
REFINAMENTO SEMIAUTOMÁTICO DO MAPA GEOMORFOLÓGICO DO RADAMBRASIL POR RECUPERAÇÃO E MODELAGEM DO CONHECIMENTO

Semiautomatic Refining of RADAMBRASIL Geomorphological Map Through Knowledge Recovery and Modeling

Pedro José Farias Fernandes
Laboratório de Geografia Física - Departamento de Geografia - UFF, Niterói.
pjf_fernandes@yahoo.com.br

Raphael e Silva Girão
Laboratório de Estudos Tectônicos - Faculdade de Geologia – UERJ.
raphaelgirao@hotmail.com

Artigo enviado para publicação em 22/05/2018 e aceito em 24/07/2018

DOI: 10.12957/tamoios.2018.34292

RESUMO

O objetivo deste trabalho é o refinamento semiautomático de um mapa geomorfológico regional de pequena escala do projeto RADAMBRASIL utilizando GEOBIA, técnicas de mineração de dados e MDEs oriundos do SRTM. Os dados SRTM provenientes do projeto TOPODATA foram preparados e tratados em ambiente SIG, o que permitiu a análise visual e quantitativa das classes RADAMBRASIL. Após as análises, o conhecimento foi estruturado para proceder com os recursos de análise de imagem baseada em objetos geográficos e de mineração de dados, que viabilizaram o mapeamento geomorfológico para cada nível taxonômico, a partir do uso do algoritmo de segmentação multirresolução. As classificações foram avaliadas com valores Kappa a partir de um mapa geomorfológico do CPRM e do próprio mapa RADAMBRASIL, e podem ser consideradas satisfatórias. Os mapas obtidos para cada nível taxonômico exibiram contornos detalhados, o que indica que a GEOBIA, as técnicas de mineração de dados e a utilização de MDEs podem contribuir na tarefa refinar contornos de mapas antigos e para o mapeamento geomorfológico regional em escalas superiores a 1:1.000.000, além de tornar possível o estabelecimento de um modelo de conhecimento geomorfológico.

Palavras-chave: Mapeamento Geomorfológico; Sensoriamento Remoto; Análise de Imagem Orientada a Objeto Geográfico.

ABSTRACT

The aim of this study is to execute semiautomatic refinement of a small-scale regional geomorphological map of the RADAMBRASIL project using GEOBIA, data mining techniques and DEMs from the SRTM. The SRTM data from the TOPODATA project were prepared and treated in a GIS environment, allowing visual and quantitative analysis of the RADAMBRASIL classes. After analysis, the knowledge was structured to proceed with geographic object based image analysis and data mining, which made the geomorphological mapping for each taxonomic level possible, by using multiresolution segmentation algorithm. The maps were evaluated with Kappa index according to geomorphological maps from CPRM and RADAMBRASIL. Classifications made can be considered satisfactory. The maps obtained for each taxonomic level exhibited detailed outlines of the polygons from the geomorphological classes, which indicated that GEOBIA, data mining techniques and DEMs can be used to refine contours of old maps at regional scales greater than 1:1,000,000. Thus, it is possible the establishment of a geomorphological knowledge model from SRTM data.

Key-Words: Geomorphological mapping; Remote Sensing; Geographic Object Based Image Analysis.

INTRODUÇÃO

O mapeamento geomorfológico particiona a superfície terrestre em unidades a partir de critérios específicos relacionados com o relevo como a morfologia, a gênese, a estrutura, a cronologia, entre outros. O mapeamento de formas de relevo possui importância, por exemplo, para o entendimento dos processos atuantes na superfície terrestre e da evolução da paisagem, levantamento de recursos naturais e identificação de riscos naturais (BISHOP et al., 2012).

Antigos mapeamentos de recursos naturais envolveram técnicas de fotointerpretação que não contavam com as ferramentas computacionais disponíveis atualmente, bem como bases de dados recentes, o que não permitia uma caracterização geomorfométrica em bases quantitativas (BARBOSA, 1973). No Brasil, o mapeamento geomorfológico do projeto RADAMBRASIL deve ser vista neste contexto. O projeto RADAMBRASIL, inicialmente chamado de RADAM (Radar na Amazônia), foi criado em 1971 com o objetivo de levantar os recursos naturais na Amazônia (ROESEL e GODOY, 1974), depois, o projeto foi ampliado para o resto do território nacional e recebeu a denominação RADAMBRASIL (ALLEVATO e PINGARILHO, 1978). O projeto utilizou, principalmente, mosaicos na escala 1:250.000 de imagens de RADAR aerotransportado de visada lateral (SLAR) na banda X, além de imagens orbitais, fotografias aéreas e trabalhos de campo. A interpretação visual desse material permitiu a elaboração de mapas de geologia, geomorfologia, pedologia, vegetação, entre outros, em escala original de 1:250.000, entretanto, publicados na escala 1:1.000.000.

Atualmente, Modelos Digitais de Elevação (MDEs) obtidos por Sensoriamento Remoto (SR) são uma importante fonte de dados topográficos cuja análise em Sistemas de Informação Geográfica (SIG) auxilia no mapeamento geomorfológico (DRAGUT e BLASCHKE, 2006). Os MDEs do Shuttle Radar Topography Mission (SRTM) e suas variáveis geomorfométricas (VG) derivadas são dados, que quando tratados em SIG, servem como auxílio ao mapeamento geomorfológico, principalmente em países de grande extensão territorial como o Brasil, e permitem a caracterização geomorfométrica das classes mapeadas. Além disso, as novas abordagens de processamento digital de imagens, como a análise de imagem baseada em objetos geográfico (*Object-based Image Analysis* - GEOBIA) e as técnicas de mineração de dados constituem em uma importante contribuição para a análise mais automatizada dos dados de sensoriamento remoto (BENZ et al., 2004; HSU et al., 2002).

A GEOBIA se difere das técnicas convencionais de classificação de imagem por não ter como foco as análises baseadas nos *pixels* das imagens, mas tem como unidades de processamento básico os objetos gerados pelos algoritmos de segmentação (BENZ et al., 2004; BAATZ e SCHÄPE, 2000). De acordo com Navulur (2007), em imagens de sensoriamento remoto, um objeto pode ser definido como um agrupamento de *pixels* com propriedades espectral e espacial homogêneas. Um objeto possui, além de valores espectrais, diversos outros atributos, incluindo forma, textura, morfologia, contexto e outros, que podem ser utilizados em análise de imagens. Além disso, há possibilidade de se gerar diferentes níveis de segmentação na abordagem multirresolução e, assim sendo, cada objeto identifica seus objetos vizinhos, os superobjetos e os subobjetos, possibilitando o respeito das bordas de cada um deles nas análises. A partir, principalmente, de tais características, é permitida a estruturação espacial e semântica do conteúdo de uma imagem pela inserção do conhecimento especialista em uma rede hierárquica (BENZ et al., 2004). Essa característica é uma das vantagens da GEOBIA no mapeamento geomorfológico, pois permite a classificação do relevo em diferentes

níveis, adequando-se bem ao mapeamento geomorfológico taxonômico (DRAGUT e BLASCHKE, 2006).

Já a mineração de dados, que, segundo Witten et al. (2011), consiste no processo automático ou semiautomático de identificação de padrões nos dados, permite a identificação dos atributos mais adequados e eficientes para distinguir classes diferentes, baseada em um conjunto amostral de objetos referentes a cada uma dessas classes. A mineração de dados é útil para explicar e revelar padrões escondidos nos dados e fazer previsões de classificação a partir deles. Entretanto, as previsões são apenas uma definição do comportamento médio das amostras fornecidas para mineração de dados, que é utilizado no processo de classificação. O comportamento médio estabelecido corresponde ao conhecimento implícito - como, por exemplo, o que está inserido no mapa - que foi estimado a partir da base de dados (WITTEN et al., 2011).

Recentemente, uma série de trabalhos de mapeamento geomorfológico foram desenvolvidos a partir da utilização de GEOBIA: Dragut e Blaschke (2006), Asselen e Seijmonsbergen (2006), Camargo et al. (2009), Anders et al. (2011), Verhagen e Dragut (2012), Dragut e Eisank (2012) e Girão et al. (2016).

Portanto, admite-se como hipótese que a GEOBIA, as técnicas de mineração de dados e a utilização de MDEs podem contribuir com o mapeamento e refinamento dos limites das classes existentes em mapas geomorfológicos antigos, de modo que estes possam ser utilizados com qualidade correspondente aos recursos disponíveis hoje e em escalas de maior detalhe. Dessa forma, o objetivo deste trabalho é o refinamento semiautomático de um mapa geomorfológico regional de pequena escala (1:1.000.000) do projeto RADAMBRASIL utilizando GEOBIA, técnicas de mineração de dados e MDEs oriundos do SRTM.

ÁREA DE ESTUDO

A área de estudo (Figura 1) está localizada nas Folhas SF.23/24 Rio de Janeiro/Vitória do Projeto RADAMBRASIL, no Estado do Rio de Janeiro, entre, aproximadamente, as latitudes -22° S e -23° S, e longitudes -43° W e -44° W. A área foi selecionada pela sua variedade topográfica e abriga seis Unidades Geomorfológicas no terceiro táxon mapeadas pelo RADAMBRASIL.

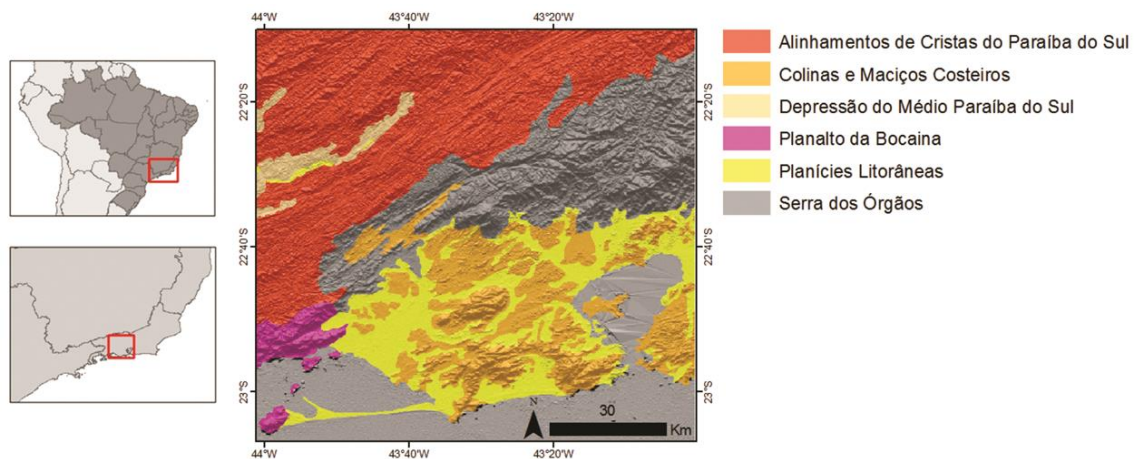


Figura 1 – Localização da área de estudo.

Fonte: RADAMBRASIL (1983)

O estado do Rio de Janeiro está localizado no extremo leste do Rift Continental do Sudeste do Brasil - RCSB (RICCOMINI, 1989; 1995), se estendendo da região de

serras e de depressões interplanálticas correspondente ao vale do rio Paraíba do Sul à região dos maciços litorâneos, colinas e terrenos quaternários da Baía de Guanabara, de Jacarepaguá, de Sepetiba, da Região dos Lagos e da planície costeira de Campos. Sua geomorfologia é caracterizada por um contraste significativo das formas de relevo, produto da complexa evolução geológica-geomorfológica meso-cenozóica do sudeste brasileiro e, ainda, do processo acentuado de alteração da paisagem resultante das condições do regime climático tropical úmido (SILVA, 2002).

Em resumo, a geologia da área de estudo é composta, em sua maior parte, por um terreno de alto grau metamórfico com intrusões graníticas referentes à Faixa Ribeira, que consiste em um complexo orógeno resultante das convergências litosféricas (Orogênese Brasileira) que atuaram entre o Neoproterozoico e o Cabro-Ordoviciano (HEILBRON et al., 2014). As orogêneses, que originaram a Faixa Ribeira, resultaram em metamorfismo, fusão parcial das rochas, deformações, cisalhamento transcorrente regional e granitogênese que originou corpos granitoides (SILVA & CUNHA, 2001). A geomorfologia do estado do Rio de Janeiro é sustentada, em sua maior parte, pelas rochas originadas por tais processos no Pré-Cambriano.

Durante o Cambriano, ocorreu um magmatismo tardi a pós-tectônico referente ao estágio pós-tectônico do orógeno do Brasileiro, tendo como produtos finais os granitoides que deram origem a maciços graníticos como o da Pedra Branca e de Mangaratiba (SILVA & CUNHA, 2001). Posteriormente, no Mesozóico, ocorreu magmatismo constituído pelas intrusões de enxames de diques básicos (diabásio) de idade Juro-Cretácea, relacionados à ruptura do Gondwana Ocidental, que antecedeu à abertura do oceano Atlântico Sul (HEILBRON et al., 2014).

A partir do Mesozóico até o Cenozóico, um período de instabilidade tectônica (evento Sul-Atlântico ou reativação Wealdeniana), referente ao mecanismo de separação das placas tectônicas Sul-Americana e Africana, resultou na formação do Oceano Atlântico e na geração de um episódio de rifteamento, produzindo um sistema de riftes, o Sistema de Riftes Cenozoicos do Sudeste do Brasil (SRCSB - RICCOMINI, 1989), resultando em depressões tectônicas, como da Guanabara e Paraíba do Sul; e em terrenos elevados, como a Serra da Mantiqueira e a Serra do Mar. Nesse período, também ocorreu um magmatismo alcalino relacionado a Reativação Wealdeniana, dando origem, por exemplo, aos Maciços do Medanha, do Tinguá e da Ilha de Cabo Frio (SILVA & CUNHA, 2001).

No SRCSB, especificamente na porção do território fluminense, estão presentes importantes bacias sedimentares paleogênicas, as bacias de Resende e Volta Redonda, que compõem o denominado Segmento Central do RCSB, e as bacias de São José de Itaboraí e Macacu, localizadas no Segmento Oriental do RCSB (HEILBRON et al., 2014).

Por fim, a configuração geológica-geomorfológica da área de estudo é complementada por depósitos neogênicos e quaternários, originados e retrabalhados em um clima predominantemente tropical, mas com inúmeras variações paleoclimáticas (umidificação e aridez), resultando, por exemplo, nas colinas suavemente onduladas que caracterizam a morfologia de “mar de morros” do Planalto Sudeste do Brasil (SILVA, 2002)

Na região da depressão da Baía da Guanabara e da área costeira no Norte Fluminense, as variações dos paleoníveis marinhos, combinadas com o aporte de materiais provindos da região serrana adjacente, resultaram em um entulhamento generalizado da paisagem do entorno da Baía da Guanabara, Baixadas de Jacarepaguá, Sepetiba, Região dos Lagos e Baixada Campista, originando as extensas planícies

flúvio-marinhas, além das morfologias de terraços marinhos, restingas e lagunas costeiras (SILVA, 2002).

MATERIAIS E MÉTODOS

Para realizar o trabalho, os seguintes *softwares* foram utilizados: ArcGIS 9.3 (ESRI, 2008) (utilizado na visualização de Planos de Informação (PI) e na confecção dos cartogramas); eCognition 8.7 (TRIMBLE, 2011) (utilizado na segmentação e na classificação); ENVI 4.7 (ITT, 2009) (utilizado para o processamento digital de imagens e análises); e WEKA 3.6 (HALL et al., 2009) (utilizado na mineração de dados).

A metodologia envolve quatro etapas gerais: (i) preparação dos PI e do banco de dados, (ii) análises e estruturação do conhecimento, (iii) classificação geomorfológica e (iv) avaliação dos resultados. A Figura 2 abaixo resume as etapas do trabalho.

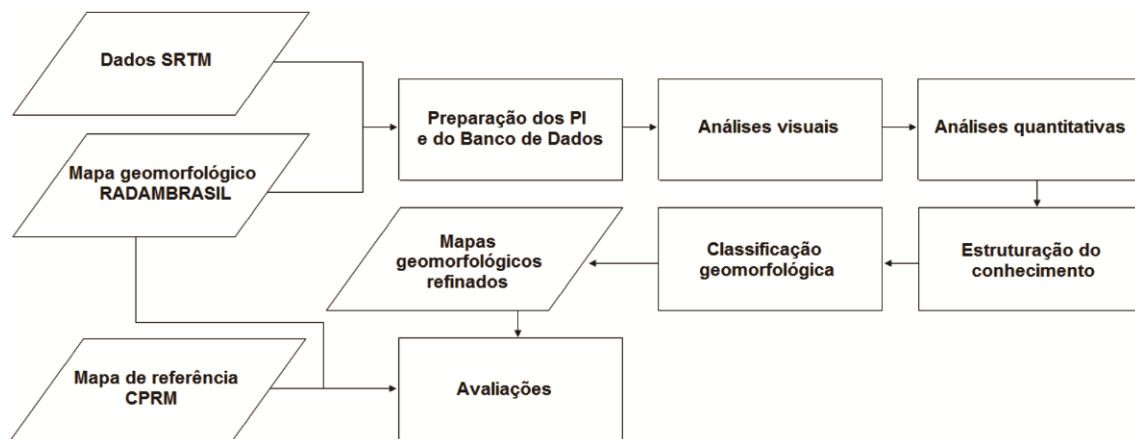


Figura 2 – Fluxograma geral das etapas.

Dados

Foram utilizados dados SRTM refinados e suas VG provenientes do projeto Topodata (<http://www.dsr.inpe.br/topodata/>) no formato TIFF, folhas 22S45, 22S435, 23S45_ e 23S435, e com tamanho de pixel de 1 arco-segundo (aproximadamente 30m): altimetria, declividade, curvatura horizontal, curvatura vertical e relevo sombreado.

Fez-se uso de duas imagens TM/Landsat-5 adquiridas no Catálogo de Imagens do INPE (<http://www.dgi.inpe.br/CDSR/>), cenas 217/76 e 217/75, dos dias 23/02/2007 e 02/08/2007, banda 4 do infravermelho próximo, com resolução espacial de 30 m. Essas imagens serviram para a separação das classes água e não água no primeiro nível de segmentação.

Foram utilizados dois mapas geomorfológicos para eventuais consultas, levantamento de dados e avaliações dos resultados: (i) mapa geomorfológico no formato vetorial para a área de estudo do Projeto RADAMBRASIL das Folhas SF.23/24 Rio de Janeiro/Vitória na escala 1:1.000.000 (RADAMBRASIL, 1983); (ii) mapa geomorfológico no formato vetorial do Projeto Rio de Janeiro realizado pela Companhia de Pesquisa de Recursos Minerais (CPRM) na escala 1:250.000 (CPRM, 2001).

Preparação dos Planos de Informação (PI) e do banco de dados

As VG do SRTM foram organizadas em ambiente SIG, onde foram aplicados processamentos matemáticos que permitiram a obtenção de PI de módulo da curvatura vertical, módulo da curvatura horizontal e declividade em graus. Os dados foram visualizados para fins de apreciação inicial. A Tabela 1 mostra os PI utilizados no projeto criado no software eCognition 8.7.

Tabela 1– PI usados no projeto do eCognition 8.7.

| Plano de Informação (PI) | Unidade |
|--------------------------------|---------|
| Altimetria | m |
| Declividade | ° |
| Curvatura vertical | °/m |
| Módulo da curvatura vertical | °/m |
| Curvatura horizontal | °/m |
| Módulo da curvatura horizontal | °/m |
| Relevo sombreado | - |
| Banda 4 TM | - |

Análises

Para a execução do presente trabalho foi feita a análise visual para comparar o mapa geomorfológico RADAMBRASIL com as VG SRTM, e, assim, identificar as relações entre os dados. Foram feitas também análises quantitativas para obter a estatística descritiva, histogramas por classes e gráficos boxplot, também para identificar relações entre as variáveis utilizadas e as classes a serem mapeadas.

Classificação Geomorfológica

As classes geomorfológicas do mapa do RADAMBRASIL da área de estudo são hierarquizadas em três táxons (Figura 3): Domínios Morfoestruturais, que são um padrão de formas grandes de relevo semelhantes em suas características estruturais; Regiões Geomorfológicas, geradas pelo clima ao longo do tempo sobre a estrutura; e Unidades Geomorfológicas, que possuem formas de relevo semelhantes na fisionomia e geomorfometria. Essas também foram as classes mapeadas no presente trabalho.

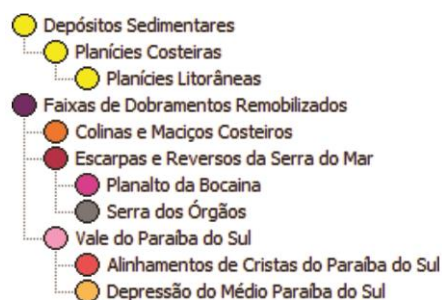


Figura 3 – Hierarquia de classes presentes na área de estudo.

Dessa forma, a classificação geomorfológica se deu com base nos passos apresentados na Figura 4 até o terceiro nível taxonômico conforme proposta de Girão et al. (2016).

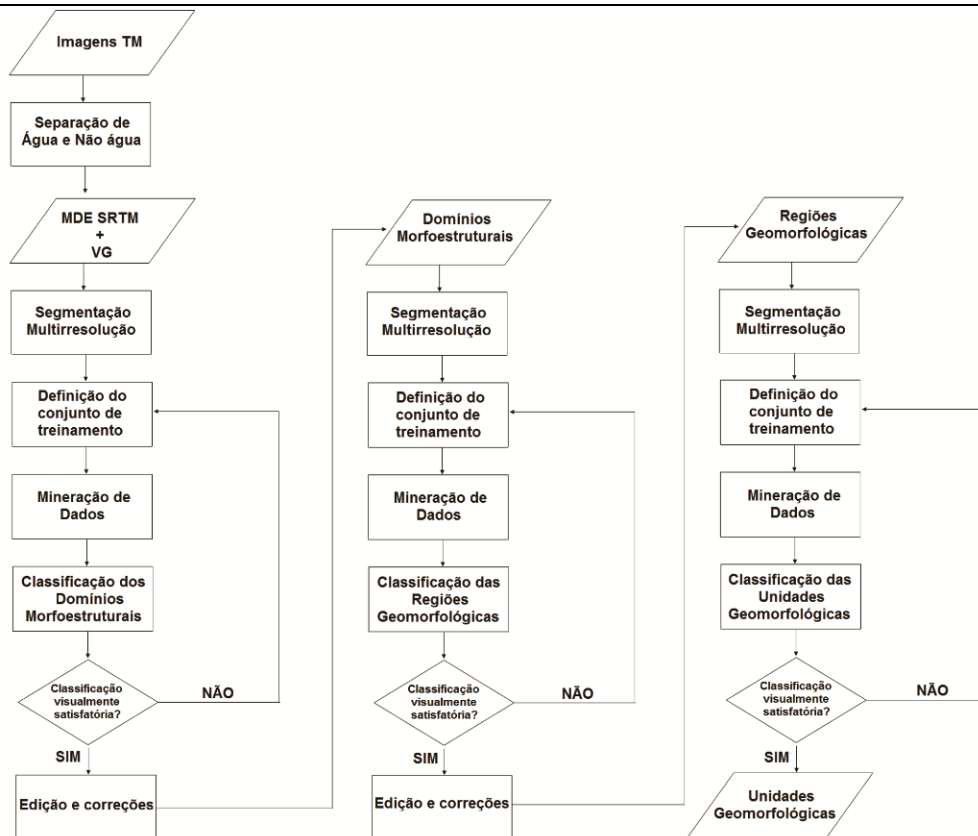


Figura 4 – Fluxograma das etapas da classificação geomorfológica.

Segmentação Multirresolução

A segmentação foi aplicada a partir do nível mais geral (de separação entre água e não água) até o nível mais detalhado, referente às Unidades Geomorfológicas. Este detalhamento foi estruturado de acordo com a hierarquia taxonômica do mapa geomorfológico do RADAMBRASIL, no qual a classificação geomorfológica deve ter os Domínios Morfoestruturais como ponto de partida, por fornecerem as características principais do relevo mapeado. Desse modo, o segundo nível de segmentação foi utilizado para classificar os Domínios Morfoestruturais, o terceiro na classificação das Regiões Geomorfológicas, e o quarto nível no mapeamento das Unidades Geomorfológicas.

Na segmentação multirresolução foi empregado vários testes, utilizando todos os PI e vários valores para os parâmetros de segmentação (*Shape*, *Compactness*, Escala) até a escolha de valores ótimos para cada segmentação com base na análise visual das VG e do relevo sombreado.

No primeiro nível de segmentação, para a separação da água e não água, as bandas 4 (IV próximo) das duas imagens TM da área de estudo foram unidas em um mosaico, que foi georreferenciado de acordo com os dados SRTM. Para reduzir ruídos e atenuar bordas que pudessem afetar a segmentação, foi aplicado um filtro de mediana 5x5, que suaviza as variações detalhadas da imagem. A perda de eventuais detalhes dos dados TM favoreceu uma generalização compatível com as VG e a simplificação de seu processamento. A imagem filtrada foi segmentada, após diversos testes para a escolha dos parâmetros de segmentação (*Shape*: 0,3; *Compactness*: 0,3; Escala: 50)

No segundo nível de segmentação, referente aos Domínios Morfoestruturais, a classe não água foi segmentada utilizando o PI de declividade como variável, já que ela

é indicadora da ocorrência dos Depósitos Sedimentares nessa área de estudo, e os seguintes parâmetros de segmentação: Shape: 0,1; Compactness: 0,1; Escala: 111.

No terceiro nível de segmentação, referente às Regiões Geomorfológicas, a classe Depósitos Sedimentares não foi segmentada, pois originou somente uma outra classe: Planícies Costeiras. A classe Faixas de Dobramentos Remobilizados foi segmentada utilizando o PI de declividade como variável e os seguintes parâmetros de segmentação: Shape: 0,4; Compactness: 0,5; Escala: 105.

No quarto nível de segmentação, referente às Unidades Geomorfológicas, a classe Planícies Costeiras não foi segmentada, pois originou somente uma outra classe: Planícies Litorâneas. A Unidade Geomorfológica de Colinas e Maciços Costeiros é idêntica a sua superclasse existente no nível superior, logo, também não houve necessidade de segmentá-la. A classe Escarpas e Reversos da Serra do Mar foi segmentada utilizando o PI de declividade como variável e os seguintes parâmetros de segmentação: Shape: 0,7; Compactness: 0,4; Escala: 700. Já a classe Vale do Paraíba do Sul foi segmentada utilizando o PI de altimetria como variável e os seguintes parâmetros de segmentação: Shape: 0,7; Compactness: 0,4; Escala: 950

Definição do conjunto de treinamento

A coleta de amostras foi feita com base na interpretação da chave de interpretação proposta nas tabelas 2, 3 e 4; nos PI de VG, de relevo sombreado; e dos mapas geomorfológicos utilizados neste trabalho (RADAMBRASIL e CPRM). Foram coletados, apenas, os objetos que se tinha certeza da classificação. Áreas duvidosas e de transição não fizeram parte do conjunto de amostras. As amostras foram selecionadas ao longo de toda a área de estudo para o conjunto amostral ser o mais representativo possível.

A partir do conjunto amostral selecionado, foram exportadas três tabelas para a realização de experimentos com a mineração de dados: (i) tabela com os valores de média e desvio padrão das VG locais por objeto, (ii) tabela com além das VG locais, VG regionais calculadas no eCognition 8.7 por objeto (Altura, Amplitude, Dissecação e Predominância), e (iii) tabela apenas com as VG indicadas pela etapa das Análises para a separação das classes.

Tabela 2 – Chaves de interpretação dos Domínios Morfoestruturais.

| Domínios Morfoestruturais | |
|-------------------------------------|--|
| Classe | Descrição |
| Depósitos Sedimentares | Áreas planas; textura lisa ou rugosidade muito baixa; próxima à costa, lagunas, vertentes e rios. |
| Faixas de Dobramentos Remobilizados | Áreas que abrigam colinas suaves a escarpas; textura com rugosidade baixa até alta; localização variada. |

Tabela 3 – Chaves de interpretação das Regiões Geomorfológicas.

| Regiões Geomorfológicas | |
|--------------------------------------|---|
| Classe | Descrição |
| Colinas e Maciços Costeiros | Áreas de declividades baixas até altas; rugosidade baixa até alta, circundadas por áreas planas de Depósitos Sedimentares; próximo à costa. |
| Escarpas e Reversos da Serra do Mara | Áreas que abrigam escarpas e relevos serranos; declividades altas e muito altas; rugosidade alta e muito alta; localizadas no contato abrupto com Depósitos Sedimentares ou com o Oceano. |
| Planícies Costeiras | Áreas planas; textura lisa ou rugosidade muito baixa; próxima à costa, lagunas, vertentes e rios. |
| Vale do Paraíba do Sul | Áreas com declividades médias e altas; textura com alta rugosidade; abriga áreas de mares de morros e algumas áreas serranas; próximas ao rio Paraíba do Sul |

Tabela 4 – Chaves de interpretação das Unidades Geomorfológicas.

| Unidades Geomorfológicas | |
|---|--|
| Classe | Descrição |
| Alinhamentos de Cristas do Paraíba do Sul | Áreas com declividades altas; textura com alta rugosidade; abriga áreas de mares de morros e algumas áreas serranas; próximas ao rio Paraíba do Sul. |
| Colinas e Maciços Costeiros | Áreas de declividades baixas até altas; rugosidade baixa até alta, circundadas por áreas planas de Depósitos Sedimentares; próximo à costa. |
| Depressão do Médio Paraíba do Sul | Áreas com declividades médias; textura de média rugosidade; altitude menor em relação aos Alinhamentos de Cristas do Paraíba do Sul; adjacentes à planície do rio Paraíba do Sul |
| Planalto da Bocaina | Área escarpada com declividade muito alta; rugosidade muito alta; localiza-se no contato com o Oceano. |
| Planícies Litorâneas | Áreas planas; textura lisa ou rugosidade muito baixa; próxima à costa, lagunas, vertentes e rios. |
| Serra dos Órgãos | Áreas que abrigam escarpas e relevos serranos; declividades altas e muito altas; rugosidade alta e muito alta; não se localiza no contato com o Oceano. |

Mineração de Dados

A mineração de dados serviu para recuperar um modelo de conhecimento a partir do mapa geomorfológico RADAMBRASIL, que pôde ser utilizado na

classificação dos objetos. As árvores de decisão obtidas na mineração de dados (do primeiro ao terceiro táxon da área de estudo do presente trabalho) expressam a generalização requerida para realizar previsões e, assim, consolidar o conhecimento obtido. Estes resultados podem ser comparados com as árvores de decisão geradas pela mineração de dados aplicada por Qi e Zhu (2003), em que os autores expressam um modelo de conhecimento recuperado a partir de um mapa de solos.

Foi decidido usar do algoritmo C4.5 (disponível no WEKA como J48, a versão em java do C4.5), devido aos bons resultados obtidos em aplicações de sensoriamento remoto por Carvalho (2011), Vieira et al. (2012), Girão et al. (2016) e outros. Basicamente, o algoritmo C4.5 examina todos os atributos previsores candidatos, seleciona aquele que maximiza a taxa de ganho para rotular o nó atual da árvore e repete o processo sucessivamente para dar continuação à construção da árvore de decisão até que todas as instâncias do conjunto de treinamento sejam classificadas (VIEIRA, 2010). As amostras exportadas no formato de tabela foram inseridas no *software* WEKA. Foi feita a mineração de dados que gerou um modelo de conhecimento utilizado na classificação nos quatro níveis.

Há parâmetros no algoritmo de mineração de dados que podem ser modificados pelo usuário. Carvalho (2011) e Vieira (2010) mostram que os seguintes parâmetros na mineração podem ser alterados (i) fator de confiança, (ii) número mínimo de instâncias por nó-folha, (iii) número de ramificações e (iv) poda ou não poda. A modificação desses parâmetros resulta em diferentes modelos de conhecimento, no caso, representados por uma árvore de decisão. Ou seja, a alteração desses parâmetros influencia diretamente no tamanho e na complexidade da árvore de decisão (CARVALHO, 2011). Por isso, em todas as etapas que houve a mineração de dados, vários testes foram feitos com a modificação desses parâmetros. Foram escolhidos os parâmetros que apresentaram resultados mais adequados, em relação ao tamanho, complexidade e interpretabilidade do modelo de conhecimento, para a construção da árvore no eCognition 8.7.

Classificação

Na classificação, as árvores de decisão resultantes da mineração de dados foram implementadas no programa eCognition 8.7. Por fim, os erros de classificação foram corrigidos com a utilização das ferramentas de classificação e edição manual.

Avaliação das classificações

Primeiro, as classificações obtidas com a metodologia proposta foram comparadas visualmente com o mapa geomorfológico RADAMBRASIL para a identificação das relações entre os dois. As avaliações quantitativas foram feitas através da matriz de confusão, índices de exatidão global e índice Kappa (COHEN, 1960; CONGALTON e GREEN, 1999).

Após isso, as classificações foram avaliadas com o mapa geomorfológico do projeto Rio de Janeiro executado pelo CPRM. Para isso, os polígonos do mapa geomorfológico do CPRM no formato vetorial foram compatibilizados e reclassificados de acordo com a classificação do projeto RADAMBRASIL na área de estudo. Assim, obteve-se um mapa base para avaliação com os polígonos mapeados pelo CPRM, mas com a classificação sugerida pelo RADAMBRASIL. A partir daí as classificações foram avaliadas quantitativamente com base no mapa CPRM compatibilizado e

reclassificado por nível taxonômico. Novamente, as avaliações quantitativas foram realizadas pela análise da matriz de confusão, índices de exatidão global e índice Kappa.

RESULTADOS

Primeiro Nível de Classificação

Os objetos foram classificados, primeiramente, em duas classes: água e não água. A classe água foi subdividida em Lagunas, Oceano e Reservatórios pela classificação automática. Após, foram efetuadas correções manuais (Figura 5). Finalmente, os objetos classificados foram unidos em objetos únicos, e este nível de segmentação foi copiado para o nível abaixo. As tabelas 5 e 6 mostram os atributos utilizados na classificação.

Tabela 5 – Atributos utilizados na separação de água e não água.

| Classes | Atributos |
|----------|------------------------------|
| Água | Média da Banda 4 \leq 48,1 |
| Não água | Not Água |

Tabela 6 – Atributos utilizados na classificação das subclasses de Água.

| Classes | Atributos |
|---------------|---|
| Lagunas | Altimetria \leq 15 Borda relativa com a classe Continente = 1 |
| Oceano | Área > 40.000 pixels Not Lagunas Not Reservatórios |
| Reservatórios | Altimetria \geq 40 Borda relativa com a classe Continente \geq 0,7 |

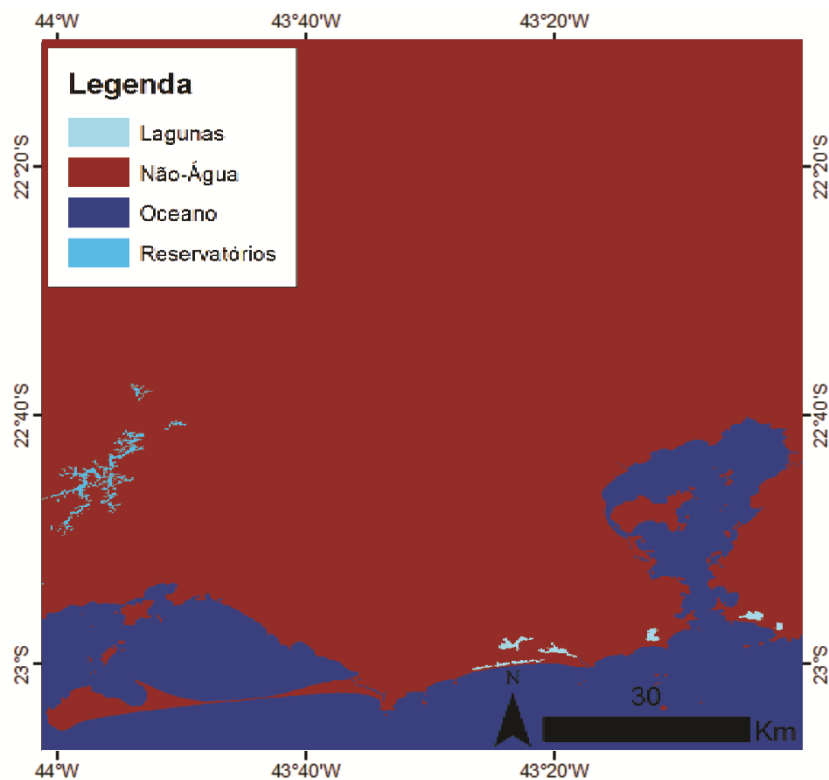


Figura 5 – Classificação da banda 4 Landsat.

Primeiro Táxon (Domínios Morfoestruturais)

No primeiro táxon, a classe não água criada no primeiro nível foi segmentada com base na declividade para dar origem aos objetos que foram utilizados na classificação dos Domínios Morfoestruturais. Vale destacar que as áreas íngremes da Serra do Mar apresentaram a criação de inúmeros objetos pequenos, o que leva a uma supersegmentação causada por esses valores altos de declividade. Mas no geral, é possível dizer que a segmentação utilizada nessa etapa permitiu garantir um compromisso na obtenção de objetos com significado geomorfológico e/ou geomorfométrico, como mencionam Dragut e Blaschke (2006). Um aumento no valor de escala permitiria aumentar o tamanho dos objetos presentes na Serra do Mar, mas isso causaria a subsegmentação de outras áreas.

A partir do conjunto amostral selecionado, obteve-se uma árvore de decisão (Figura 6) caracterizada por expressar apenas um limiar de média de declividade para separar as Faixas de Dobramentos Remobilizados dos Depósitos Sedimentares. Dessa forma, os Depósitos Sedimentares são os objetos com média de declividade menor que $3,913358^\circ$, enquanto que as Faixas de Dobramentos Remobilizados são os objetos com média de declividade superior a este valor, ou seja, áreas que possuem ondulações suaves até escarpas serranas íngremes.

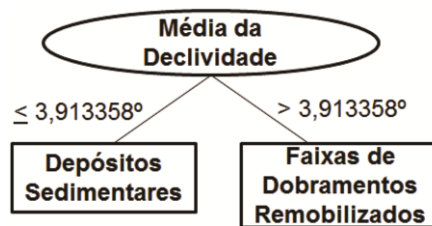


Figura 6 – Árvore de decisão obtida para o primeiro nível taxonômico.

O relatório gerado pelo WEKA mostra a avaliação por validação cruzada da árvore de decisão (Figura 7). 99% das instâncias foram classificadas corretamente, além de ter sido alcançado um índice Kappa de 0,9688. O relatório indica a adequação da árvore de decisão para expressar o conhecimento geomorfométrico e/ou geomorfológico para o primeiro nível taxonômico.

| | | |
|---|--------|-----|
| Instâncias classificadas corretamente | 198 | 99% |
| Instâncias classificadas incorretamente | 2 | 1% |
| Índice Kappa | 0,9688 | |

=== Matriz de Confusão ===

| | | |
|----|-----|---|
| a | b | ← classificados como |
| 39 | 1 | a = Depósitos Sedimentares |
| 1 | 159 | b = Faixas de Dobramentos Remobilizados |

Figura 7 – Relatório da mineração de dados da árvore do primeiro táxon.

A árvore de decisão da Figura 6 foi implementada no software eCognition 8.7, e foi utilizada na classificação dos objetos originados pela segmentação nas classes do primeiro nível taxonômico do mapa RADAMBRASIL. A Figura 8 mostra as classificações obtidas antes e depois da edição manual. Poucos objetos de Depósitos Sedimentares foram classificados como Faixas de Dobramentos Remobilizados. Esses erros de classificação estão em áreas costeiras densamente ocupadas na parte sul da

Baía de Guanabara. As construções presentes podem ter interagido com as ondas de RADAR de tal forma que aumentou a declividade do local, uma vez que os prédios são formações de declividades abruptas. Assim, os objetos criados pela segmentação possuem médias de declividade altas, o que causa confusão com a classe Faixas de Dobramentos Remobilizados. Os contornos desses poucos objetos tiveram que ser editados e classificados manualmente pelo uso das ferramentas de edição disponíveis no eCognition 8.7. A edição é importante para não haver propagação de erros, como menciona Camargo (2009).

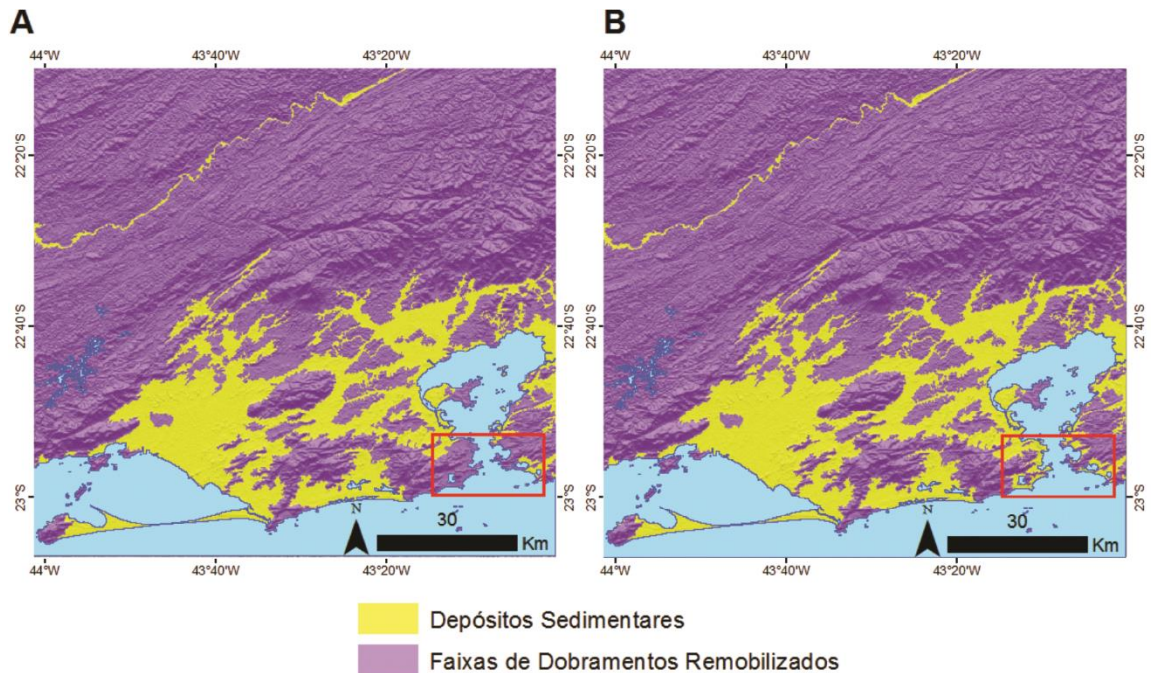


Figura 8– Classificações para o primeiro táxon: (A) antes da edição e (B) depois da edição.

Pela comparação visual entre o mapa RADAMBRASIL original e o obtido no presente trabalho para o primeiro nível taxonômico, é possível notar que a metodologia empregada forneceu um resultado com os limites das classes refinados, já que os dados SRTM estão em uma escala maior que 1:1.000.000 (Figura 9). Além disso, a comparação mostra que há correspondência entre os dois mapas.

A Tabela 7 mostra a matriz de confusão, a exatidão global o índice Kappa obtidos com base no mapa original RADAMBRASIL para a avaliação da classificação editada. Foram obtidos um índice Kappa de 0,6529 e uma exatidão global de 90,5225%. Esses valores, junto com a análise da matriz de confusão, indicam que há concordância. Obviamente, a concordância não será alta, pois além dos dois dados estarem em escalas diferentes, há a subjetividade da interpretação geomorfológica, que foi realizada por autores diferentes.

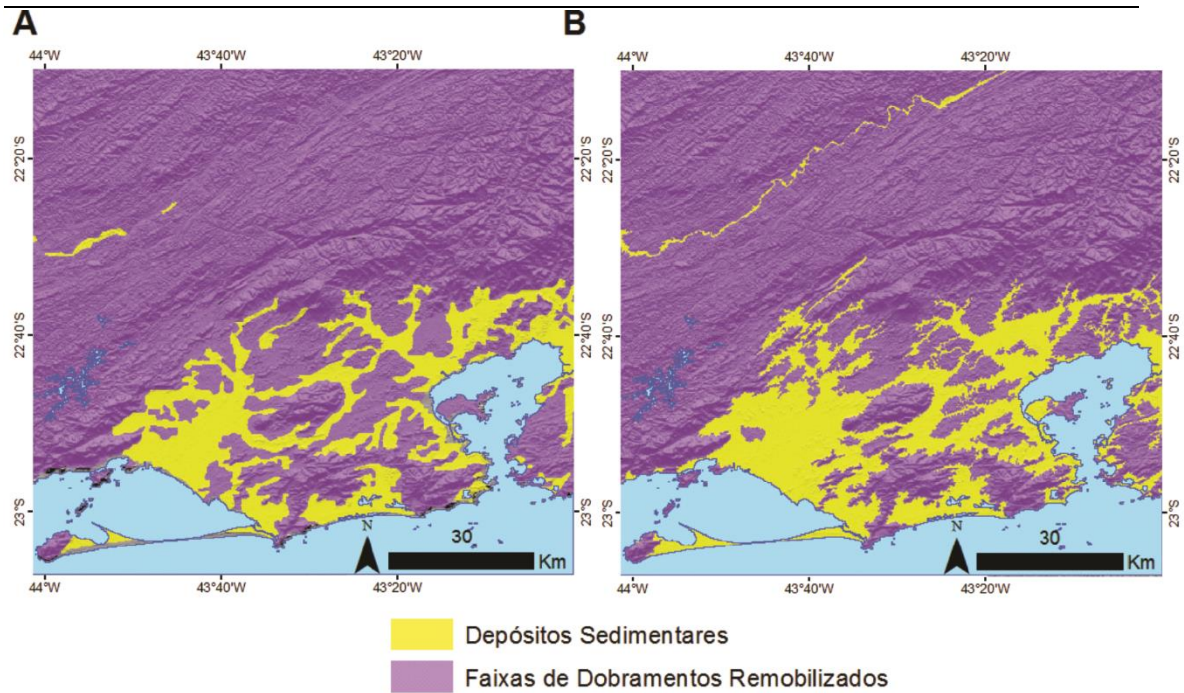


Figura 9 – Comparação para o primeiro táxon: (A) Mapa original RADAMBRASIL. (B) Classificação a partir do SRTM.

Tabela 7 – Matriz de confusão para a classificação editada com base no mapa RADAMBRASIL para o primeiro táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | |
|---|--------------------------|-------------|-----------------------------|
| | DS | FDR | Total |
| Depósitos Sedimentares (DS) | 78,86% | 7,48% | 17,92% |
| Faixas de Dobramentos Remobilizados (FDR) | 21,14% | 92,52% | 82,08% |
| Total | 100% | 100% | 100% |
| Exatidão Global: 90,5225% | | | Índice Kappa: 0,6529 |

As Tabelas 8 e 9 mostram as matrizes de confusão, as exatidões globais e os índices Kappa para as classificações sem e com edição manual, com base no mapa do CPRM compatibilizado. O índice Kappa obtido para o mapa sem edição é de 0,7660, enquanto que o valor obtido para o mapa editado foi de 0,7751. Esses valores indicam que, como constatado anteriormente, poucos objetos foram classificados incorretamente devido aos ruídos no SRTM nas áreas costeiras densamente povoadas. A intervenção manual permitiu uma pequena melhora, ou seja, aumentou ligeiramente o valor do índice Kappa.

Tabela 8 – Matriz de confusão para a classificação sem edição avaliada com base no mapa CPRM para o primeiro táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | |
|---|--------------------------|-------------|-----------------------------|
| | DS | FDR | Total |
| Depósitos Sedimentares (DS) | 81,43% | 4,31% | 17,80% |
| Faixas de Dobramentos Remobilizados (FDR) | 18,57% | 95,69% | 82,20% |
| Total | 100% | 100% | 100% |
| Exatidão Global: 93,1995% | | | Índice Kappa: 0,7660 |

Tabela 9 – Matriz de confusão para a classificação editada avaliada com base no mapa CPRM para o primeiro táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | |
|---|--------------------------|-------------|-----------------------------|
| | DS | FDR | Total |
| Depósitos Sedimentares (DS) | 83,09% | 4,40% | 18,17 % |
| Faixas de Dobramentos Remobilizados (FDR) | 16,91% | 95,60% | 81,83% |
| Total | 100% | 100% | 100% |
| Exatidão Global: 93,4108% | | | Índice Kappa: 0,7751 |

Segundo Táxon (Regiões Geomorfológicas)

O Domínio das Faixas de Dobramentos Remobilizados foi segmentado com base na declividade para ser classificado em três classes de Regiões Geomorfológicas: Colinas e Maciços Costeiros, Escarpas e Reversos da Serra do Mar e Vale do Paraíba do Sul. Os objetos criados, dessa vez, possuem tamanhos apropriados. Portanto, visualmente, a segmentação pôde ser considerada satisfatória, já que foi útil para gerar os compartimentos topográficos a serem classificados.

A partir do conjunto amostral selecionado, obteve-se uma árvore de decisão (Figura 10) caracterizada por três divisões para a classificação. Na primeira divisão, as instâncias com média de altimetria menor que 207,672026 m foram classificadas como Colinas e Maciços Costeiros, se maior do que esse valor, o teste seguinte é feito com a média da declividade. Então, as instâncias com média da altimetria maior que 207,672026 m e com a média da declividade maior que 20,694207° são classificadas como Escarpas e Reversos da Serra do Mar, caso contrário, o teste é feito com o desvio padrão da declividade. A divisão final é feita, portanto, com o desvio padrão da declividade, em que as instâncias com média de declividade menor ou igual a 20,694207° são separadas em Escarpas e Reversos da Serra do Mar (desvio padrão da declividade maior que 7,153971°) e em Vale do Paraíba do Sul (desvio padrão da declividade menor ou igual a 7,153971°).

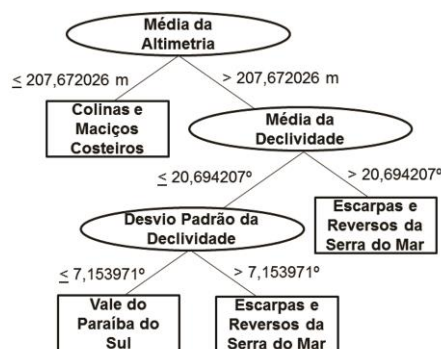


Figura 10 – Árvore de decisão obtida para o segundo nível taxonômico.

A Figura 11 mostra o relatório de avaliação da árvore de decisão. 83,3% das amostras foram classificadas corretamente e, obteve-se um índice Kappa de 0,75. Pode-se considerar este resultado satisfatório para expressar o conhecimento geomorfométrico/geomorfológico.

| | | |
|--|-------------|------------------|
| Instâncias classificadas corretamente | 100 | 83,3333 % |
| Instâncias classificadas incorretamente | 20 | 16,6667 % |
| Índice Kappa | 0,75 | |

=== Matriz de Confusão ===

| a | b | c | ← classificados como |
|----|----|----|---|
| 35 | 5 | 0 | a = Colinas e Maciços Costeiros |
| 3 | 31 | 6 | b = Escarpas e Reversos da Serra do Mar |
| 0 | 6 | 34 | c = Vale do Paraíba do Sul |

Figura 11 – Relatório da mineração de dados da árvore do segundo táxon.

A Figura 12 mostra os mapas sem e com a edição manual. Neste nível taxonômico, a confusão entre classes foi um pouco maior, portanto, a edição dos objetos e classificação manual dos erros, apesar de rápida, foi um pouco mais trabalhosa em relação à etapa anterior. Objetos de declividade e altitude altas presentes nos Maciços da Tijuca, Mendanha e da Pedra Branca foram classificados como Escarpas e Reversos da Serra do Mar. Alguns alinhamentos serranos presentes no Vale do Paraíba do Sul também foram confundidos com as Escarpas e Reversos da Serra do Mar. Além, disso, as Escarpas e Reversos da Serra do Mar possuem trechos de declividades mais suaves que foram confundidas com a classe Vale do Paraíba do Sul. No geral, pode-se considerar que a árvore de decisão utilizada e que a classificação sem edição foi satisfatórias, pois as poucas intervenções manuais foram rápidas de serem executadas.

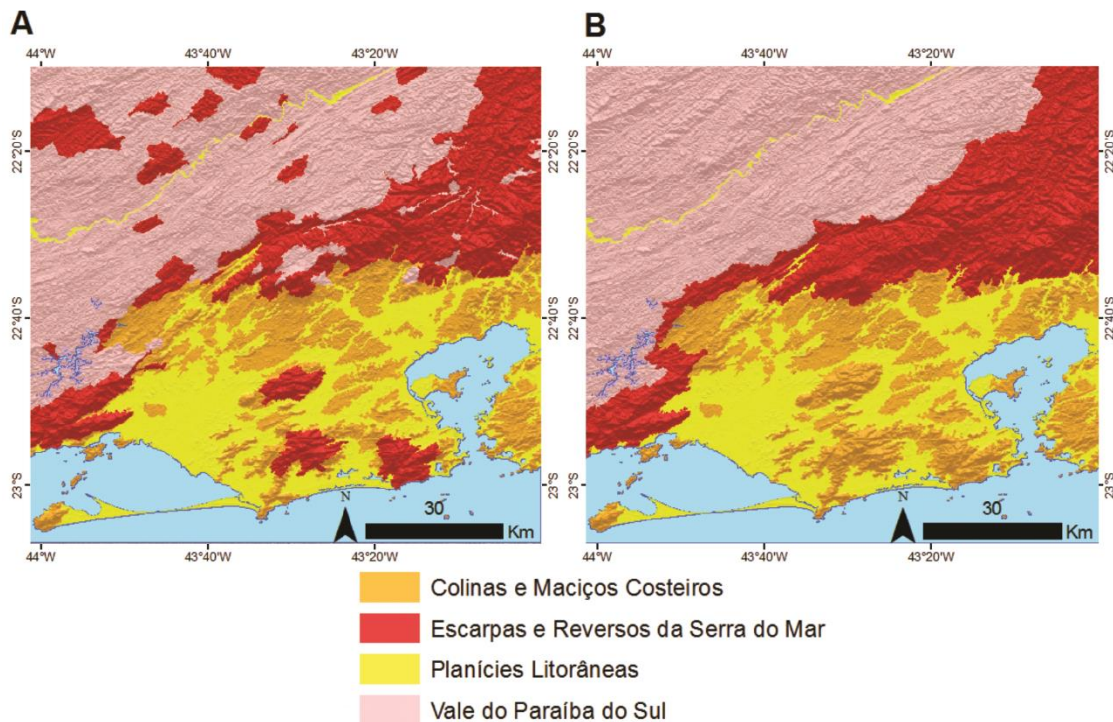


Figura 12 – Classificações para o segundo táxon: (A) antes da edição e (B) depois da edição.

Pela análise visual da Figura 13, percebe-se que há correspondência entre os dois mapas. Os contornos da classificação pelos dados SRTM estão refinados em relação aos do mapa em escala menor (1:1.000.000). O SRTM permitiu, então, trazer as Regiões Geomorfológicas do projeto RADAMBRASIL para uma escala superior a 1:1.000.000. Outro fato notório é a possibilidade de se alterar, redesenhar, atualizar e corrigir classes do mapa original. Por exemplo, a partir do segundo nível taxonômico é possível notar que as ilhas existentes no 3º quadrante do mapa original foram classificadas como Escarpas e Reversos da Serra do Mar devido a uma semelhança litológica (RADAMBRASIL, 1983). Entretanto, com base em consultas ao mapa geomorfológico do CPRM, foi decidido classificá-las como integrantes da Região das Colinas e Maciços Costeiros.

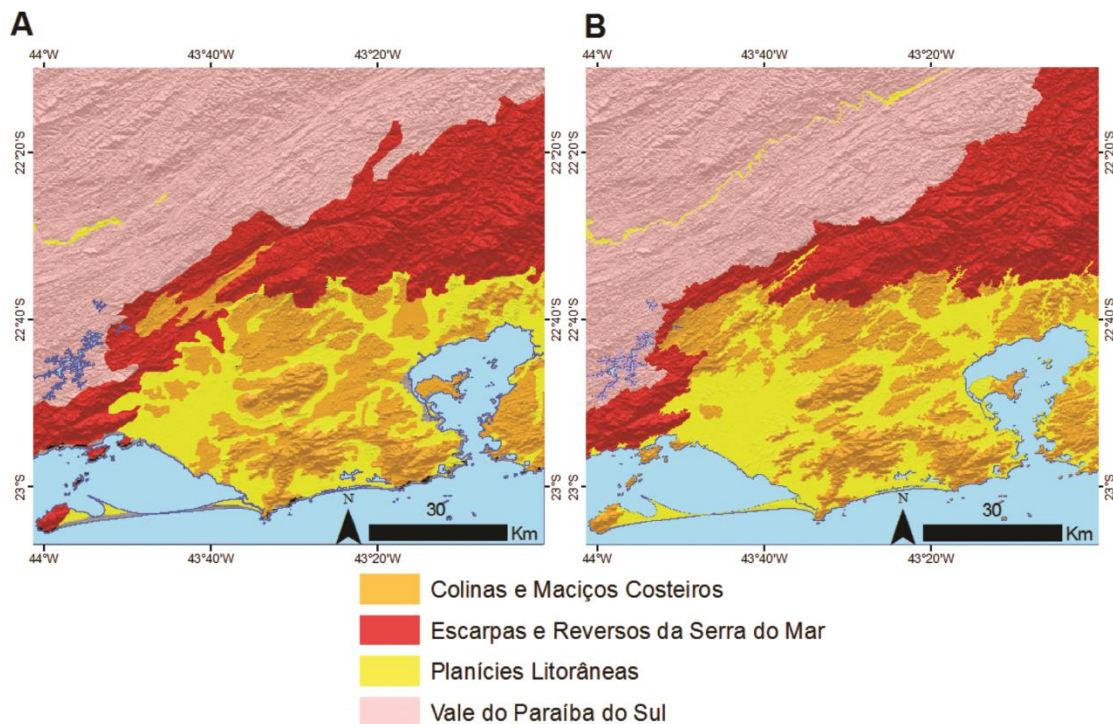


Figura 13 – Comparação para o segundo táxon: (A) Mapa original RADAMBRASIL. (B) Classificação a partir do SRTM.

A Tabela 10 mostra a matriz de confusão, a exatidão global e o índice Kappa obtidos com base no mapa original RADAMBRASIL para a avaliação da classificação editada. Obteve-se exatidão global de 83,4681% e índice Kappa de 0,7688. Tais valores mostram que há concordância entre os dados SRTM e os originais, apesar de estarem em escalas diferentes, e de haver mudanças nas classificações de alguns polígonos de ilhas, como tratado anteriormente.

Tabela 10 – Matriz de confusão para a classificação editada com base no mapa RADAMBRASIL para o segundo táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | | | |
|--|--------------------------|-------------|-------------|-------------|-----------------------------|
| | CMC | ERSM | PC | VPS | Total |
| Colinas e Maciços Costeiros (CMC) | 70,06% | 9,78% | 18,23% | 0,02% | 18,45% |
| Escarpas e Reversos da Serra do Mar (ERSM) | 1,79% | 79,73% | 1,82% | 5,14% | 23,34% |
| Planícies Costeiras (PC) | 28,14% | 2,43% | 78,94% | 1,11% | 17,95% |
| Vale do Paraíba do Sul (VPS) | 0% | 8,06% | 1,01% | 93,73% | 40,25% |
| Total | 100% | 100% | 100% | 100% | 100% |
| Exatidão Global: 83,4646% | | | | | Índice Kappa: 0,7688 |

As Tabelas 11 e 12 mostram as matrizes de confusão obtidas para as classificações sem e com edição, geradas com base no mapa CPRM. Obteve-se um índice Kappa de 0,6987 para a classificação sem edição, e um índice Kappa de 0,8410 para a classificação editada. No geral, as concordâncias foram inferiores aos do primeiro táxon devido ao maior número de classes existentes no segundo nível taxonômico.

O valor do índice Kappa da classificação sem edição é consideravelmente menor que o do mapa editado. Isto mostra a necessidade de intervenções manuais para não haver propagações de erros quando a classificação do terceiro nível taxonômico for realizada. Como discutido no item anterior, as confusões entre classes foram causadas devido às semelhanças geomorfométricas entre as classes. Entretanto, pode-se dizer que a árvore de decisão foi eficaz na classificação automática, uma vez que a maioria dos objetos foi classificada corretamente, e que as edições manuais foram rapidamente realizadas.

Tabela 11 – Matriz de confusão para a classificação sem edição avaliada com base no mapa CPRM para o segundo táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | | | |
|--|--------------------------|-------------|-------------|-------------|-----------------------------|
| | CMC | ERSM | PC | VPS | Total |
| Colinas e Maciços Costeiros (CMC) | 65,24% | 8,62% | 15,54% | 0 % | 17,48% |
| Escarpas e Reversos da Serra do Mar (ERSM) | 17,12% | 73,37% | 0,85% | 14,01% | 25,70% |
| Planícies Costeiras (PC) | 15,66% | 0,34% | 83,27% | 0,72% | 18,20% |
| Vale do Paraíba do Sul (VPS) | 0,97% | 17,66% | 0,34% | 85,27% | 38,62% |
| Total | 100% | 100% | 100% | 100% | 100% |
| Exatidão Global: 78,3050% | | | | | Índice Kappa: 0,6987 |

Tabela 12 – Matriz de confusão para a classificação editada avaliada com base no mapa CPRM para o segundo táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | | | |
|--|--------------------------|-------------|-------------|-------------|-----------------------------|
| | CMC | ERSM | PC | VPS | Total |
| Colinas e Maciços Costeiros (CMC) | 77,45% | 3,90% | 14,93% | 0 % | 18,71% |
| Escarpas e Reversos da Serra do Mar (ERSM) | 5,89% | 90,31% | 1,46% | 3,95% | 23,36% |
| Planícies Costeiras (PC) | 16,66% | 0,34% | 83,27% | 0,72% | 18,20% |
| Vale do Paraíba do Sul (VPS) | 0% | 5,44% | 0,33% | 95,33% | 39,74% |
| Total | 100% | 100% | 100% | 100% | 100% |
| Exatidão Global: 88,5789% | | | | | Índice Kappa: 0,8410 |

Terceiro Táxon (Unidades Geomorfológicas)

Para a classificação das Unidades Geomorfológicas, a segmentação foi feita em duas etapas. Primeiro, a Região Geomorfológica do Vale do Paraíba do Sul foi segmentada, depois, as Escarpas e Reversos da Serra do Mar.

A partir do conjunto amostral selecionado, obteve-se uma árvore de decisão para a classificação das unidades do Vale do Paraíba do Sul (Figura 14) caracterizada por apresentar um limiar de altimetria entre a Depressão do Médio Paraíba do Sul e os Alinhamentos de Cristas do Paraíba do Sul. O limiar indica que a depressão possui objetos com média de altitude menor ou igual a 411,430608 m.



Figura 14 - Árvore de decisão obtida para a classificação das unidades do Vale do Paraíba do Sul.

O resumo do relatório (Figura 15) da mineração de dados para a árvore da Figura 14 indica que ela é satisfatória para a classificação das unidades geomorfológicas do Vale do Paraíba do Sul, uma vez que 97,1429% das instâncias foram classificadas corretamente. Além disso, foi obtido um índice Kappa de 0,7853.

| | | |
|---|--------|-----------|
| Instâncias classificadas corretamente | 34 | 97,1429 % |
| Instâncias classificadas incorretamente | 1 | 2,8571 % |
| Índice Kappa | 0,7853 | |

=== Matriz de Confusão ===

| | | |
|----|---|---|
| a | b | ← classificados como |
| 32 | 0 | a = Alinhamentos de Cristas do Paraíba do Sul |
| 1 | 2 | b = Depressão do Médio Paraíba do Sul |

Figura 15 – Relatório da mineração de dados da árvore das unidades do Vale do Paraíba do Sul.

Os testes de mineração de dados, que utilizam apenas geomorfometria, para a classificação das unidades das Escarpas e Reversos da Serra do Mar geraram uma mesma árvore de decisão que exibiu alta confusão entre as classes. Esta confusão foi esperada, devido à semelhança geomorfométrica entre o Planalto da Bocaina e a Serra dos Órgãos.

O Planalto da Bocaina possui contato com o oceano, enquanto a Serra dos Órgãos não. Por isso, foi decidido realizar um experimento de mineração de dados com o atributo borda ao oceano. Assim, a árvore de decisão gerada (Figura 16) é caracterizada por utilizar apenas este atributo. Apesar do resumo do relatório da mineração (Figura 17) indicar resultados ruins, a classificação no eCognition 8.7 não gerou erros, como será visto adiante.

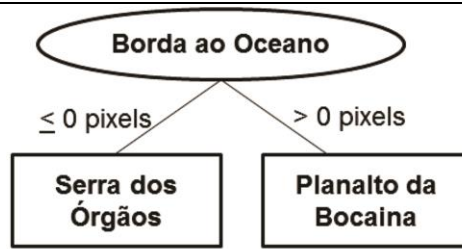


Figura 16- Árvore de decisão obtida para a classificação das unidades das Escarpas e Reversos da Serra do Mar.

| | | |
|--|----------|------------|
| Instâncias classificadas corretamente | 8 | 80% |
| Instâncias classificadas incorretamente | 2 | 20% |
| Índice Kappa | 0 | |

=== Matriz de Confusão ===

| | | |
|----------|----------|-----------------------------|
| a | b | ← classificados como |
| 0 | 2 | a = Planalto da Bocaina |
| 0 | 8 | b = Serra dos Órgãos |

Figura 17 – Resumo do relatório da mineração de dados da árvore das unidades do Escarpas e Reversos da Serra do Mar.

As duas árvores geradas foram implementadas no eCogniton 8.7 e as classificações foram realizadas (Figura 18). Como os erros foram corrigidos anteriormente (nos níveis taxonômicos acima), a segmentação neste nível não gerou objetos incorretos. Dessa forma, não houve necessidade de intervenções manuais no terceiro nível taxonômico. Todas as Unidades Geomorfológicas foram classificadas corretamente.

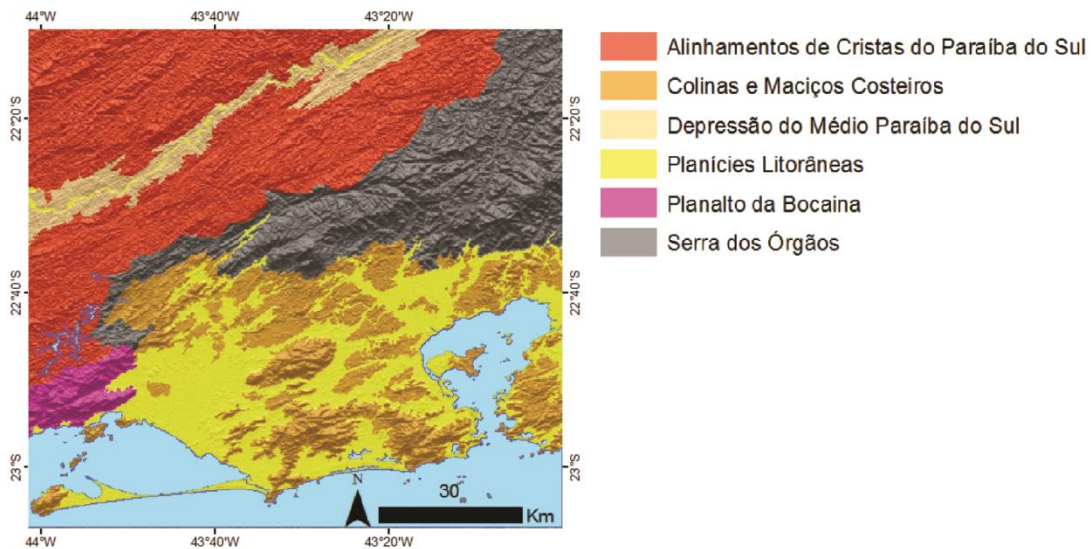


Figura 18 – Unidades Geomorfológicas mapeadas a partir dos dados SRTM

A comparação visual com os dados RADAMBRASIL permite verificar que há correspondência entre os dois mapas, e que os limites dos contornos originais ficaram refinados devido ao uso dos dados SRTM, que está em escala superior (Figura 19). As pequenas mudanças temáticas feitas no táxon superior (classificação das ilhas como Colinas e Maciços Costeiros) se mantiveram no terceiro nível taxonômico. A Tabela 13 mostra a matriz de confusão gerada com base no mapa RADAMBRASIL. Obteve-se

exatidão global 79,2155% e índice Kappa de 0,7240. Estes resultados indicam que há concordância entre os dois mapas.

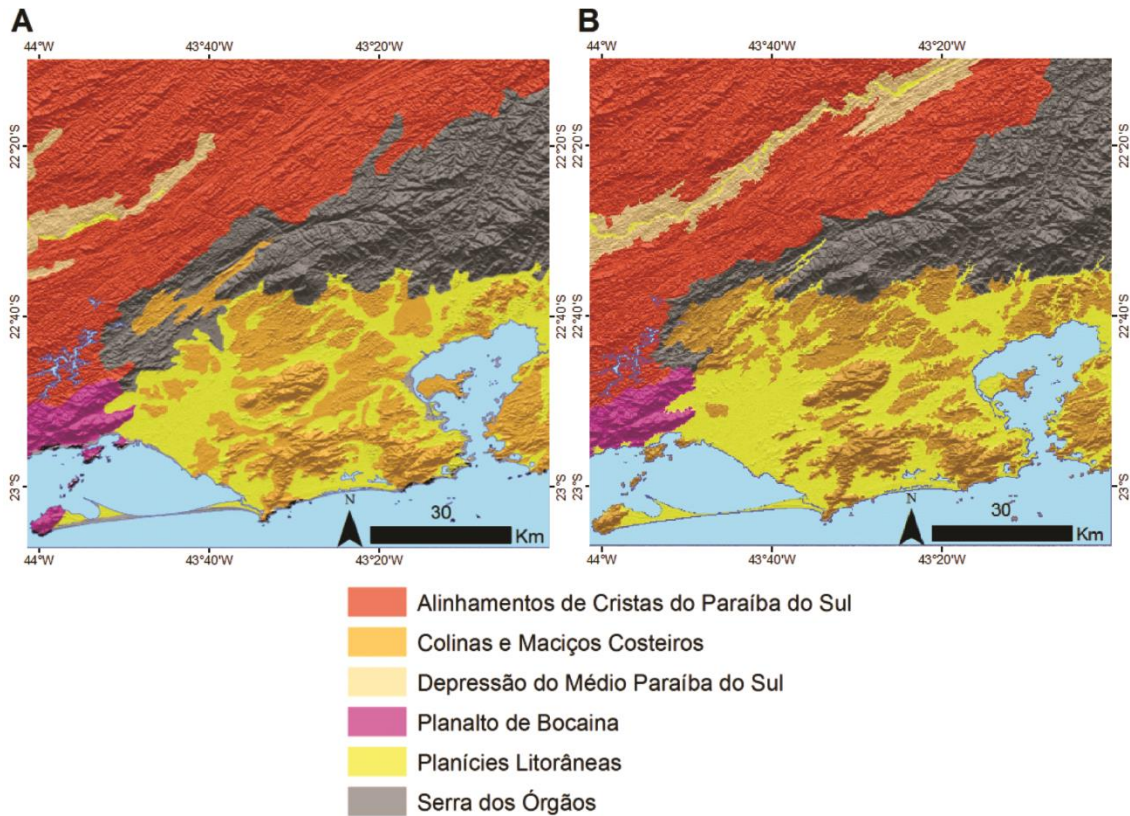


Figura 19 – Comparação para o terceiro táxon: (A) Mapa original RADAMBRASIL. (B) Classificação a partir do SRTM.

Tabela 13 – Matriz de confusão para a classificação editada avaliada com base no mapa RADAMBRASIL para o terceiro táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | | | | | Total |
|--|--------------------------|-------------|-------------|-------------|-------------|-----------------------------|-------------|
| | ACPS | CMC | DMPS | PB | PL | SO | |
| Alinhamentos de Cristas do Paraíba do Sul (ACPS) | 84,83% | 0% | 30,28% | 2,67% | 0% | 8,74% | 35,28% |
| Colinas e Maciços Costeiros (CMC) | 0,02% | 70,04% | 0,06% | 12,46% | 18,26% | 9,43% | 18,44% |
| Depressão do Médio Paraíba do Sul (DMPS) | 9,08% | 0% | 61,13% | 0% | 1,01% | 0% | 5% |
| Planalto da Bocaina (PB) | 0,51% | 0% | 0% | 80,71% | 0,28% | 0,39% | 2,41% |
| Planícies Litorâneas (PL) | 0,68% | 28,14% | 8,52% | 4,16% | 78,86% | 2,20% | 17,93% |
| Serra dos Órgãos (SO) | 4,88% | 1,83% | 0% | 0% | 1,60% | 79,24% | 20,94% |
| Total | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Exatidão Global: 79,2155% | | | | | | Índice Kappa: 0,7240 | |

A Tabela 14 mostra a matriz de confusão obtida com base no mapa CPRM. Foram obtidos exatidão global de 84,6769% e índice Kappa de 0,7985. Os resultados indicam que há concordância entre o mapa de Unidades Geomorfológicas RADAMBRASIL obtidas a partir do SRTM e o mapa do CPRM, o que mostra o potencial da metodologia

desenvolvida nesse estudo para o refinamento e recuperação de um modelo de conhecimento a partir de mapas geomorfológicos.

Tabela 14 – Matriz de confusão para a classificação editada avaliada com base no mapa CPRM para o terceiro táxon.

| CLASSES | REFERÊNCIA (Porcentagem) | | | | | | Total |
|--|--------------------------|-------------|-------------|-------------|-------------|-----------------------------|-------------|
| | ACPS | CMC | DMPS | PB | PL | SO | |
| Alinhamentos de Cristas do Paraíba do Sul (ACPS) | 89,05% | 0% | 35,16% | 1,97% | 0% | 5,93% | 34,90% |
| Colinas e Maciços Costeiros (CMC) | 0% | 77,39% | 0% | 0,21% | 14,98% | 4,29% | 18,70% |
| Depressão do Médio Paraíba do Sul (DMPS) | 6,29% | 0% | 60,59% | 0% | 0,33% | 0% | 4,86% |
| Planalto da Bocaina (PB) | 0,16% | 0% | 0% | 96,27% | 0,37% | 0,70% | 2,50% |
| Planícies Litorâneas (PL) | 0,30% | 16,68% | 4,25% | 1,43% | 83,18% | 0,18% | 18,18% |
| Serra dos Órgãos (SO) | 4,20% | 5,94% | 0% | 0,12% | 1,13% | 88,90% | 20,85% |
| Total | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| Exatidão Global: 84,6769% | | | | | | Índice Kappa: 0,7985 | |

CONCLUSÃO

Confirma-se a hipótese que a GEOBIA, as técnicas de mineração de dados e a utilização de MDEs podem contribuir com o mapeamento e refinamento dos limites das classes existentes em mapas geomorfológicos antigos, de modo que estes possam ser utilizados com qualidade correspondente aos recursos disponíveis hoje e em escalas superiores. Os mapas refinados permitem que o projeto RADAMBRASIL seja utilizado de acordo com os recursos computacionais da atualidade e em escalas superiores a 1:1.000.000. O presente trabalho serve como uma fonte de conhecimento para o desenvolvimento de metodologias que objetivem o refinamento de mapas geomorfológicos regionais antigos, bem como a obtenção de um modelo de conhecimento.

A análise dos resultados permite algumas outras conclusões:

- Os dados SRTM permitem, no geral, a caracterização geomorfométrica regional, tanto visual quanto quantitativa, de classes de relevo em escalas superiores a 1:1.000.000. Por tais caracterizações, foi possível indicar as VG que permitem a separação entre classes até o terceiro táxon. A declividade serviu para a separação do primeiro táxon, a altimetria e declividade para o segundo táxon, e altimetria para a separação da Depressão do Médio Paraíba do Sul dos Alinhamentos de Cristas do Paraíba do Sul. Entretanto, para classes semelhantes geomorfometricamente, como o Planalto da Bocaina e Serra dos Órgãos no terceiro táxon, não foi possível separá-las com as VG adotadas pela metodologia do presente trabalho.
- A segmentação é sensível às descontinuidades geomorfológicas do relevo expressas nos dados SRTM, o que permite a obtenção dos contornos refinados (em escala superior a 1:1.000.000) das classes mapeadas. A segmentação da declividade forneceu bons resultados, pois esta variável realça os limites entre os compartimentos. Entretanto, para definir os limites de classes em que há um desnível altimétrico entre elas, como no caso dos Alinhamentos de Cristas do

Paraíba do Sul e da Depressão do Médio Paraíba do Sul, a altimetria foi utilizada na segmentação.

- Cada classe possui configurações de segmentação ideais. Por isso, é necessário configurá-la, depois da realização de testes, de acordo com os parâmetros que permitem o delineamento dos compartimentos de interesse. Vale lembrar que diversas configurações do algoritmo de segmentação foram utilizadas até se chegar à definição da configuração ótima. Uma pequena variação nos valores dos parâmetros do algoritmo, já muda sensivelmente a compartimentação topográfica.
- A mineração de dados forneceu resultados que estão de acordo com os limiares de diferenciação entre classes, estabelecidos nas etapas de análises. Portanto, os algoritmos de mineração de dados foram úteis para diminuir o tempo gasto na seleção de atributos.
- A edição manual dos poucos erros de segmentação e de classificação foi rápida e não trabalhosa, o que mostra a utilidade desse procedimento na obtenção de bons resultados. As edições foram feitas com base na interpretação visual do relevo sombreado, que é útil para tal tarefa.
- A escolha da metodologia utilizada, como a configuração do algoritmo de segmentação multirresolução e critérios de interpretação foi definida após diversos testes, e são específicos para a área de estudo deste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDERS, N. S.; SEIJMONSBERGEN, A. C.; BOUTTEN, W. Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. **Remote Sensing of Environment**, v. 115, n. 12, p. 2976-2985, 2011.

ASSELEN, S.; SEIJMONSBERGEN, A. C. Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM. **Geomorphology**, v. 78, n. 3.4, p. 309-320, 2006.

ALLEVATO, S. R.; PINGARILHO, M. N. F. Análise e tratamento de material fotográfico no Projeto RADAMBRASIL. **Revista Brasileira de Cartografia**, n. 24, p. 19-24, 1978.

BARBOSA, G. V. **Cartografia geomorfológica utilizada pelo projeto RADAM**. In: In: CONGRESSO BRASILEIRO DE GEOLOGIA, 27., 1973, Aracaju, , Anais... Aracaju: SBG, v.1, p. 427-433, 1973.

BAATZ, M.; SCHAPE, A. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation In: Strobl, J., Blaschke, T. (eds.). **Angewandte Geographische Informationsverarbeitung XII**. Beiträge zum AGIT Symposium Salzburg 2000. Karlsruhe: Herbert Wichmann Verlag, 2000. p. 12-23.

BENZ, U. C.; HOFMANN, P.; WIKKHAUCK, G.; LINGENFELDER, I.; HEYNEN, M. Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. **Photogrammetry & Remote Sensing**, v. 58, p. 239-258, 2004.

BISHOP, M. P.; JAMES, L. A.; SHRODER, JR; J. F.; WALSH, S. J. Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. **Geomorphology**, v. 137, n. 1, p. 5-26, 2012.

CAMARGO, F. F.; FLORENZANO, T. G.; ALMEIDA, C. M.; OLIVEIRA, C. G. Abordagens cognitivas para a identificação de unidades geomorfológicas por meio de dados ASTER/ Terra. **Revista Brasileira de Geociências**, v. 39, n. 2, p. 276-288, 2009.

CARVALHO, M. V. A. **Classificação da cobertura e uso do solo com imagens Worldview-2 utilizando mineração de dados e análise baseada em objetos**. 2011. 201 p. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2011.

CONGALTON, R. G.; GREEN, K. **Assessing the accuracy of remotely sensed data: principles and practices**. Boca Raton: Lewis Publishers, 1999. 137p.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, v. 20, n. 1, p. 37-40, 1960.

CPRM - SERVIÇO GEOLÓGICO BRASILEIRO. **Projeto Rio de Janeiro**. Brasília, 2001. 417 p.

DRAGUT, L.; BLASCHKE, T. Automated classification of landform elements using object-based image analysis. **Geomorphology**, v. 81, n. 3-4, p. 330-344, 2006.

DRAGUT, L.; EISANK, C. Automated object-based classification of topography from SRTM data. **Geomorphology**, v. 141-142, n. 1, p. 21-33, 2012.

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE (ESRI). **ArcGIS**. Version 9.3. Software. 2008.

GIRÃO, R. S.; MELLO, C. L.; FERNANDES, P. J. F.; OLIVEIRA, C. G. Mapeamento Geomorfológico da Bacia Hidrográfica do rio São João por Análise de Imagens Orientada a Objeto e Mineração de Dados. **Revista Brasileira de Geomorfologia**, v. 17, n. 1, p. 3-21, 2016.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA data mining software: an update. **SIGKDD Explorations**, v. 11, n. 1, p. 10-18, 2009.

HEILBRON, M., EEIRADO, L.G. & ALMEIDA, J. (Org). **Mapa Geológico e de Recursos Minerais do Estado do Rio de Janeiro. Escala 1:400.000**. Programa Geologia do Brasil (PGB), Mapas Geológicos Estaduais. CPRM-Serviço Geológico do Brasil, Superintendência Regional de Belo Horizonte. 2014.

HSU, W.; LEE, L. M.; ZHANG, J. Image mining: trends and developments. **Journal of Intelligent Information Systems**, v. 19, n. 1, p. 7-23, 2002.

ITT Visual Information Solutions. **ENVI 4.7**. Version 4.7. Boulder, Colorado, EUA, 2009.

NAVULUR, K. **Multispectral image analysis using the object-oriented paradigm**. Boca Raton, FL: Taylor & Francis Group, 2007. 206p.

QI, F.; ZHU, A. Knowledge Discovery from soil maps using inductive learning. **International Journal of Geographical Information Science**, v. 17, n. 8, p.771-795, 2003.

RADAMBRASIL. **Folhas SF.23/24 Rio de Janeiro/Vitória**. Rio de Janeiro, 1983. 780 p. (Levantamento de Recursos Naturais, 32).

RICCOMINI, C.. **O Rift Continental do Sudeste do Brasil**. Tese de Doutorado, Depto Geologia, IG/USP, São Paulo. 256 p., 1989.

RICCOMINI, C.. **Tectonismo gerador e reformador dos depósitos sedimentares Pós-Gondwânicos da porção centro-ocidental do Estado de São Paulo e áreas vizinhas**. Tese de Livre Docência, Depto Geologia, IG/USP. São Paulo. 198p., 1995.

ROESEL, J. W.; GODOY, R. C. SLAR mosaics for Project RADAM. **Photogrammetric Engineering**, v. 77, n. 5, p.583-595, 1974.

SILVA, L.C.; CUNHA, H.C. **Geologia do Estado do Rio de Janeiro**. In: CPRM – Serviço Geológico do Brasil. Projeto Rio de Janeiro. Brasília: CPRM. 2001.

SILVA, T. M. **A Estruturação geomorfológica do Planalto Atlântico no Estado do Rio de Janeiro**. 2002. 265 p. Tese (Doutorado em Geografia) - Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2002.

TRIMBLE. **eCognition**. Version 8.7. Software. 2011.

VERHAGEN, P.; DRAGUT, L. Object-based landform delineation and classification from DEMs for archaeological predictive mapping. **Journal of Archaeological Science**, v. 39, n. 3, p. 608-703, 2012.

VIEIRA, M. A. **Análise de imagem orientada a objeto e mineração de dados aplicadas ao mapeamento da cultura da cana-de-açúcar**. 2010. 110 p. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.

VIEIRA, M.; FORMAGGIO, A. R.; RENNÓ, C. D.; ATZBERGER, C.; AGUIAR, D. A.; MELLO, M. P. Object based image analysis and data mining applied to a remotely sensed Landsat time-series to map sugarcane over large áreas. **Remote Sensing of Environment**, v.123, p. 533-562, 2012.

WITTEN. H.; FRANK, E.; HALL, M. A. **Data mining practical machine learning tools and techniques**. Morgan Kaufmann Publishers: Burlington. 2011. 629 p.