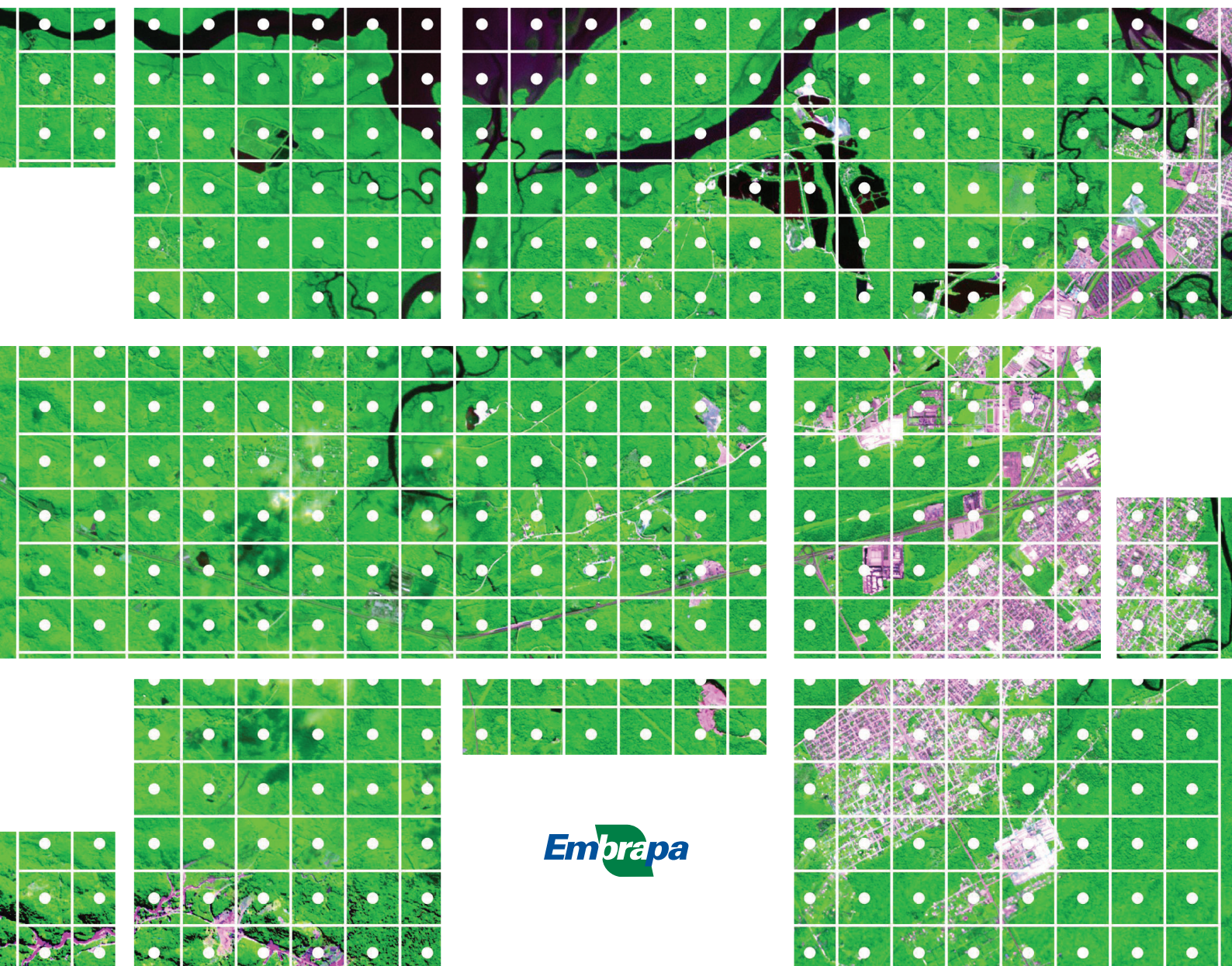


Manual de Análise de Paisagem volume 2

Procedimentos para a execução do
controle de qualidade do mapeamento de uso
e cobertura da terra



**Empresa Brasileira de Pesquisa Agropecuária
Embrapa Florestas
Ministério da Agricultura, Pecuária e Abastecimento**

DOCUMENTOS 317

**Manual de Análise de Paisagem
volume 2**

**Procedimentos para a execução do
controle de qualidade do mapeamento de uso e
cobertura da terra**

*Naissa Batista da Luz
Jéssica Caroline Maran
Maria Augusta Doetzer Rosot
Marilice Cordeiro Garrastazu
Luziane Franciscon
Yeda Maria Malheiros de Oliveira
Joberto Veloso de Freitas*

Exemplares desta publicação podem ser adquiridos na:

Embrapa Florestas

Estrada da Ribeira, km 111, Guaraituba,
Caixa Postal 319
83411-000, Colombo, PR, Brasil
Fone: (41) 3675-5600
www.embrapa.br/florestas
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações da
Embrapa Florestas

Presidente
Patrícia Póvoa de Mattos

Vice-Presidente
José Elidney Pinto Júnior

Secretária-Executivo
Neide Makiko Furukawa

Membros
*Álvaro Figueredo dos Santos, Gizelda Maia Rego,
Guilherme Schnell e Schühli, Ivar Wendling, Luis
Cláudio Maranhão Froufe, Maria Izabel Radomski,
Marilice Cordeiro Garrastazu, Valderés Aparecida
de Sousa*

Supervisão editorial
José Elidney Pinto Júnior

Revisão de texto
José Elidney Pinto Júnior

Normalização bibliográfica
Francisca Rasche

Projeto gráfico da coleção
Carlos Eduardo Felice Barbeiro

Editoração eletrônica
Neide Makiko Furukawa

1ª edição
Versão digital (2018)

Todos os direitos reservados.

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

Dados Internacionais de Catalogação na Publicação (CIP)

Nome da unidade catalogadora

Manual de análise da paisagem: vol. 2: procedimento para a execução do
mapeamento de uso e cobertura da terra. [recurso eletrônico] / Naissa
Batista da Luz ... [et al.]. - Colombo : Embrapa Florestas, 2018.
v. 2 : 37 p. : il. color. - (Documentos / Embrapa Florestas, ISSN 1980-3958
; 317)

Modo de acesso:

<<http://www.infoteca.cnptia.embrapa.br/handle/item/221>>

1. Sistemas de informação geográfica. 2. Análise de imagens. 3.
Monitoramento ambiental. I. Luz, Naissa Bastista da. II. Maran, Jéssica
Caroline. III. Rosot, Maria Augusta Doetzer. IV. Garrastazu, Marilice Cordeiro.
V. Franciscan, Luziane. VI. Oliveira, Yeda Maria Malheiros de. VII. Freitas,
Joberto Veloso de. X. Série.

CDD (21. ed.) 630.25158162

Autores

Naissa Batista da Luz

Engenheira Florestal, doutora em Ciências Geodésicas, consultora da FAO, Colombo, PR

Jéssica Caroline Maran

Engenheira Florestal, mestre em Engenharia Florestal, consultora da FAO, Colombo, PR

Maria Augusta Doetzer Rosot

Engenheira Florestal, doutora em Engenharia Florestal, pesquisadora da Embrapa Florestas, Colombo, PR

Marilice Cordeiro Garrastazu

Engenheira Florestal, mestre em Engenharia Agrícola, pesquisadora da Embrapa Florestas, Colombo, PR

Luziane Franciscon

Estatística, mestre em Estatística e Experimentação Agronômica, analista da Embrapa Florestas, Colombo, PR

Yeda Maria Malheiros de Oliveira

Engenheira Florestal, doutora em Ciências Florestais, pesquisadora da Embrapa Florestas, Colombo, PR

Joberto Veloso de Freitas

Engenheiro Florestal, doutor em Manejo Florestal, Serviço Florestal Brasileiro, Brasília, DF

Apresentação

O Inventário Florestal Nacional (IFN-BR) visa fundamentar a formulação e implementação de políticas públicas de desenvolvimento, uso e conservação dos recursos florestais, bem como a gestão desses recursos, por meio de informações suficientes, confiáveis e periodicamente atualizadas, coletadas no campo. Estudos em escala de paisagem, por sua vez, também têm adquirido importância fundamental dentro desse contexto, por permitirem a avaliação da dinâmica de uso e cobertura da terra e das interações entre padrões espaciais e processos ecológicos, envolvendo florestas e outros usos da terra, como a agricultura.

Por sua vez, a Política Nacional de Meio Ambiente no Brasil tem por objetivo a preservação, melhoria e recuperação da qualidade ambiental propícia à vida, visando assegurar, no país, condições ao desenvolvimento socioeconômico, aos interesses da segurança nacional e à proteção da dignidade da vida humana. Dessa maneira, o objetivo geral da Política Nacional do Meio Ambiente está dividido em preservação, melhoramento e recuperação do meio ambiente. Para poder implementar ações relativas a tais objetivos, requer-se instrumentos de diagnóstico e análise que permitam mapear e avaliar ecossistemas e os respectivos serviços que proveem em seus territórios, ambos espacialmente explícitos. Daí advém a necessidade de incluir dados e indicadores espaciais na análise. Nessa avaliação, torna-se também necessário traduzir o resultado de abordagens técnico-científicas em informação compreensível para a implementação de políticas públicas e tomada de decisão, o que pode ser feito por meio de mapas, indicadores, relatórios e gráficos.

Assim, a Embrapa Florestas desenvolveu, em conjunto com o Serviço Florestal Brasileiro (SFB) e com o suporte da Food and Agriculture Organization of the United Nations (FAO) e do BID (Banco Interamericano de Desenvolvimento), uma metodologia para a análise espacial da estrutura da paisagem no contexto do IFN-BR. O objetivo do denominado Componente Geoespacial dentro do projeto do IFN-BR, ao longo das diferentes edições, é observar a dinâmica de uso das florestas por meio de imagens orbitais, em escalas compatíveis com os interesses nacionais e estaduais e utilizando, para tanto, indicadores como mudanças no uso da terra e fragmentação florestal. De tal maneira, pretende-se produzir informações sobre a importância e qualidade dos recursos florestais em relação a outros usos da terra, em escala de paisagem, ressaltando suas funções, qualidade e pressões incidentes, de modo a subsidiar a formulação de políticas públicas que sejam apropriadas à região e à sua escala de abordagem.

A análise da paisagem complementa outros dois componentes do IFN-BR, a coleta de dados em campo e o levantamento socioeconômico, além de ser destinada ao monitoramento do componente florestal em escala de paisagem, ao longo do tempo. Nesse contexto estão inseridas as chamadas Unidades Amostrais de Paisagem (UAPs), por meio das quais é implementado o Componente Geoespacial do IFN-BR, concebidas para oferecer um ferramental que permita ao usuário a visualização de aspectos da paisagem conjugados sob a forma de indicadores e seus respectivos índices. As UAPs são áreas amostrais permanentes de 100 km², distribuídas sistematicamente em um grid de 40 km x 40 km sobre todo o território nacional, perfazendo um total de aproximadamente 5.500 unidades. Todas contêm uma Unidade Amostral de Campo do IFN-BR, na forma de um conglomerado, localizada em seu centro geométrico. Dessa maneira, a estratégia adotada foi desenvolver a metodologia de todos os componentes do IFN-BR com vistas à sua integração e posterior análise conjunta.

Uma vez que proporcionam a possibilidade tanto de análises estáticas, ou seja, em apenas uma data, quanto dinâmicas - quando os índices são calculados para sucessivas ocasiões -, as UAPs constituem unidades de diagnóstico e de monitoramento. A base para o cálculo dos índices de paisagem e posteriores análises é o mapa de uso e cobertura da terra, obtido por meio de processos de classificação e análise de imagens orientada a objetos. Assim, os indicadores e índices de paisagem permitem estabelecer um diagnóstico integrado de cada UAP que, por sua vez, reflete uma determinada combinação de características biogeoclimáticas (classe territorial ou ecorregião), associadas a fatores de influência antrópicos ou naturais, ocorrentes naquela localidade.

Apesar da utilização de técnicas sofisticadas para o mapeamento temático, algumas limitações ainda são encontradas, de forma que, nas classificações decorrentes de processos automatizados, como é o caso da metodologia utilizada na elaboração dos mapas de uso e cobertura da terra nas UAPs do IFN-BR, erros resultantes de interações complexas entre as estruturas espaciais da paisagem, a resolução do sensor, os algoritmos de pré-processamento e os procedimentos da própria classificação podem ocorrer (Campbell, 1996). Assim, a qualidade dos mapas de uso e cobertura da terra, utilizados para a análise de paisagem do IFN-BR, deve ser conhecida, tendo em vista que o cálculo dos indicadores e índices de paisagem deriva diretamente deste produto. A qualidade dos resultados gerados pelos índices de paisagem está diretamente relacionada à qualidade dos mapas de uso e cobertura da terra.

O presente manual se atém à descrição detalhada dos procedimentos relativos à avaliação da acuracidade temática dos mapas de uso e cobertura da terra do INF-BR.

Sérgio Gaiad

Chefe Adjunto de Pesquisa e Desenvolvimento

Lista de Figuras

Figura 1. Janelas do software ArcGIS (versão 10.4.1) indicando a localização da ferramenta <i>Create Fishnet</i> no <i>ArcToolbox</i> (a), bem como os <i>inputs</i> para a criação da rede de pontos (b).	20
Figura 2. Janela do software ArcGIS (versão 10.4.1) exibindo a grade vetorial e os pontos de referência criados pela ferramenta <i>Create Fishnet</i> para a UAP1214, utilizando a extensão do mosaico <i>RapidEye</i> em composição NIR/RE/R.....	22
Figura 3. Janela do software ArcGIS exibindo pontos de referência que contemplam a classe 08 (“Influência Urbana”), selecionados em azul e destacados na tabela de atributos do arquivo de pontos.	23
Figura 4. Exemplos de pontos de referência ocorrentes na UAP1444, localizada na Bahia, ilustrando áreas de “Floresta Plantada”, “Agricultura e Pastagem”, e “Floresta Natural”, no mosaico <i>RapidEye</i> (5R, 4G, 3B), e na imagem do Google Earth, relativa ao ano de 2014 (b).	24
Figura 5. Exemplos de diferentes áreas pertencentes à classe “Floresta Plantada”, ilustrando a resposta espectral da imagem <i>RapidEye</i> (5R, 4G, 3B) em talhões temporariamente não estocados devido ao recente corte raso (b), plantios jovens (a) e plantios adultos (c), bem como sua correspondência nas imagens do Google Earth (e) (d) (f).....	25
Figura 6. Exemplos de diferentes áreas pertencentes à classe “Gramíneas e Herbáceas”, ilustrando a resposta espectral da imagem <i>RapidEye</i> (5R, 4G, 3B) em localidades com presença de vegetação herbácea, com ocorrência esparsa de elementos arbóreo/arbustivos, devido à presença de solo raso (a), e em áreas úmidas (b) (c), e sua correspondência nas imagens do Google Earth (d) (e) (f).....	25
Figura 7. Janelas do software ArcGIS (versão 10.4.1) ilustrando a ferramenta <i>Split By Attribute</i> e os arquivos <i>geodatabase</i> criados por ela.	26
Figura 8. Janelas do software ENVI Classic ilustrando o processo de preparo de uma imagem classificada para posterior cálculo da matriz de erro.....	27
Figura 9. Janelas do software ENVI Classic ilustrando o processo de preparo dos pontos de referência representando a verdade de campo para posterior cálculo da matriz de erro.....	28
Figura 10. Janelas do software ENVI Classic ilustrando o processo de cálculo de uma matriz de erro a partir de pontos da verdade de campo.....	29
Figura 11. Matrizes de erro geradas por meio do software ENVI, referentes à avaliação da acuracidade do mapa de uso e cobertura da terra para a UAP1214, localizada no Paraná, em valores absolutos (a) e em valores percentuais (b).....	30
Figura 12. Coeficientes de concordância gerados pelo software ENVI, referentes aos erros de comissão e omissão das classes de uso e cobertura da terra (a), bem como os valores de acuracidade do produtor e do usuário (b), em valores percentuais e em pixels.	30
Figura 13. Janela do software ArcGIS (versão 10.4.1) ilustrando a escala que deve ser utilizada para a avaliação qualitativa do mapa de uso e cobertura, bem como um erro de atribuição entre as classes, apontado pela seta em vermelho, onde um polígono que deveria ser classificado como pertencente à classe “Floresta Natural” (01), recebeu atribuição da classe “Floresta Plantada” (05).....	35

Lista de Tabelas

Tabela 1. Matriz de erro, onde os elementos em destaque representam a diagonal principal da matriz que contém os casos em que a classe representada na classificação da imagem e os pontos de referência representando a verdade de campo coincidem, enquanto os elementos fora desta diagonal contemplam os casos em que existe divergência entre esses.	16
Tabela 2. Matriz de erro para quatro classes, onde as entradas das células (p_{ij}) representam a proporção da área da população (amostra) que tem classe da imagem classificada i e classe da referência j	17
Tabela 3. Descrição sucinta e codificação das classes de uso e cobertura da terra das UAPs do IFN-BR.	21
Tabela 4. Número de polígonos e superfícies absoluta e percentual ocupadas pelas classes de uso e cobertura da terra, calculadas a partir do mapeamento na área de abrangência de 11 km x 11 km, na projeção UTM, Fuso 22S, WGS84, para a UAP1214.....	31
Tabela 5. Matriz de confusão referente à avaliação da acuracidade do mapa de uso e cobertura da terra em pixels, para a UAP1214.	31
Tabela 6. Matriz de erros ponderada pela proporção de área estimada para cada classe de uso e cobertura da terra, para a UAP1214.....	32
Tabela 7. Acurácias do produtor e do usuário ponderadas pela área, por classe de uso e cobertura da terra.	33
Tabela 8. Valores de referência para a acurácia da classificação do uso e cobertura da terra nas Unidades Amostrais de Paisagem do IFN-BR.	34

Sumário

1	Validação dos mapas de uso e cobertura da terra	13
1.1	Desenho amostral	14
1.2	Estratégia de verificação	14
1.3	Análise e estimativa dos resultados.....	16
1.4	Criação de uma rede de pontos.....	19
1.5	Classificação dos pontos de referência	22
1.6	Cálculo da matriz de erro	27
1.7	Critérios quantitativos e qualitativos para a avaliação do produto	34
	Referências	36

1 Validação dos mapas de uso e cobertura da terra

A validação de um mapa de uso e cobertura da terra refere-se a um conjunto de técnicas necessárias à determinação da sua qualidade, que se divide em três pontos principais: o desenho amostral, a estratégia de verificação e o protocolo de análise e estimativa dos resultados (Stehman; Czaplewski, 1998). Inclui a avaliação da acurácia por meio da comparação entre os resultados da imagem classificada e outras fontes de informação, em observações baseadas em um desenho amostral probabilístico, permitindo a estimativa da acurácia global, da acurácia do produtor e do usuário e dos erros de omissão e comissão (Strahler et al., 2006). A avaliação da acurácia é um requisito básico no processo de elaboração e distribuição de mapas temáticos e uma condição necessária para a comparação de resultados e o uso dos produtos de um mapeamento (Stehman; Wickham, 2011; Radoux; Bogaert, 2017). Pontos de referência verificados em campo ou em imagens de melhor escala espacial permitem quantificar a qualidade dos resultados gerados (Stehman; Czaplewski, 1998; Stehman, 2009).

Muitas são as estratégias recomendadas para a avaliação da qualidade de classificações temáticas, sobretudo quando a análise orientada a objetos é adotada. De acordo com Radoux e Bogaert (2017), as boas práticas de avaliação da acuracidade temática envolvendo a avaliação de matrizes de erro descritas por Olofsson et al. (2014) e Strahler et al. (2006) podem ser adotadas em muitos casos, mesmo quando a segmentação de imagens e a classificação orientada a objetos é utilizada. As boas práticas de Olofsson et al. (2014) também são adotadas no guia prático elaborado pela Organização das Nações Unidas para Alimentação e Agricultura (FAO) (2016), e serão seguidas neste trabalho.

Olofsson et al. (2014) recomendam que a avaliação da acurácia leve em consideração a proporção de área com base na amostragem dos pontos de referência. As boas práticas incluem:

- i. implementação de um desenho amostral probabilístico selecionado para atingir os objetivos prioritários da estimativa de área e de acurácia ao mesmo tempo em que existem restrições práticas como os custos e a disponibilidade de fontes de referências;
- ii. implementação de um protocolo de verificação baseado em fontes de dados de referência que forneçam informações suficientemente detalhadas, em termos de resolução espacial e temporal, para a correta rotulagem de cada ponto de referência ;
- iii. implementação de uma forma de análise consistente com o desenho amostral e o protocolo de verificação;
- iv. sumarização da avaliação da acurácia por meio da matriz de erros, em termos da proporção de área e estimativa da acurácia global, acurácia do usuário (ou erro de comissão) e acurácia do produtor (ou erro de omissão);
- v. estimativa da área por classe baseada na classificação dos pontos de referência;
- vi. quantificação da incerteza pela inclusão do intervalo de confiança, para os parâmetros de acurácia e área;
- vii. avaliação da variabilidade e do erro potencial dos pontos de referência da classificação;
- viii. documentação dos desvios das boas práticas que podem substancialmente afetar os resultados.

A seguir são apresentados os conceitos básicos adotados em consonância com as boas práticas propostas por Olofsson et al. (2014) e os procedimentos necessários para a validação do mapeamento de uso e cobertura da terra.

1.1 Desenho amostral

O desenho amostral consiste no protocolo para a seleção de pontos de referência, sendo que a recomendação mais crítica é a de que as amostras sejam selecionadas de forma randômica, onde a probabilidade de inclusão de cada amostra seja conhecida e maior que zero para cada uma delas (Olofsson et al., 2014).

No caso das unidades amostrais de paisagem do IFN-BR, foi selecionada a amostragem sistemática por meio de um grid de pontos equidistantes entre si em 500 m. O ponto de referência corresponde ao pixel onde o ponto se encontra, tendo área igual àquela do pixel, ou seja, 25 m². Considera-se este formato apropriado ao componente Geoespacial do IFN-BR, pois confere padronização, bastante desejável em um inventário nacional dada a necessidade de atuação de diferentes profissionais na elaboração dos produtos e a sistematização dos procedimentos. A implementação do grid e a interpretação dos pontos neste modelo são práticas e podem ser realizadas em tempo inferior àquele quando amostras são distribuídas de forma aleatória ou estratificada.

A representatividade amostral deste tipo de grid foi comprovada quando experimentos elaborados com diferentes densidades de pontos foram avaliados, incluindo redes de pontos equidistantes com 100 m, 250 m, 500 m, 750 m e 1.000 m, em comparação com uma interpretação visual e avaliação de todos os pixels da imagem, ou seja, um censo. A partir de 500 m, percebeu-se que o adensamento do grid não mais contribui para a melhoria na estimativa da acurácia da classificação. Outro aspecto observado é que, para um grid com pontos equidistantes a 500 m, a proporção da área das classes na UAP reflete-se na proporção da área por pontos de referência por classe.

A seleção sistemática de unidades amostrais é capaz de fornecer estimativas mais precisas que a seleção randômica simples e o desenho amostral sistemático é bastante apropriado para inventários florestais nacionais, já que as amostras de campo seguem a mesma configuração, o que permitiria a coleta de informações durante as campanhas de campo para a validação da imagem classificada (Olofsson et al., 2014).

1.2 Estratégia de verificação

No contexto da avaliação de acurácia, a estratégia de verificação compreende todos os aspectos do protocolo que determinam a concordância entre a classe atribuída à unidade amostral na imagem classificada e a mesma área na referência. Os principais elementos da estratégia de verificação são as dimensões das amostras de referência, a fonte de informações utilizada como referência, o protocolo de rotulagem e as definições de concordância entre as classificações da imagem e da referência (Olofsson et al., 2014). As características desses elementos no âmbito do Componente Geoespacial no IFN-BR serão apresentadas a seguir.

A menor dimensão para um ponto de referência e a mais comumente adotada na maioria das estratégias para avaliação da acuracidade temática de imagens classificadas é o pixel, sendo bastante conhecidos os protocolos e as formas de análise associadas com a implementação de uma amostragem baseada em pixels (Stehman; Wickham, 2011). A utilização de polígonos como pontos de referência apresenta alguns problemas práticos. A possibilidade de ocorrência de múltiplas classes

de uso e cobertura em um mesmo polígono, por exemplo, leva à subjetividade na definição do percentual de ocorrência de cada classe. A necessidade de avaliação das diferenças entre os limites do polígono de referência e da imagem por meio do delineamento manual também leva à subjetividade. Por se tratar de um inventário florestal nacional, a sistematização é necessária, buscando-se protocolos que permitam a operacionalização e a padronização de processos. Considerando-se que a segmentação da imagem é um processo padronizado por um algoritmo que define automaticamente, para cada imagem, o fator de escala a ser adotado e que as imagens utilizadas são ortorretificadas e tiveram sua acurácia posicional avaliada (Costa et al., 2015), não serão realizadas avaliações dos limites dos polígonos. O algoritmo implementado define o fator de escala pelo incremento deste parâmetro, de forma iterativa, otimizando a homogeneidade espectral e o tamanho dos polígonos, até que a maioria absoluta dos polígonos (90%) tenha área maior ou igual à unidade mínima de mapeamento (500 m²), levando em conta a similaridade espectral, conforme definido no Manual de análise de paisagem: volume 1: procedimentos para a execução do mapeamento de uso e cobertura da terra (Luz et al., 2018).

A impossibilidade de visitas ao campo, com maior densidade do que a já praticada pelo levantamento realizado pelo IFN-BR e pela necessidade de intensificação dos pontos amostrais para a avaliação da acurácia da classificação das Unidades Amostrais de Paisagem, torna necessária a seleção de fonte alternativa de referência. A primeira fonte de dados de referência consiste no próprio mosaico das imagens *RapidEye* (RE), utilizado para a elaboração do mapa de uso e cobertura. Porém, dados de maior resolução espacial e temporal devem ser utilizados como fonte de referência complementar às imagens *RapidEye*, tendo-se selecionado as imagens do Google Earth para tal finalidade. Desta forma, é válido o tipo de uso e cobertura interpretado a partir das imagens *RapidEye*, utilizando-se as imagens Google Earth como um apoio na interpretação do contexto, no qual a unidade amostral em avaliação está inserida. Além de estarem disponíveis para praticamente todo o país, têm melhor qualidade para a interpretação visual, pois a resolução espacial é maior e a resolução temporal permite a análise de informações pretéritas, essenciais em algumas classes de uso do solo, como os plantios florestais, por exemplo. Desta forma, uma densidade de aproximadamente 500 pontos pode ser levantada rapidamente e com grande riqueza de informações auxiliares. A cobertura, no entanto, não está disponível em alta resolução e com séries históricas para todo o país, porém segue sendo a melhor alternativa, já que apenas uma pequena percentagem tem cobertura deficiente. A utilização dessa fonte de informação tem sido amplamente adotada nos últimos anos e consta das boas práticas citadas por Olofsson et al. (2014). Uma prática recomendada pelos autores quando imagens do Google Earth são utilizadas é a data de aquisição das cenas, que consta na parte inferior da tela. A data mais próxima possível da data de aquisição das imagens *RapidEye* utilizadas para a elaboração do mosaico deve ser preferida.

O ponto central de cada UAP contém uma unidade amostral de campo do IFN-BR, onde valiosas informações são coletadas, incluindo fotografias da unidade amostral de campo e arredores. Estes dados devem ser utilizados para a interpretação do uso e cobertura da terra, principalmente os tipos de vegetação que ocorrem na área. A estratégia de rotulagem das classes de uso e cobertura da terra deve seguir a classificação e a descrição fornecida pelo IFN-BR e completamente descrita em Luz et al. (2018). O sistema de classificação de uso e cobertura da terra da FAO foi adotado e compatibilizado com o sistema oficial de classificação da vegetação brasileira do IBGE (IBGE, 2012). O processo de rotulagem da referência e a sua comparação com os resultados da classificação não é tarefa trivial, já que muitas definições para os diferentes tipos de uso e cobertura podem ser adotadas, e o sistema da FAO facilita esta transição (Ahlqvist, 2008).

1.3 Análise e estimativa dos resultados

Esta etapa inclui os protocolos de como quantificar a acurácia, a formulação e a abordagem de inferência a serem adotadas para a estimativa da acurácia e da área e a quantificação da incerteza dessas estimativas. A matriz de erros tem papel principal nesta etapa (Olofsson et al., 2014).

As matrizes de erro são o meio mais comum de se expressar a acuracidade de uma classificação temática (Rosot, 2001; Foody, 2002), fornecendo a base para descrever a precisão da classificação e caracterizar os erros, ajudando a refinar a classificação ou estimativas que possam derivar dessa (Foody, 2002; Figueiredo; Vieira, 2007). Tais matrizes comparam, classe a classe, a relação entre os dados de referência conhecidos e os respectivos resultados de uma classificação automática (Lillesand; Kiefer, 1994).

A matriz de erro é uma matriz quadrada, de dimensão $q \times q$, onde q é o número de classes. Os elementos da matriz apresentam o número de pontos que representam uma classe da imagem classificada “ i ” e da verdade de campo “ j ” (n_{ij}), amostrados para cada classe. Nas linhas se encontram o número de pontos amostrais de cada classe segundo a classificação automática e nas colunas aqueles da referência. Dessa forma, a diagonal principal da matriz conterá o número de pixels classificados corretamente para cada categoria, enquanto as células fora da diagonal apresentam os erros de omissão e comissão (Tabela 1). A partir dessa matriz, a avaliação da acurácia pode ser obtida por meio de coeficientes de concordância, que podem ser expressos como concordância total ou para classes individuais (Lillesand; Kiefer, 1994).

Tabela 1. Matriz de erro, onde os elementos em destaque representam a diagonal principal da matriz que contém os casos em que a classe representada na classificação da imagem e os pontos de referência representando a verdade de campo coincidem, enquanto os elementos fora desta diagonal contemplam os casos em que existe divergência entre esses.

		Referência				
		A	B	C	D	Σ
Classificação	A	n_{AA}	n_{AB}	n_{AC}	n_{AD}	n_{A+}
	B	n_{BA}	n_{BB}	n_{BC}	n_{BD}	n_{B+}
	C	n_{CA}	n_{CB}	n_{CC}	n_{CD}	n_{C+}
	D	n_{DA}	n_{DB}	n_{DC}	n_{DD}	n_{D+}
	Σ	n_{+A}	n_{+B}	n_{+C}	n_{+D}	n

Fonte: Adaptado de Foody (2002).

A contagem absoluta do número de pixels na amostra n_{ij} (Figura 11) pode ser convertida na proporção de área estimada \hat{p}_{ij} (Tabela 2) com a Equação 1, quando o estrato corresponde às classes do mapa quando a amostragem aleatória simples, sistemática ou estratificada aleatória é utilizada. Dessa forma, os valores na matriz de erros representam a proporção da área e não a contagem de pixels (Tabela 2).

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_i} \quad \text{Equação 1}$$

Em que:

\hat{p}_{ij} = proporção de área estimada para a classe da imagem classificada i e a classe de referência j ;

W_i = proporção de área da classe da imagem classificada i , calculada pela divisão da área ocupada pela classe i sobre o total de área da imagem classificada;

n_{ij} = número de pixels para a classe da imagem classificada i e a classe de referência j ;

n_i = número total de pixels para a classe i .

Tabela 2. Matriz de erro para quatro classes, onde as entradas das células (\hat{p}_{ij}) representam a proporção da área da população (amostra) que tem classe da imagem classificada i e classe da referência j .

		Referência				Total
		Classe 1	Classe 2	Classe 3	Classe 4	
Classificação	Classe 1	p_{11}	p_{12}	p_{13}	p_{14}	$p_{1.}$
	Classe 2	p_{21}	p_{22}	p_{23}	p_{24}	$p_{2.}$
	Classe 3	p_{31}	p_{32}	p_{33}	p_{34}	$p_{3.}$
	Classe 4	p_{41}	p_{42}	p_{43}	p_{44}	$p_{4.}$
	Total	$p_{.1}$	$p_{.2}$	$p_{.3}$	$p_{.4}$	1

A acurácia global, a do produtor (por classe) e do usuário (por classe) podem ser estimadas, respectivamente, pelas Equações 2, 3 e 4 a seguir.

a. **Acurácia global:** proporção de área classificada corretamente.

$$\hat{A} = \sum_{j=1}^q \hat{p}_{jj}$$

Equação 2

Em que:

\hat{A} = estimativa de acurácia global para cada classe j , onde j varia de 1 a q ;

\hat{p}_{jj} = proporção estimada para a classe j ponderada pela proporção de área estimada.

b. **Acuracidade do produtor para a classe j :** proporção da área da referência para a classe j que é mapeada como a classe j na imagem; tem como medida complementar o **erro de omissão da classe j** , $1 - \hat{p}_{jj}/\hat{p}_{.j}$:

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{.j}}$$

Equação 3

Em que:

\hat{P}_j = estimativa da acuracidade do produtor para a classe j ;

\hat{p}_{ij} = proporção estimada para a classe j ponderada pela proporção de área estimada;

\hat{p}_j = total da proporção estimada para a classe j ponderada pela proporção de área estimada.

- c. **Acuracidade do usuário para a classe i** : proporção da área mapeada como classe i que tem na referência a classe i ; tem como medida complementar o **erro de comissão da classe i** , $1 - \hat{p}_{ii}/\hat{p}_i$:

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_i} \quad \text{Equação 4}$$

Em que:

\hat{U}_i = estimativa da acuracidade do usuário para a classe i ;

\hat{p}_{ii} = proporção estimada para a classe i ponderada pela proporção de área estimada (W_i);

\hat{p}_i = total da proporção estimada para a classe i ponderada pela proporção de área estimada;

O **coeficiente de Kappa** apresenta diversos problemas (Pontius Junior; Millones, 2011). Sua utilização não é recomendada por Olofsson et al. (2014) e, portanto, não deverá ser reportado.

Com relação à acurácia global, deve ser informado seu valor e o respectivo intervalo de confiança, dado pela Equação 5 (Olofsson et al., 2014).

d. Intervalo de confiança para a acurácia global

$$IC = \hat{A} \pm t_{tab} * \sqrt{\hat{V}(\hat{A})} \quad \text{Equação 5}$$

Em que:

IC = intervalo de confiança para a acurácia global;

\hat{A} = estimativa da acurácia global;

t_{tab} = valor de teste t para 95% de probabilidade encontrado na tabela t de Student ($t_{tab} = 1,96$);

$\hat{V}(\hat{A})$ = variância estimada para a acurácia global, dada pela Equação 6.

e. Variância estimada para a acurácia global

$$\hat{V}(\hat{A}) = \sum_{i=1}^q W_i^2 \hat{U}_i (1 - \hat{U}_i) / (n_i - 1) \quad \text{Equação 6}$$

Em que:

$\hat{V}(\hat{A})$ = variância estimada para a acurácia global;

W_i = proporção de área da classe da imagem classificada i ;

\hat{U}_i = estimativa da acuracidade do usuário para a classe i ;

n_i = número total de pixels para a classe i .

1.4 Criação de uma rede de pontos

A avaliação do mapa de uso e cobertura da terra deve ser realizada por meio de um grid de pontos distribuídos a cada 500 metros, estabelecido de acordo com os limites da UAP. Considerando que os mapas de uso e cobertura da terra das unidades amostrais de paisagem serão processados para uma área de 11 km x 11 km, essa mesma área deve ser adotada na criação da rede de pontos. Para a criação dessa rede de pontos pode-se utilizar a ferramenta “*Create Fishnet*” do software ArcGIS, localizada no diretório *ArcToolbox>Data>Management Tools>Sampling>Create Fishnet* (Figura 1 (a)). O arquivo *shapefile* contendo os limites da UAP corresponde aos limites do grid de pontos (“*Template Extent*”) (Figura 1 (b)). Na janela da ferramenta “*Create Fishnet*” também é necessário indicar o tamanho da célula do grid a ser criada (“*Cell Size Width*” e “*Cell Size Height*”), ou seja, a distância de 500 metros entre os pontos de referência amostrais, e marcar a opção para a criação dos pontos (“*Create Label Points*”) (Figura 1 (b)).

No software QGIS é possível criar a mesma rede de pontos de forma similar ao processo previamente exposto, utilizando-se primeiro a ferramenta “Grade vetorial”, localizada na barra de ferramentas principal do programa, em *Vetor>Investigar>Grade Vetorial* e, posteriormente, aplicando-se a ferramenta “Centroide do Polígono”, com base na grade vetorial recém-criada, localizada também na barra de ferramentas principal, em *Vetor>Geometrias>Centroide do Polígono*.

O processo descrito gera dois arquivos *shapefile*, um com feições do tipo polígono, representando a grade vetorial composta por quadrados de 500 m por 500 m, e o outro com feições do tipo ponto, localizadas ao centro de cada polígono. Os 484 pontos de referência gerados devem ser classificados de acordo com a legenda pré-definida (Tabela 3), para avaliação e validação do mapa de uso e cobertura, representando a verdade de campo a ser utilizada no cálculo da matriz de erro e demais coeficientes de concordância. Na Figura 2 pode-se observar os arquivos criados pela ferramenta “*Create Fishnet*”, para uma das unidades amostrais de paisagem do estado do Paraná.

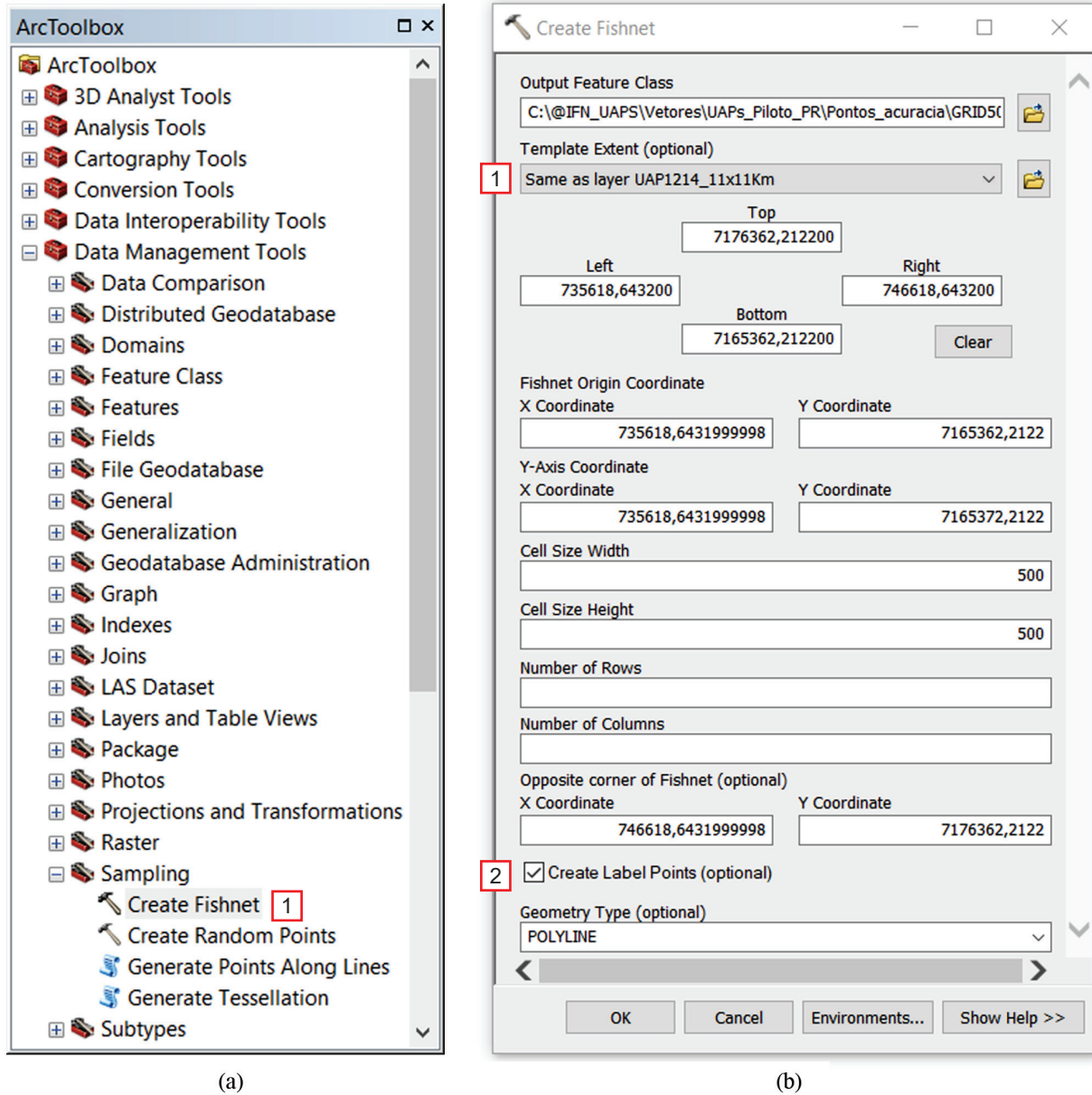


Figura 1. Janelas do software ArcGIS (versão 10.4.1) indicando a localização da ferramenta *Create Fishnet* no *ArcToolbox* (a), bem como os *inputs* para a criação da rede de pontos (b).

Tabela 3. Descrição sucinta e codificação das classes de uso e cobertura da terra das UAPs do IFN-BR.

Classe de uso e cobertura	Código	Descrição
Floresta Natural	01	Florestas são definidas, de acordo com FAO (2016a), como áreas estendendo-se a mais do que 0,5 ha com árvores de mais de 5 m de altura e uma cobertura de copa maior que 10%, ou árvores capazes de atingir estes limiares in situ. Não incluem áreas de uso predominantemente agrícola ou urbano. Assim, florestas são definidas tanto pela presença de árvores quanto pela ausência de outros usos da terra predominantes.
Outras Terras com Vegetação Lenhosa	02	Conceitualmente, "Outras Terras com Vegetação Lenhosa" são áreas não classificadas como florestas, que se estendem por mais de 0,5 ha, com árvores maiores que 5 m e cobertura de copas entre 5% e 10%; ou árvores capazes de atingir estes limiares in situ; ou com uma combinação de indivíduos arbóreos, arbustivos e de árvores com mais de 10% de cobertura do dossel.
Outras Terras com Árvores	03	Árvores isoladas, ou agrupadas, que não se enquadram na definição de floresta. Estão associadas a outros usos, como áreas de agricultura e pastagem, ou influência urbana. Contemplam polígonos que tenham área superior a 0,05 ha (500 m ²) e inferior a 0,5 ha (5.000 m ²).
Gramíneas e Herbáceas	04	Corresponde a áreas naturais caracterizadas por um estrato predominantemente herbáceo, com eventual ocorrência de elementos arbóreo/arbustivo, desde que esparsamente distribuídos, sobre um tapete gramíneo-lenhoso. Encontram-se disseminadas por diferentes regiões fitogeográficas, compreendendo diferentes tipologias, com diversos graus de antropização (IBGE, 2013).
Floresta Plantada	05	Classe relativa aos plantios ou formações de maciços com espécies florestais nativas ou exóticas, puros ou consorciados. Nesta definição não se considera se o plantio é realizado em áreas anteriormente povoadas com espécies florestais ou não; considera-se reflorestamento todas as áreas povoadas com essências florestais, independentemente do ambiente.
Solo Exposto	06	Classe que inclui áreas sem cobertura vegetal, podendo ser ocupada por atividades de mineração, exploração de jazidas, lavras, bem como extração de areia. Apresentam, em geral, coloração clara, textura lisa e forma variável.
Agricultura e Pastagem	07	Engloba áreas utilizadas para a produção de alimentos, fibras e commodities do agronegócio, incluindo todas as terras cultivadas, caracterizadas pelo delineamento de áreas cultivadas ou em descanso, podendo também compreender áreas alagadas (IBGE, 2013) que sejam utilizadas para essa mesma finalidade. Classe constituída por culturas temporárias, cíclicas e permanentes. Inclui, ainda, as plantas hortícolas, floríferas, medicinais, aromáticas e condimentares de pequeno porte. Inclui lavouras semipermanentes e plantios lenhosos perenes, como os de frutíferas.
Influência Urbana	08	A essa nomenclatura estão associados os tipos de uso da terra de natureza antrópica intensiva, estruturadas por edificações e sistema viário, onde predominam as superfícies artificiais não agrícolas. Estão incluídas nesta categoria as metrópoles, cidades, vilas, áreas de rodovias, serviços e transporte, energia, comunicações e terrenos associados, bem como edificações presentes no meio rural (casas, galpões, estábulos, entre outras)
Superfície com Água	09	Áreas compostas pelo oceano, por rios, lagos, lagoas, barragens, represas, canais naturais ou artificiais, tanques d'água, com área igual ou superior a um hectare.
Dunas e Afloramentos Rochosos	10	Áreas referentes a ambientes naturais, sem cobertura vegetal, como afloramentos rochosos, praias, dunas e extensões de areia ou seixos (IBGE, 2013).
Área não Observada	00	Áreas não observadas por impedimentos tais como erro no sistema de imageamento, obstruções como nuvens e sombra de nuvens e demais condições de observação desfavoráveis.

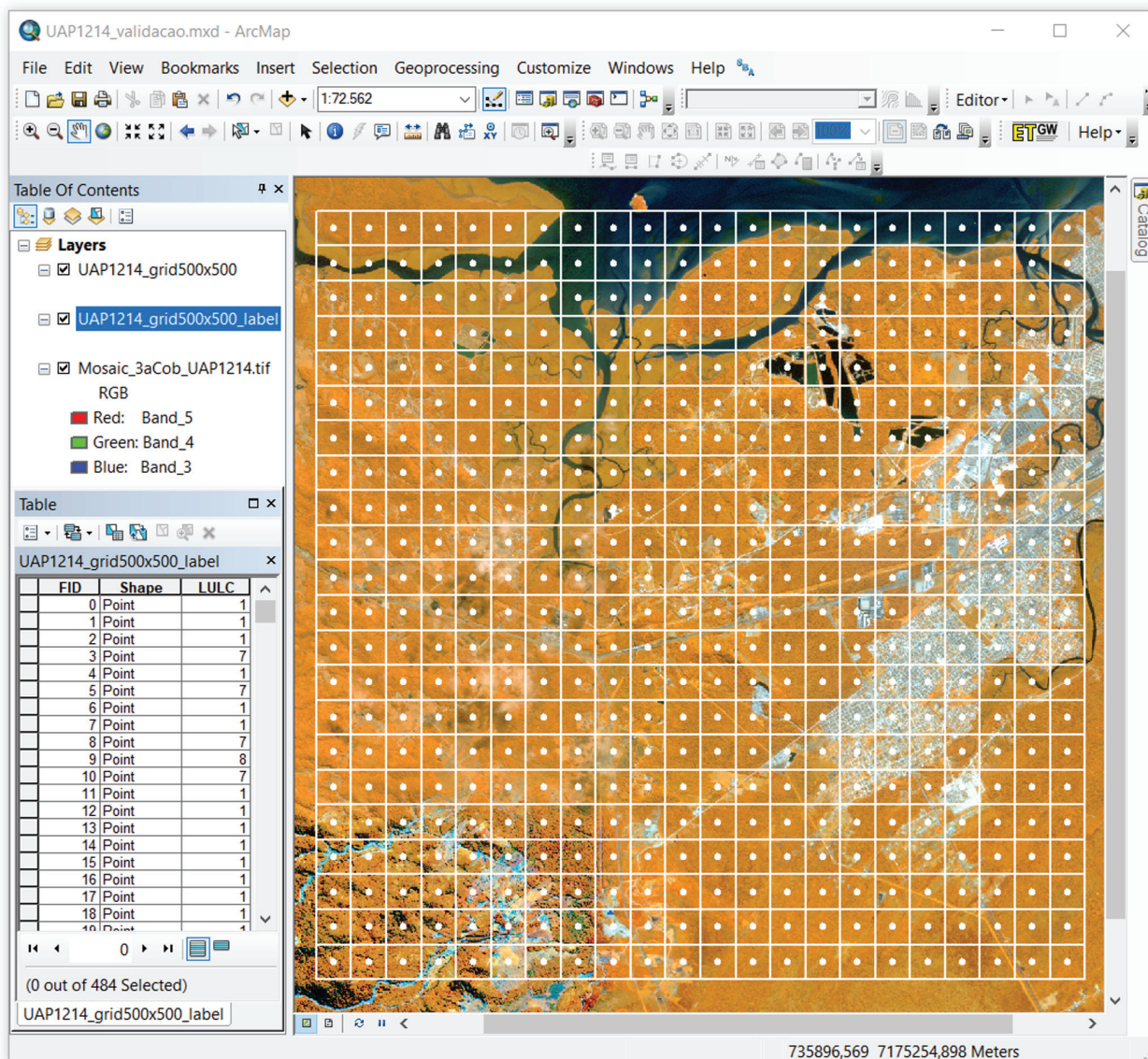


Figura 2. Janela do software ArcGIS (versão 10.4.1) exibindo a grade vetorial e os pontos de referência criados pela ferramenta *Create Fishnet* para a UAP1214, utilizando a extensão do mosaico *RapidEye* em composição NIR/RE/R.

1.5 Classificação dos pontos de referência

Na coleta dos pontos de referência para a avaliação do mapeamento, deve-se utilizar o arquivo de pontos gerado no processo previamente descrito. Nesse arquivo deve ser criado um novo campo, do tipo numérico, para receber o código da classe correspondente ao ponto (Figura 3). Após a criação desse campo, deve-se colocar a camada em edição, clicando com o botão direito sobre a camada correspondente na Tabela de Conteúdos (*Table of Contents*, Figura 3) e selecionando a opção *Edit Features>Start Editing*, iniciando a classificação dos 484 pontos de referência.

A primeira fonte de dados de referência consiste no próprio mosaico das imagens *RapidEye* (RE), utilizado para a elaboração do mapa de uso e cobertura. Para a coleta, deve-se identificar a classe contemplada por cada ponto de referência, cuja descrição pode ser consultada na Tabela 3. Na Figura 3, por exemplo, estão selecionados, em azul, cinco pontos de referência localizados em área de “Influência Urbana”, devendo, então, receber o código 08.

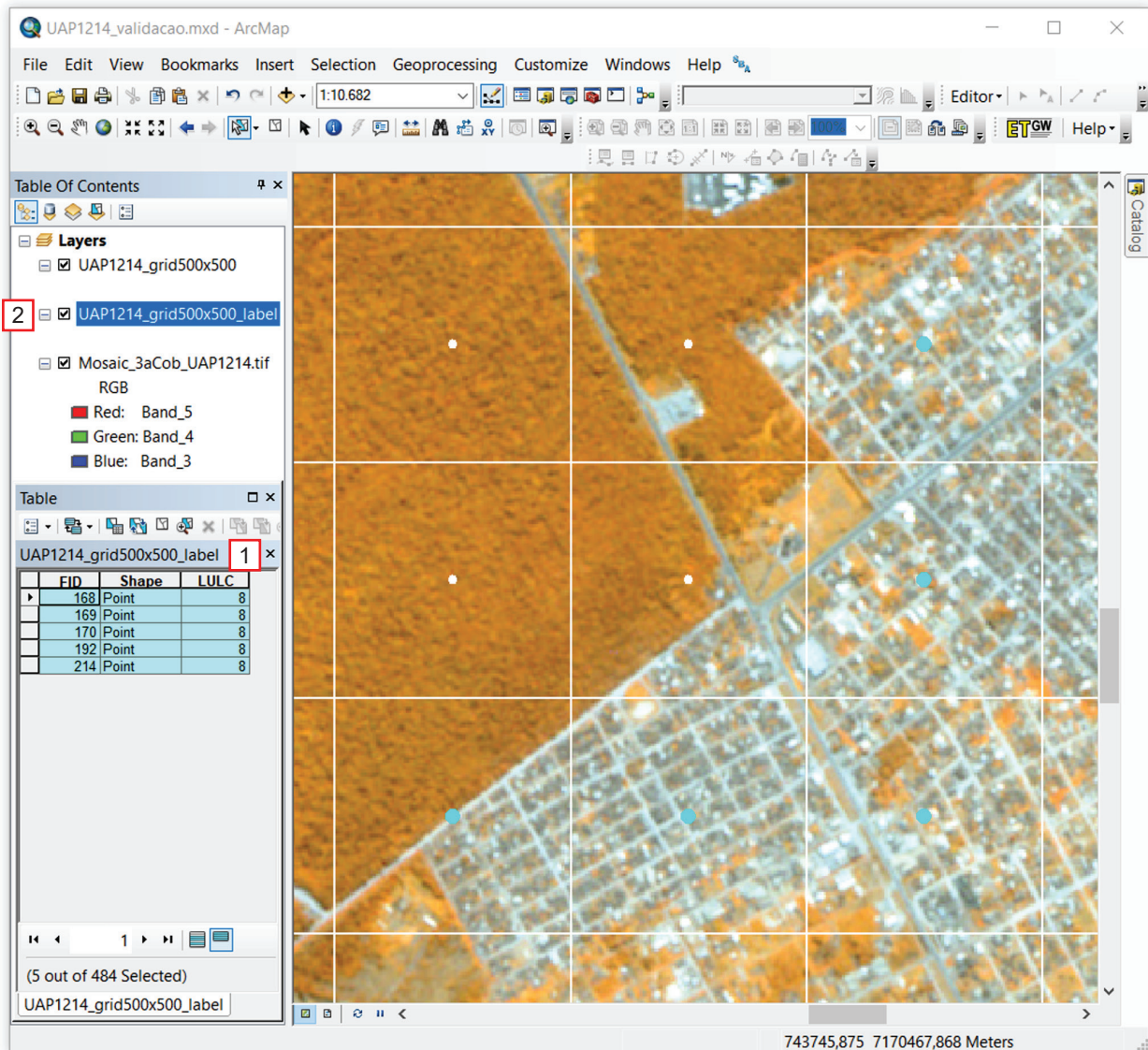


Figura 3. Janela do software ArcGIS exibindo pontos de referência que contemplam a classe 08 (“Influência Urbana”), selecionados em azul e destacados na tabela de atributos do arquivo de pontos.

Os códigos e uma descrição sucinta de cada classe de uso e cobertura da terra podem ser consultados na Tabela 3. A descrição completa das classes de uso e cobertura da terra foi disponibilizada no manual que descreve os procedimentos para a execução do mapeamento de uso e cobertura da terra das Unidades Amostrais de Paisagem do IFN-BR e deve ser consultada previamente à avaliação da classificação de uso e cobertura da terra (Luz et al., 2018).

Como ferramenta auxiliar na classificação dos pontos, pode-se fazer uso das imagens disponíveis no Google Earth, muito úteis principalmente devido à alta resolução espacial, além de conter séries históricas, que permitem uma análise mais completa da dinâmica do uso e cobertura da terra.

Um exemplo que ocorre com frequência diz respeito à classe “Floresta plantada” (05), na qual áreas temporariamente não estocadas devido ao corte raso devem ser incluídas, podendo ser confundidas com a classe “Solo Exposto” (06), ou mesmo “Agricultura e Pastagem” (07). Na Figura 4 ilustra-se a seguinte situação: a UAP1444 contém predominância de áreas de reflorestamento, de forma que um dos pontos ilustrado na Figura 4, quando observado somente no mosaico *RapidEye* (Figura 4 (a), item (2)), pode ser erroneamente classificado como pertencente à classe “Floresta plantada”,

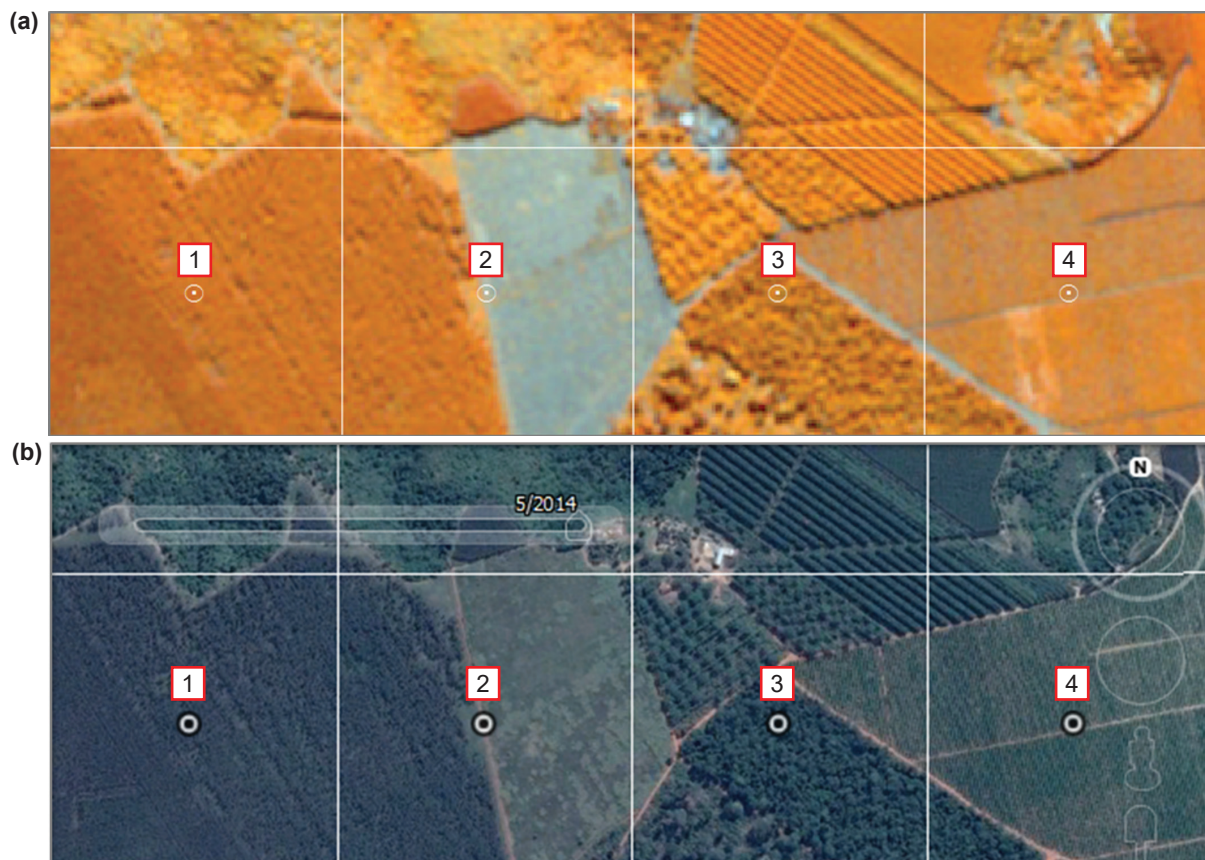


Figura 4. Exemplos de pontos de referência ocorrentes na UAP1444, localizada na Bahia, ilustrando áreas de “Floresta Plantada”, “Agricultura e Pastagem”, e “Floresta Natural”, no mosaico *RapidEye* (5R, 4G, 3B), e na imagem do Google Earth, relativa ao ano de 2014 (b).

quando, na verdade, corresponde a uma área de “Agricultura e Pastagem”, facilmente identificada na imagem do Google Earth (Figura 4 (b), item (2)).

Outras situações que podem ocorrer em áreas de “Floresta Plantada” (05) estão ilustradas na Figura 5, que mostra a resposta espectral de talhões cujo corte raso foi realizado recentemente em relação à data de aquisição das imagens (Figura 5 (a)). A análise da série histórica das imagens do Google Earth revela que essa mesma área vem sendo usada há anos para este propósito (Figura 5 (d)), apresentando plantios ainda jovens (b) (e) e plantios maduros (c) (f). Dessa forma, ainda que tenha apresentado cobertura formada por solo nu, a área deve ser classificada como “Floresta Plantada”.

Por sua vez, áreas pertencentes à classe “Gramíneas e Herbáceas” (04) são particularmente de difícil identificação e classificação, devendo-se utilizar as imagens oriundas do Google Earth para elucidar quaisquer dúvidas. Na Figura 6 (a), por exemplo, a resposta espectral da vegetação presente nessa área pode ser classificada erroneamente como pertencente à classe “Agricultura e Pastagem” (07), devido ao seu porte herbáceo, com poucos elementos arbóreo-arbustivos presentes. A consulta às imagens do Google Earth (Figura 6 (d)) permite a correta classificação do ponto, caracterizado por uma vegetação herbácea, com ocorrência de elementos arbóreo/arbustivos de forma esparsa, devido à presença de solo raso. A classe denominada “Gramíneas e Herbáceas” também pode ocorrer em áreas úmidas, com vegetação natural de porte herbáceo, localizada às margens de rios, alagadas sazonalmente, como ilustrado nas Figuras 6 (b) e (e), bem como em áreas de várzeas (Figura 6 (c) e (f)).

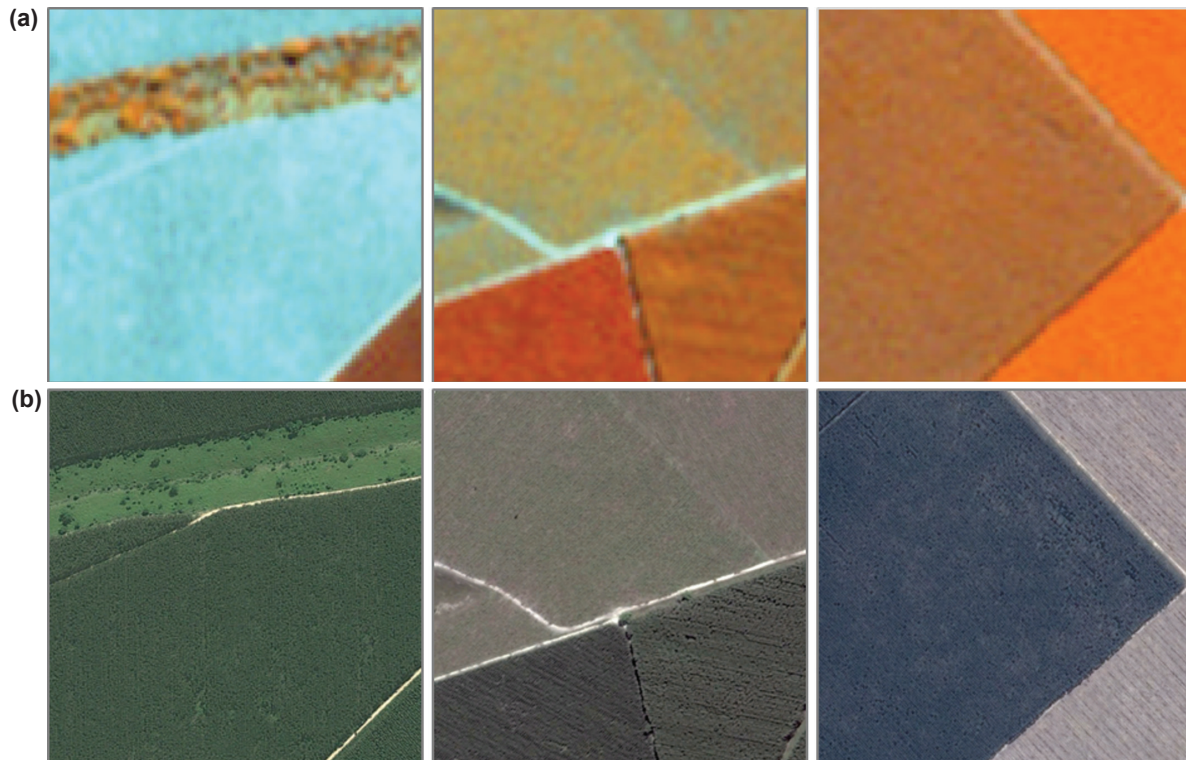


Figura 5. Exemplos de diferentes áreas pertencentes à classe “Floresta Plantada”, ilustrando a resposta espectral da imagem *RapidEye* (5R, 4G, 3B) em talhões temporariamente não estocados devido ao recente corte raso (b), plantios jovens (a) e plantios adultos (c), bem como sua correspondência nas imagens do Google Earth (e) (d) (f).

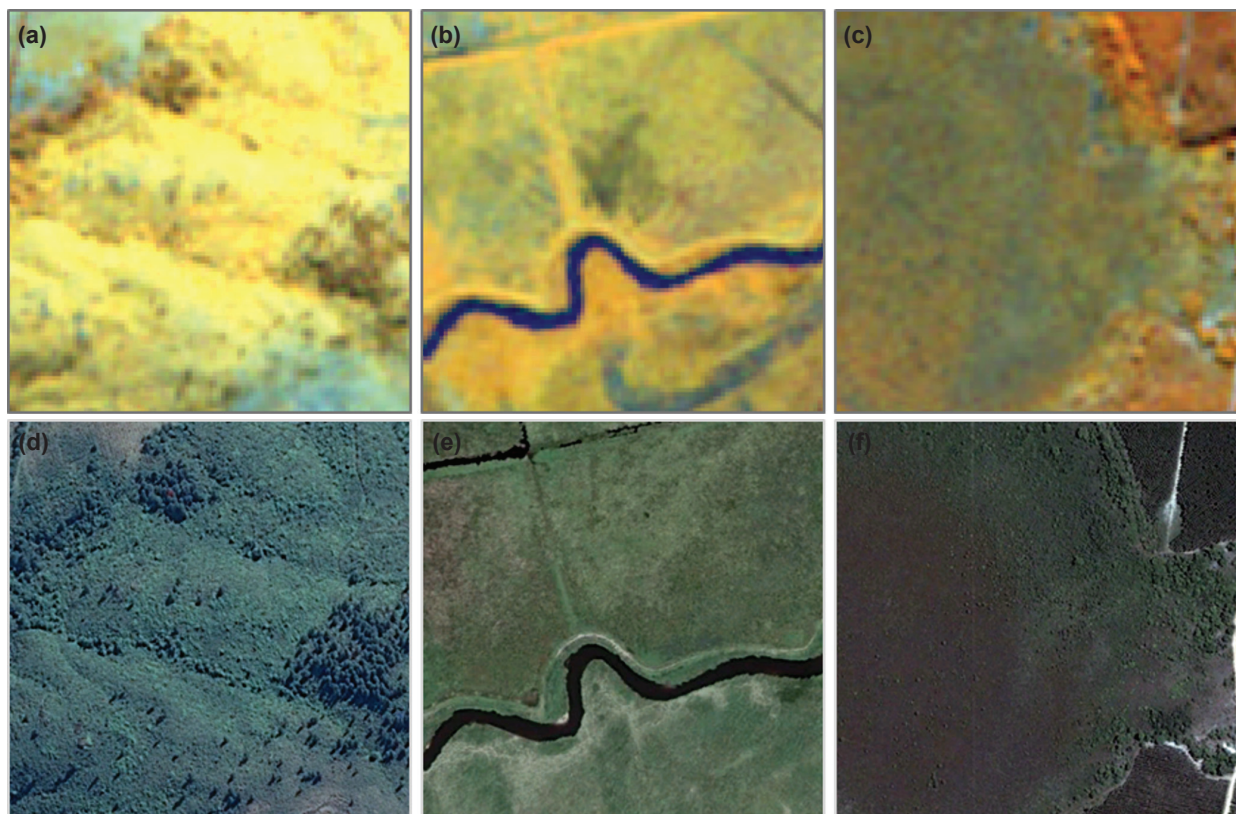


Figura 6. Exemplos de diferentes áreas pertencentes à classe “Gramíneas e Herbáceas”, ilustrando a resposta espectral da imagem *RapidEye* (5R, 4G, 3B) em localidades com presença de vegetação herbácea, com ocorrência esparsa de elementos arbóreo/arbustivos, devido à presença de solo raso (a), e em áreas úmidas (b) (c), e sua correspondência nas imagens do Google Earth (d) (e) (f).

Após a verificação e classificação de todos os pontos da verdade de campo deve-se utilizar a ferramenta “*Split by Attribute*” para separar este arquivo vetorial em vários arquivos, com base no campo contendo o código da classificação. Assim, a ferramenta irá separar todos os pontos classificados como “01”, por exemplo, em um *shapefile* contendo somente os pontos pertencentes a essa classe, e assim sucessivamente, para todas as classes, de forma que ao fim do processo existirá um arquivo *shapefile* para cada classe presente no mapeamento. A ferramenta “*Split by Attribute*” pode ser encontrada para *download* no endereço eletrônico: <http://www.umesc.usgs.gov/management/dss/split_by_attribute_tool.html>.

Conforme ilustrado na Figura 7 (a) deve-se selecionar como camada de entrada (“*Feature Layer*”) o arquivo vetorial contendo a classificação dos pontos da verdade de campo, e escolher o campo a ser utilizado para a divisão dessa camada (“*Split Field*”), bem como o prefixo para a nomenclatura dos arquivos que serão criados. Após a execução, um arquivo do tipo *geodatabase* (.gdb) será criado (Figura 7 (b)), contendo tantas camadas quantas forem as classes de uso e cobertura da terra identificadas. Estas camadas de informação devem ser abertas no projeto do ArcGIS e exportadas para o formato *shapefile*, uma a uma. Esses arquivos correspondem às entradas dos pontos de referência utilizados no software ENVI para a avaliação da acurácia dos mapas de uso e cobertura da terra. Para realizar a exportação do arquivo de cada classe referente aos pontos de referência coletados, basta clicar com o botão direito na camada que se deseja exportar e selecionar a opção “*Data>Export Data*”.

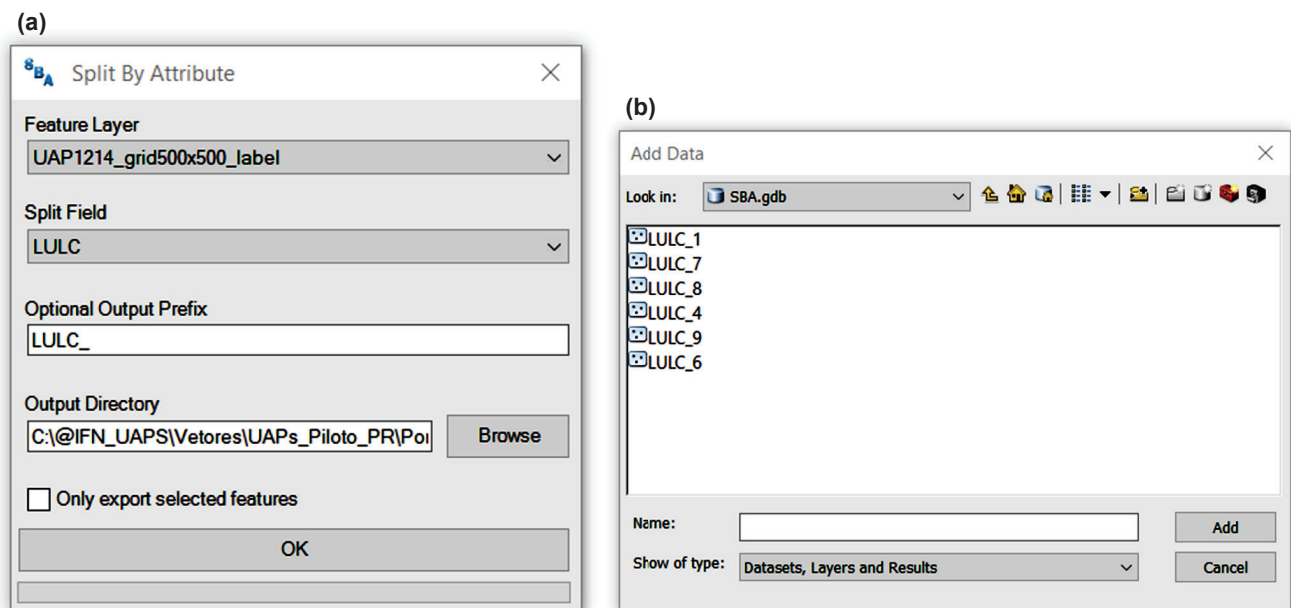


Figura 7. Janelas do software ArcGIS (versão 10.4.1) ilustrando a ferramenta *Split By Attribute* e os arquivos *geodatabase* criados por ela.

No software QuantumGIS é possível realizar a mesma operação, com a ferramenta “*Dividir Camada Vetorial*”, localizada na barra de ferramentas principal, em: “*Vetor>Gerenciar Dados>Dividir Camada Vetorial*”. Basta selecionar o arquivo *shapefile* com os pontos classificados como a camada vetorial de entrada e escolher o campo de identificação exclusivo como sendo a coluna contendo a codificação da classificação, bem como uma pasta de saída para salvar os arquivos no formato *shapefile*.

1.6 Cálculo da matriz de erro

O cálculo da matriz de erro para a avaliação da acurácia do mapeamento pode ser realizado por meio do software ENVI, na versão Classic. Por meio da barra de ferramentas principal deve-se abrir o arquivo *raster* correspondente ao mapa de uso e cobertura da terra para o qual se pretende calcular a acurácia, clicando em “File>Open Image File”.

Convém salientar que, conforme indicado por Luz et al. (2018), após a edição do mapa de uso e cobertura da terra, deve ser criada nova coluna na tabela de atributos, que conterà os códigos variando de 00 a 10, correspondentes às classes descritas na Tabela 3. Com base neste código será gerado o arquivo *raster* correspondente ao mapa de uso e cobertura da terra, utilizado para a verificação da acurácia.

Depois de carregada no software, a imagem deve ter sua legenda ajustada de acordo com as classes abrangidas pelo mapeamento. Para isso, deve-se acessar a ferramenta “Density Slice”, disponível na barra de ferramentas da imagem, em “Tools>Color Mapping>Density Slice” (Figura 8 (a)).

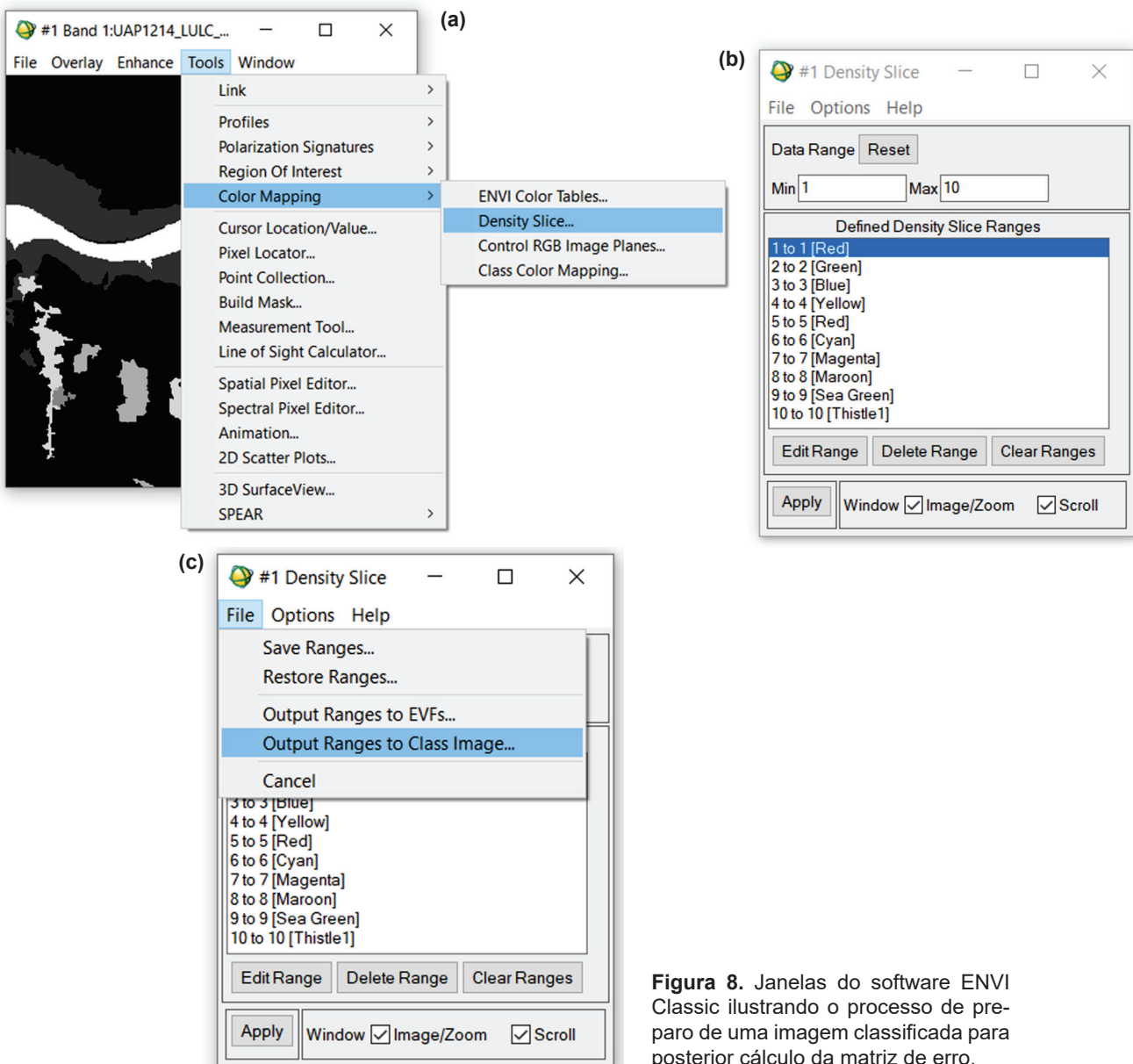


Figura 8. Janelas do software ENVI Classic ilustrando o processo de preparo de uma imagem classificada para posterior cálculo da matriz de erro.

Na janela da ferramenta “*Density Slice*” deve-se configurar a distribuição das classes na legenda conforme exposto na Figura 8 (b). O intervalo de dados (*Data Range*) deve abranger todas as classes contempladas na imagem, de forma que seu intervalo de dados varia de 01 (*Min*) a 10 (*Max*). Para cada classe deve-se definir a amplitude de valores dos pixels (“*Defined Density Slice Ranges*” na opção “*Edit Range*”), de acordo com os códigos de classes, variando de 1 a 1 para a primeira classe, de 2 a 2 para a segunda classe e assim por diante. Após a configuração correta de todas as classes, deve-se aplicar as edições (“*Apply*”) e exportar a imagem classificada com o novo intervalo de dados em “*File>Output Ranges to Class Image*” (Figura 8 (c)). Para facilitar a visualização de cada classe, e posterior análise da matriz de erro que será gerada, os nomes das camadas presentes na nova imagem podem ser editados, representando as diferentes classes do mapeamento, que variam de 01 a 10. Para isso, por meio da barra de ferramentas da imagem, pode-se acessar a janela “*Class Color Mapping*”, em “*Tools>Color Mapping>Class Color Mapping*”, e editar o nome das camadas do *raster*.

Após preparada a imagem para a qual será calculada a matriz de erro, deve-se adicionar os arquivos de pontos previamente classificados que representarão a verdade de campo (referência) na análise. Na barra de ferramentas principal deve-se clicar em “*Vector>Open Vector File*” e abrir o conjunto de arquivos *shapfile* contendo os pontos. Esses pontos também podem ter seus nomes editados para melhor visualização e análise da matriz de erro, em “*Options>Edit Layer Names*”, na janela que será aberta quando os arquivos vetoriais forem carregados. Em seguida, deve-se exportar os arquivos vetoriais, em “*File>Export Layers to ROI*” (Figura 9 (a)), selecionando-se a opção para conversão de todos os registros da camada para *ROI* (*Convert all records of an EVF layer to one ROI*) (Figura 9 (b)).

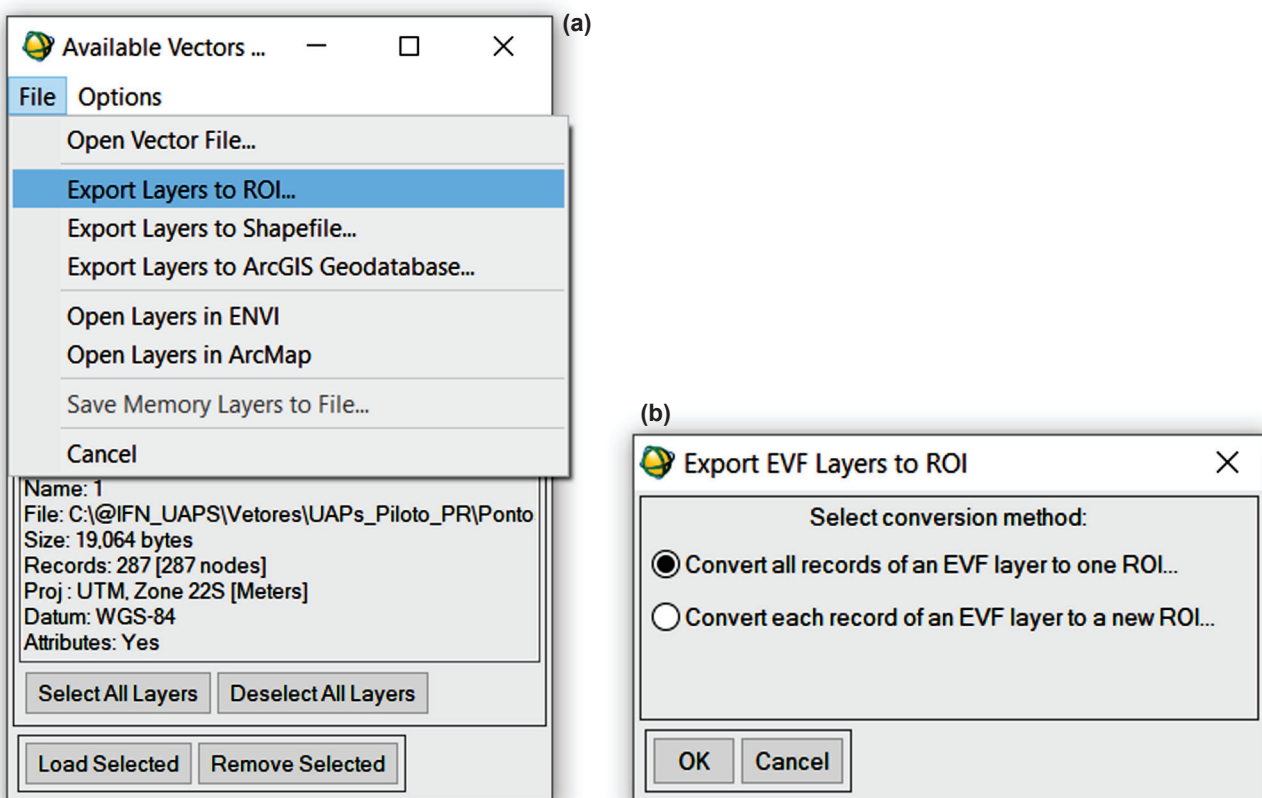


Figura 9. Janelas do software ENVI Classic ilustrando o processo de preparo dos pontos de referência representando a verdade de campo para posterior cálculo da matriz de erro.

Com os dados preparados pode-se proceder à elaboração da matriz de erro. Na barra de ferramentas principal do software clicar em “Classification>Post Classification>Confusion Matrix>Using Ground Truth ROIs” (Figura 10 (a)). Na janela que se abre (“Match Classes Parameters”) devem ser selecionados os pontos de verdade de campo (“Ground Truth ROI”) e da imagem correspondente (“Classification Image”) a cada classe e clicar em “Add Combination”, da mesma maneira como ilustrado na Figura 10 (b). Na janela “Confusion Matrix Parameters”, selecionar as opções para que a matriz contenha dados tanto em pixels quanto em percentuais, bem como a opção “Report Accuracy Assessment” (Figura 10 (c)).

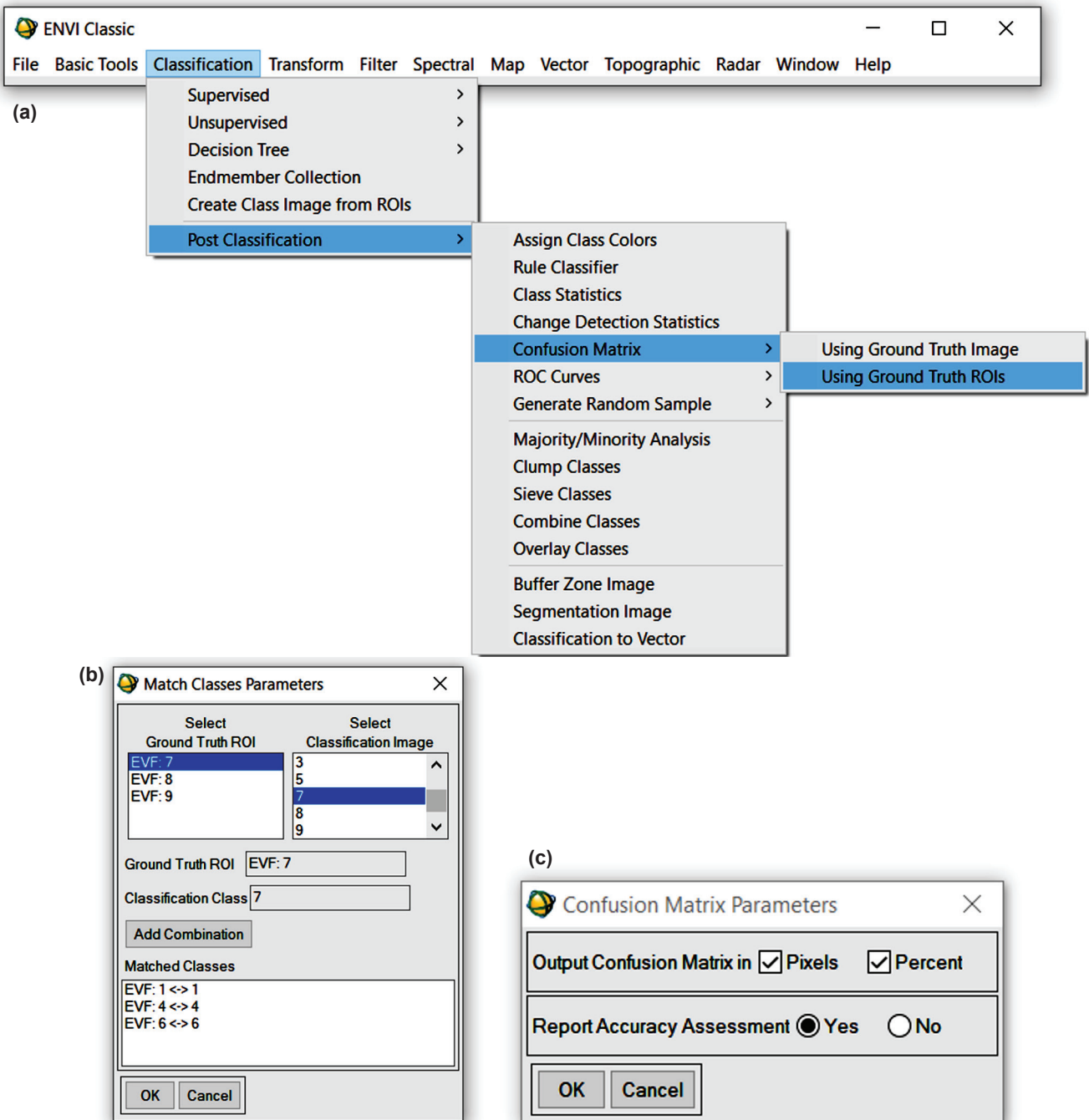


Figura 10. Janelas do software ENVI Classic ilustrando o processo de cálculo de uma matriz de erro a partir de pontos da verdade de campo.

Desta forma, o arquivo de saída gerado apresenta a acurácia global, o coeficiente Kappa, a matriz de erro em termos de pixels e também em termos percentuais, bem como os erros de comissão e omissão de cada classe, a acuracidade do produtor e a acuracidade do usuário de cada classe. Na Figura 11 são apresentadas as matrizes de erro geradas para uma determinada UAP, com resultados tanto em pixels (a) quanto em valores percentuais (b), bem como a acurácia global e o valor do coeficiente Kappa.

Overall Accuracy = (418/484) 86.3636%
Kappa Coefficient = 0.7681

Class	Ground Truth (Pixels)						Total
	EVF: 1	EVF: 4	EVF: 6	EVF: 7	EVF: 8	EVF: 9	
Unclassified	0	0	0	0	0	0	0
1	272	15	0	3	11	0	301
4	4	22	0	0	1	0	27
6	0	0	1	0	2	0	3
7	6	0	0	26	9	1	42
8	3	0	0	3	64	2	72
9	2	4	0	0	0	33	39
Total	287	41	1	32	87	36	484

(a)

Class	Ground Truth (Percent)						Total
	EVF: 1	EVF: 4	EVF: 6	EVF: 7	EVF: 8	EVF: 9	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	94.77	36.59	0.00	9.38	12.64	0.00	62.19
4	1.39	53.66	0.00	0.00	1.15	0.00	5.58
6	0.00	0.00	100.00	0.00	2.30	0.00	0.62
7	2.09	0.00	0.00	81.25	10.34	2.78	8.68
8	1.05	0.00	0.00	9.38	73.56	5.56	14.88
9	0.70	9.76	0.00	0.00	0.00	91.67	8.06
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

(b)

Figura 11. Matrizes de erro geradas por meio do software ENVI, referentes à avaliação da acuracidade do mapa de uso e cobertura da terra para a UAP1214, localizada no Paraná, em valores absolutos (a) e em valores percentuais (b).

Por sua vez, a distribuição percentual de pixels classificados de forma correta com relação ao número total de pixels de referência é representada pela acurácia global que, no caso da UAP exemplificada, foi aproximadamente 86%, de forma que de um total de 484 pixels, 418 foram classificados corretamente. Na Figura 12 (a) são apresentados os valores de comissão e omissão em pixels e em percentuais e, na Figura 12 (b), a acuracidade do produtor e do usuário por classe, também em pixels e em percentuais.

Class	Commission	Omission	Commission (Pixels)	Omission (Pixels)	Class	Prod. Acc.	User Acc.	Prod. Acc. (Pixels)	User Acc. (Pixels)
	(Percent)	(Percent)				(Percent)	(Percent)		
1	9.63	5.23	29/301	15/287	1	94.77	90.37	272/287	272/301
4	18.52	46.34	5/27	19/41	4	53.66	81.48	22/41	22/27
6	66.67	0.00	2/3	0/1	6	100.00	33.33	1/1	1/3
7	38.10	18.75	16/42	6/32	7	81.25	61.90	26/32	26/42
8	11.11	26.44	8/72	23/87	8	73.56	88.89	64/87	64/72
9	15.38	8.33	6/39	3/36	9	91.67	84.62	33/36	33/39

(a)

(b)

Figura 12. Coeficientes de concordância gerados pelo software ENVI, referentes aos erros de comissão e omissão das classes de uso e cobertura da terra (a), bem como os valores de acuracidade do produtor e do usuário (b), em valores percentuais e em pixels.

A matriz de erros gerada pelo software ENVI servirá de base para a construção da matriz de erros em termos de proporção de área (Tabela 2), a partir da qual deverão ser calculados e apresentados os valores da acurácia global proporcional à área e sua respectiva variância estimada, de acordo com a formulação apresentada na Seção 1.3, Análise e Estimativa dos Resultados.

Apresenta-se a seguir um exemplo do cálculo da matriz de acurácia proporcional à área, contendo os valores das classes, a acurácia global, a acurácia do produtor e do usuário por classe, bem como o intervalo de confiança a 95% de probabilidade para a acurácia global (proporcional à área) e sua respectiva variância estimada, utilizando os dados da UAP1214, localizada no município de Paranaguá, PR.

Para o cálculo dos valores da matriz de erros ponderada pela proporção de área para cada classe de uso e cobertura da terra, deve-se utilizar a Equação 1. A sequência de cálculos apresentada a seguir exemplifica a obtenção de \hat{p}_{11} para a Classe 1. Os valores utilizados no cálculo aparecem destacados em verde nas Tabelas 4 e 5.

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_i} \rightarrow \hat{p}_{11} = 0,6581 * 272/301 \rightarrow \hat{p}_{11} = 0,5947$$

Tabela 4. Número de polígonos e superfícies absoluta e percentual ocupadas pelas classes de uso e cobertura da terra, calculadas a partir do mapeamento na área de abrangência de 11 km x 11 km, na projeção UTM, Fuso 22S, WGS84, para a UAP1214.

Classe	Número de polígonos	Área (ha)	Área (%)	W_i
1	138	7962,64	65,81	0,6581
2	0	0,00	0,00	0,0000
3	59	14,40	0,12	0,0012
4	156	764,02	6,31	0,0631
5	0	0,00	0,00	0,0000
6	36	36,99	0,31	0,0031
7	225	854,04	7,06	0,0706
8	90	1538,18	12,71	0,1271
9	82	922,80	7,63	0,0763
10	7	6,93	0,06	0,0006
Total	793	12100,00	100,00	1,0000

Tabela 5. Matriz de confusão referente à avaliação da acuracidade do mapa de uso e cobertura da terra em pixels, para a UAP1214.

Classe	Dados de referência (“verdade de campo”) em pixels						Total (n_i)
	1	4	6	7	8	9	
1	272	15	0	3	11	0	301
4	4	22	0	0	1	0	27
6	0	0	1	0	2	0	3
7	6	0	0	26	9	1	42
8	3	0	0	3	64	2	72
9	2	4	0	0	0	33	39
Total (n_j)	287	41	1	32	87	36	484

Na Tabela 6 são apresentados os resultados do cálculo dos valores da matriz de \hat{p}_{ij} , bem como o somatório das linhas e colunas.

Tabela 6. Matriz de erros ponderada pela proporção de área estimada para cada classe de uso e cobertura da terra, para a UAP1214.

Classe	Dados de referência (“verdade de campo”) ponderados pela proporção de área da classe						Total ($p_{.j}$)
	1	4	6	7	8	9	
1	0,5947	0,0328	0,0000	0,0066	0,0240	0,0000	0,6581
4	0,0094	0,0514	0,0000	0,0000	0,0023	0,0000	0,0631
6	0,0000	0,0000	0,0010	0,0000	0,0020	0,0000	0,0031
7	0,0101	0,0000	0,0000	0,0437	0,0151	0,0017	0,0706
8	0,0053	0,0000	0,0000	0,0053	0,1130	0,0035	0,1271
9	0,0039	0,0078	0,0000	0,0000	0,0000	0,0645	0,0763
Total ($p_{.j}$)	0,6233	0,0921	0,0010	0,0555	0,1565	0,0697	

A partir dessa matriz deve-se calcular a Acurácia Global (\hat{A}), as acurácias do Produtor (\hat{P}_j) e do Usuário (\hat{U}_i) por classe, proporcionais à área. A Acurácia Global é dada pela Equação 2, e corresponde ao somatório da diagonal principal da matriz de erros representada pela Tabela 6, onde os valores foram destacados em azul, conforme demonstrado a seguir:

$$\hat{A} = \sum_{i=1}^q p_{ii} \rightarrow \hat{A} = 0,5947 + 0,0514 + 0,0010 + 0,0437 + 0,1130 + 0,0645$$

$$\hat{A} = 0,8684$$

O cálculo da acurácia do Produtor (Equação 3) e do Usuário (Equação 4) para a Classe 1 está exemplificada na sequência:

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{.j}} \hat{p}_{.j} = 0,5947/0,6233 \rightarrow \hat{P}_1 = 0,9540$$

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{.i}} \hat{p}_{.i} = 0,5947/0,6581 \rightarrow \hat{U}_1 = 0,9037$$

Os resultados do cálculo dos valores de acurácia do produtor e do usuário, proporcionais à área, estão apresentados na Tabela 7.

Tabela 7. Acurácias do produtor e do usuário ponderadas pela área, por classe de uso e cobertura da terra.

Classe	Acurácia do produtor (\hat{P}_i)	Acurácia do usuário (\hat{U}_i)
1	0,9540	0,9037
4	0,5588	0,8148
6	1,0000	0,3333
7	0,7866	0,6190
8	0,7218	0,8889
9	0,9253	0,8462

Também deve ser apresentada Variância Estimada ($\hat{V}(\hat{A})$) da Acurácia Global ponderada pela proporção de área de cada classe. O cálculo para a Variância, obtida a partir da Equação 6, está demonstrado na sequência:

$$\hat{V}(\hat{A}) = \sum_{i=1}^q \frac{W_i^2 \hat{U}_i (1 - \hat{U}_i)}{n_i - 1}$$

$$\hat{V}(\hat{A}) = \frac{0,6581^2 * 0,9037(1 - 0,9037)}{301 - 1} + 0,00002314 + 0,000001038 + 0,00002866 + 0,00002247 + 0,00001992$$

$$\hat{V}(\hat{A}) = 0,000220913$$

A partir da Variância deve ser calculado o Intervalo de Confiança (I.C.) (Equação 5) de 95% de probabilidade para a Acurácia Global ponderada pela proporção de área de cada classe.

$$IC = \hat{A} \pm t_{tab} * \sqrt{\hat{V}(\hat{A})} \rightarrow IC = 0,8684 \pm 1,96 * \sqrt{0,0002209} \rightarrow IC = 0,8684 \pm 0,02913$$

Desta forma, o Intervalo de confiança da acurácia global (proporcional à área) do mapeamento de uso e cobertura da terra da UAP 1214 resulta em:

$$IC_{acurácia\ global} = [83,92; 89,75]$$

Assim, o relatório de acuracidade, para cada UAP, deve ser constituído por:

- Matriz de acurácia em pixels, contendo os valores das classes, a acurácia global, a acurácia do produtor e do usuário por classe.
- Matriz de acurácia em valores percentuais do número de pixels, contendo os valores das classes, a acurácia global, a acurácia do produtor e do usuário por classe.

- Matriz de acurácia proporcional à área, contendo os valores das classes, a acurácia global, a acurácia do produtor e do usuário por classe.
- O intervalo de confiança a 95% de probabilidade para a acurácia global (proporcional à área) e sua respectiva variância estimada.

1.7 Critérios quantitativos e qualitativos para a avaliação do produto

Especificar uma única medida para todos os fins de avaliação da acurácia da classificação, apesar de ser o desejado, não é possível, pois um único coeficiente não é capaz de fornecer uma avaliação satisfatória quanto à qualidade de uma classificação (FOODY, 2002). Assim, recomenda-se que, ao invés de simplesmente fornecer um coeficiente básico para avaliação da acurácia da classificação de imagens (tal como, o índice percentual de acurácia global de cada classe), sejam adotadas mais medidas de acurácia, além de fornecer a matriz de erro, como uma descrição mais completa da qualidade da classificação.

Além disso, nas avaliações clássicas de acurácia de mapas temáticos, todas as atribuições incorretas têm o mesmo peso. Muitas vezes, no entanto, alguns erros são mais importantes, ou prejudiciais, do que outros. Em muitos casos, erros observados em uma classificação ocorrem entre classes similares e, por vezes, podem não ter importância para o produto final, enquanto outros podem ser muito significativos (FOODY, 2002). Considerando que as classes temáticas mais importantes no IFN-BR, tanto para o mapa de uso e cobertura da terra quanto para as análises de paisagem a serem realizadas com esse produto, são as classes com cobertura arbórea/arbustiva (“Floresta Natural”, “Floresta Plantada”, “Outras Terras com Vegetação Lenhosa” e “Outras Terras com Árvores”) e as áreas de ocorrência de campos naturais (“Gramíneas e Herbáceas”), percentuais de acerto mais elevados para essas classes são desejados. De menor relevância, por exemplo, é o erro entre as classes de “Influência Urbana” e “Solo Exposto”.

Dessa forma, é adotado um limiar mínimo de 80% de acurácia global e limiares de 60% na acurácia do produtor e usuário para as classes de cobertura arbórea/arbustiva (“Floresta Natural”, “Floresta Plantada”, “Outras Terras com Vegetação Lenhosa” e “Outras Terras com Árvores”), dada a sua importância, e de 50% na acurácia do produtor e usuário para a classe de campo natural (“Gramíneas e Herbáceas”), dada a dificuldade em seu reconhecimento e classificação, salvo nos casos em que a ocorrência da classe na UAP for inferior a 5%. Um resumo desses valores é apresentado na Tabela 8, a seguir.

Tabela 8. Valores de referência para a acurácia da classificação do uso e cobertura da terra nas Unidades Amostrais de Paisagem do IFN-BR.

Acurácia	(%)
Geral	80
Floresta Natural (produtor e usuário)	60
Floresta Plantada (produtor e usuário)	
Outras Terras com Vegetação Lenhosa (produtor e usuário)	
Outras Terras com Árvores (produtor e usuário)	
Gramíneas e Herbáceas (produtor e usuário)	50

Além dos coeficientes acima expostos, que correspondem aos critérios quantitativos para a avaliação do produto, recomenda-se uma avaliação também qualitativa, baseada na interpretação visual do mapeamento realizado com base no mosaico *RapidEye*, pois, muitas vezes, apesar de os coeficientes apresentarem resultados aparentemente satisfatórios, basta uma conferência visual da classificação gerada para que erros sejam encontrados e facilmente corrigidos, como exposto na Figura 13. Essa etapa de análise visual também garante que áreas que não correspondem às amostradas pelo *grid* de pontos de referência sejam controladas.

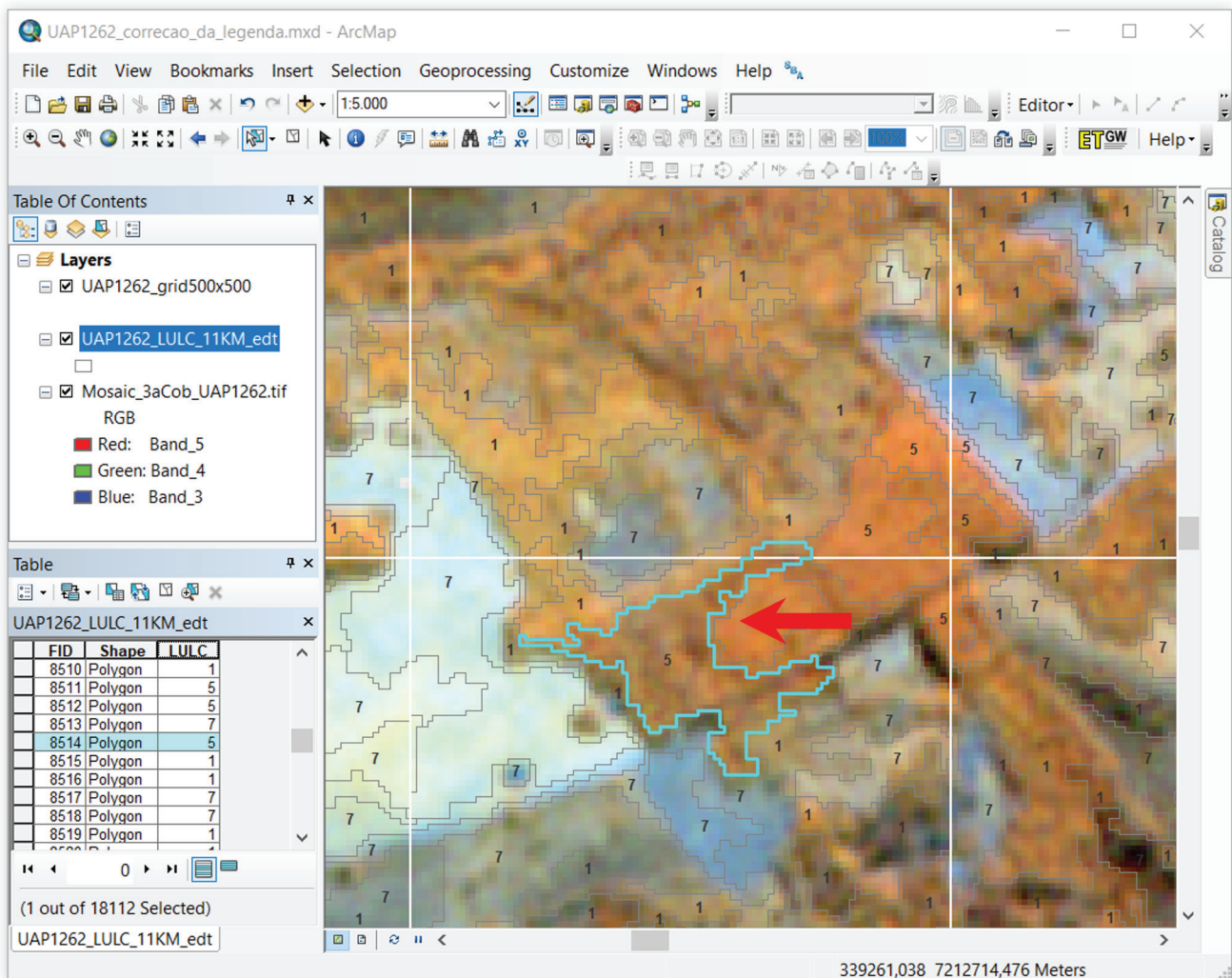


Figura 13. Janela do software ArcGIS (versão 10.4.1) ilustrando a escala que deve ser utilizada para a avaliação qualitativa do mapa de uso e cobertura, bem como um erro de atribuição entre as classes, apontado pela seta em vermelho, onde um polígono que deveria ser classificado como pertencente à classe “Floresta Natural” (01), recebeu atribuição da classe “Floresta Plantada” (05).

Para execução de tal conferência, o Google Earth também pode ser utilizado como ferramenta auxiliar. Recomenda-se, ainda, considerando a resolução espacial do mosaico *RapidEye*, uma escala de trabalho de 1:5.000, e a utilização da grade vetorial de 500 m x 500 m criada previamente para a geração dos pontos de referência, como guia para garantir que toda a cobertura da imagem passe por essa avaliação qualitativa, na forma de uma varredura sistemática.

Referências

- AHLQVIST, O. In search of classification that supports the dynamics of science: The FAO Land Cover Classification System and proposed modifications. **Environment and Planning B: Planning and Design**, v. 35, n. 1, p. 169-186, 2008.
- CAMPBELL, J.B. **Introduction to remote sensing**. 2 ed. New York: The Guildford Press, 1996. 622 p.
- COSTA, C. R.; LUZ, N. B.; ARAKI, H.; OLIVEIRA, Y. M. M.; ROSOT, M. A. D.; GARRASTAZU, M. C.; KRUEGER, C. P. Análise da exatidão cartográfica das imagens *RapidEye* adotadas no Inventário Florestal Nacional do Brasil (IFN-BR). In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 17., 2015, João Pessoa. **Anais**. São José dos Campos: INPE, 2015. p. 3289-3296.
- FAO. **Global Forest Resources Assessment 2015**: how have the world's forests changed? Rome, 2016a. 54 p.
- FAO. **Map accuracy assessment and area estimation**: a practical guide. Rome, 2016b. 69 p. (FAO. National forest monitoring assessment working paper n. 46/E). Disponível em: <<http://www.fao.org/3/a-i5601e.pdf>>. Acesso em: 20 nov. 2017
- FIGUEIREDO, G. C.; VIEIRA, C. A. O. Estudo do comportamento dos índices de Exatidão Global, Kappa e Tau, comumente usados para avaliar a classificação de imagens do sensoriamento remoto. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 13., 2007, Florianópolis. **Anais**. São José dos Campos: INPE, 2007. p. 5755-5762.
- FOODY, G. M. Status of land cover classification accuracy assessment. **Remote Sensing of Environment**, v. 80, n. 1, p. 185-201, 2002.
- IBGE. **Manual Técnico da vegetação brasileira**. 2 ed. Rio de Janeiro: IBGE, 2012. 271 p. (IBGE. Manuais Técnicos em Geociências n. 1). Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv63011.pdf>>. Acesso em: 19 jan. 2018.
- IBGE. **Manual Técnico de uso da terra**. 3 ed. Rio de Janeiro: IBGE, 2013. 171 p. (IBGE. Manuais Técnicos em Geociências, n. 7). Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv81615.pdf>>. Acesso em: 19 jan. 2018.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.
- LILLESAND, T. M.; KIEFER, R. W. **Remote sensing and image interpretation**. 3 ed. Crawfordsville: John Wiley and Sons, 1994.
- LUZ, N. B.; MARAN, J. C.; GARRASTAZU, M. C.; ROSOT, M. A. D.; FRANCISCON, L.; HOLLER, W. A.; GAIAD, N. P.; OLIVEIRA, Y. M. M.; FREITAS, J. V. **Manual de análise de paisagem: volume 1**: procedimentos para a execução do mapeamento de uso e cobertura da terra. Colombo: Embrapa Florestas, 2018. (Embrapa Florestas. Documentos, 316).
- OLOFSSON, P.; FOODY, G. M.; HEROLD, M.; STEHMAN, S. V.; WOODCOCK, C. E.; WULDER, M. A. Good practices for estimating area and assessing accuracy of land change. **Remote Sensing of Environment**, v. 148, n. 1, p. 42-57, 2014.
- PONTIUS JUNIOR, R. G.; MILLONES, M. Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. **International Journal of Remote Sensing**, v. 32, n. 15, p. 4407-4429, 2011.
- RADOUX, J.; BOGAERT, P. Good practices for object-based accuracy assessment. **Remote Sensing**, v. 9, n. 7, p. 646-669, 2017.
- ROSOT, N. C. **Integração de imagens de sensores de microondas e ópticos para fins de mapeamento e classificação de reflorestamentos no sul do Brasil**. 2001. 188 f. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis.
- STEHMAN, S. V.; CZAPLEWSKI, R. L. Design and analysis for thematic map accuracy assessment: fundamental principles. **Remote Sensing of Environment**, v. 64, n. 3, p. 331-344, 1998.
- STEHMAN, S. V.; WICKHAM, J. D. Pixels, blocks of pixels, and polygons: choosing a spatial unit for thematic accuracy assessment. **Remote Sensing of Environment**, v. 115, n. 12, p. 3044-3055, 2011.
- STEHMAN, S. V. Sampling designs for accuracy assessment of land cover. **International Journal of Remote Sensing**, v. 30, n. 20, p. 5243-5272, 2009.

STRAHLER, A. H.; BOSCHETTI, L.; FOODY, G. M.; FRIEDL, M. A.; HANSEN, M. C.; HEROLD, M.; MAYAUX, P.; MORISETTE, J. T.; STEHMAN, S. V.; WOODCOCK, C. E. **Global land cover validation**: Recommendations for evaluation and accuracy assessment of global land cover maps. Luxembourg, 2006, v. 51, n. 4, 58 p. (EUR 22156 EN). Disponível em: <<https://landval.gsfc.nasa.gov/pdf/GlobalLandCoverValidation.pdf>>. Acesso em: 19 jan. 2018.

Embrapa

Florestas