



Impacto da Correção Atmosférica na Classificação da Cobertura da Terra da Carta Topográfica SC-24-V-C (Petrolina) Utilizando o Google Earth Engine e Imagens Landsat-8
Impact of Atmospheric Correction on Land Cover Classification of the Topographic Chart SC-24-V-C (Petrolina) Using Google Earth Engine and Landsat-8 images

Khalil Ali Ganem¹; Gustavo Macedo de Mello Baptista^{1,2};
Washington de Jesus Sant'Anna da Franca Rocha²; Rodrigo Nogueira Vasconcellos²;
Marcos Reis Rosa³ & Deorgia Tayane Mendes de Souza^{1,2}

¹Universidade de Brasília – UnB – Instituto de Geociências. Campus Universitário Darcy Ribeiro, 70910-900, Caixa Postal 4465, Brasília, DF – Brasil

²Universidade Estadual de Feira de Santana – UEFS

Programa de Pós-Graduação em Modelagem em Ciências da Terra e do Ambiente.

Km 3, BR 116, Módulo 5 – Campus Universitário, 44031-460, Feira de Santana, BA – Brasil

³Universidade de São Paulo – USP. Departamento de Geografia.

Avenida Professor Lineu Prestes, 338, Caixa Postal 72042, 05508-000, São Paulo, SP – Brasil

E-mails: khalilganem@hotmail.com; gmbaptista@unb.br; wrocha@uefs.br;

rnuviesppgm@gmail.com; marcosrosa@usp.br; deorgiasouza.geo@gmail.com

Recebido em: 12/09/2018 Aprovado em: 01/04/2019

DOI: http://dx.doi.org/10.11137/2019_2_100_116

Resumo

Sabe-se que mapear a cobertura da terra usando técnicas de sensoriamento remoto não é um procedimento trivial e requer um alto nível de conhecimento das técnicas de processamento adotadas. O Mapeamento Anual da Cobertura e Uso do Solo no Brasil – MapBiomias – é um projeto focado na classificação da cobertura da terra dos biomas brasileiros por meio de imagens dos satélites da série Landsat. No entanto, o fato de os dados disponibilizados em reflectância de superfície para o Landsat-8 não estarem ortorretificados até a conclusão deste trabalho, levou ao uso do dado em reflectância no topo da atmosfera (TOA) pelo projeto na geração das classificações. Tendo em vista que os efeitos atmosféricos podem interferir na qualidade do dado, torna-se necessário investigar qual dado melhor se ajusta às classificações. Assim, após a seleção da carta SC-24-V-C (Petrolina), pertencente ao bioma Caatinga, foram geradas três classificações: uma em reflectância TOA e as outras com correção atmosférica a partir dos algoritmos FLAASH e QUAC. As classificações foram geradas no *Code Editor* do *Google Earth Engine* (GEE) com base em uma árvore de decisão definida empiricamente. A avaliação da acurácia para as classificações advindas do dado sem correção atmosférica (TOA) e dos dados corrigidos pelo FLAASH e QUAC mostrou, respectivamente, valores *Tau*, de, 54,13%, 39,13% e 58,10%; *Kappa* de 49,24%, 34,08% e 52,97%, e, exatidão global de 61,78%, 49,28% e 65,08%. Isso mostrou um melhor ajuste do dado corrigido pelo QUAC ao contexto da carta em comparação com os demais dados. O dado TOA diferiu em aproximadamente 4% menos em relação ao QUAC, evidenciando a indispensabilidade do dado, desde que sejam feitos novos ajustes nos parâmetros da árvore de decisão. Ao serem comparados por classe, os coeficientes apresentaram elevada variação, muito embora os dados que tiveram a atmosfera corrigida tenham apresentado, de modo geral, valores de exatidão mais elevados frente ao dado sem correção atmosférica.

Palavras-chave: Caatinga; QUAC; FLAASH; MapBiomias; Reflectância TOA

Abstract

It is known that mapping land cover through remote sensing techniques is not a trivial procedure and it requires a high level of knowledge of the processing techniques adopted. The Brazilian Annual Land Use and Land Cover Mapping - MapBiomias - is a project focused on the generation of land use and land cover classifications of the Brazilian biomes through Landsat satellite images. However, the fact that surface reflectance data provided for Landsat-8 are not orthorectified until the conclusion of this paper lead to the use of top-of-the-atmosphere (TOA) reflectance data by the project to proceed with the classifications. Since atmospheric effects may interfere with the quality of the data, it's necessary to investigate which data best fit the classifications. Thus, after the selection of the SC-24-V-C (Petrolina) topographic chart, belonging to the Caatinga biome, three classifications were generated: one in TOA reflectance and the others with atmospheric correction from FLAASH and QUAC algorithms. The classifications were generated in the Code Editor of Google Earth Engine (GEE) based in an empirically defined decision tree. The accuracy assessment for the TOA classifications as well as those originated by FLAASH and QUAC showed, respectively, Tau values of 54.13%, 39.13% and 58.10%, Kappa values of 49.24%, 34.08% and 52.97%, and overall accuracy of 61.78%, 49.28% and 65.08%. This showed a better adjustment of QUAC-corrected data to the context of the chart compared to the other data. The TOA data differed by approximately 4% less than the data corrected by QUAC, which highlights its indispensability, as long as new adjustments are made in the decision tree parameters. When compared by class, the coefficients showed a high variation, although the data atmospherically corrected presented, generally, higher values of accuracy compared to the data without atmospheric correction.

Keywords: Caatinga; QUAC; FLAASH; MapBiomias; TOA Reflectance



1 Introdução

Mapear a superfície terrestre por meio do uso de técnicas de sensoriamento remoto requer um alto nível de conhecimento, tanto com relação às técnicas de processamento adotadas, como do contexto real da paisagem da área a ser representada. Isso configura-se como uma informação primordial para projetos de gestão ambiental, avaliação da biodiversidade e, conseqüentemente, como suporte à adoção de políticas públicas. Além disso, outro fator bastante discutido por especialistas em sensoriamento remoto relaciona-se com a necessidade de corrigir ou não os efeitos atmosféricos na imagem antes de processá-la.

É sabido que os efeitos produzidos pela atmosfera podem interferir nos dados capturados pelo sensor, o que pode requerer a correção desses efeitos (Richards & Jia, 2006). Para tal, no campo da radiação solar refletida, os dados devem ser convertidos de radiância para reflectância de superfície ao invés de reflectância no topo da atmosfera (TOA), ao passo que, na faixa térmica do espectro, devem ser convertidos para temperatura de superfície ao invés de temperatura de brilho. Além disso, a correção atmosférica pode alterar os valores de reflectância em algumas bandas de diferentes formas (Fonseca & Glerani, 2005), pois os efeitos de absorção e espalhamento atuam em menor ou maior escala, e este fato depende do comprimento de onda ou das bandas em que o sensor registra a imagem.

Para a reflectância de superfície, a precisão do método de transferência radiativa e o modelo atmosférico têm uma influência significativa (Rojas, 2005). Atualmente, existem vários modelos para corrigir os efeitos atmosféricos em imagens multiespectrais. Dentre os principais modelos para proceder com a correção atmosférica dos dados, os mais conhecidos e consolidados na literatura são o *Fast Line-of-sight Atmospheric Analysis of Hypercubes* (FLAASH) (Adler-Golden *et al.*, 1998) e o *Quick Atmosphere Correction* (QUAC) (Bernstein *et al.*, 2004).

Mandanici *et al.* (2015) compararam modelos de correção atmosférica, baseados em modelos empíricos e físicos, e relataram não ser possível afirmar precisamente qual método é o mais apropriado,

muito embora tenham destacado que modelos como o FLAASH e QUAC destacam melhor a forma das assinaturas espectrais dos alvos. Guo & Zeng (2012) também constataram que ambos os métodos, mesmo com suas especificidades, podem eliminar efeitos atmosféricos e recuperar características típicas de diferentes feições superficiais de maneira mais eficiente, incluindo alvos vegetais.

Um dos projetos nacionais no âmbito da classificação da cobertura da superfície terrestre é o Mapeamento Anual da Cobertura e Uso do Solo no Brasil – MapBiomias (Projeto MapBiomias, 2017a). O projeto nasceu de um seminário organizado pelo Sistema de Estimativas e Emissões de Gases de Efeito Estufa (SEEG), iniciativa do Observatório do Clima (OC), em março de 2015, na cidade de São Paulo, e envolve uma rede colaborativa com especialistas de diversos ramos da ciência que utiliza processamento em nuvem e classificadores automatizados desenvolvidos e operados a partir da plataforma *Google Earth Engine* (GEE) para gerar uma série histórica de mapas anuais de cobertura e uso da superfície terrestre do Brasil.

Os primeiros trabalhos desenvolvidos pelo MapBiomias objetivaram, basicamente, diferenciar áreas vegetadas de não vegetadas. Essa etapa também é referida como Coleção 1. A Coleção 2, definida em 2016 e já incorporada por este trabalho, busca aumentar o número de classes, para especificar melhor as áreas vegetadas e não vegetadas. As diversas classes mapeadas pelo projeto foram identificadas com base em imagens de toda a série Landsat, disponibilizadas pelo Serviço Geológico dos Estados Unidos (USGS) no catálogo de imagens do *Google Earth Engine* (GEE). Informações mais detalhadas sobre o projeto podem ser obtidas na página do MapBiomias na internet (mapbiomias.org).

Para o satélite Landsat-8, os dados do sensor *Operational Land Imager* (OLI) disponibilizados em reflectância de superfície não apresentaram um bom ajuste, por não estarem ortorretificados até a conclusão deste trabalho, não sendo, portanto, apropriados para a classificação. Isso justificou a adoção do dado em reflectância no topo da atmosfera (TOA) pelo MapBiomias na geração de suas classificações

da cobertura da superfície terrestre. No entanto, a necessidade de se constatar qual dado de reflectância (TOA ou de superfície) melhor se adequa à geração dessas classificações em ambientes da Caatinga, com características biofísicas semelhantes ao da área de estudo deste trabalho, é um aspecto relevante e que carece de investigação.

Diante disso, o objetivo do presente trabalho é comparar diferentes classificações da cobertura da superfície terrestre, tanto oriundas do dado em reflectância TOA como dos dados corrigidos pelo FLAASH e pelo QUAC (reflectância de superfície), na área abrangida pela carta topográfica SC-24-V-C (Petrolina), pertencente ao bioma Caatinga. As classificações deste trabalho foram geradas em imagens do sensor OLI, do satélite Landsat-8, e seguiram a lógica de processamento proposta pelo MapBiomas/Caatinga, com a incorporação dos dados corrigidos pelo FLAASH e pelo QUAC. Ao final, as classificações tiveram sua acurácia avaliada, a nível global e por classe, a fim de identificar o dado que melhor representa a superfície terrestre da área da carta topográfica em questão, com base no conjunto de classes definido.

2 Material e Métodos

O fluxograma da Figura 1 descreve as etapas metodológicas seguidas para o desenvolvimento

deste trabalho, com ênfase em cada software utilizado para o cumprimento de cada etapa. Como a unidade de mapeamento adotada neste trabalho é a Carta Internacional do Mundo ao Milionésimo na escala de 1:250.000, logo, a caracterização climático-temporal revela-se uma etapa inicial fundamental para identificar os melhores meses do ano para a seleção da carta e posterior geração do mosaico temporal.

O mosaico temporal foi gerado tanto em reflectância TOA como em número digital. O dado TOA seguiu a lógica de processamento adotada na coleção 2.1 do MapBiomas, com aplicação do modelo espectral de mistura e posterior geração da classificação. Já o mosaico em número digital foi submetido à correção atmosférica (FLAASH e QUAC), fora do ambiente de processamento do *Google Earth Engine*. Trata-se de uma etapa proposta por este trabalho. Em seguida, os mosaicos em reflectância de superfície foram reinseridos no GEE para aplicação do modelo espectral de mistura e posterior geração das classificações da cobertura da terra.

É importante destacar que a geração do mosaico temporal, a aplicação do modelo espectral de mistura e a geração da classificação da cobertura da terra foram desenvolvidas a partir de scripts pré-definidos (<http://mapbiomas.org/pages/methodology>). Ao final, na etapa de pós-classificação, os dados gerados foram submetidos a testes estatísticos para avaliar a acurácia de cada classificação.

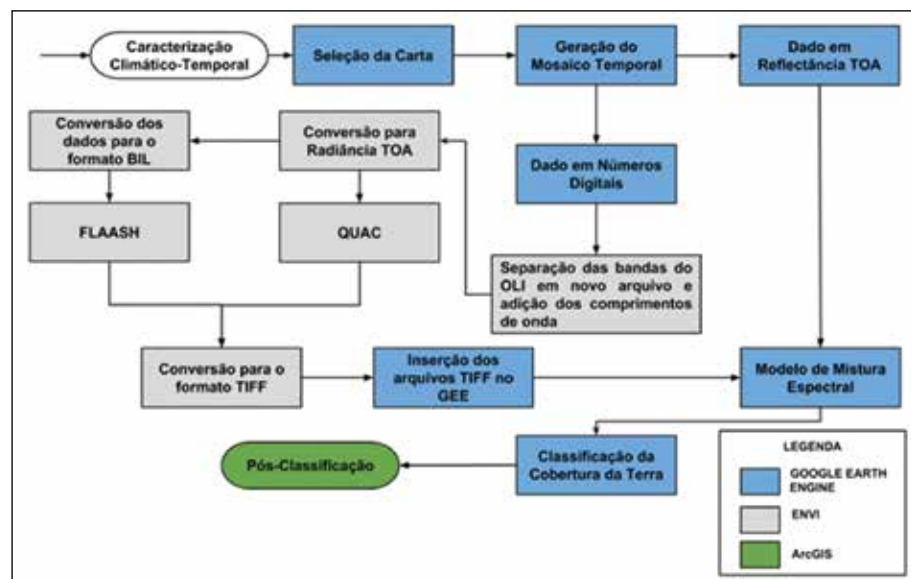


Figura 1 Fluxograma detalhando os processos realizados para a correção atmosférica dos dados, seu papel dentro da metodologia do MapBiomas e as ferramentas utilizadas para produzir cada dado.

2.1 Google Earth Engine (GEE)

O *Google Earth Engine* – GEE é uma plataforma utilizada em ampla escala com o objetivo principal de analisar dados ambientais. Uma das grandes vantagens dessa plataforma é o fato dela reunir mais de 40 anos de imagens globais de satélites, de décadas anteriores e atuais (Gorelick *et al.*, 2017). Essa plataforma também fornece ferramentas e uma vasta capacidade computacional para a realização de análises e a exploração de uma elevada amplitude de dados. Dentre suas principais aplicações destacam-se a detecção de desmatamento, a classificação da cobertura da terra e suas alterações correspondentes, assim como as estimativas de biomassa florestal e de carbono.

A ferramenta integra um conjunto de dados geoespaciais com cobertura global, capacidade de armazenamento e processamento altamente elevados, correspondentes ao ambiente de computação da nuvem Google, e um ambiente de desenvolvimento integrado suportando a implementação de algoritmos de análise e processamento de dados geoespaciais nas linguagens *Java Script* e *Python* (Gorelick *et al.*, 2017).

2.2 Landsat-8

O Landsat-8 foi lançado em 11 de fevereiro de 2013 e tem como objetivo principal coletar, arquivar, processar e distribuir dados científicos de forma consistente com o registro de dados existente dos demais sensores da série Landsat (Irons *et al.*, 2012). O satélite é referenciado ao sistema WRS-2 e opera em uma órbita polar heliosíncrona próxima de 705 km de distância da superfície terrestre e cruzando a linha do Equador aproximadamente às 10:13, horário médio local durante o nó descendente de cada órbita.

O Landsat-8 possui resolução radiométrica de 16 bits, resolução temporal de 16 dias e 185 km de faixa imageada e destaca-se pela presença de dois instrumentos imageadores: um óptico, *Operational Land Imager* (OLI), e outro abrangendo a faixa do infravermelho termal, *Thermal Infrared Sensor* (TIRS). O OLI possui 9 bandas espectrais: azul costal (banda 1), azul (banda 2), verde (banda 3), vermelho (banda 4), infravermelho próximo (banda 5), infravermelho de ondas curtas 1 (banda 6) e infravermelho de ondas curtas 2 (banda 7). Já o TIRS apresenta duas bandas espectrais de ondas longas (bandas 10

e 11) que equivalem ao comprimento de onda antes coberto por uma única banda do sensor TM, a banda 6 (USGS, 2012). A resolução espacial do OLI é de 30 metros e a do TIRS, que é originalmente de 100 metros, foi reamostrada pelo USGS para 30 metros para fins de comparação com dados do OLI.

2.3 Caracterização Climático-Temporal

Essa etapa caracterizou-se pela avaliação dos períodos de chuva a fim de estabelecer a janela temporal mais adequada para a geração dos mosaicos no contexto da Caatinga, visto que é importante desconsiderar pixels com informações do período da seca a fim de evitar maior confusão espectral entre as classes, especialmente as que se referem à vegetação. Para isso, foi feita uma avaliação prévia da média mensal dos dados fornecidos pelas estações pluviométricas do site do Instituto Nacional de Meteorologia (INMET, www.inmet.gov.br) a fim de identificar os meses que garantem uma boa representatividade do período úmido no bioma.

A partir das observações feitas aos dados pluviométricos do INMET nos limites da Caatinga, definiu-se que a janela temporal mais apropriada é de seis meses. Esse compromisso com o período de seis meses se dá, uma vez que esse intervalo temporal garante uma quantidade representativa de imagens para compor o mosaico temporal, visto que, uma parcela considerável das imagens do período úmido apresenta elevada quantidade de nuvens, o que comprometeria a análise das classes em mosaicos gerados a partir de uma janela temporal menor. Além disso, a janela temporal para a área em questão deve abranger, prioritariamente, os meses de janeiro a junho, fevereiro a julho e, em último caso, março a agosto. Isso ocorre porque o primeiro semestre sofre menor interferência das variações sazonais, especialmente no que se refere ao período de seca.

2.4 Seleção da Área de Estudo (Carta SC-24-V-C)

Como o sensor utilizado foi o *Operational Land Imager* (OLI) do satélite Landsat-8, só há imagens disponíveis a partir de 2013. A definição da área de estudo se baseou em uma avaliação da qualidade dos mosaicos das 68 cartas que abrangem o bioma Caatinga, tendo como parâmetro o percentual mínimo de nuvens presentes em cada carta. Essa

Impacto da Correção Atmosférica na Classificação da Cobertura da Terra da Carta Topográfica SC-24-V-C (Petrolina) Utilizando o Google Earth Engine e Imagens Landsat-8

Khalil Ali Ganem; Gustavo Macedo de Mello Baptista;

Washington de Jesus Sant'Anna da Franca Rocha; Rodrigo Nogueira Vasconcellos; Marcos Reis Rosa & Deorgia Tayane Mendes de Souza

avaliação considerou três categorias de percentual de cobertura de nuvens: boa (<10%), regular (entre 10 e 30%) e ruim (> 30%).

Nos anos considerados (2013 a 2016), a carta SC-24-V-C (Petrolina – PE) apresentou qualidade boa, dada a janela temporal de, no máximo, seis meses. O ano de 2013, em especial, apresentou presença quase nula de nuvens. Isso também pode ser explicado pelo fato desse ano abranger uma menor quantidade de cenas, já que o Landsat-8 foi colocado em órbita em fevereiro de 2013.

A carta SC-24-V-C está limitada pelas coordenadas 42°00' a 43°30' de longitude oeste e 13° a 14° de latitude sul, e está localizada na região norte da Bahia, totalizando uma área de cerca de 18 mil km².

Além disso, possui grande relevância econômica e ambiental para o bioma, por abranger o reservatório de Sobradinho e a cidade de Petrolina. O mapa da vegetação produzido pelo Ministério do Meio Ambiente (MMA), em parceria com diversos órgãos no âmbito do PROBIO (Figura 2) apresenta as formações vegetais e não vegetais presentes na carta.

2.5 Geração do Mosaico Temporal

O período considerado para a geração do mosaico temporal foi de 01/01/2013 a 01/07/2013 e o percentual de nuvens considerado foi de 4%. No entanto, cabe considerar que como as cenas do Landsat-8 começaram a ser disponibilizadas de março em diante para essa região, logo o intervalo do mosaico em questão foi de 3 meses, aproximadamente.

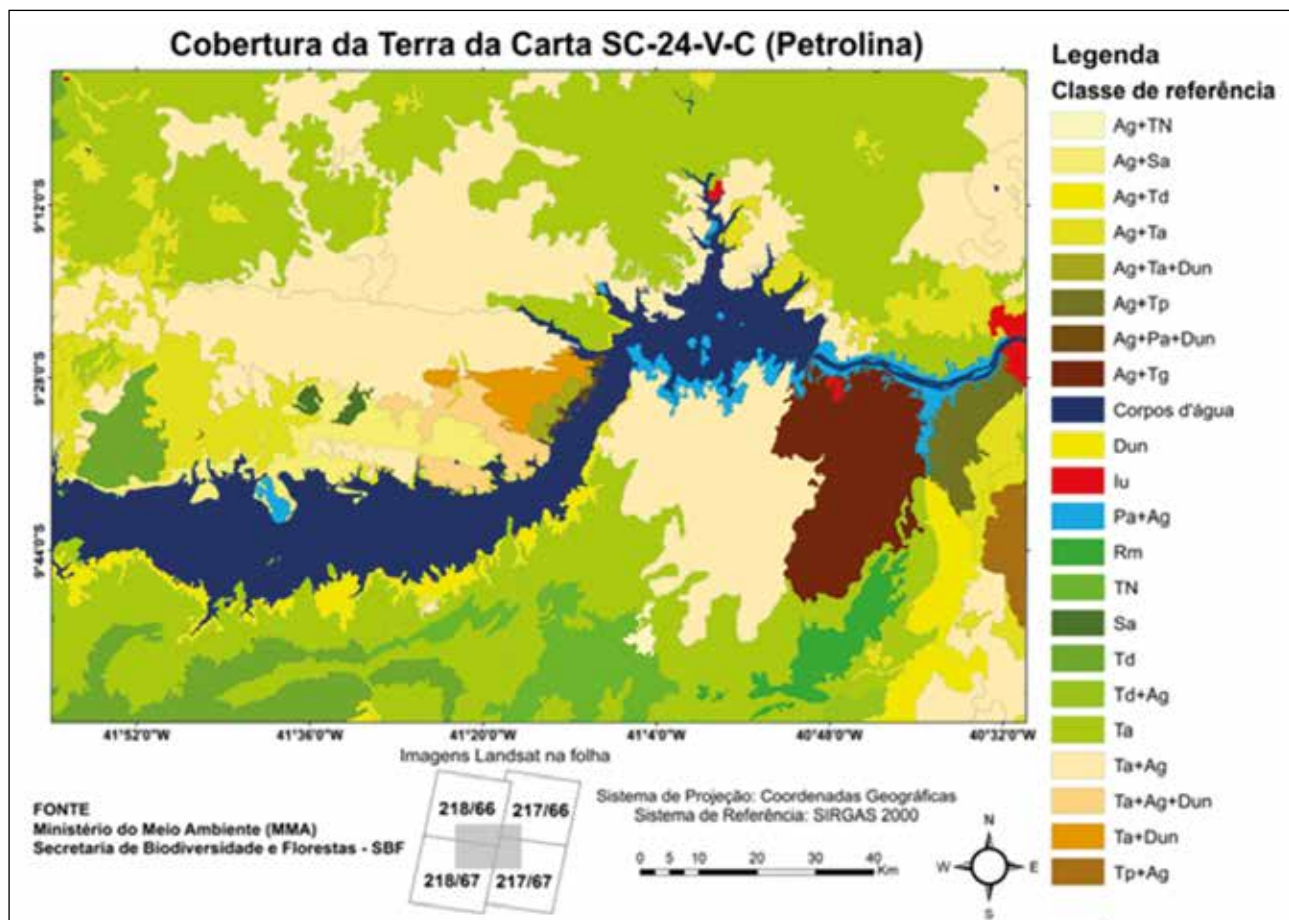


Figura 2 Mapa de Referência da vegetação a carta SC-24-V-C. Legenda: Áreas de Tensão Ecológica (TN), caracterizadas por savana estépica/floresta estacional; Savana Arborizada (Sa); Savana-Estépica Florestada (Td), Savana-Estépica Arborizada (Ta); Savana-Estépica Parque (Tp); Savana-estépica gramíneo-lenhosa (Tg); Dunas (Dun); Influência urbana (Iu); Formação pioneira com influência fluvial e/ou lacustre (Pa); Agropecuária (Ag); e Refúgio Montano (Rm). Fonte: Adaptado de MMA (2006).

O mosaico temporal em questão foi composto por um total de 9 cenas dos seguintes pontos/órbitas: 217/66 (1 cena, datada de 30/05/2013), 217/67 (1 cena, datada de 30/05/2013); 218/66 (3 cenas, datadas de 05/05/2013, 06/06/2013 e 22/06/2013) e 218/67 (4 cenas, datadas de 05/05/2013, 21/05/2013, 06/06/2013 e 22/06/2013). Para obter o mosaico, inicialmente, removeu-se a borda da cena Landsat para suprimir pixels que não possuem dados em todas as bandas (ruídos).

O próximo passo caracterizou-se pela aplicação de um algoritmo de detecção e remoção das nuvens e das sombras provocadas pelas mesmas em imagens Landsat, o *Function of Mask* (FMASK), proposto por Zhu & Woodcock (2012). O algoritmo requer como entrada, dados de reflectância no topo da atmosfera e de temperatura de brilho. A partir de critérios baseados nas propriedades físicas das nuvens, os pixels considerados como nuvens potenciais são separados dos pixels de céu claro.

Em seguida, por meio da combinação das probabilidades dos parâmetros temperatura, variabilidade espectral e brilho dos alvos, é produzida uma “máscara” para nuvens sobre superfícies terrestres e aquáticas, separadamente. Os valores da “máscara” criada pelo algoritmo são: 0 – céu claro; 1 – água; 2 – sombra; 3 – neve; 4 – nuvem; e 255 – não identificado.

Após a aplicação do FMASK, os pixels marcados como nuvem ou sombra de nuvem são removidos da imagem original, mantendo apenas pixels com valores válidos. Por fim, uma redução das cenas consideradas no intervalo temporal pré-definido pelo valor da mediana dos pixels válidos gera o mosaico da imagem final da carta. O filtro da mediana é importante por remover os valores extremos em cada pixel, que podem ser ruídos do sensor ou contaminação por névoa.

2.6 Correção Atmosférica

Primeiramente, os dados brutos foram convertidos para radiância no topo da atmosfera (TOA). De acordo com Chander *et al.* (2009), durante a calibração radiométrica, os valores dos pixels das ima-

gens brutas são convertidos para valores de radiância absoluta. Para a conversão para radiância TOA (ou radiância a nível do sensor), é preciso conhecer os fatores de reescalonamento originais, disponibilizados nos metadados de cada imagem.

A conversão dos dados de radiância para reflectância de superfície foi feita por meio dos métodos de correção atmosférica FLAASH e QUAC, ambos presentes no software ENVI 5.1. Para ambos os métodos de correção atmosférica, considerou-se o dado de radiância como entrada. No caso do FLAASH, utilizou-se, basicamente, fator de escala (*single scale factor*) 10 e taxa de CO₂ de 396 ppm. A altura de escala do aerossol (*aerosol scale height*) é definida, de acordo com Turner *et al.* (2001), como a altura inferior pela qual a profundidade óptica do aerossol atinge 63% da sua camada inteira. Assim, tendo em vista que os seus valores variam entre 1 e 2 km, optou-se por manter o valor recomendado pelo FLAASH, ou seja, 1,5 km.

Idealmente, deve-se utilizar informações atmosféricas relativas ao momento da passagem do satélite (aquisição de dados) para se proceder com a correção atmosférica em um maior nível de precisão. No entanto, como tais informações são de difícil obtenção, utiliza-se uma aproximação. O FLAASH disponibiliza seis modelos atmosféricos (perfis atmosféricos padrão), cada um com uma quantidade padrão de vapor d'água e de temperatura de superfície. Diante disso, a escolha do modelo atmosférico para processar o FLAASH foi o Tropical, uma vez que a área de estudo possui temperaturas bastante elevadas ao longo do ano. Além disso, o módulo de correção atmosférica do guia do usuário do ENVI (2009), afirma que a posição latitudinal da área de estudo em função da sazonalidade também embasa a escolha do modelo atmosférico mais apropriado. Com isso, como a carta SC-24-V-C está situada entre 9° e 10°, o modelo Tropical mostra-se ideal em qualquer período do ano. Esse modelo atmosférico assume uma quantidade de vapor d'água de 4,11 g/cm² presente na coluna de ar atmosférico (ENVI, 2009). O parâmetro visibilidade, por sua vez, é estimado pela própria imagem, quando nela são detectados pixels escuros. Entretanto, antes de processar o FLAASH, é importante definir este parâmetro. Para isso, utilizou-se o valor de 36,31 km.

Para a maioria dos sensores multiespectrais, incluindo o Landsat, o campo *Water Retrieval* no FLAASH é automaticamente ajustado para a opção “não”, o que quer dizer que a estimativa da presença de água não é restaurada em cada pixel. Quando isso acontece, é atribuído um valor constante para o valor da coluna de vapor d’água em todos os pixels da imagem, que deve ser preenchido no campo *Water Column Multiplier*. O valor padrão adotado neste trabalho foi 1, o que se justifica pelo fato de o Landsat não possuir uma banda própria para corrigir o vapor d’água. Alguns estudos procederam da mesma forma ao corrigirem os efeitos atmosféricos com o FLAASH em imagens Landsat (Nascimento & Sano, 2010; Mitchell *et al.*, 2013; Els *et al.*, 2015).

Além disso, o modelo de aerossol escolhido para a área de estudo foi o Rural, uma vez que a área da carta SC-24-V-C não sofre forte influência de áreas urbano-industriais. Esse modelo de aerossol é composto pela mistura de dois padrões de partícula, uma maior e outra menor. É importante destacar que a relevância desse parâmetro para o FLAASH se torna maior caso a visibilidade seja menor que 40 km, como foi o caso da cena em questão. Para determinar a quantidade de aerossol, no campo *Aerosol Retrieval* foi selecionada a opção 2-Band (K-T), a qual utiliza as bandas 4 (655 nm) e 7 (2200 nm) do sensor OLI do Landsat-8, assim como uma razão quase constante de reflectância entre essas bandas para identificar pixels escuros. Ao selecionar esta opção, caso os pixels escuros adequados não sejam encontrados para o dado de entrada, o valor preenchido no campo *Initial Visibility*, que se trata da visibilidade inicial, passa a ser considerado. O valor padrão de visibilidade mantido pelo FLAASH é de 40 km. Entretanto, é recomendável que este parâmetro seja estimado à máxima exatidão, uma vez que pode variar de cena para cena. Os valores utilizados para os demais parâmetros foram os pré-definidos pelo FLAASH (default).

2.7 Modelo de Mistura Espectral

O Modelo de Mistura Espectral (MME), proposto por Smith *et al.* (1985), é um modelo linear baseado na mistura espectral de dois ou mais espectros de reflectância de componentes puros, ou *end-*

members, que permite representar a variabilidade na composição e iluminação de uma imagem.

O MME aplicado no MapBiomas utiliza uma biblioteca espectral definida por quatro componentes puros: Vegetação (GV), Vegetação Não-Fotosinteticamente Ativa (NPV), Solo (SOIL) e Nuvem (CLOUD). Essa biblioteca espectral foi definida originalmente para o bioma Amazônia, de acordo com estudo desenvolvido por Souza Jr. *et al.* (2005). No caso deste trabalho, foram utilizadas as bandas de 2 a 7 do Landsat-8.

O MME assume que a reflectância de cada pixel, em cada banda da imagem Landsat, resulta da combinação linear do produto do percentual do componente puro (leia-se, a proporção no pixel), pelo seu respectivo valor na banda (Equação 1).

$$Rb = F_i R_{i,b} + \epsilon b \quad (1)$$

onde:

Rb é a reflectância na banda b;

R_{i,b} é a reflectância do componente puro i, na banda b;

F_i a fração (ou proporção) do componente puro i, e;

εb é o erro residual de cada banda.

O MME calcula uma fração para cada alvo. Assim, os valores dos componentes puros GV, NPV, SOIL e CLOUD podem variar de 0 a 1 (ou 0 a 100%). O efeito da sombra (SHADE), conforme mostra a Equação 2, é calculado pela subtração de 1 pela soma das frações dos componentes puros (GV + NPV + SOIL + CLOUD), podendo também variar de 0 a 1. A sombra aqui referida está relacionada ao alvo (por exemplo, sombra da rugosidade da copa das árvores) e não possui ligação com a sombra removida pelo FMASK, que é oriunda apenas de nuvens.

$$SHADE = 1 - \sum F_i \quad (2)$$

A Equação 3 demonstra como normalizar a vegetação (GV) pela sombra (SHADE). As frações são necessárias para o cálculo do Índice Normalizado por Diferença de Fração (NDFI), proposto por Souza Jr. *et al.* (2005), o qual é obtido pela diferença

normalizada da imagem fração Vegetação normalizada pela sombra (GVs) e da soma das imagens fração Vegetação Não- Fotossinteticamente Ativa (NPV) e Solo (SOIL), conforme demonstrado pela Equação 4. O NDFI, por ser um índice normalizado, varia de -1 a 1.

$$GVs = \frac{GV}{(100 - SHADE)} \quad (3)$$

$$NDFI = \frac{GVs - (NPV + SOIL)}{GVs + NPV + SOIL} \quad (4)$$

As imagens fração dos componentes puros mencionados também foram necessárias para calcular o *Savannah Ecosystem Fraction Index* (SEFI), desenvolvido pela equipe do MapBiomias, do bioma Pantanal (Projeto MapBiomias, 2017b). O SEFI foi desenvolvido a partir de um ajuste do índice NDFI e também é utilizado na Caatinga, conforme apresentado pela Equação 5. Já a Equação 6 apresenta o cálculo da variável GVNPs.

$$SEFI = \frac{GVNPs - (SOIL + CLOUD)}{GVNPs + SOIL + CLOUD} \quad (5)$$

$$GVNPs = \frac{(GV + NPV)}{(100 - SHADE)} \quad (6)$$

Esse ajuste foi feito com o objetivo de minimizar o efeito da sazonalidade da Caatinga no NDFI, uma vez que o período seco afeta a vegetação, no sentido de promover uma maior predominância da vegetação seca em detrimento da vegetação verde, diminuindo os resultados que indicam maior predominância de solo exposto em períodos de seca, ou seja, reduzindo a variação do NDFI.

2.8 Classificação da Cobertura da Terra

As imagens fração obtidas a partir do MME foram utilizadas na classificação da cobertura da terra para as seguintes classes: Agropecuária, Água, Áreas Não Vegetadas, Área Úmida (não identificada na carta SC-24-V-C), Floresta Aberta, Floresta Den-

sa, e Outras Formações Nativas Não Florestais. As classes de vegetação foram definidas de acordo com o Manual Técnico de Vegetação do Instituto Brasileiro de Geografia e Estatística (IBGE, 2012). A Floresta Densa equivale à Floresta Estacional Decidual e Semidecidual, assim como a Savana Estépica-Florestada. Já a Floresta Aberta, se refere à Savana-Estépica Arborizada.

A classe Outras Formações Nativas Não Florestais, refere-se à Savana-Estépica Parque e Savana-Estépica gramíneo-lenhosa, ou seja, com um padrão maior de espaçamento entre a vegetação. Na legenda do MapBiomias e nas figuras com os resultados obtidos nesse trabalho, essa classe foi nomeada, genericamente, como Savana. Por fim, cabe ressaltar que a classe Áreas Não Vegetadas se refere às áreas urbanas, afloramentos rochosos e solo exposto.

Com a caracterização das classes, foi aplicada uma “máscara” para mapear áreas cobertas por nuvens que não puderam ser filtradas durante o processo de geração do mosaico temporal da carta. Em seguida, foi aplicada uma árvore de decisão (Figura 3), definida empiricamente para mapear cada classe (Projeto MapBiomias, 2017a).

O ponto de partida para a interpretação da árvore é o algoritmo FMASK. A escolha desse parâmetro como raiz parte do princípio de que a água é um alvo facilmente identificado e que se confunde menos com os principais alvos presentes na cobertura da terra.

Para a árvore de decisão apresentada, foi feito um reescalonamento dos valores do SEFI, os quais passaram a variar de 0 a 200, sendo que, quanto maior a presença de vegetação verde e sombra, mais próximos de 200 ficaram os valores. O reescalonamento foi feito por meio da soma do pixel após a aplicação do referido índice a 1, seguida da sua multiplicação por 100. Esse nó é responsável pela distinção entre um pixel que pode ser classificado como Floresta Densa ou Floresta Aberta, de um pixel que pode ser classificado como outra formação nativa não florestal, como o caso da Savana.

O parâmetro SHADE aparece em dois nós, sendo importante deixar claro que o valor do nó que

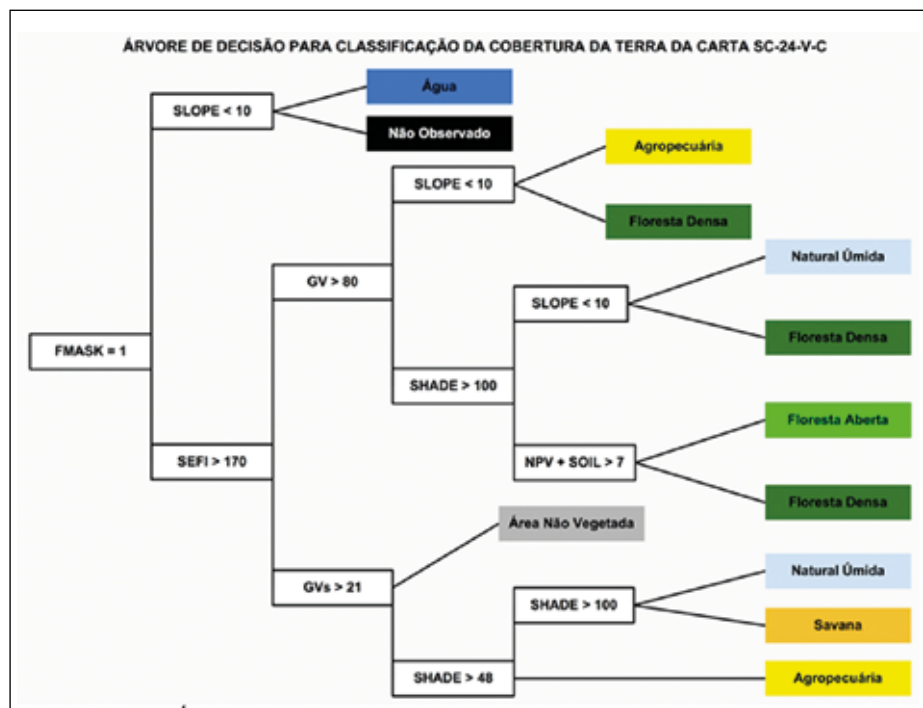


Figura 3 Árvore de decisão para classificação da cobertura da terra na Caatinga.

caracteriza a classe Savana deve ser sempre maior que o valor do nó anterior, que caracteriza a classe Agropecuária. Isso ocorre porque esse último nó classifica apenas pixels com SEFI baixo e, consequentemente, acaba sendo mais eficiente na diferenciação de áreas com pouca sombra, que, geralmente, trata-se de pastagem e áreas com aspectos vegetais gramíneo-herbáceo. Já a classe Natural Úmida (áreas alagadas) não foi identificada na carta em questão.

A informação com relação à declividade (SLOPE) do terreno foi proveniente dos dados *Shuttle Radar Topography Mission* (SRTM), disponibilizados pelo USGS, e foi utilizada para fazer correções mais simples, como vegetações com maior padrão de espaçamento entre si. Áreas com alto GV costumam ter pouca sombra e por isso são, normalmente, classificadas como áreas de agricultura, conforme observado em testes de parametrização realizados. Entretanto, cabe fazer uma ressalva. Vertentes de alta declividade podem ter muita iluminação solar e, por isso, tendem a ser classificadas como agricultura, mesmo que se trate de floresta densa. Isso faz com que a interferência da iluminação solar aliada à declividade seja determinante para diferenciar ambas as classes.

2.9 Pós-Classificação

A avaliação da acurácia temática para as classificações foi feita a partir das seguintes etapas: definição do número de amostras mínimas para garantir um bom desempenho da avaliação; distribuição aleatória dessas amostras nas imagens Landsat em composição colorida; checagem por interpretação visual; verificação da exatidão por meio do uso da matriz de confusão; e cálculo de índices de concordância de forma geral e por classe.

Para o cálculo do número de amostras, foi utilizado o método de distribuição multinomial (Congalton & Green, 2008), que leva em consideração o número de classes e a proporção da classe com maior área na classificação. Além disso, a unidade de amostragem utilizada para este trabalho foi o pixel (30 metros para o sensor OLI do Landsat-8), uma vez que a similaridade entre as informações espectrais de cada um serviu como embasamento para gerar as classificações. Essa forma de validação, com base na estrutura de imagem de satélite, é bastante utilizada, por permitir um levantamento ágil e de boa precisão (Epiphanyo *et al.*, 2002), especialmente se comparado aos levantamentos de campo.

Definido o tamanho da amostra para a carta, realizou-se uma amostragem aleatória simples para a imagem Landsat, por meio da ferramenta *Hawths Analysis Tools* disponível no ArcGIS 10.2, o qual utilizou, como espaço amostral, um arquivo de pontos do centroide de cada pixel das imagens de cada carta.

O passo seguinte caracterizou-se pela atribuição das classes correspondentes a cada amostra por meio de interpretação visual na tela do computador, utilizando imagem de alta resolução espacial do aplicativo *Google Earth*. Esses pontos foram checados e classificados com rigor, para então serem utilizados como dados de referência (verdade de campo) para construir a matriz de confusão, a qual permitiu computar a exatidão global e os coeficientes de concordância total e por classe.

O coeficiente de concordância total mais utilizado, segundo Levada (2010), é a taxa de acerto ou Exatidão Global (Hellden & Stern, 1980). Porém, apenas o seu cálculo não é suficiente, havendo a necessidade de índices que levam em conta a matriz por completo, como os coeficientes de concordância *Tau* (Klecka, 1980) e *Kappa* (Cohen, 1960). A estatística *Kappa* inclui todos os pixels da imagem, sejam aqueles classificados corretamente, ou aqueles classificados erroneamente (Moreira, 2001). Já o *Tau* baseia-se na probabilidade a priori, permitindo, assim, a obtenção do valor da concordância esperada, mesmo antes de efetuar a checagem dos resultados da simulação de mapeamento.

3 Resultados e Discussão

A informação derivada dos dados de sensores remotos, de acordo com Jensen (2009), é geralmente apresentada como uma imagem realçada, um mapa-imagem, mapas temáticos, arquivos espaciais em banco de dados, estatísticas ou gráficos. No caso deste trabalho, os resultados foram apresentados em mapas e análises estatísticas.

3.1 Mapas de Classificação da Cobertura da Terra dos Dados com e sem Correção Atmosférica

Analisando visualmente o mapa da Figura 4, é possível observar maior predomínio da classe Áreas Não Vegetadas no resultado oriundo do dado TOA,

seguida da classe Floresta Aberta. As classes Savana e Agropecuária, por sua vez, aparecem em menor proporção. A Água foi classificada adequadamente e praticamente não foi detectada Floresta Densa.

A classificação exibida no mapa da Figura 5, gerada a partir do dado em reflectância de superfície, gerado pelo FLAASH, mostra um grande predomínio de Áreas Não Vegetadas. Em comparação com o dado TOA, nota-se que a Floresta Aberta foi subestimada. Já em comparação com o dado corrigido pelo QUAC, além da Floresta Aberta, observa-se também uma subestimação da classe Savana. Com relação às classes Agropecuária e Floresta Densa, verifica-se uma maior semelhança desse dado com o dado TOA. A Água foi bem representada na classificação.

Já no mapa da Figura 6 (QUAC), verifica-se que as classes Savana e Áreas Não Vegetadas são as mais abrangentes na imagem. Em comparação com a classificação sem correção atmosférica (TOA), observa-se que a classe Floresta Aberta foi subestimada, enquanto a classe Floresta Densa foi melhor evidenciada. A classe Água foi classificada adequadamente, enquanto a classe Agropecuária não foi bem detectada.

3.2 Matriz de Confusão das Classificações da Cobertura da Terra com e sem Correção Atmosférica

As matrizes de confusão permitem que seja feita uma avaliação do grau de confusão entre as classes de referência (verdade de campo) e das classes de cada mapa resultante. A partir da matriz de confusão é possível calcular índices globais e por classe (Congalton & Green, 2008). A Tabela 1 reúne as matrizes de confusão das classificações oriundas dos dados sem correção (TOA) e com correção atmosférica (FLAASH e QUAC).

Observando a Tabela 1 é possível constatar que, de um total de 696 pontos distribuídos aleatoriamente sobre a carta e que representam os dados de referência (verdade de campo), 430 se mostraram concordantes com a classificação oriunda dos dados em reflectância TOA, 343 com a classificação oriunda dos dados corrigidos pelo FLAASH e 453

Impacto da Correção Atmosférica na Classificação da Cobertura da Terra da Carta Topográfica SC-24-V-C (Petrolina) Utilizando o Google Earth Engine e Imagens Landsat-8

Khalil Ali Ganem; Gustavo Macedo de Mello Baptista;

Washington de Jesus Sant'Anna da Franca Rocha; Rodrigo Nogueira Vasconcellos; Marcos Reis Rosa & Deorgia Tayane Mendes de Souza

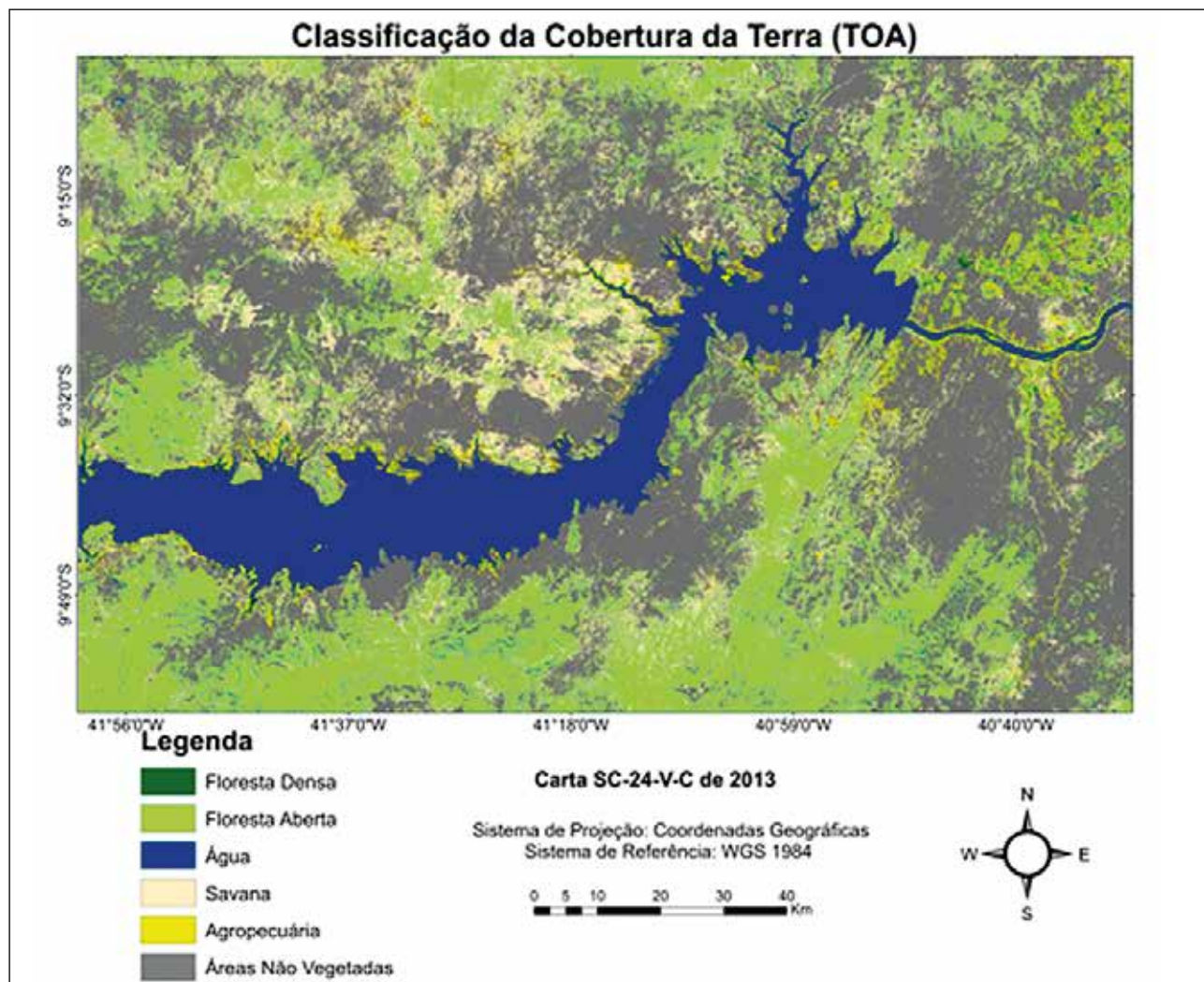


Figura 4 Mapa da Classificação da Cobertura da Terra a partir do dado em reflectância no topo da atmosfera da carta SC-24-V-C processado pelo Google Earth Engine.

		REFERÊNCIA (VERDADE)						
		AGROPECUÁRIA	ÁGUA	ÁREAS NÃO VEGETADAS	FLORESTA ABERTA	FLORESTA DENSE	SAVANA	TOTAL
CLASSIFICAÇÃO TOA	Agropecuária	6	-	3	-	-	5	14
	Água	-	65	-	-	-	-	65
	Áreas Não Vegetadas	9	-	133	48	-	116	306
	Floresta Aberta	6	-	1	169	11	36	223
	Floresta Densa	1	-	-	-	-	-	1
	Savana	2	-	2	25	1	57	87
	Total	24	65	139	242	12	214	696
CLASSIFICAÇÃO FLAASH	Agropecuária	7	-	3	1	-	6	17
	Água	-	65	-	-	-	1	66
	Áreas Não Vegetadas	10	-	133	69	4	132	348
	Floresta Aberta	4	-	1	67	6	4	82
	Floresta Densa	-	-	1	1	-	-	2
	Savana	3	-	1	104	2	71	181
	Total	24	65	139	242	12	214	696
CLASSIFICAÇÃO QUAC	Agropecuária	-	-	-	-	-	-	-
	Água	1	65	1	1	-	1	69
	Áreas Não Vegetadas	8	-	121	21	1	67	218
	Floresta Aberta	3	-	-	128	6	9	146
	Floresta Densa	4	-	-	2	2	-	8
	Savana	8	-	17	90	3	137	255
	Total	24	65	139	242	12	214	696

Tabela 1 Matriz de Confusão das classificações da cobertura da terra oriundas dos dados sem correção atmosférica (TOA) e corrigidos pelo FLAASH e QUAC para a carta SC-24-V-C do ano de 2013.

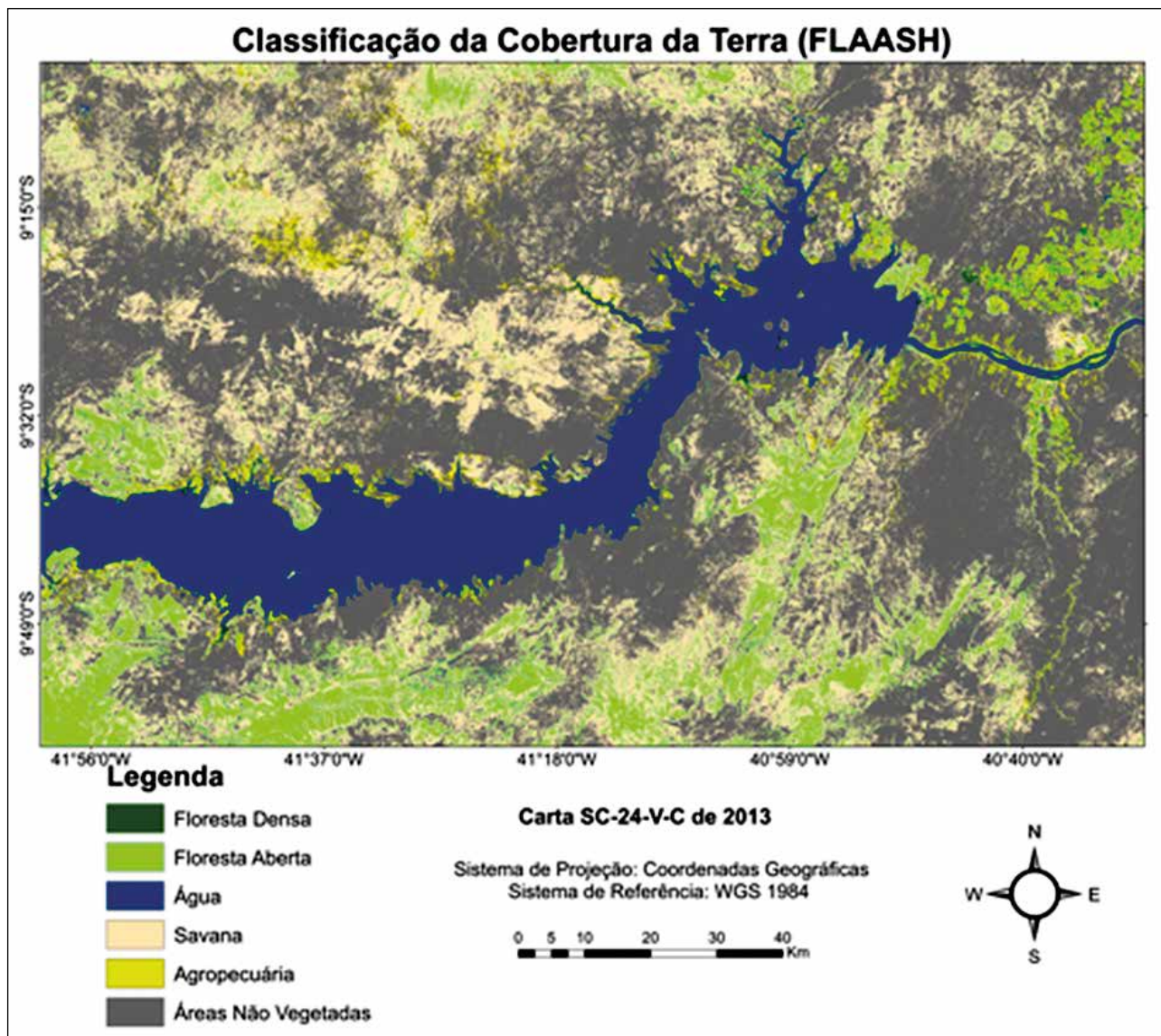


Figura 5 Mapa da Classificação da Cobertura da Terra processado pelo Google Earth Engine a partir do dado de reflectância de superfície gerado pelo FLAASH.

com a classificação oriunda dos dados corrigidos pelo QUAC.

Ainda de acordo com a Tabela 1, a Água foi a classe que apresentou os melhores resultados em todas as classificações, seguida da classe Áreas Não Vegetadas, que também foi muito bem representada, especialmente nas classificações oriundas do dado TOA e do dado corrigido pelo FLAASH.

Já os piores resultados variaram em razão de cada classificação. A classe Floresta Densa, por

exemplo, foi muito mal representada em todas as classificações. No entanto, houveram menos de 2% da totalidade dos pontos amostrais para a classe, o que indica sua baixa representatividade. Nas três classificações, a classe Agropecuária apresentou resultados muito baixos, chegando a zero no caso do dado corrigido pelo QUAC.

Dentre as classes de vegetação, a Floresta Aberta foi a mais bem representada na classificação que utilizou o dado TOA, a Savana melhor repre-

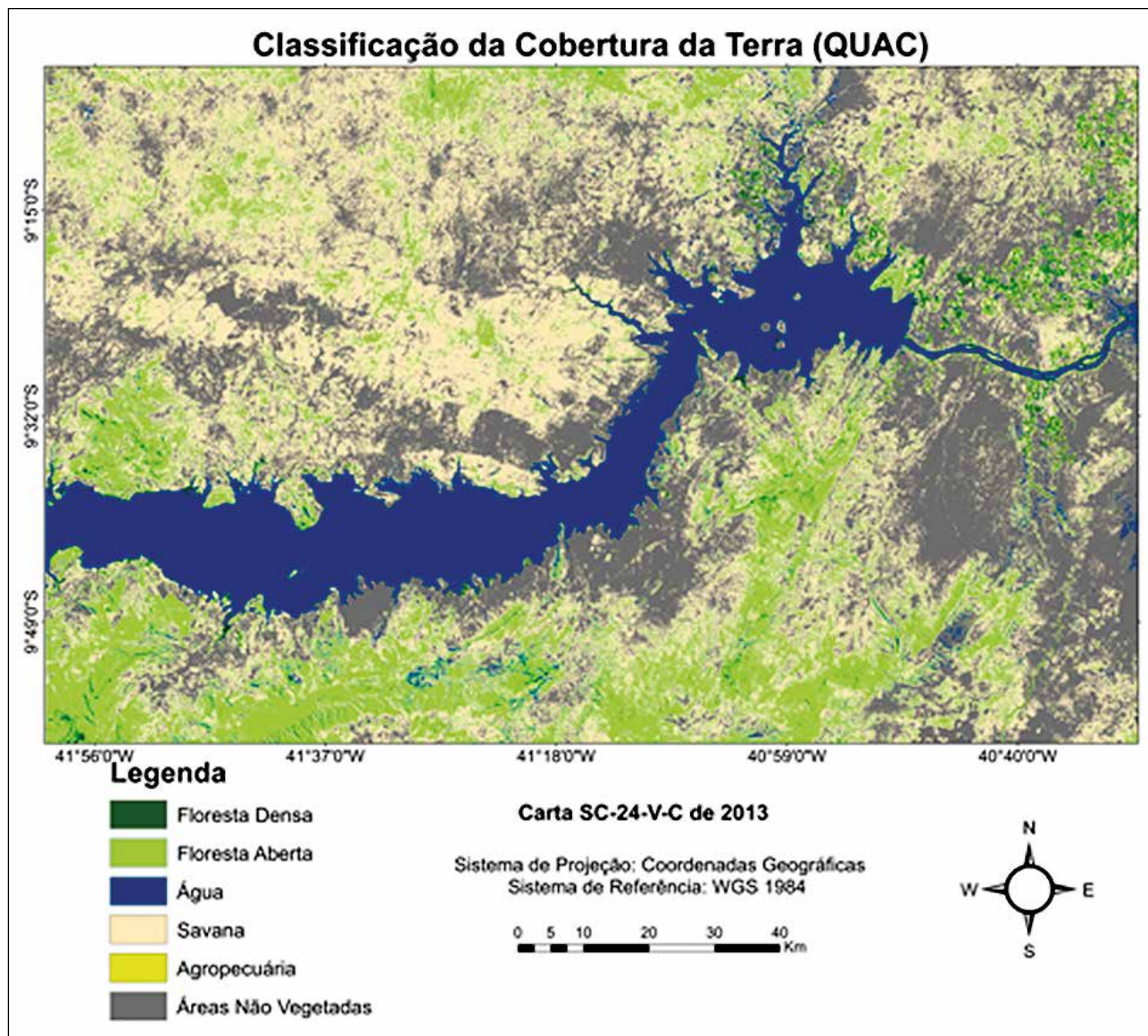


Figura 6 Mapa da Classificação da Cobertura da Terra processado pelo Google Earth Engine a partir do dado de reflectância de superfície gerado pelo QUAC.

sentada nas classificações que utilizaram dados atmosféricamente corrigidos, tendo o dado corrigido pelo QUAC apresentado resultados superiores em comparação com o dado corrigido pelo FLAASH.

No entanto, é importante ressaltar que a classe Savana foi muito confundida com as Áreas Não Vegetadas em todas as três classificações, o que mostra a necessidade de se aprimorar os parâmetros da árvore de decisão, de modo que seja possível melhor distinguir essas classes.

3.3 Comparação entre os Resultados Obtidos a partir dos Coeficientes de Concordância Globais e por Classe

A Tabela 2 apresenta uma síntese dos resultados estatísticos a partir do uso dos índices de concordância globais para as classificações baseadas nos três diferentes dados de reflectância utilizados.

As diferenças encontradas entre a exatidão global e o índice *Kappa* é normal (Mather, 2004), visto que o segundo resume a informação provenien-

te da matriz de erro; enquanto o primeiro é a média da acurácia individual das classes. Brites *et al.* (1996) afirma que na formulação do índice *Tau*, o valor da concordância casual é estabelecido *a priori*, evitando falhas que incorrem da exatidão global e do índice *Kappa*. Isso faz com que o *Tau*, de acordo com o autor, talvez seja o índice que represente com maior fidedignidade a exatidão dos processos de classificação.

	TOA	FLAASH	QUAC
Exatidão Global	61,78%	49,28%	65,08%
<i>Kappa</i>	49,24%	34,08%	52,97%
<i>Tau</i>	54,13%	39,13%	58,10%

Tabela 2 Coeficientes de Concordância Global para cada método.

A Tabela 2 mostra uma exatidão global de 61,78%, *Kappa* de 49,24%, considerado bom, e *Tau* de 54,13%, para a classificação TOA, exatidão global de 49,28%, *Kappa* de 34,08%, considerado ruim, e *Tau* de 39,13%, para a classificação oriunda dos dados corrigidos pelo FLAASH, e exatidão global de 65,08%, *Kappa* de 52,97%, considerado bom, e *Tau* de 58,10%, para a classificação cujo dado foi corrigido pelo QUAC. Assim, analisando as classificações em função dos dados de reflectância utilizados, verifica-se que a classificação oriunda de dados corrigidos pelo QUAC obteve o melhor resultado. Isso representa uma grande vantagem para o método, considerando sua maior rapidez e menor complexidade para corrigir os efeitos atmosféricos.

Muito embora os coeficientes de concordância global forneçam um bom panorama com relação ao impacto do dado de reflectância utilizado na classificação, faz-se necessário avaliar também os valores dos coeficientes de concordância por classe,

tendo em vista o fato de permitirem a identificação das classes que mais se confundiram entre si, além do seu grau de subestimação ou superestimação em cada classificação. Assim, a Tabela 3 apresenta os resultados das exatidões do produtor e do usuário para cada classe avaliada. Enquanto a exatidão do produtor indica a probabilidade de acerto da classificação (erro de omissão), a exatidão do usuário indica a probabilidade de um pixel classificado na imagem, de fato, corresponder com o dado de referência (verdade de campo), estando, portanto, associada ao erro de comissão.

Para todos os dados de reflectância, com atmosfera corrigida ou não, a Água foi a classe que apresentou melhor resultado, com 100% de exatidão do produtor. A Floresta Densa apresentou o pior resultado. As Áreas Não Vegetadas apresentaram valores de exatidão do produtor bastante elevados, o que reforça a probabilidade de acerto na classificação para esta classe.

A exatidão do usuário também foi alta para a Água nos três métodos, seguida da classe Floresta Aberta, com destaque para o QUAC, com 87%. O dado corrigido por esse método foi o único que apresentou algum grau de probabilidade de pixels da classe Floresta Densa de fato corresponderem com o dado de referência, já que os outros dois dados de reflectância apresentaram resultados nulos para a classe. O QUAC também apresentou maior exatidão do usuário para a classe Áreas Não Vegetadas, com 55%. Já o dado gerado pelo FLAASH não se destacou por ser o melhor para nenhuma classe, comportamento esse já indicado pelos coeficientes de concordância global.

A Tabela 3 permite concluir que a Agropecuária foi melhor classificada pelo dado TOA, mesmo tendo apresentado exatidões inferiores a 50%.

Tabela 3 Coeficientes de concordância global por classe das classificações geradas a partir dos dados sem correção atmosférica (TOA) e corrigidos pelo FLAASH e QUAC.

CLASSES	EXATIDÃO DO PRODUTOR			EXATIDÃO DO USUÁRIO		
	TOA	FLAASH	QUAC	TOA	FLAASH	QUAC
Agropecuária	0,25	0,29	-	0,42	-	-
Água	1	1	1	1	0,98	0,94
Áreas Não Vegetadas	0,95	0,95	0,87	0,43	0,38	0,55
Savana	0,26	0,33	0,64	0,65	0,39	0,53
Floresta Aberta	0,70	0,27	0,52	0,75	0,81	0,87
Floresta Densa	-	-	0,16	-	-	0,25

A Água foi classificada com alto nível de exatidão, tendo apresentado valores superiores a 90% tanto para a exatidão do produtor como para a exatidão do usuário.

Para a classe Áreas Não Vegetadas, o dado TOA apresentou exatidão do produtor de 95%. Entretanto, o dado corrigido pelo QUAC apresentou 55% de exatidão do usuário. A Savana apresentou valores mais baixos de exatidão do produtor para o dado TOA. Cabe ressaltar ainda, que para a classe Floresta Densa, apesar dos dados muito baixos para os coeficientes utilizados, apenas o dado corrigido pelo QUAC apresentou resultado diferente de zero de exatidão do produtor.

Sano *et al.* (2007) enfatizaram que a ocorrência de confusão espectral entre certas classes vegetais representa grandes desafios no mapeamento temático para o bioma Cerrado, o qual mais se aproxima da Caatinga dentre os biomas brasileiros. Os autores também destacaram a sazonalidade marcante da vegetação natural como sendo outro fator que exerce grande influência no processo de mapeamento do bioma. A resolução espacial de 30 metros do Landsat é outro fator que pode ter contribuído para as confusões entre as classes, já que um pixel pode vir a abranger mais de uma classe, especialmente considerando que os pontos amostrais não compreendem toda a dimensão do pixel. Galletti & Myint (2014) também alegaram que a confusão espectral é um aspecto inevitável em regiões áridas e semiáridas, especialmente entre classes de agropecuária, com solo exposto, ou áreas pouco povoadas.

Isso pode explicar a presença marcante da classe Áreas Não Vegetadas nos resultados das classificações, especialmente tendo em vista o fato de o mosaico temporal ter considerado imagens de períodos de menor ocorrência de chuva, o que aumentou a semelhança da vegetação seca com solo exposto. Ainda nesse sentido, é importante destacar o fato de a cobertura vegetal da Caatinga estar passando por uma grande transição, com aumento do desmatamento pela retirada da vegetação do bioma e substituição rápida e progressiva por áreas com pastagem e agricultura, conforme constatado por Coelho *et al.* (2014), em uma bacia do semiárido no estado de Pernambuco.

Um aspecto que merece ser destacado na análise dos resultados das classificações geradas é em relação à interferência de questões climáticas aliadas ao fator sazonalidade. O ano de 2013 registrou a maior seca dos últimos 50 anos que o antecederam, conforme mostra o relatório “Declaração sobre o Estado do Clima”, divulgado pela Organização Meteorológica Mundial, agência das Nações Unidas especializada em monitorar eventos climáticos (WMO, 2014). O evento foi tão forte que resultou na morte de cerca de 4 milhões de animais, além da baixa dos reservatórios de água das hidrelétricas na região.

Em relação aos dados utilizados para gerar cada classificação, o pior desempenho do dado corrigido pelo FLAASH pode estar associado diretamente às incertezas oriundas do uso de um perfil atmosférico constante, diferente das condições da atmosfera no momento da passagem do satélite. No entanto, há casos em que essa aproximação pode gerar um resultado satisfatório, como no estudo desenvolvido por Ganem *et al.* (2017), onde a correção atmosférica do dado pelo FLAASH mostrou melhor desempenho em comparação com as classificações geradas a partir de dados em reflectância TOA e corrigidos pelo QUAC, com diferença de aproximadamente 4% em comparação com este último dado, também em uma área da Caatinga.

O bom desempenho do QUAC em relação ao FLAASH pode ser explicado devido ao fato de o modelo determinar parâmetros de compensação atmosférica diretamente das informações contidas na cena inserida como entrada, não havendo necessidade de se preencher informações auxiliares, como no caso do FLAASH. Isso pode explicar a maior estabilidade do QUAC ao realizar a correção atmosférica, já que a inserção de parâmetros de modo equivocado no FLAASH, pode comprometer significativamente o desempenho do algoritmo.

Pacifici (2018) afirmou que o QUAC realiza uma correção atmosférica mais aproximada que o FLAASH ou outros métodos baseados em princípios físicos. Isso ocorre, pois, de acordo com Bernstein *et al.* (2012), considerando que a precisão desses métodos está diretamente ligada à precisão da geometria de calibração e medição do sensor, o desem-

penho do QUAC não se degradará de maneira significativa. Isso não ocorre com métodos baseados em princípios físicos, onde as incertezas provenientes do sensor em si e suas medições aumentam. Além disso, esses métodos exigem a presença de bandas específicas para corrigir a absorção de água e o espalhamento de aerossóis, enquanto o QUAC funciona com qualquer coleção de bandas que abrangem o intervalo do infravermelho de ondas curtas e infravermelho próximo.

Em contrapartida, Carvalho & Baptista (2013), verificaram em um estudo desenvolvido em uma área do bioma Cerrado, que o dado em reflectância TOA forneceu melhores resultados em comparação com o dado corrigido pelo FLAASH. Os autores apontaram a correção atmosférica como sendo dispensável para o caso avaliado. Isso faz com que o FLAASH, por apresentar uma elevada variabilidade em seus resultados, ora satisfatórios, ora insatisfatórios, não seja recomendável para análises da natureza deste trabalho.

4 Conclusões

A classificação da cobertura da superfície da terra que utilizou o dado corrigido pelo QUAC obteve o melhor resultado estatístico para a carta SC-24-V-C, tendo sido 19% maior que o dado corrigido pelo FLAASH, cujo resultado foi o pior. Além de ter apresentado maior acurácia, o QUAC tem a vantagem de ser mais rápido e menos complexo, sendo, dessa forma, mais apropriado para o mapeamento do bioma.

O panorama de incertezas no que diz respeito à aproximação feita pelo FLAASH em relação ao estado da atmosfera, pode justificar a sua dispensabilidade para se proceder com a correção atmosférica, visto que há uma probabilidade elevada de o dado resultante ser distinto da realidade do contexto real da área a ser classificada. Isso sem falar na quantidade de parâmetros requeridos para rodar o algoritmo, o que resultaria em maior tempo de processamento gasto.

Por outro lado, apesar de o dado sem correção atmosférica (TOA) ter superestimado algumas

classes em detrimento de outras na classificação, esse dado pode apresentar maior consistência, caso a parametrização da árvore de decisão seja aprimorada. Além disso, cabe salientar que o trabalho de executar o pré-processamento fora do ambiente em nuvem para depois inserir os dados atmosféricamente corrigidos na plataforma talvez não se justifique, já que o dado em reflectância TOA diferiu em apenas 4% a menos em comparação com o dado corrigido pelo QUAC.

Como sugestão para trabalhos futuros, destaca-se o uso de novos algoritmos para classificação, como o *Random Forest* (Breiman, 2001), e o uso de dados em reflectância de superfície disponibilizados pela plataforma Earth Explorer do USGS, visto que se trata de um modelo desenvolvido para o próprio Landsat e que, potencialmente, pode se adequar melhor a estudos voltados à classificação da cobertura da superfície terrestre. Além disso, a realização de testes envolvendo outros índices de vegetação, como o SAVI, que considera a interferência do solo na detecção da vegetação, também é recomendável.

5 Agradecimentos

Os autores agradecem à Coordenação Geral e Técnica do MapBiomias, Tasso Azevedo e Carlos Souza, a João Victor Siqueira (Terras App.) e César Diniz (Solved). Finalmente, ao Laboratório de Sensoriamento Remoto e Análise Espacial, do Instituto de Geociências (IG) da Universidade de Brasília, por ter disponibilizado os computadores com a licença dos softwares ENVI 5.0 e ArcGIS 10.2, bem como ao Laboratório de Processamento de Imagens da Universidade Estadual de Feira de Santana pelo suporte na utilização do GEE.

6 Referências

- Adler-Golden, S.; Berk, A.; Bernstein, L.S.; Richtsmeier, S.; Acharya, P.K.; Matthew, M.W. & Chetwynd, J.H. 1998. FLAASH, a MODTRAN4 atmospheric correction package for hyperspectral data retrievals and simulations. In: ANNUAL JPL AIRBORNE EARTH SCIENCE WORKSHOP, 7, Pasadena. *Anais*, Pasadena, JPL, p. 9-14.
- Bernstein, S.M.; Adler-Golden, R.L.; Sundberg, R.Y.; Levine, T.C.; Perkins, A.; Berk, A.; Ratkowski, A. & Hoke, M.L. 2004. A new method for atmospheric correction and aerosol optical property retrieval for VIS-SWIR multi and hyperspectral imaging

Impacto da Correção Atmosférica na Classificação da Cobertura da Terra da Carta Topográfica SC-24-V-C (Petrolina) Utilizando o Google Earth Engine e Imagens Landsat-8

Khalil Ali Ganem; Gustavo Macedo de Mello Baptista;

Washington de Jesus Sant'Anna da Franca Rocha; Rodrigo Nogueira Vasconcellos; Marcos Reis Rosa & Deorgia Tayane Mendes de Souza

- sensors: QUAC (Quick atmospheric correction). In: AVIRIS WORKSHOP, 5, Pasadena, 2004. *Anais*, Pasadena, JPL, p. 3549-3552.
- Bernstein, L.S.; Jin, X.; Gregor, B.; & Adler-Golden, S.M. 2012. Quick atmospheric correction code: algorithm description and recent upgrades. *Optical engineering*, 51(11):111719.
- Breiman, L. 2001. Random forests. *Machine learning*, 45(1): 5-32.
- Brites, R.S.; Soares, P.S. & Ribeiro, C.A.A.S. 1996. Comparação de desempenho entre três índices de exatidão aplicados a classificação de imagens orbitais. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8, Salvador, 1996. *Anais*, São José dos Campos, INPE, p. 813-821.
- Carvalho, A.A. & Baptista, G.M.M. 2013. Influência da correção atmosférica na conversão dos dados de radiancia para reflectância para o cálculo do EVI em imagem Rapideye. *Revista Brasileira de Cartografia*, 67(7): 1371-1390.
- Chander, G.; Markham, B.L. & Helder, D.L. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5): 893-903.
- Coelho, V.H.R.; Montenegro, S.M.G.L.; Almeida, C.N.; Lima, E.R.V.; Ribeiro, A.N. & Moura, G.S.S. 2014. Dinâmica do uso e ocupação do solo em uma bacia hidrográfica do semiárido brasileiro. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 18: 64-72.
- Cohen, J.A. 1960. Coefficient of Agreement for Nominal Scales. *Educational and Measurement*, 20(1): 37-46.
- Congalton, R.G. & Green, K. 2008. *Assessing the accuracy of remotely sensed data: principles and applications*. Boca Raton, Lewis Publishers. 200 p.
- Els, A.; Merlo, S. & Knight, J. 2015. Comparison of two Satellite Imaging Platforms for Evaluating Sand Dune Migration in the Ubari Sand Sea (Libyan Fazzan). *Remote Sensing and Spatial Information Sciences*, 40(7): 1375.
- ENVI. 2009. User's Guide: Atmospheric Correction Module: QUAC and FLAASH. ESRI. 44 p. Acessado em 30 de Agosto de 2017 através do link: http://www.harrisgeospatial.com/portals/0/pdfs/envi/flaash_module.pdf.
- Epiphanyo, J.C.N.; Luiz, A.J.B. & Formaggio, A.R. 2002. Estimativa de áreas agrícolas municipais, utilizando sistema de amostragem simples sobre imagens de satélite. *Bragantia*, 61(2): 187-197.
- Fonseca, E.L. & Glerani, J.M. 2005. Avaliação do efeito da correção atmosférica e retificação radiométrica no cálculo da transformação Tasseled Cap. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12, Goiânia. *Anais*, Goiânia, INPE, p. 513-519.
- Galletti, C.S. & Myint, S.W. 2014. Land-use mapping in a mixed urban-agricultural arid landscape using object-based image analysis: A case study from Maricopa, Arizona. *Remote Sensing*, 6(7): 6089-6110.
- Ganem, K.A.; Baptista, G.M.M.; Franca Rocha, W.J.S.; Vasconcellos, R.N.; Rosa, M.R. & Souza, D.T.M. 2017. Comparação entre dados com e sem correção atmosférica na classificação da cobertura da terra de uma área da Caatinga utilizando o Google Earth Engine. *Revista Brasileira Cartografia*, 69(6): 1053-1074.
- Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D. & Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202: 18-27.
- Guo, Y. & Zeng, F. 2012. Atmospheric correction comparison of Spot-5 Image based on model FLAASH and model QUAC. *ISPRS Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci.*, 39(B7): 7-11.
- Helliden, U. & Stern, M. 1980. Evaluation of Landsat Imagery and Digital Data for Monitoring Desertification Indicators in Tunisia. In: SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT, 14. *Anais*, p. 1601-1611.
- IBGE – Instituto Brasileiro de Geografia e Estatística. 2012. *Manual Técnico da Vegetação Brasileira*. Rio de Janeiro, IBGE, 271p.
- Irons, J.R.; Dwyer, J.L. & Barsi, J.A. 2012. The next Landsat satellite: The Landsat data continuity mission. *Remote Sensing of Environment*, 122: 11-21.
- Jensen, J.R. 2009. *Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres*. São José dos Campos, Parêntese. 598 p.
- Klecka, W.R. 1980. *Discriminant Analysis*. Beverly Hills, SAGE Publications, Inc. 71 p.
- Levada, A.L.M. 2010. *Combinação de modelos de campos aleatórios markovianos para classificação contextual de imagens multiespectrais*. Programa de Pós-Graduação do Instituto de Física de São Carlos – USP, Tese de Doutorado, 146p.
- Mandanici, E.; Franci, F.; Bitelli, G.; Agapiou, A.; Alexakis, D. & Hadjimitsis, D.G. 2015. Comparison between empirical and physically based models of atmospheric correction. In: INTERNATIONAL CONFERENCE ON REMOTE SENSING AND GEOINFORMATION OF THE ENVIRONMENT, 3, 95350E-1-95350E-10.
- Mather, P.M. 2004. *Computer processing of remotely-sensed images: an introduction*. Chichester, John Wiley & Sons. 324 p.
- Mitchell, J.J.; Shrestha, R.; Moore-Ellison, C.A. & Glenn, N.F. 2013. Single and multi-date Landsat classifications of basalt to support soil survey efforts. *Remote Sensing*, 5(10): 4857-4876.
- MMA – Ministério do Meio Ambiente. 2006. Mapas de Cobertura Vegetal dos Biomas Brasileiros. Acessado em 20 de julho de 2017 através do link: <http://mapas.mma.gov.br/mapas/aplic/probio/datadownload.htm?/>.
- Moreira, M.A. 2001. *Fundamentos do sensoriamento remoto e metodologias de aplicação*. São José dos Campos, INPE. 250 p.
- Nascimento, E.R.P. & Sano, E.E. 2010. Identificação de Cerrado Rupetre por meio de imagens multitemporais do Landsat: proposta metodológica. *Sociedade e Natureza*, 23(1): 93-106.
- Pacifici, F. 2018. Atmospheric compensation in satellite imagery. *U.S. Patent No. 9,990,705*. Washington, Patent and Trademark Office.
- Projeto MapBiomas. 2017a – Coleção 2 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. Acessado em 30 de abril de 2017 através do link: <http://mapbiomas.org>.
- Projeto MapBiomas. 2017b – Coleção 2 da Série Anual de Mapas de Cobertura e Uso de Solo do Brasil. ATDB do Bioma Pantanal. Acessado em 05 de maio de 2017 através do link: <http://mapbiomas.org>.
- Richards, J.A. & JIA, X. 2006. *Remote sensing digital image analysis-hardback*. New York, Springer. 464 p.
- Rojas, E.H.M. 2005. *Síntese Genética de Redes Neurais Artificiais ART2 na classificação de imagens ASTER para mapeamento de uso e cobertura da terra na região Norte de Mato Grosso*. Programa de Pós-Graduação em Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais – INPE, Dissertação de Mestrado, 126p.
- Sano, E.E.; Ferreira, L.G.; Asner, G.P. & Steinke, E.T. 2007. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *International Journal of Remote Sensing*, 28: 2739-2752.
- Smith, M.O.; Johnson, P.E. & Adams, J.J. 1985. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. *Journal of Geophysical Research*, 90: 797-804.
- Souza Jr. C.M.; Roberts, D. & Cochrane, M.A. 2005. Combining Spectral and Spatial Information to Map Canopy Damages from Selective Logging and Forest Fires. *Remote Sensing of Environment*, 98: 329-343.
- Turner, D.D.; Ferrare, R.A. & Brasseur, L.A. 2001. Average aerosol extinction and water vapor profiles over the Southern Great Plains. *Geophysical research letters*, 28(23): 4441-4444.
- WMO – World Meteorological Organization. 2014. *WMO statement on the status of global climate*. Geneva, WMO. 290 p.
- Zhu, Z. & Woodcock, C.E. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118: 83-94.