheme = GTheme.Make(p3)
TheView.AddTheme(outTheme)
outTheme.SetName("p3")
'# Adds P4 to the View #
outTheme = GTheme.Make(p4)
TheView.AddTheme(outTheme)
outTheme.SetName("p4")
'# Adds Psoma to the View #
outTheme = GTheme.Make(Psoma)
TheView.AddTheme(outTheme)

```

Figura 12. Algoritmo do script empregado para executar as regressões logísticas múltiplas em ambiente ArcView 3.2.

No mesmo *script* foram escritas programações em Avenue para construção de mapas de probabilidade de ocorrência e distribuição de cada classe de solo, com base nas variáveis selecionadas para construção do modelo através das regressões logísticas.

A partir deste ponto, através das funções *Map Query* e *Map Calculator* do Arcview 3.2, foram produzidos mapas individuais mostrando em cada *célula* a UM com maior probabilidade de ocorrência, definindo esta UM como o solo ocorrente naquela *célula*.

A partir da reclassificação e sobreposição dos mapas individuais, que mostram cada UM separadamente, juntaram-se os mesmos para geração do mapa estimado de solos.

4.5. Estatística da acurácia

Para a avaliação da qualidade dos mapas estimados usou-se como referência a comparação de correspondências entre este com o mapa original de solos. A metodologia empregada é baseada na construção de matrizes de erro, relacionando o mapa estimado com o mapa original. Através da matriz de erros pode-se calcular a acurácia geral (AG), a acurácia de usuário do mapa (AU), a acurácia do produtor de mapa (AP) e o coeficiente Kappa de Cohen.

Quando os índices de acurácia mostraram baixa concordância entre o mapa original e o mapa estimado, repetiu-se o procedimento selecionando-se outras equações logísticas, através da variação das variáveis independentes utilizadas.

Quando, mesmo após este processo, não se obteve resultados satisfatórios, optou-se pela repetição do mesmo procedimento a partir de uma compilação do mapa original (de campo), simplificando-se a legenda através do agrupamento de UM com classes de solos similares, o que de certa forma representaria um mapa em escala menor. Nesses casos, são apresentadas ambas as situações, com e sem simplificação da legenda.

5. RESULTADOS E DISCUSSÃO

5.1. Resultados e discussão por Caso de Estudo

5.1.1. Caso de Estudo 1: Região de Ibirubá e Quinze de Novembro

A distribuição espacial dos diversos atributos do terreno calculados através da modelagem geomorfológica e hidrológica da área do Caso de Estudo 1, é apresentada na Figura 13, que permite visualizar padrões de distribuição mais ou menos evidenciados. Devido à sua relação direta com o MNT e deste com o fluxo de água, para diversos destes parâmetros percebe-se clara relação entre a distribuição dos solos e as principais drenagens da área.

As variáveis selecionadas para uso nas regressões logísticas múltiplas por apresentarem melhores índices de explicação da ocorrência de solos na área foram: a elevação, índice de umidade topográfica (IUT), índice de poder de córrego (IPC), curvatura do terreno, declividade, distância dos rios e comprimento de fluxo. Os parâmetros e coeficientes das equações logísticas determinados e posteriormente empregados no modelo de predição podem ser verificados na Tabela 5. Embora não incluídos nestes resultados, tentativas de regressões logísticas com menores números de variáveis para explicar a distribuição dos solos resultaram baixa concordância com o mapa original de solos.

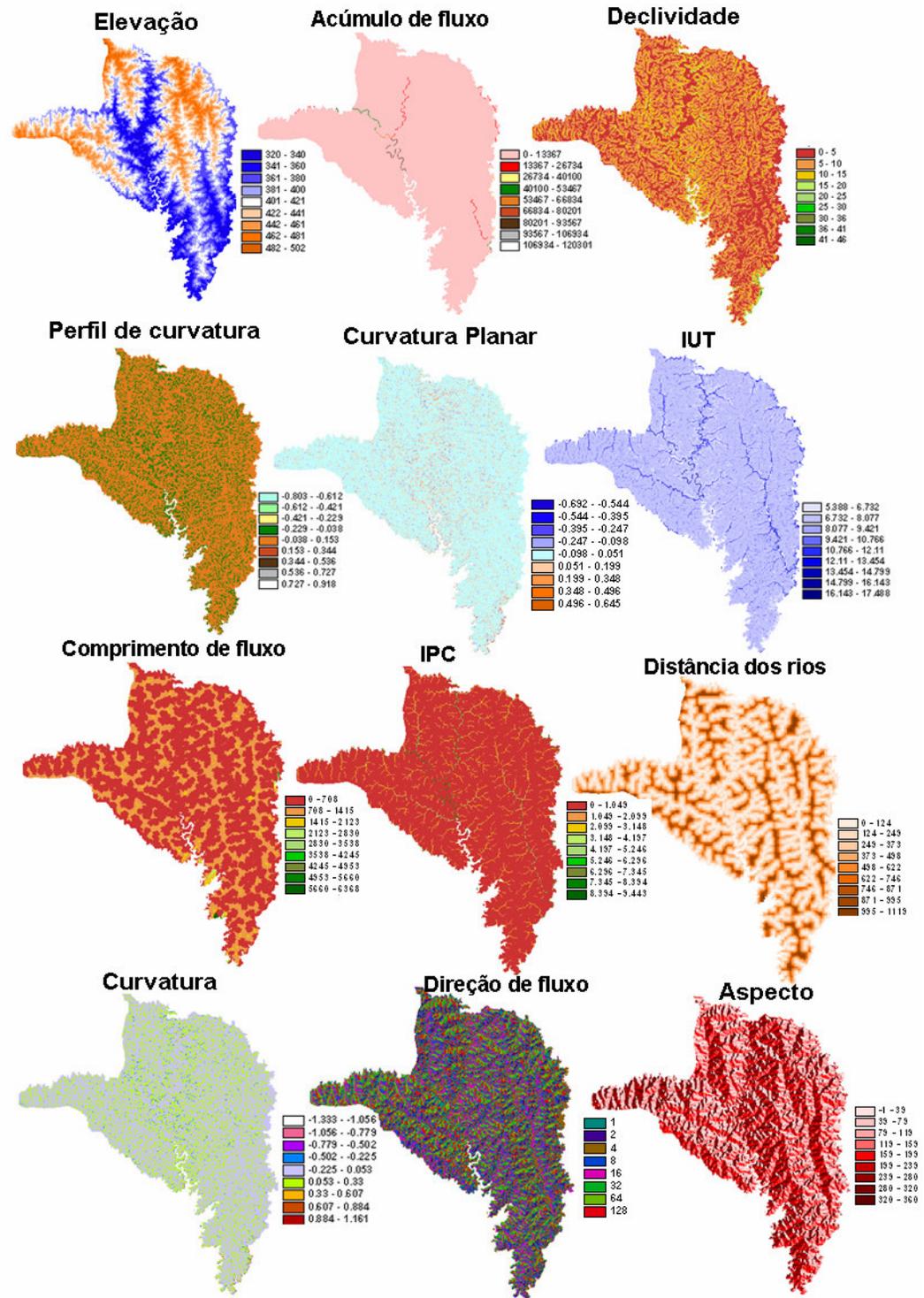


Figura 13. Variáveis oriundas da modelagem topográfica e hidrográfica a partir de MNT na área correspondente ao Município de Ibirubá e parte do Município de Quinze de Novembro, RS.

Tabela 5. Parâmetros e coeficientes de modelo (*função logit*) empregado na predição da ocorrência de classes de solos na região do Município de Ibirubá e parte do Município de Quinze de Novembro, RS.

Parâmetros do modelo	Modelo de probabilidade				
	MTf-RDe/ LVa(o)	MTf-RDe/ LVa(o-fo)	MTf-RDe/ LVd1	MTf-RDe/ Aluvial	MTf-RDe/ LVd2
<i>Constante</i>	-11,4379	-13,9980	-51,368	-15,616	-7,150
Poder de córrego	-0,8784	-2,2960	-1,5968	-2,9258	-4,5557
Comprimento de fluxo, m	-0,000069	0,0000831	-0,0016153	0,0004944	-0,000995
Índice de umidade topográfica	0,36941	0,30454	1,6272	1,1602	0,7301
Curvatura, m	-10,039	-8,331	-6,08	-11,81	-17,77
Declividade, %	-0,06995	-0,08138	-0,09233	-0,21753	-0,12153
Elevação, m	0,027598	0,034852	0,095525	0,020473	0,003217
Distância dos rios, m	-0,000403	0,0011404	0,0000313	-0,003173	0,0029561
Log-likelihood = -8566.523, G = 3908.959, P-Value = 0.000					

As probabilidades de ocorrência de cada UM com base nas equações da Tabela 5 são apresentados na Figura 14. Observaram-se probabilidades baixas em reação a LVd1 e Aluvial, médias para LVaf (o) e LVaf (o-fo) e altas para algumas partes de MTf-RDe e LVd2. Esta distribuição de probabilidades aparentemente semelhantes principalmente entre LVaf (o) e LVaf (o-fo), é uma grande fonte de erros principalmente nas zonas de limitações entre um e outra, de fato, este conflito está baseado no fato das duas UM pertencerem à mesma classe de solo com diferenças apenas nas fases de relevo.

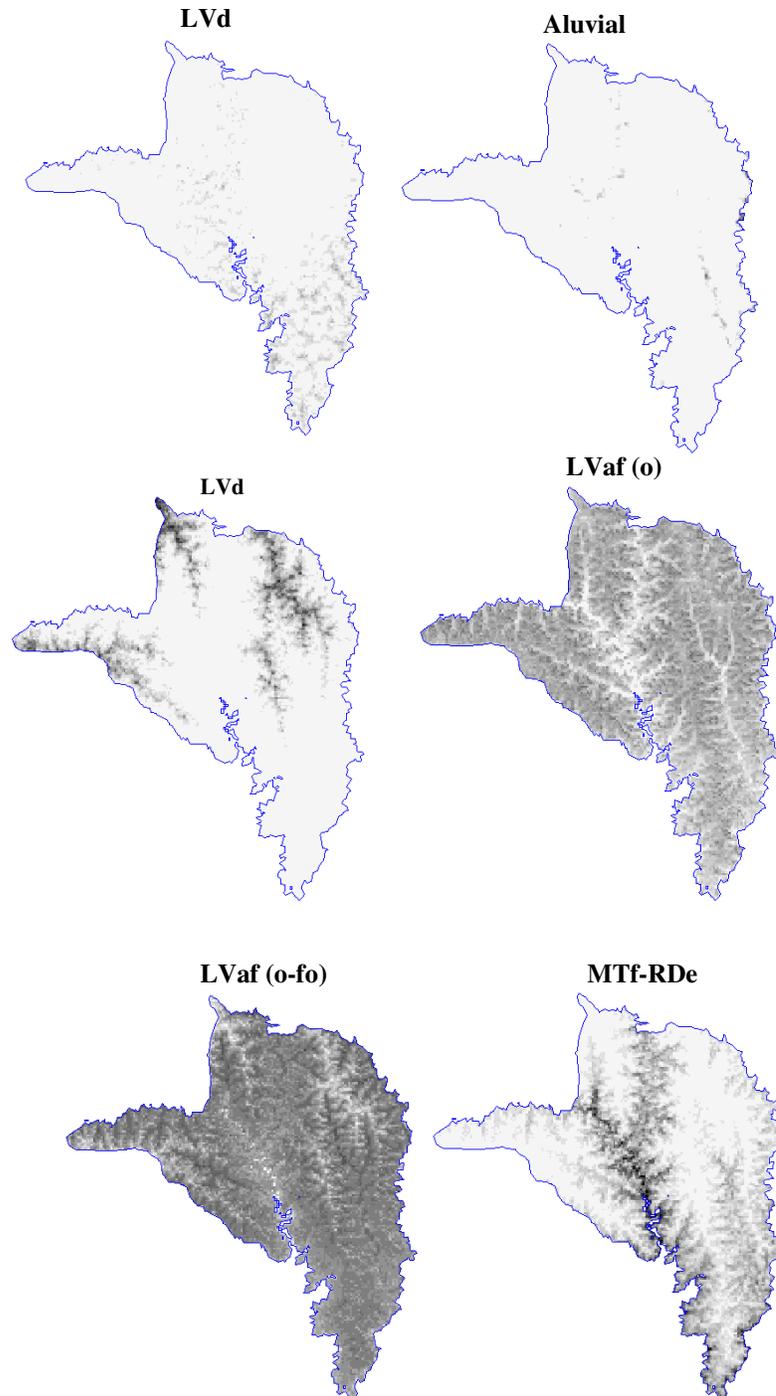


Figura 14. Probabilidades de ocorrência de unidades de mapeamento na região do Município de Ibirubá e parte do Município de Quinze de Novembro, RS. A escala apresenta-se em tons de cinza, variando entre probabilidade de 0% (branco) a probabilidade de 100% (preto).

Ao estabelecer-se uma comparação visual entre o mapa estimado com o mapa original (Figura 15), verifica-se um elevado grau de semelhança visual, que embora não seja uma medida de acurácia, indica que há uma correspondência entre os dois mapas. O fato de ocorrer essa correspondência aparentemente elevada no aspecto visual, indica que o modelo adequou-se em algum grau (a ser medido a seguir) na predição das pedoformas originais.

Ao analisar-se a matriz de erros (Tabela 6) comparando o mapa original com o mapa estimado, verificou-se que as UM LVd1 e LVd2 apresentaram os maiores valores de AU, sendo de 91,5% e 96,7%, respectivamente. Os menores valores de AU foram verificados na UM Aluvial, com 37,4% de correspondência. As unidades mais representativas em termos de extensão da área, LVaf (o-fo) e LVaf (o) tiveram uma AU média variando de 49,7% a 69,8%, respectivamente.

Na avaliação da AM, observou-se que o maior percentual de AM foi verificado para a unidade MTf-Rde com 69,6%, seguida de LVa (o) com 60,8%, sendo que as menores AM foram de 26,0% e 23,2% para LVd1 e LVd2, respectivamente. Neste caso, uma AM de 69,6% significa que, ao consultar no mapa a unidade de mapeamento MTf-Rde, 69,6% das vezes a indicação da ocorrência desta UM coincidirá com a indicação de ocorrência no mapa original.

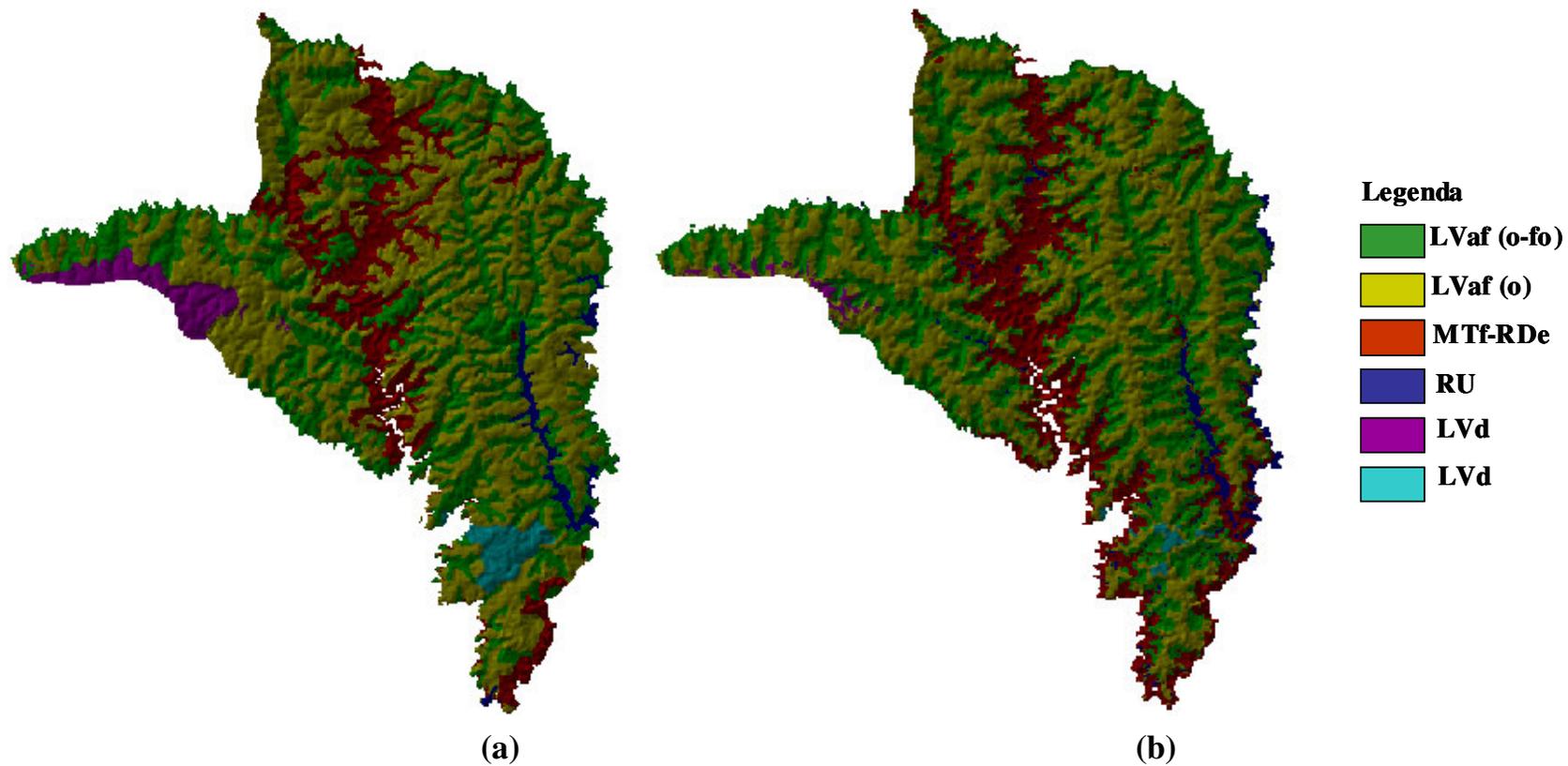


Figura 15. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas na região do município de Ibirubá e parte do município de Quinze de Novembro.

Tabela 6. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na região de Ibirubá e Quinze de Novembro, RS

Mapa estimado	Mapa original						Total	AU
	LVa (o)	LVd1	LVd2	Aluvial	MTf-RDe	LVa (fo)		
LVa (o)	28,0	1,0	0,5	0,1	1,5	9,0	40,2	69,8
LVd1	0,0	0,7	0,0	0,0	0,0	0,1	0,8	91,5
LVd2	0,0	0,0	0,5	0,0	0,0	0,0	0,5	96,7
Aluvial	0,3	0,0	0,0	0,9	0,2	1,2	2,5	35,4
MDTf-RDe	3,6	0,0	0,1	0,4	9,4	4,9	18,4	51,4
LVa (fo)	14,2	1,1	0,8	0,3	2,4	18,7	37,6	49,7
Total	46,1	2,9	2,0	1,7	13,5	33,8	100	
AM	60,8	26,0	23,2	54,0	69,6	55,3		

AG = 58,3%; Kappa = 36,6

Estas diferenças verificadas entre as duas acurácias, provavelmente contribuiriam para a redução da acurácia geral do mapa estimado (AG = 58,3% e Kappa = 36,6%). A maioria das áreas LVd1 e LVd2 foram estimadas corretamente no mapa estimado, porém o modelo estimado está superestimando a ocorrência destas duas unidades de mapeamento, prevendo sua ocorrência onde na verdade ocorrem outros solos, em detrimento das demais UM, principalmente MTf-RDe e LVa (fo).

As diferenças entre as acurácias demonstra a dificuldade do método preditivo em identificar e distinguir as UM LVd1 e LVd2 em relação às demais se baseando apenas em parâmetros derivados do MNT. De fato o fator diferencial entre essas unidades e as unidades LVaf (o) e LVaf (o-fo), de textura argilosa, é a classe textural, sendo arenosa para LVd2 e média para LVd1. Estas diferenças de textura estão relacionados com o material de origem, já que a UM LVd2 desenvolveu-se a partir de rochas da formação Tupaciretã e LVd1 recebeu interposição de fração areia, provavelmente de processos erosionais ocorridos em morros de arenito, enquanto as solos da UM LVaf desenvolveram-se a partir de rochas básicas extrusivas. A inclusão da variável material de origem ao modelo

poderia melhorar a capacidade do mapa estimado em reproduzir o mapa original, entretanto, para esta área de estudo não existe levantamento geológico disponível em escala adequada.

A fim de verificar se o uso de uma legenda simplificada poderia aumentar a acurácia do mapa estimado (embora com redução da precisão) em relação ao mapa quando com a legenda original, esta foi reduzida a três UM ((LVaf (o-fo), LVaf (o) e MTf-RDe). As variáveis utilizadas foram as mesmas quando usando a legenda original. A comparação entre o mapa estimado e o mapa original com a legenda simplificada por ser feita visualmente na Figura 16 e na Tabela 7.

Não foram observadas alterações significativas na AG, que foi de 61,4% (Kappa = 38,3%). Isto provavelmente é devido ao fato de haver um grande percentual de área indefinida entre LVaf (o-fo) e LVaf (o), como pode ser verificado na Tabela 7, onde, por exemplo 22% da área mapeada como UM LVaf (o) está estimada como sendo LVaf (o-fo) e 26% da área original da UM LVaf (o-fo) está estimada como sendo LVaf(o).

Grande parte deste problema provavelmente está associado ao fato de que estas duas unidades de mapeamento pertencem à mesma classe de solo com diferenças em relação apenas às fases de relevo. Na época do mapeamento, em 1970, o relatório descreve a classe de relevo para LVaf (o-fo) como sendo parte ondulada que não se consegue visualmente dissociar da parte forte ondulada que figura em declividades entre 8% a 15%, e a parte ondulada de LVaf (o), está em declividades de até 8%. Ao se analisar o mapa de declividade, verifica-se que dentro da área LVaf (o-fo) há uma grande diferença entre o que está mapeado com relevo forte ondulado e o que possui declividade entre 8% e 15% com base no MNT.

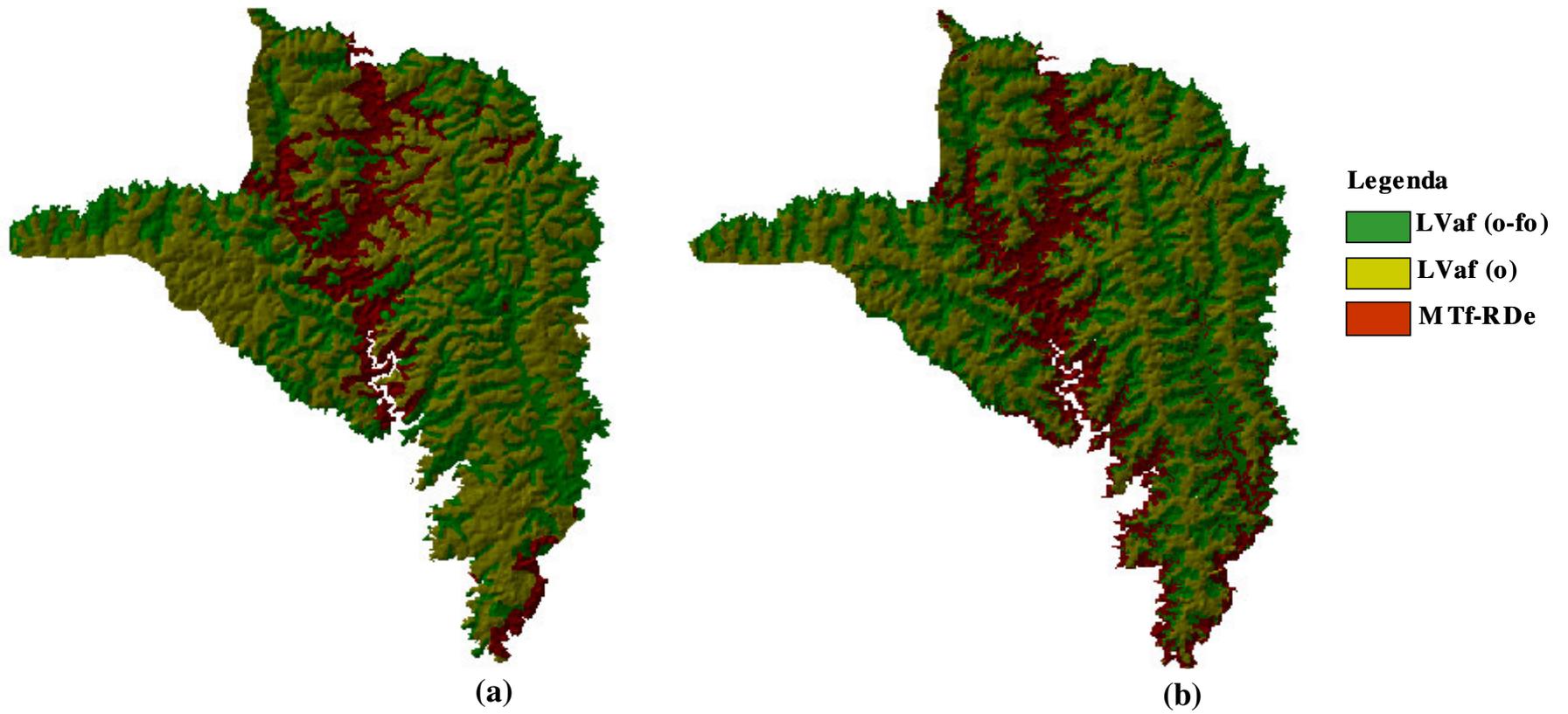


Figura 16. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas com legenda simplificada na região do município de Ibirubá e parte do município de Quinze de Novembro, RS.

Esta indefinição nas zonas de limites entre as duas UM provavelmente está contribuindo com a redução da acurácia do mapa estimado, tanto na escala original quanto na escala simplificada. Este pode ser um revés em se utilizar áreas piloto de mapeamentos anteriores muito antigos pelos métodos tradicionais de levantamentos de solos. Uma alternativa que provavelmente elevaria a acurácia dos mapas de solos tradicionais, assim como a concordância destes com os mapas estimados, seria desenvolverem-se levantamentos de solos através de métodos tradicionais tutelados por mapas digitais com base no MNT, assim como a inclusão de informações detalhadas sobre a geologia da área, que poderia compartimentar a área em zonas de influência de diversos materiais de origem.

Tabela 7. Matriz de erro entre o mapa original com legenda simplificada e o respectivo mapa estimado por regressões logísticas múltiplas na região de Ibirubá e Quinze de Novembro, RS

Mapa estimado	Mapa original			Total	AU
	LVaf (o)	MTf-RDe	Lvaf (o-fo)		
LVaf (o)	30,9	1,5	9,2	41,6	74,3
MTf-RDe	3,7	9,5	5,2	18,4	51,4
Lvaf (o-fo)	16,6	2,6	20,8	39,9	52,0
Total	51,2	13,6	35,2	100,0	
AM	60,4	69,6	59,1		
AG = 61,4%, Kappa = 38,3%					

5.1.2. Caso de Estudo 2: Bacia Hidrográfica do Arroio Portão

A Figura 17 apresenta a distribuição espacial dos valores de cada uma das 12 variáveis topográficas e hidrográficas empregadas neste estudo. Visualmente verifica-se correspondência entre as pedoformas do mapa de solos com as geoformas da elevação, declividade e curvatura. De fato, as melhores regressões logísticas múltiplas foram obtidas com as variáveis elevação, declividade, curvatura e distância dos rios (Figura 17 e Tabela 8).

Por ocorrerem grandes extensões da área com a presença de Gleissolos, esperava-se que mais variáveis hidrológicas relacionadas à identificação de zonas de acúmulo de água poderiam conceber regressões mais apuradas, o que de fato não ocorreu.

Todos os atributos selecionados para compor a equação de regressão logística possuem grande variação espacial dentro da área. Como, por exemplo, em relação à elevação este atributo varia entre 7m nas cotas mais baixas a 304 m nos morros mais altos, sendo os valores de elevação entre estes dois extremos bem distribuídos na área. Vemos áreas de baixada ao Sul, elevações suaves ao centro e morros mais íngremes ao Norte. O mesmo pode ser observado para as curvaturas e declividade. Em relação à rede de drenagem, exemplificada pelos rios, estes percorrem estendendo-se por toda a área, sendo a distância destes uma variável que adquire importância pela sua forte relação com o relevo e a dinâmica de águas superficiais e subsuperficiais, dessa forma uma importante variável para predição de UM de solos.

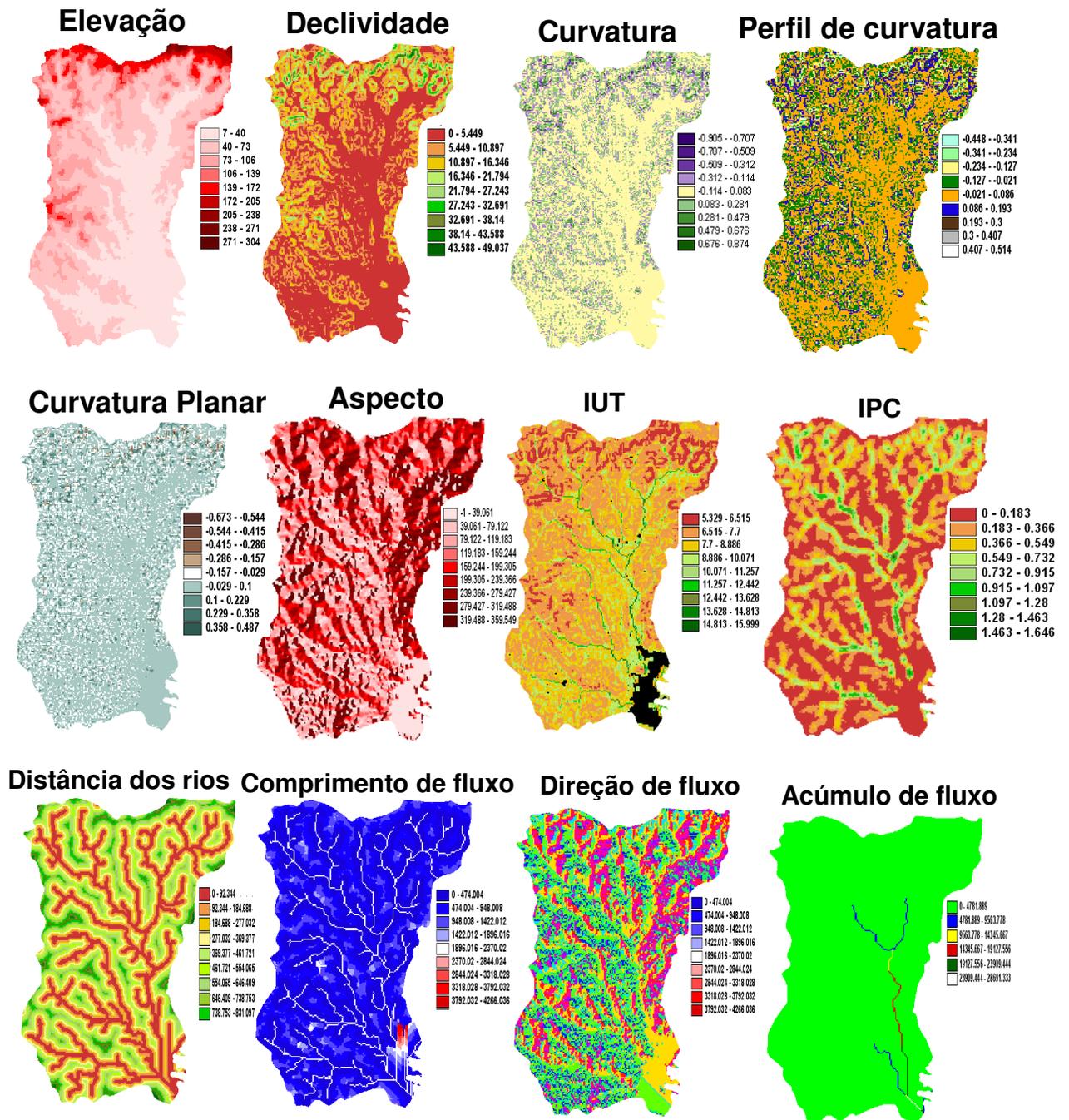


Figura 17. Variáveis oriundas da modelagem topográfica e hidrográfica a partir de MNT na Bacia Hidrográfica do Arroio Portão, RS.

Tabela 8. Parâmetros e coeficientes de modelo (*função logit*) empregado na predição da ocorrência de classes de solos na região da Bacia Hidrográfica de Arroio Portão, RS

Parâmetros do modelo	Modelos de probabilidade							
	PVa-GMe/Aluvial	PVa-GMe/PVd	PVa-GMe/PVa	PVa-GMe/MXo/RLe	PVa-GMe/Sxd	PVa-GMe/CHa	PVa-GMe/Cxf/RLe	PVa-GMe/GMd
Constante	3,5692	-11.5215	-0.6267	-1.7017	-14.796	-3.5780	-2.8916	-4,475
Declividade, %	-0,0030	0.16304	-0.2060	-0.15806	0.10917	0.10565	0.03564	-0,067
Elevação, m	-0,2947	0.09178	-0.1396	-0.06038	0.11691	0.02064	0.01486	0,0125
Dist. Rios, m	0,0021	0.00019	-0.0011	0.000040	0.00171	0.00254	0.00400	0,0007
Curvatura, m	8,726	-2.4469	10.523	2.094	-1.9574	0.5659	1.0420	1,314
Log-likelihood = 11.635, G = 2.893, P-Value = 0.000								

Na Figura 18, observamos os mapas temáticos de distribuição espacial das probabilidades de ocorrência de cada classe de solo com base nas regressões logísticas múltiplas. Nesta figura, observamos mapas com áreas de elevada probabilidade de ocorrência nas unidades MXo/RLe, GMd e CXf/RLe, de fato, ao comparar-se o mapa de solos com os mapas de probabilidade, verifica-se correspondência visual entre as pedoformas com as “formas regressionalis”. Uma probabilidade média de ocorrência foi verificada na unidade PVd. Já as unidades CHa, Aluvial e SXd tiveram uma probabilidade baixa de ocorrência. Estas diferenças de intensidade de probabilidade irão influenciar nas feições finais do mapa estimado, pois “formas regressionalis” de mapas com probabilidade média ou baixa podem se assemelhar com sua pedoforma correspondente, entretanto podem ser preteridas em função de uma “forma regressional” de outra UM porém com probabilidade maior de ocorrência, sobrepondo-se a UM a qual pretendia-se representar diminuindo a acurácia da classe de solo correspondente. Isto significa que quanto maior é a probabilidade de ocorrência de determinada UM maiores são as chances da UM estimada assemelhar-se a UM original.

As áreas que apresentaram baixa probabilidade de ocorrência foram as que não se pode estimar após a finalização do mapa estimado, estas áreas foram as unidades Aluvial, Sxd e CHa.

À semelhança do que ocorreu no Caso de Estudo 1, verificou-se visualmente uma grande correspondência entre o mapa de solos final estimado com o mapa original (Figura 19). Esta correspondência foi matematicamente confirmada na estatística da matriz de erros da área (Tabela 9). Nesta, verifica-se uma correspondência geral pela AG de 58,7%, maior, portanto, do que verificado no Caso de Estudo 1. A acurácia pelo coeficiente Kappa de Cohen foi de 41,7%, também superior ao Caso de Estudo 1. Este maior valor nas duas acurácias provavelmente reflete o fato de haver classes de solo que ocupam grandes extensões da área e estes estarem ocupando posições totalmente distintas na paisagem, como topos de morro versus áreas planas, por exemplo.

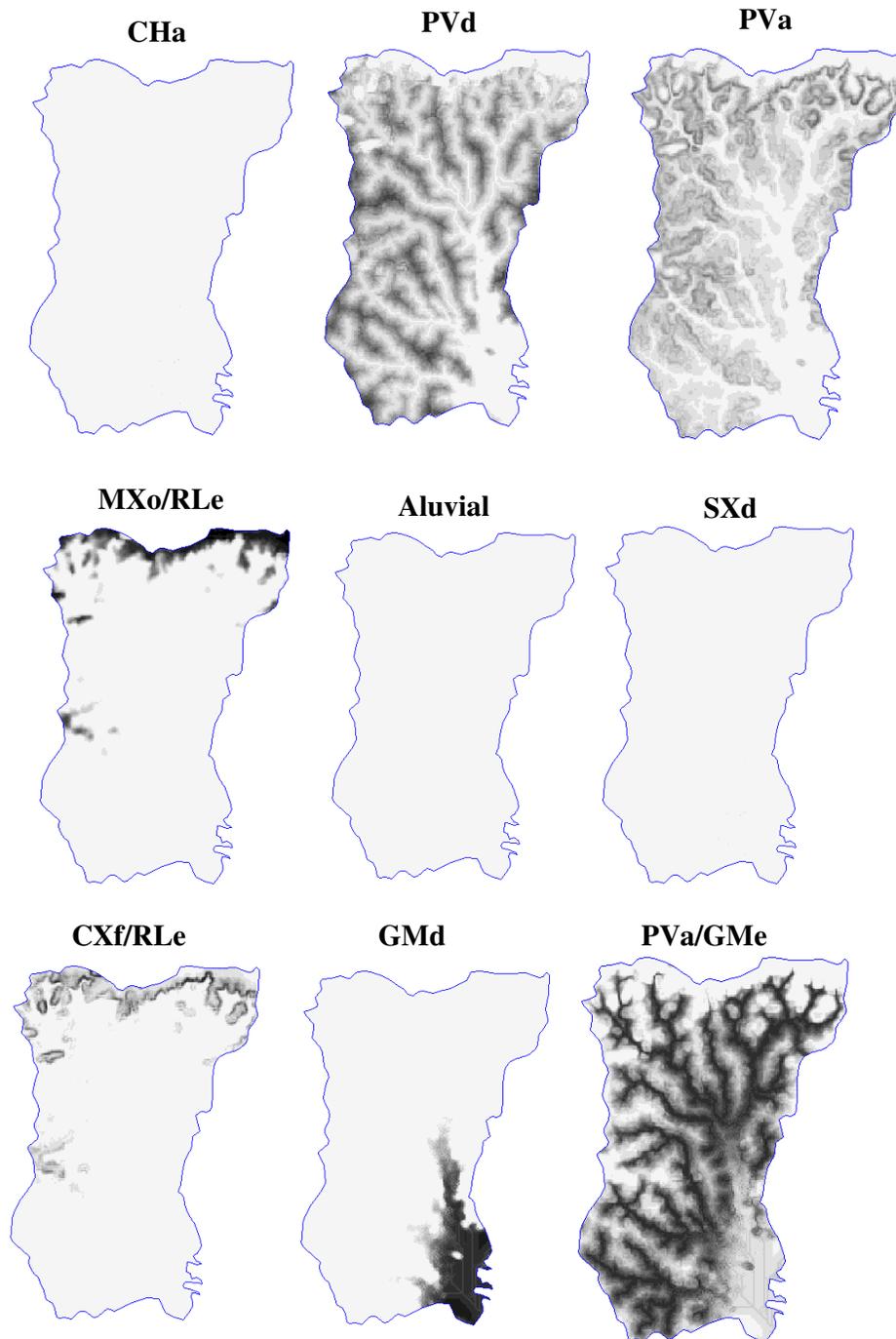


Figura 18. Probabilidade de ocorrência das unidades de mapeamento de solos na região da Bacia Hidrográfica do Arroio Portão, RS. A escala apresenta-se em tons de cinza, variando de 0% (branco) a probabilidade de 100% (preto).

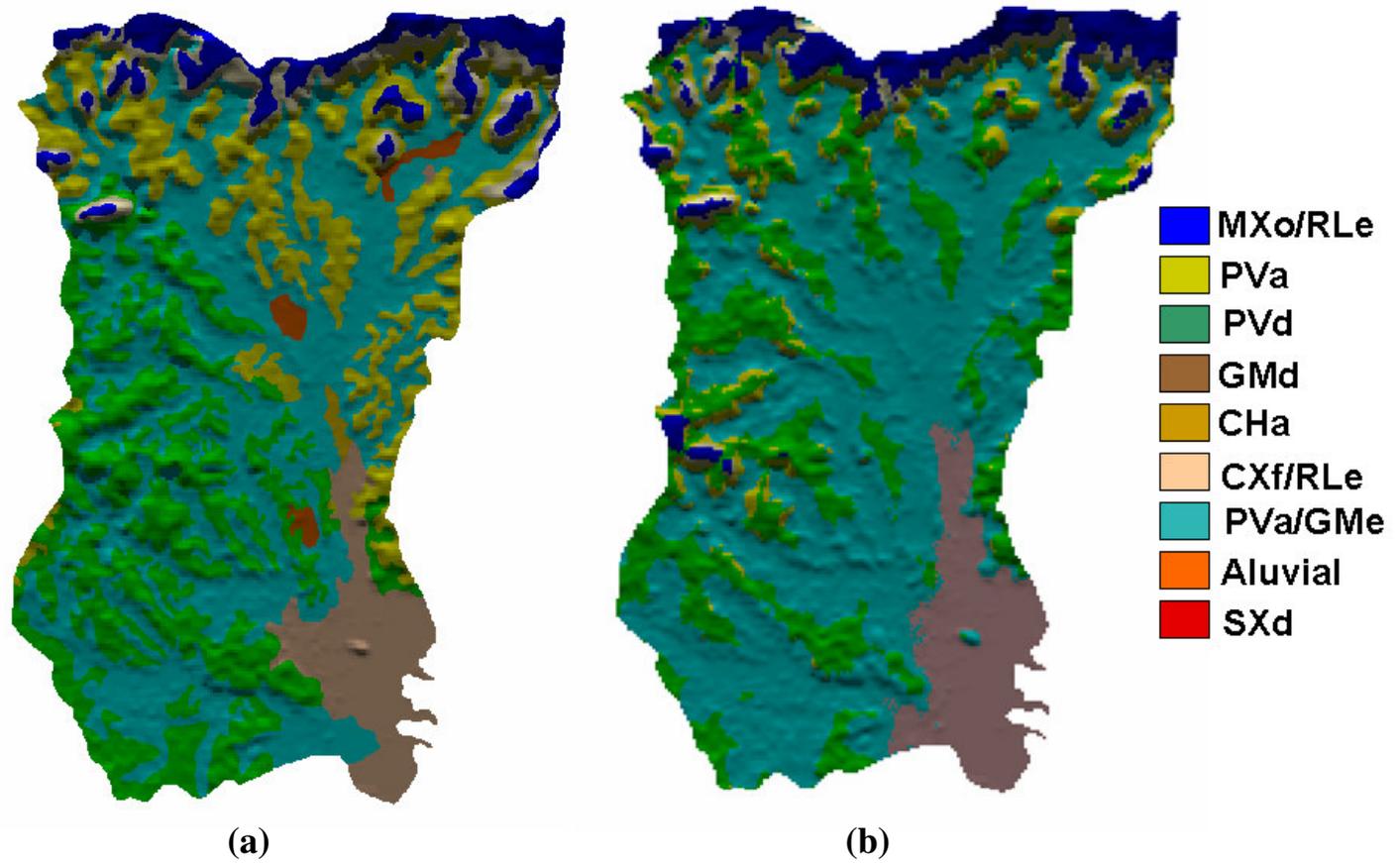


Figura 19. Mapas de solos original (a) e estimado (b) por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS.

Tabela 9. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS

Mapa estimado	Mapa original									Total	AU
	Aluvial	PVd	PVa	MXo/RLe	SXd	CHa	CXf/RLe	GMd	PVa/GMe		
Aluvial	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
PVd	0,0	10,1	6,3	0,3	0,0	0,0	0,3	0,0	4,6	21,7	46,7
PVa	0,0	1,5	2,6	0,4	0,0	0,0	1,0	0,0	1,4	6,9	37,8
MXo/RLe	0,0	0,4	0,2	4,6	0,0	0,0	1,3	0,0	0,0	6,6	70,1
SXd	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CHa	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CXf/RLe	0,0	0,2	0,6	0,6	0,0	0,0	1,5	0,0	0,2	3,0	50,6
GMd	0,0	0,3	0,1	0,0	0,0	0,1	0,0	6,7	1,9	9,1	74,3
PVa/GMe	0,4	10,8	6,6	0,0	0,6	0,2	0,1	0,8	33,1	52,7	62,8
Total	0,4	23,4	16,3	5,9	0,6	0,3	4,3	7,6	41,2	100,0	
AM	0,0	43,4	16,1	78,7	0,0	0,0	34,9	88,8	80,3		

Acurácia geral = 58,7%, Kappa = 41,7%

A unidade GMd, que ocupa 8,4% da área total, foi a que apresentou os maiores valores de AM (88,8%) seguida da unidade MXo/RLe (78,7%) sendo que esta ocupa 6,0% da área total. Os menores valores de AM foram observados em PVa (16%) e CXf/RLe (39,9%), estas UM ocupam respectivamente áreas correspondentes a 15,7% e a 5,8% da área total. Em relação à AU, novamente nas unidades GMd e MXo/RLe foram observados os maiores valores sendo de 74,3% e 70,1%, respectivamente. Da mesma forma, os menores valores de AU foram observados em PVa (37,8%), ocorrendo uma das menores AU também para PVd (46,7%) que ocupa 21% da área.

Observa-se que ocorre uma grande dificuldade de separação entre as unidades PVa e PVd, sendo que a probabilidade de ocorrência da UM PVd suplanta em muito UM PVa em quase toda a sua extensão, este conflito provavelmente justifica-se pelo fato que ambas as UM pertencem a classe dos Argissolos, tem o mesmo tipo de relevo e mesmo material de origem. Portanto são bastante homogêneas entre si em relação à variação dos atributos do terreno empregados neste estudo. O fato de haver uma prevalência da UM PVd, provavelmente está no fato desta ter uma extensão maior ocupando 21% da área enquanto a UM PVa tem 15,7%, ao ter uma extensão maior PVd recebeu mais pontos de amostragem na distribuição randômica (porém proporcional à área), tendo maior esforço de amostragem, sua probabilidade aumentou em relação a PVa.

Novamente utilizando-se da estratégia de simplificação de legenda para verificar seus efeitos sobre a acurácia do mapa estimado, as UM PVa e PVd foram agrupadas em uma única UM, as unidades Aluvial, SXd e CHa foram incorporadas como inclusões à unidade PVa/GMe, sendo que a nova legenda ficou composta pelas UM MXo/RLe, PVd/PVa, GMd, CXf/RLe e PVa/GMe. Na Figura 20, verificamos uma maior correspondência visual entre o mapa original simplificado e o mapa estimado, do que nos mapas sem simplificação de legenda. Ao analisarmos a matriz de erros da Tabela 8, verificamos uma AG de 67,7% e um Kappa de 50,3%, em relação à situação original houve um incremento de 15% e 20% em relação a AG e ao Kappa, respectivamente. Em relação à AU e a AM,

verifica-se que quando separadas as unidades PVd e PVa tem AU de 46,7% e 37,8%, respectivamente, ao juntar as unidades sua AU passa a ser de 71,7%. As unidades PVd e PVa tiveram isoladamente uma AM de 43,4% e 16,1%, respectivamente, e 51,8% ao serem agrupadas. Isso significa que apesar de perderem detalhamento, ao serem agrupadas, estas duas unidades elevaram significativamente a AU.

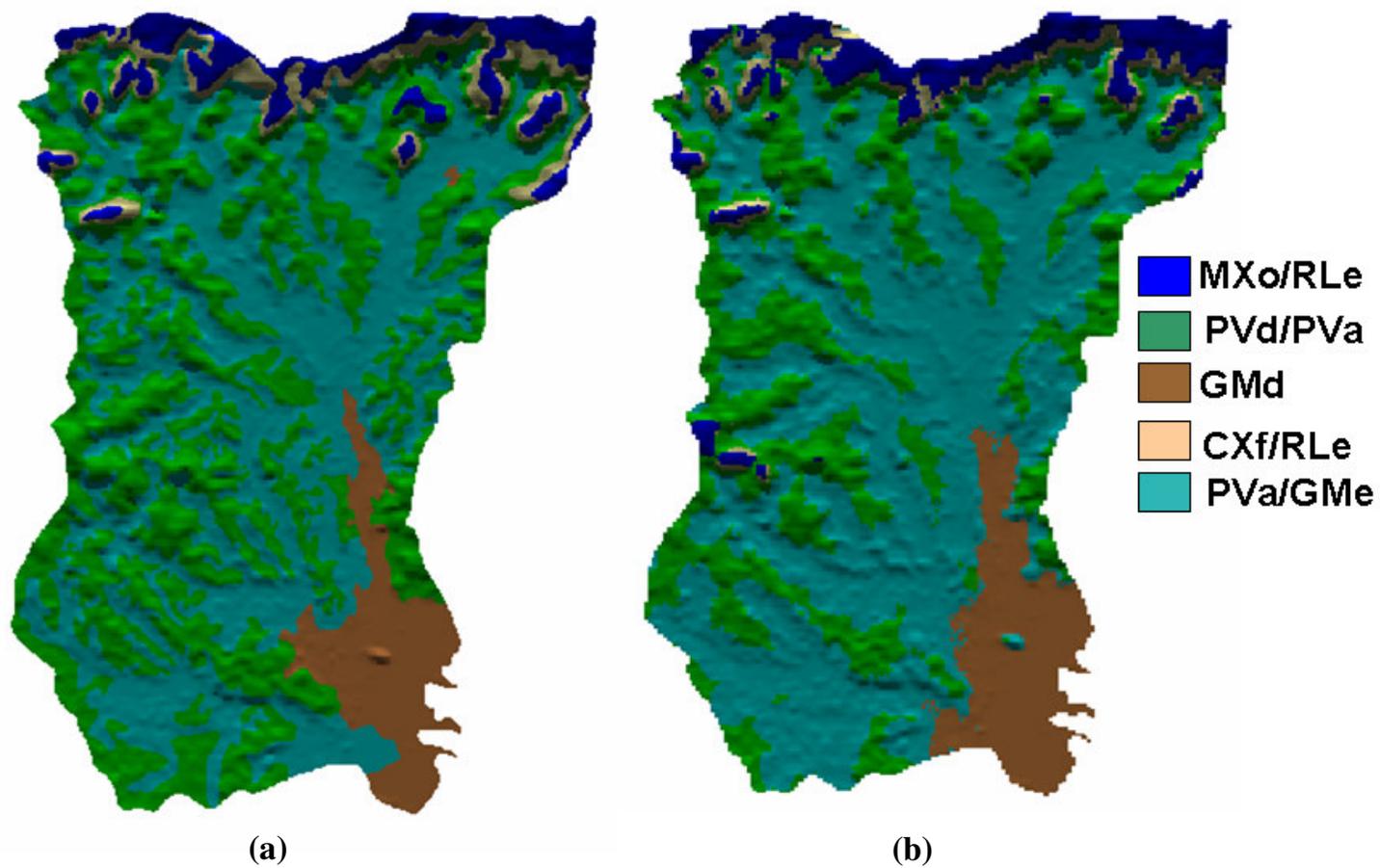


Figura 20. Mapa de solos, com legenda simplificada, original (a) e estimado (b) por regressões logísticas múltiplas na região da Bacia Hidrográfica do Arroio Portão, RS.

Tabela 10. Matriz de erro entre o mapa original e o mapa estimado por regressões logísticas múltiplas na Bacia Hidrográfica do Arroio Portão, RS, utilizando legenda simplificada

Mapa estimado	Mapa original					Total	AU
	PVd/PVa	MXo/RLe	CXf/RLe	GMd	PVa/GMe		
PVd/PVa	20,5	0,7	1,3	0,0	6,0	28,5	71,7
MXo/RLe	0,6	4,6	1,4	0,0	0,0	6,6	70,2
CXf/RLe	0,7	0,6	1,5	0,0	0,2	3,0	50,7
GMd	0,4	0,0	0,0	6,7	1,9	9,0	74,9
PVa/GMe	17,4	0,0	0,1	1,0	34,3	52,9	64,9
Total	39,5	5,9	4,3	7,8	42,4	100,0	
AM	51,8	78,9	34,9	86,0	80,8		

Acurácia geral = 67,7%, Kappa = 50,3%

5.1.3. Caso de Estudo 3: Área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas, Município de Viamão, RS

A distribuição espacial dos atributos de terreno para o Caso de Estudo 3, pode ser verificada na Figura 21.

As variáveis que melhor explicaram a distribuição e ocorrência através do cálculo de probabilidades por regressões logísticas múltiplas foram a elevação, a declividade, o IUT e a curvatura, cujos parâmetros da equação podem ser visualizados na Tabela 11. Como verificado anteriormente com o Caso de Estudo 1, o IUT volta a compor o conjunto de regressões logísticas. Valores elevados de IUT representam áreas com convergência de água, caracterizadas pela estagnação de água por longos períodos, isto é um indicativo de áreas de baixada e planas.

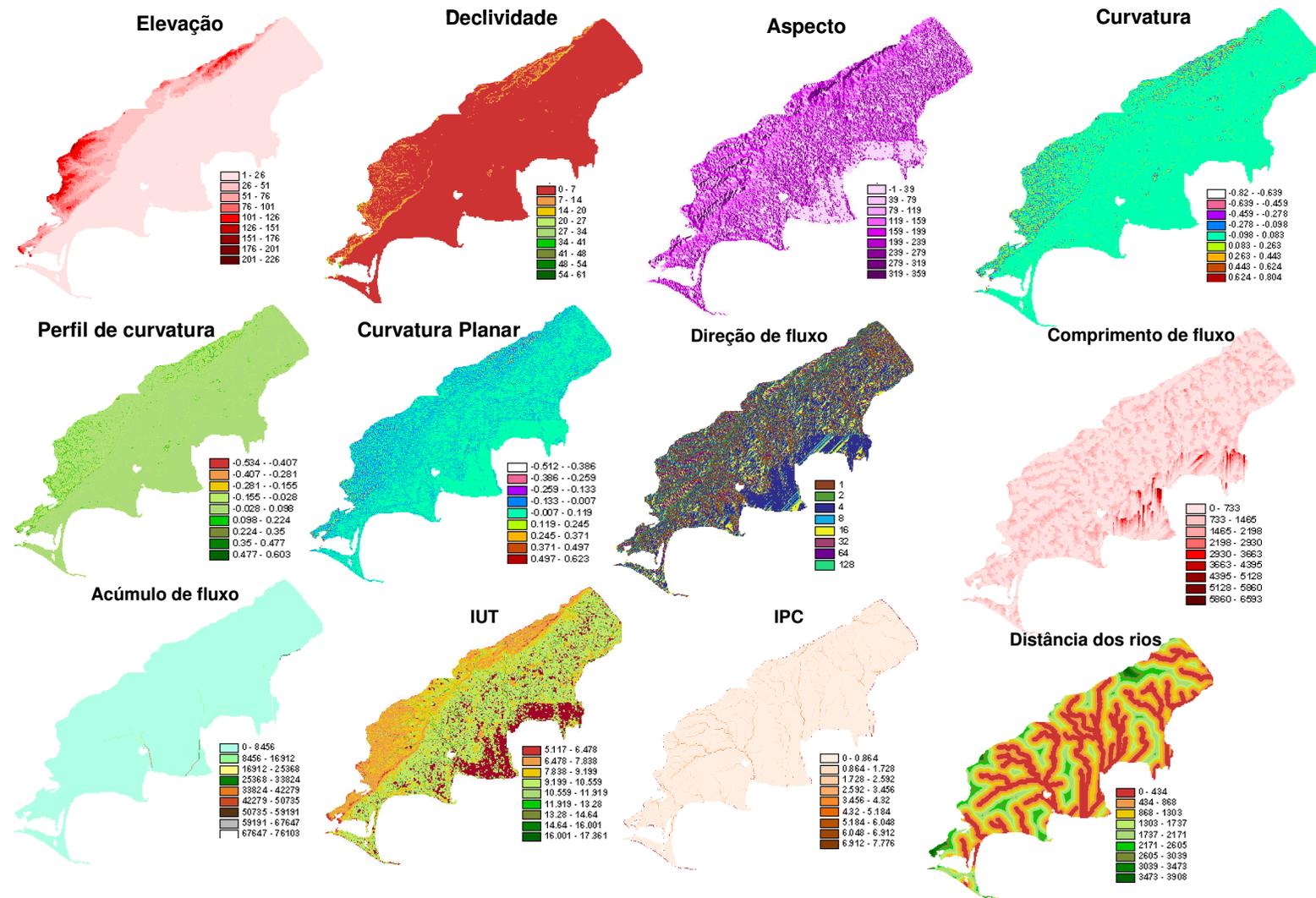


Figura 21. Variação espacial dos atributos de terreno do Caso de Estudo 3, região de Viamão, RS.

Tabela 11. Parâmetros e coeficientes de modelo (*função logit*) empregado na predição da ocorrência de classes de solos na região de Viamão, RS

Atributos do modelo	Modelos de probabilidade								
	PVAd2/ SGe	PVAd2/ OXy	PVAd2/ RQo	PVAd2/ GMd	PVAd2/ SGd- OXy	PVAd2/ PVAd1	PVAd2/ FXa- OXy- RQg	PVAd2/ OXy- RQg	PVAd2/ SGd- FXd
Constante	-3.936	-0,288	-3.620	-0,268	-4.699	-1.998	-3,628	-2,328	-7,836
Índice de Umidade Topográfica	-0,165	-0,095	-0,001	-0,057	0,017	0,031	-0,392	0,012	0,3493
Curvatura, m	-13.169	0,630	-8,921	0,729	-0,948	-5,690	-14,565	8,305	-8,832
Declividade, %	0,246	0,172	0,268	0,766	0,765	0,189	0,4026	0,446	0,4077
Elevação, m	0,454	0,053	0,339	0,097	0,028	0,250	0,4860	0,333	-0,053
Log-likelihood = 6.325, G = 4..632, P-Value = 0.000									

Na figura 22 observa-se a distribuição da probabilidade de ocorrência de cada UM, verifica-se que houve uma baixa probabilidade em todas as UM e certa homogeneidade entre as UM, provavelmente em virtude da homogeneidade das características de estudo. Neste Caso de Estudo, já na comparação visual verifica-se uma baixa correspondência entre o mapa estimado e o mapa original (Figura 23), o que de fato é verificado na matriz de erros apresentada na Tabela 12.

Na Tabela 12, observa-se que as unidades SGd, OXy, GMd, OXy-RQg e SGd-FXd não foram estimadas, provavelmente por apresentaram algum tipo de semelhança geomorfológica com outras unidades do modelo mais representativas em termos de área ocupada o que fez com que a probabilidade de ocorrência destas unidades fossem reduzidas. Outrossim, ao verificar-se os atributos de terreno, observa-se certa divisão em duas grandes partes, uma plana e outra de elevação moderada, e dentro destas áreas observa-se certo grau de homogeneidade dos atributos, o que complica qualquer modelo de estimativa que se baseia, essencialmente, na variabilidade destes e nas correspondências entre geoformas de terreno com as pedoformas do mapa de solos, o que neste caso de estudo não ocorreu.

Estas constatações refletiram-se no baixo valor de AG, que foi de 40,8% com um coeficiente de Kappa de 31,8%, dessa forma o modelo regressional com base nos atributos de terreno selecionados, que foi o melhor dentre a combinação dos 12 atributos, não se adequou para explicar a distribuição e ocorrência das unidades de mapeamento, o que provavelmente, como explicado anteriormente, decorreu-se de uma semelhança geomorfológica entre as UM.

Entretanto, valores relativamente altos de acurácia foram verificados para algumas UM, na AU verificou-se que as UM PVAd2 que ocupa 6,3% da área e PVAd1 que ocupa 17,2% da área tiveram valores de AU de 71,8% e 77,6%, respectivamente. Para a AM, observou-se altos valores para SGd-OXy (92,7%), que representa 18,7% da área total e PVAd1 (84,3%), que ocupa 17,2% da área. Estes valores elevados para as UM mencionadas indica haver um alto grau de dependência da classe de solo formada com as variáveis de modelagem geomorfológica e hidrológica selecionadas na área para compor o modelo.

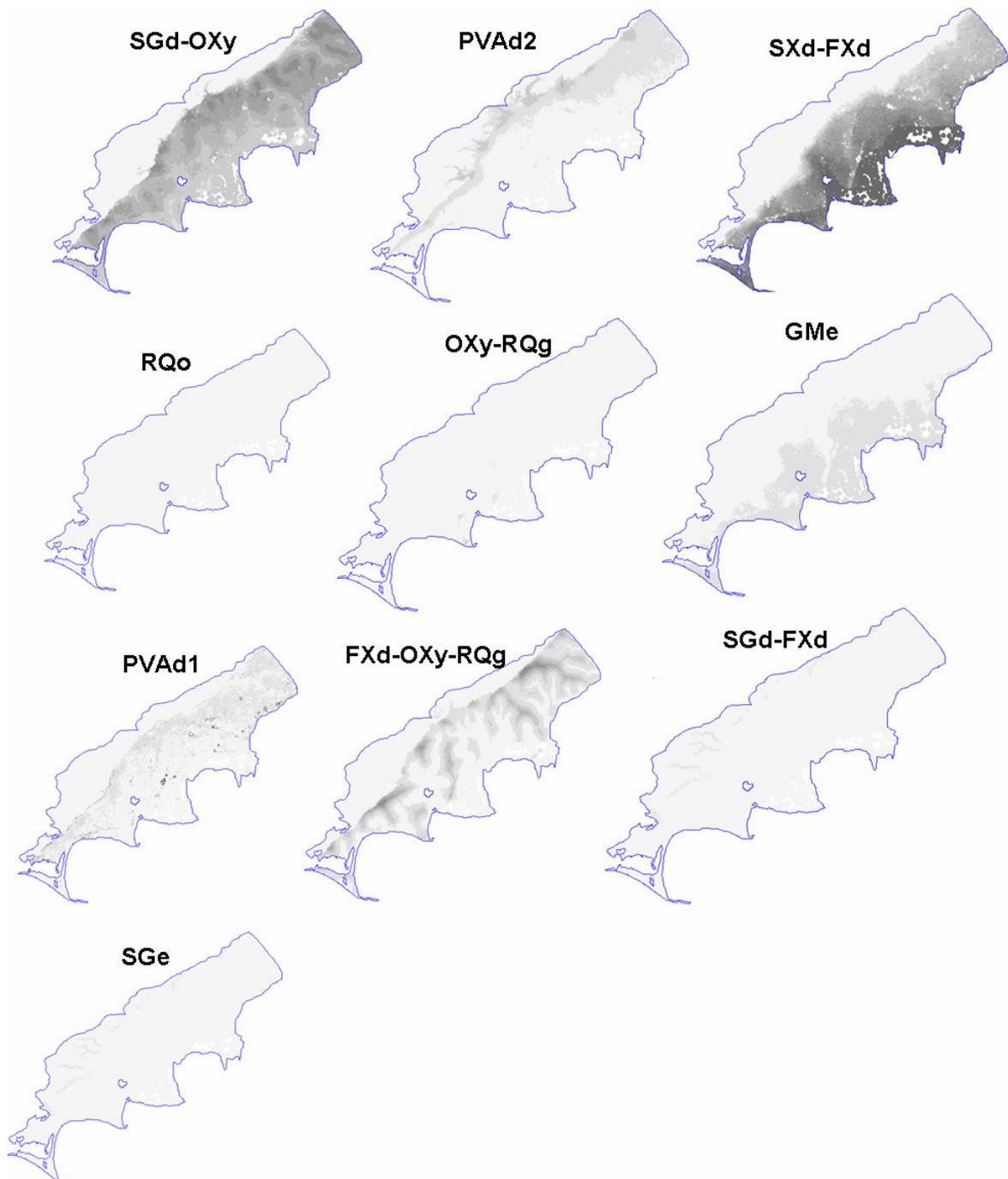


Figura 22. Probabilidade de ocorrência das unidades de mapeamento de solos na região de Viamão, RS. A escala apresenta-se em tons de cinza, variando de 0% (branco) a probabilidade de 100% (preto).

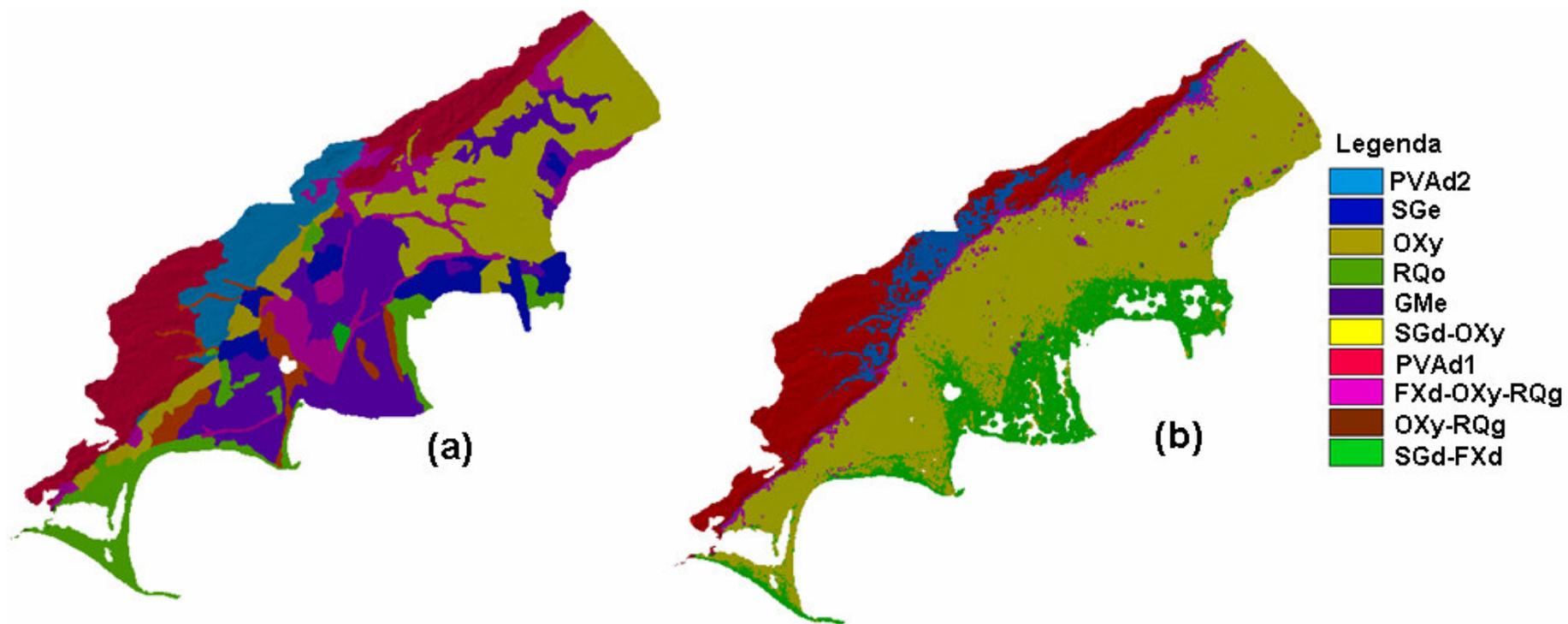


Figura 23. Mapas de solos com legenda simplificada, original (a) e estimado (b) por regressões logísticas múltiplas na região de Viamão, RS.

Tabela 12. Matriz de erro entre a legenda do mapa original e a legenda do mapa estimado por regressões logísticas múltiplas na região de Viamão, RS

Mapa estimado	Mapa original										Total	AU
	PVAd2	SGd	OXY	RQo	GMd	SGd-OXY	PVAd1	FXa-OXY-RQg	OXY-RQg	SGd-FXd		
PVAd2	3,9	0,0	0,1	0,0	0,0	0,0	1,0	0,4	0,1	0,0	5,5	71,8
SGd	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
OXY	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
RQo	0,0	2,9	0,0	3,1	5,3	1,0	0,0	1,5	1,5	0,0	15,4	34,5
GMd	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
SGd-OXY	0,4	2,3	5,3	5,0	14,1	18,3	0,8	6,5	1,7	0,3	54,7	33,5
PVAd1	3,3	0,0	0,1	0,1	0,0	0,0	15,8	0,7	0,4	0,0	20,4	77,6
FXa-OXY-RQg	0,7	0,1	0,5	0,2	0,1	0,4	1,2	0,9	0,0	0,0	4,0	22,7
OXY-RQg	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
SGd-FXd	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Total	8,3	5,3	6,0	8,4	19,5	19,8	18,8	10,0	3,7	0,3	100,0	
AM	47,7	0,0	0,0	36,8	0,0	92,7	84,3	9,1	0,0	0,0		

Acurácia geral =40,8%, Kappa = 31,8%

5.2. Resultados e discussão geral

5.2.1. Atributos do terreno empregados no estudo

Dos 12 atributos utilizados para compor os modelos de regressões logísticas, verificamos que alguns poucos são selecionados para compor os modelos dentre os diferentes Casos de Estudo, também verificamos que alguns destes tendem a se repetir. Estas variáveis mais freqüentes nos modelos foram a elevação, as curvaturas do declive e o IUT. Isto vem em concordância com vários casos citados na literatura, onde a seleção de variáveis como a declividade, curvatura do declive e elevação estão presentes em vários modelos de predição de classes de solos (Zhu et al., 2001; McBratney et al., 2003). A elevação está relacionada com a profundidade dos horizontes do solo bem como com o acúmulo de matéria orgânica; já a declividade se relaciona com a suscetibilidade erosional do solo, o que invariavelmente está envolvido com a profundidade do horizonte A e com as concentrações de matéria orgânica; por sua vez, os padrões de curvaturas do terreno estão relacionadas com a drenagem do solo e com a profundidade dos solos bem como com os processos erosionais e, conseqüentemente, com transporte de sedimentos para as cotas mais baixas. Fatores como suscetibilidade à erosão, espessura do horizonte A, conteúdo de matéria orgânica e profundidade do solo são parâmetros chave no processo de identificação e classificação de classes de solos (Embrapa, 1999).

Recentemente o IUT começou a surgir em alguns modelos de predição da ocorrência e distribuição de solos (Hengl & Rossiter, 2003), sendo que este atributo reflete a tendência de acúmulo de água em qualquer ponto da paisagem. Este índice, apesar de adimensional, reflete importantes propriedades relacionadas com a interação entre variáveis hidrológicas e topográficas. Algumas relações têm sido estabelecidas entre o IUT com fatores importantes na pedogênese como profundidade do horizonte A, conteúdo de matéria orgânica, textura do solo e suscetibilidade à erosão (Florinsky et al., 2002).

5.2.2. Acurácia dos mapas

Segundo Hengl & Rossiter (2003), dentro da teoria das estatísticas de mapas, um modelo considerado ótimo deveria apresentar uma boa correspondência com a realidade, o que seria representado por uma AG = 80% e Kappa = 70%. De fato, até o presente momento, em literatura, tais valores não foram satisfatoriamente alcançados através do mapeamento supervisionado de solos quanto se compara um mapa de solos pré-existente, mapeado tradicionalmente, com um mapa estimado baseado neste sendo a variável resposta.

Alguns autores têm alcançado níveis próximos ou superiores a estes valores quando realizam inspeções a campo, ou quanto fazem correlações entre as classes estimada com a original através apenas dos pontos de amostragem selecionados e não de toda a área, como o caso de Hengl & Rossiter (2003) que observaram AG = 90,2% (Kappa = 89,3%) utilizando-se apenas dos pontos de amostragem com uma simplificação de legenda. Quando consideraram toda a área do mapa, os autores verificaram AG = 45,3% (Kappa = 42,6%), enquanto que ao dividirem a área de estudo entre montanhosa e plana, os autores verificaram valores de AG = 58,1% (Kappa = 51,6%) para a área montanhosa e AG = 39,1% (Kappa = 34,4%) para área plana.

Outrossim, conforme visto no trabalho de Hengl & Rossiter (2003), áreas de terreno acidentado, montanhosas, de paisagem heterogênea e, conseqüentemente, contempladas com classes de solos com processos pedogenéticos bastante distintos têm valores de acurácia mais altos que áreas planas e de características mais homogêneas quanto aos atributos do terreno e pedogênese de solos.

As observações de acurácia, presentes neste trabalho de pesquisa, estão na mesma escala de magnitude que outras observações quando se empregou regressões logísticas tais como, por exemplo, Campling et al. (2002), que empregaram regressões logísticas múltiplas na predição da probabilidade

espacial da ocorrência de uma única característica (drenagem) do solo, usando atributos derivados do modelo numérico de terreno e índices de vegetação através de imagens LANDSAT-5 em região tropical úmida da Nigéria, com concordância entre os mapas de solos (original e estimado) variando entre 95% e 65% à medida que se aumentava o grau de detalhamento em função da separação das classes. Da mesma forma, Bailey et al. (2003) usaram regressões logísticas para predição de classes de solos com acurácias variando entre 47% e 65%.

Estas metodologias de estatísticas de mapas em que se comparam mapas de solos tradicionais com estimados são poucas utilizadas, em parte por acreditar-se que o modelo deve adequar-se mais a realidade de campo do que ao mapa de solos tradicional a que pretende representar e que possui invariavelmente certo grau de erro que lhe é inerente, e em parte por que os valores de AG e Kappa apresentam-se sempre relativamente baixos. Em trabalhos, por exemplo, de Zhu et al. (2001) e Scull et al. (2005), os autores omitem as informações referentes à AG e Kappa entre o mapa tradicional e o mapa estimado, entretanto apresentam ambos os mapas lado a lado para ilustrar a semelhança. Entretanto, uma simples comparação visual destes mapas é suficiente para verificar que estas acurácias são baixas. Uma hipótese é que isto deva-se: a (i) erros inerentes no mapa de solos tradicional; (ii) erros decorrentes da fonte de dados, principalmente o MNT; (iii) inadequação do método preditivo e/ou (iv) um fator aleatoriedade complicador da racionalização do mapeamento.

5.2.3. Vantagens da metodologia

A utilização de regressões logísticas múltiplas têm a vantagem de fácil compreensão por remeter ao conceito de regressão já amplamente popularizado na comunidade científica. De fácil compreensão e também operação, as regressões logísticas têm produzido resultados em termos de acurácia semelhante a outros métodos como já discutido anteriormente.

Outra grande vantagem da metodologia é o fato de estar baseada no cálculo de probabilidades de ocorrência e distribuição, dessa forma a acurácia pode ser relacionada com a probabilidade.

5.2.4. Limitações da metodologia

A primeira grande limitação da metodologia é que esta é dependente da pré-existência e disponibilidade de um mapa de solos produzido de forma tradicional em escala adequada com a fonte de dados ambientais empregados, já que o método não tem a pretensão de gerar mapas sem ser a partir da consideração de uma realidade de campo já definida (mapeada). Um dos grandes objetivos do mapeamento supervisionado de solos é que, ao se estabelecer relações matemáticas entre as classes de solo e os atributos ambientais no qual este está inserido, poder se extrapolar a informação sobre a distribuição espacial dos solos para além dos limites da área do mapa de solos original, para as áreas vizinhas, que poderiam ser de tamanho de até 10 vezes o tamanho da área mapeada tradicionalmente. Dessa forma, pelo menos 1/10 da área mapeada, incluindo-se aquelas mapeadas de forma tradicional e aquelas mapeadas com mapeamento supervisionado, deve ser previamente mapeada de forma tradicional (Hengl & Rossiter, 2003).

Outra limitação está na resolução e qualidade do MNT, que para a América Latina é de 90m (pixel = 0,81 ha) enquanto nos EUA e Europa chega a 30m (0,09 ha) e até a 10m (0,01 ha) conferindo um maior grau de detalhamento do terreno e, obviamente, das variáveis do terreno dele derivadas. Provavelmente, num futuro próximo, novos levantamentos serão realizados na América Latina objetivando produzir MNT mais detalhados. Entretanto, mesmo MNT detalhados possuem erros inerentes, como por exemplo não conseguirem diferenciar a elevação natural do terreno da elevação de construções e de árvores. Uma alternativa promissora para MNT pouco detalhados é a interpolação de pontos vizinhos por krigagem, gerando maior número de pontos com capacidade para detalhar o MNT em até 10 vezes mais através de um estudo geoestatístico da variabilidade espacial da elevação da área.

Por fim, a acurácia atualmente alcançada por este e por vários outros métodos de mapeamento supervisionado de solos, ainda limita o uso desta

metodologia para uso do mapeador de solos e do usuário do mapa de solos onde reside aplicação mais útil do mapa supervisionado.

5.2.5. Perspectivas da metodologia

O objetivo final da metodologia, tanto de regressões logísticas múltiplas quanto de qualquer outro método de mapeamento supervisionado de solos, é produzir um mapa de solos digital a partir de um mapa de solos pré-existente mapeado de forma tradicional e de atributos ambientais, no caso deste estudo, com atributos topográficos e hidrográficos do terreno. Com base nisso, apesar das limitações inerentes, as perspectivas da metodologia são promissoras, com chances de extrapolar os resultados de levantamentos de solos pelo método tradicional.

O suporte dos modelos preditivos com maiores fontes de atributos ambientais, embora adequado por contemplar mais fatores da equação de Jenny (1941), não tem refletido grandes efeitos no sentido de aumentar a acurácia dos mapas estimados. Entretanto, a possibilidade de empregar fontes de dados como geologia, índices de vegetação e clima quando disponível não pode ser descartada em futuros trabalhos envolvendo o mapeamento supervisionado de solos.

Avaliando os resultados de acurácia deste trabalho e confrontando com as observações de literatura, observa-se que o cerne da questão não reside em números de atributos ou em diversas fontes destes atributos, mas sim na metodologia empregada. Segundo McBratney et al. (2003), o sucesso do mapeamento supervisionado de solos residirá no estabelecimento de uma harmoniosa interação entre o conhecimento subjetivo pedológico, que deve ser matematicamente racionalizado, com ferramentas matemáticas complexas como as regressões logísticas múltiplas. Nesta interação reside o grande desafio dos pedométricos, desafio este que ao ser alcançado irá popularizar o mapeamento de solos e expandir a informação sobre os tipos de solos para proporções nunca antes alcançadas.

6. CONCLUSÕES

As regressões logísticas múltiplas têm potencial em prever a distribuição e ocorrência de classes de solos em estudos de mapeamento supervisionado de solos.

A eficiência do método preditivo depende da qualidade e do tipo das fontes de dados utilizadas no modelo, sendo que estas devem ter interação com a pedogênese destes solos e ter correspondência morfológica com as formas das delimitações da unidades de mapeamento em um mapa de solos.

Com base neste estudo verifica-se uma melhor adequação do método com o mapa de escala 1:50.000 (Bacia Hidrográfica do Arroio Portão), subdividido em UM que ocupam grandes extensões do mapa e são bastante diferenciadas entre si quanto aos seus atributos de terreno geomorfológico e hidrológico.

A estratégia de simplificar a legenda não surtiu efeitos significativos quanto a incrementar a acurácia geral e a acurácia pelo coeficiente Kappa. Entretanto, mostrou-se eficiente em elevar a acurácia do usuário e acurácia do mapeador quando unidades de mapeamento conflitantes entre si quanto à probabilidade de distribuição e ocorrência e morfológicamente semelhantes são agrupadas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

BAILEY, N.; CLEMENTS, T.; LEE, J.T.; THOMPSON, S. Modelling soil series data to facilitate targeted habitat restoration: a polytomous logistic regression approach. **Journal of Environmental Management**, Amsterdam, v. 67, pp. 395-407, 2003.

BISHOP, T.F.A.; MCBRATNEY, A.B. A comparison of prediction methods for the creation of field-extent soil property maps. **Geoderma**, Amsterdam, v. 103, p.149–160, 2001.

BÖNISCH, S.; LOPES ASSAD, M. L.; MONTEIRO, A.M.V. ; CÂMARA, G. Representação e propagação de incertezas em dados de solo. II – Atributos numéricos. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 28, pp. 33-47, 2004.

BRASIL, Ministério da Agricultura. **Levantamento de reconhecimento dos solos do estado do Rio Grande do Sul**. Recife, 1973. 431p. (Boletim Técnico, 30).

BUI, E.N. Soil survey as a knowledge system. **Geoderma**, Amsterdam, v. 120, p. 17-26, 2004.

BUI, E.N.; MORAN, C.J.. A strategy to fill gaps in soil survey over large spatial extents: an example from the Murray– Darling basin of Australia. **Geoderma**, Amsterdam, v.111, p. 21–44, 2003.

CAMPLING, P.; GOBIN, A.; FEYEN, J. Logistic Modeling to Spatially Predict the Probability of Soil Drainage Classes. **Soil Science Society of America Journal**, Madison, v. 66, pp.1390–1401, 2002.

CHANG, D.H.; ISLAM, S. Estimation of soil physical properties using remote sensing and artificial neural network. **Remote Sensing of Environment**, v. 74, p.534– 544, 2000.

COMLEY, J.W.; DOWE, D.L.. Minimum Message Length, MDL and Generalised Bayesian Networks with Asymmetric Languages. In: GRUNWALD, P.; PITT, M.A.; MYUNG, I.J. (eds). **Advances in Minimum Description Length: Theory and Applications**, Ohio, M.I.T. Press, 2005, p. 265-294.

COOK, S.E.; CORNER, R.; GREALISH, G.J.; GESSLER, P.E.; CHARTRES, C.J. A rule-based system to map soil properties. **Soil Science Society of America Journal**, Madison, v.60, p.1893–1900, 1996.

DALMOLIN, R. S. D.; KLAMT, E.; PEDRON, F.A.; AZEVEDO, A.C. Relação entre as características e o uso das informações de levantamentos de solos de diferentes escalas. **Ciência Rural**, Santa Maria, v.34, n.5, p. 1479-1486, 2004.

DEMATTÊ, J.A.M.; GENÚ, A.M.; FIORIO, P.R.; ORTIZ, J.L.; MAZZA, J.A.; HUDSON, C.A. Comparação entre mapas de solos obtidos por sensoriamento remoto espectral e pelo método convencional. **Pesquisa agropecuária brasileira**, Brasília, v.39, n.12, p.1219-1229, dez. 2004.

DOBERMANN, A.; OBERTHUR, T. Fuzzy mapping of soil fertility—a case study on irrigated riceland in the Philippines. **Geoderma**, Amsterdam, v.77, p.317– 339, 1997.

EMBRAPA. Centro Nacional de Pesquisa em Solos. **Procedimentos Normativos de Levantamentos Pedológicos**. Brasília : Embrapa, 1995. 116p.

ESRI. Arcview, versão 3.2. Environmental Research Institute: Redlands, CA/EUA, 1995.

FLORINSKY, I.V.; EILERS, R.G.; MANNING, G.; FULLER, L.G. Prediction of soil properties by digital terrain modelling. **Environmental Modelling and Software**, v. 17, p. 295–311, 2002.

GERSHENFELD, N. The Nature of Mathematical Modelling. Cambridge: Cambridge University Press, UK, 1999.

GESSLER, P.E.; MOORE, I.D.; MCKENZIE, N.J.; RYAN, P.J. Soil–landscape modelling and spatial prediction of soil attributes. **International Journal of Geographical Information Systems**, v. 9, p. 421– 432, 1995.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. The elements of statistical learning: data mining, inference and prediction. Springer Series in Statistics. New York: Springer-Verlag, 2001.

HENGL, T. **Pedometric mapping bridging the gaps between conventional and pedometric approaches**, Wageningen: International Institute for Geo-Information Science and Earth Observation, 2003. 214 f. (Ph.D. thesis).

HENGL, T.; ROSSITER, D.G. Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. **Soil Science Society of America Journal**, v. 67, p. 1810–1822, 2003.

HENGL, T.; ROSSITER, D.G.; HUSNJAK, S. Mapping soil properties from an existing national soil data set using freely available ancillary data. WORLD CONGRESS OF SOIL SCIENCE, 17, Bangkok, Tailândia, 14- 21 de Agosto, 2002.

HOSMER, D. W.; LEMESHOW, S. **Applied logistic regression**. New York: Wiley: New York, USA, 1989.

IPPOLITI R., G.A.; COSTA, L. M.; SCHAEFER, C. E. G. R.; FERNANDES FILHO, E. I.; GAGGERO, M. R. Análise digital do terreno: ferramenta na identificação de pedoformas em microbacia na região de "Mar de Morros" (MG) **Revista Brasileira de Ciência do Solo**, Viçosa, v.29, n.2, p. 269-276, 2005.

JENNY, H. **Factors of Soil Formation: A System of Quantitative Pedology**. McGraw-Hill: New York, 1941.

JENSON S. K. ; DOMINGUE, J. O. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. **Photogrammetric Engineering and Remote Sensing**, v. 54, n. 11, p. 1593-1600, 1988.

KING, D. ; JAMAGNE, M. ; ARROUAYS, D. ; BORNAND, D. ; FAVROT, J.C. ; HARDY, R. ; LE BAS ; C. STENGEL, P. Inventaire cartographique et surveillance des sols en France. Etat d'avancement et exemples d'utilisation. Montpellier : [s.n.], 1999. p. 215– 228 (Estude et Gestion des Sols 6).

KIRKBY, M.J. Hydrograph modeling strategies. In: PEEL et al. (Ed) Progress in physical and human geography. London: Heinemann, 1975, p. 69-90.

KLAMT, E.; SCHNEIDER, P.; AZEVEDO, A.C.; SABADIN, E.M.; MOBUS, G.; SOARES, J.M.D.; CARPENEDO, V. Solos da Bacia do Arroio Portão: Características, Distribuição e Aptidão de Uso. Porto Alegre: Departamento de Solos da UFRGS, 1992.

KOHONEN, T. Self-organized formation of topologically correct feature maps. **Biological Cybernetics**. V. 43, p.59–69, 1982.

LIMA, Z. M. C.; RIBEIRO, M.R.; LIMA, A.T.O. Utilização de imagens TM/LANDSAT-5 como ferramenta no mapeamento de solos. **Revista brasileira de engenharia agrícola e ambiental**. Campina Grande, v.5, n.3, p. 425-430, 2001.

LAGACHERIE, P.; HOLMES, S. Addressing geographical data errors in a classification tree soil unit prediction. **International Journal of Geographical Information Science**, v.11, p.183–198, 1997.

McBRATNEY, A.B.; MENDONÇA SANTOS, M.L.; MINASNY, B. On digital soil mapping. **Geoderma**, v. 117, p. 3-52, 2003.

MCBRATNEY, A.B.. A theoretical framework for land evaluation. **Geoderma**, v. 72, p. 198-199, 1996.

MINASNY, B., MCBRATNEY, A.B. The neuro-m method for fitting neural network parametric pedotransfer functions. **Soil Science Society of America Journal**, v.66, p.352–361, 2002.

MINITAB Inc. Minitab for Windows, statistical software, 1996.

MOORE, D.M.; LEES, B.G.; DAVEY, S.M. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. **Environmental Management** , v.15, p.59–71, 1991.

MOORE, I.D.; GESSLER, P.E.; NIELSEN, G.A.; PETERSON, G.A. Soil attribute prediction using terrain analysis. **Soil Science Society of America Journal**, v. 57, p. 443-452, 1993.

MORAN, C.J.; BUI, E. Spatial data mining for enhanced soil map modelling. **International Journal of Geographical Information Science**, v. 16, p. 533– 549, 2002.

O'LOUGHLIN, E.M. Saturation regions in catchments and their relations to soil and topographic properties. **Journal of Hydrology**, v. 53, p. 229-246, 1981.

PARK, S.J.; MCSWEENEY, K.; LOWERY, B. Identification of the spatial distribution of soils using a process-based terrain characterization. **Geoderma**, v. 103, p. 249–272, 2001.

REIS, T.E.S.; BARROS, O.N.F.; REIS, L.C. Utilização de sistema de informações geográficas para obtenção das cartas de solo e de declividade do município de Bandeirantes-PR. **Geografia**, Londrina, v. 13, p. 1-17, 2004.

ROSSITER, D.G. **Technical Note: Statistical methods for accuracy assesment of classified thematic maps**. 2004.

SANTOS, M.C.L.; KLAMT, E.; KAMPF, N.; ABRÃO, P.U.R. Levantamento e utilização agrícola dos solos do município de Ibirubá. Porto Alegre: Departamento de Solos da UFRGS, 1970.

SCHNEIDER, P.; MENDES D'ÁVILA, A.L.; BRAIDA, J. A.; SPINELLI, L. F.; JUNGBLUT, M. Levantamento de reconhecimento de média intensidade dos solos da área abrangida entre a margem norte da Lagoa dos Patos e o divisor de águas da Coxilha das Lombas (município de Viamão). Porto Alegre, Departamento de Solos da UFRGS, 1992.

SCULL, P.; FRANKLIN, J.; CHADWICK, O.A.. The application of classification tree analysis to soil type prediction in a desert landscape. **Ecological modeling**, 181, p. 1-15, 2005.

SOMARATNE, S.; SENEVIRATNE, G.; COOMARASWAM, U. Prediction of Soil Organic Carbon across Different Land-use Patterns: A Neural Network Approach. **Soil Science Society for America Journal**, Madison, v. 69, p.1580-1589, 2005.

STEHMAN, S.V.; CZAPLEWSKI, R.L. Design and analysis of thematic map accuracy assessment: fundamental principles. **Remote Sensing of Environment**, v.64, p.331-344, 1998.

TAN, P.J.; DOWE, D.L. MML Inference of Oblique Decision Trees. **Lecture Notes in Artificial Intelligence**, New York, p. 1082-1088, 2004.

WILSON, J.P.; GALLANT, J.C. Secondary topographic attributes. In: WILSON, J.P.; GALLANT, J.C. (Eds.) *Terrain Analysis: Principles and Applications*. Wiley, New York, pp. 87–132, 2000.

WONG, G.Y.; MASON W.M. The hierarchical logistic regression model for multilevel analysis. **Journal of the American Statistical Association**, v. 80, pp. 513-524, 1985.

ZEVENBERGEN, L.W.; THORNE, C.R. Quantitative analysis of land surface topography. Earth Surfer Process Landforms, Ohio, v. 12, pp. 47-56, 1987.

ZHU, A.X.; BAND, L.E.; VERTSEY, R.; DUTTON, B. Derivation of soil properties using a soil land inference model (SoLIM). **Soil Science Society of America Journal**, v.61, pp. 523-533, 1997a.

ZHU, A.X.; BAND, L.E.; DUTTON, B.; NIMLOS, T.J. Automated soil inference under fuzzy logic. **Ecological Modelling**, v.90, p. 123–145.,1997b.

ZHU, A.X. Mapping soil landscape as spatial continua: the neural network approach. **Water Resources Research**, v. 36, p. 663– 677, 2000.

ZHU, A.X.; Hudson, B.; Burt, J.; Lubich, K.; Simonson, D. Soil mapping using GIS, expert knowledge, and fuzzy logic. **Soil Science Society of America Journal**, v. 65, pp. 1463-1472, 2001.