

Análise Espaço-Temporal de Áreas de Queimadas no Estado do Maranhão a partir de Imagens MODIS e Classificação Random Forest

Spatio-temporal Analysis of Burning Areas in the State of Maranhão based on MODIS Images and Random Forest Classification

Admilson da Penha Pacheco¹  & Juarez Antonio da Silva Junior² 

¹Universidade Federal de Pernambuco, Instituto de Geociências, Departamento de Engenharia Cartográfica e de Agrimensura. Recife, PE, Brasil.

²Universidade Federal de Pernambuco, Curso de Engenharia Cartográfica e de Agrimensura. Recife, PE, Brasil.

E-mails: pacheco3p@gmail.com; juarez.ufpe@gmail.com

Resumo

O mapeamento de áreas queimadas através de imagens de sensoriamento remoto apresenta uma série de questões importantes no gerenciamento espacial de estudos de detecção de focos de calor, análise de risco de incêndio, avaliação de danos e gerenciamento de processos de regeneração florestal. Neste estudo foi apresentada uma abordagem metodológica para mapeamento de áreas de queimadas no estado do Maranhão de 2001 a 2019 a partir de dados do satélite Terra/MODIS e do Algoritmo de Classificação Binária *Random Forest*. A avaliação da qualidade dos mapas gerados foi realizada a partir do produto padrão NASA MCD64A1 de áreas queimadas do sensor MODIS, que resultou num Coeficiente de Determinação geral (R^2) de 0,55 e Correlação de Spearman de 0,78. O modelo *Random Forest* com 400 árvores permitiu avaliar a banda espectral de maior contribuição na classificação, bem como os erros relacionados ao número de árvores empregado. Algumas feições de áreas queimadas foram superestimadas, apresentando elevados erros de comissão. Os resultados mostraram que esta abordagem é útil para determinar áreas queimadas derivadas de dados de satélite de órbita polar. O modelo *Random Forest* mostrou-se aplicável em áreas de transição entre biomas. Constatou-se uma relação não linear entre a variabilidade espaço-temporal de incêndio e o clima em ecossistemas temperados. A Metodologia estabelecida e validada neste estudo poderá ser aplicada em outras regiões de clima temperado.

Palavras-chave: Vegetação; Sensoriamento Remoto; Incêndios; Mapeamentos

Abstract

The mapping of burnt areas through remote sensing images presents a series of important issues in the spatial management of studies for the detection of hot spots, fire risk analysis, damage assessment and management of forest regeneration processes. In this study, a methodological approach was presented for mapping the burning areas in the state of Maranhão from 2001 to 2019 using data from the Terra / MODIS satellite and the Random Forest Binary Classification Algorithm. The quality assessment of the generated maps was carried out using the standard product NASA MCD64A1 of burned areas of the MODIS sensor, which resulted in a general Coefficient of Determination (R^2) of 0.55 and Spearman Correlation of 0.78. The Random Forest model with 400 trees allowed to evaluate the spectral band with the greatest contribution in the classification, as well as the errors related to the number of trees used. Some features of burnt areas were over estimated, showing high commission errors. The results showed that this approach is useful to determine burnt areas derived from polar orbit satellite data. The Random Forest model proved to be applicable in areas of transition between biomes. There was a non-linear relationship between the space-time variability of fire and the climate in temperate ecosystems. The methodology established and validated in this study can be applied in other regions of temperate climate.

Keywords: Vegetation; Remote Sensing; Fires; Mappings

1 Introdução

O incêndio florestal é um fenômeno natural que afeta muitos ecossistemas em todo o mundo, envolvendo diversas consequências ambientais tais como emissão de gases do efeito estufa, degradação do solo e destruição de espécies, biomassa e biodiversidade (Carmenta *et al.*, 2011; Oertel *et al.*, 2016). A frequência e severidade dos incêndios, juntamente com as mudanças de temperatura e precipitação afetam severamente as características dos padrões das florestas (Chu & Guo, 2014).

A detecção de Área Queimada (AQ) e a previsão das condições de perigo de incêndios são tópicos de pesquisa ativos, que são estudados em vários ecossistemas. A relação do fogo com a vegetação é contraditória: como fator ecológico; contribui para manter a dinâmica do ecossistema, a produtividade e a biodiversidade e, como ferramenta de gerenciamento da terra; é amplamente empregada em áreas de cultivo, pastagens e florestas em todo o mundo (Pereira *et al.*, 2017). As alterações nos regimes naturais de incêndio podem ter impactos negativos sobre a biodiversidade, a estrutura e a resiliência da floresta, principalmente, em regiões equatoriais onde florestas sempre verdes tornam-se vulneráveis ao fogo (Chuvieco *et al.*, 2018; Chuvieco *et al.*, 2019).

O Cerrado brasileiro apresenta 2.000.000 km² de área, ocupando quase um quarto do território do país, e é considerada a savana mais biodiversa do mundo (Melchiori *et al.*, 2014). O desmatamento já reduziu esse bioma pelo menos 50 %, e se for considerada a degradação, esse valor aumenta significativamente (Klink & Machado, 2005).

No Maranhão, o fogo atinge até mesmo as Unidades de Conservação (UCs). Este fato é crítico, já que estas áreas são legalmente utilizadas para a proteção e conservação de ecossistemas únicos. Entre 2008 e 2012 foram registrados 19.048 focos de calor nas (UCs), correspondendo a 19,5% de todos os focos identificados no período para todo o Maranhão (Bezerra *et al.*, 2018).

Os meses de agosto e setembro são os mais críticos em ocorrências de queimadas, por serem os meses mais secos do ano, e a vegetação está mais suscetível a ocorrência de queimada. Assim, nesse período, a probabilidade de ocorrência de queimada é bem maior em comparação aos outros meses. As queimadas nessa região do Maranhão são influenciadas diretamente pela quantidade de chuva mensal, dias seguidos de chuva, umidade e temperatura do ar.

O sensoriamento remoto orbital devido suas características espaciais, espectral e temporal, tem se apresentado com uma tecnologia eficaz para prevenir e monitorar incêndios florestais. Calle & Casanova (2015) afirmam que é possível distinguir alguns tópicos de sensoriamento remoto aplicado diretamente ao tema dos incêndios florestais: risco de propagação do incêndio,

detecção de pontos quentes, estabelecimento de parâmetros térmicos do fogo e cartografia das áreas afetadas.

Nas últimas décadas, o uso do sensoriamento remoto permitiu avanços no mapeamento da dinâmica do fogo, especialmente para localizar a ocorrência do fogo no tempo e no espaço e quantificar a extensão total da área queimada Pereira *et al.*, (2017). Vários estudos foram realizados usando imagens de sensoriamento remoto para mapeamento de áreas queimadas em escala regional e global (Chuvieco *et al.*, 2008; Pereira *et al.*, 2016; Lizundia-Loiola *et al.*, 2020). Imagens de média resolução geradas pelo Espectrorradiômetro MODIS (Moderate-Resolution Imaging Spectroradiometer) a bordo dos satélites de Observação da Terra (EOS) “Terra” e “Aqua”, tem sido muito utilizadas no monitoramento de incêndios florestais em diversas partes do mundo.

Os dados do sensor MODIS fornece desde fevereiro de 2000 o produto MOD09A1 (Surface Reflectance Daily L2G Global 1 km and 500 m). Dentro desse conjunto está o produto padrão de áreas queimadas MCD64A1, que mapeia a extensão espacial e a data aproximada da queima de biomassa em todo o mundo, com uma resolução espacial de 500 m (Giglio *et al.*, 2006).

Random Forest (RF) é um algoritmo de aprendizagem de máquina flexível, que produz várias árvores de decisão. *Random Forest* permite integrar dados de diferentes escalas e fontes, o que explica o seu amplo uso em muitas aplicações de mapeamentos a partir de imagens de satélites (Ramo & Chuvieco, 2017).

Vários estudos têm mostrado o potencial do RF aplicado em imagens de satélites para a detecção de incêndios em florestas. Os trabalhos de Belgiu & Drăguț (2016), Collins *et al.*, (2018), Yao *et al.*, (2018) são referências recentes de aplicação do algoritmo RF em estudos de incêndios florestais por sensoriamento remoto.

Neste contexto, este artigo tem como objetivo apresentar uma abordagem metodológica para mapeamento de áreas de queimadas no estado do Maranhão de 2001 a 2019, a partir de dados do Satélite Terra/MODIS e do algoritmo de Classificação Binária *Random Forest*.

2 Materiais e Métodos

2.1 Área de Estudo

O estado do Maranhão (Figura 1) possui uma área superficial de 331 984 km², e está localizado entre os paralelos 1° 01' e 10° 21' sul e os meridianos 41° 48' e 48° 50' oeste. Ao norte limita-se com o Oceano Atlântico (639 km), ao sul e sudoeste com o Tocantins (1060 km), a oeste com o Pará (798 km), a leste e sudeste limita-se com o Piauí (1365 km) (MARANHÃO, 2011). O Maranhão

apresenta uma grande variedade de ecossistemas devido à sua localização entre o Nordeste e o Norte, ou seja, entre as condições do semiárido brasileiro e os biomas Cerrado e Amazônia. Destacando-se como um verdadeiro conjunto de ambientes transacionais, onde predominam as seguintes formações vegetais: Floresta Ombrófila Densa, Savana, Savana Estépica, Floresta Estacional Decídua, Floresta

Estacional e formações com influência marinha (Bezerra *et al.*, 2018).

No geral, o Estado do Maranhão é submetido a um clima tropical úmido, com um verão seco e um inverno seco, e a precipitação média anual é da ordem de 1500 mm. Devido ao desmatamento na região, o estado representa cerca de 20% das áreas antropizadas do Cerrado brasileiro (MMA, 2018).

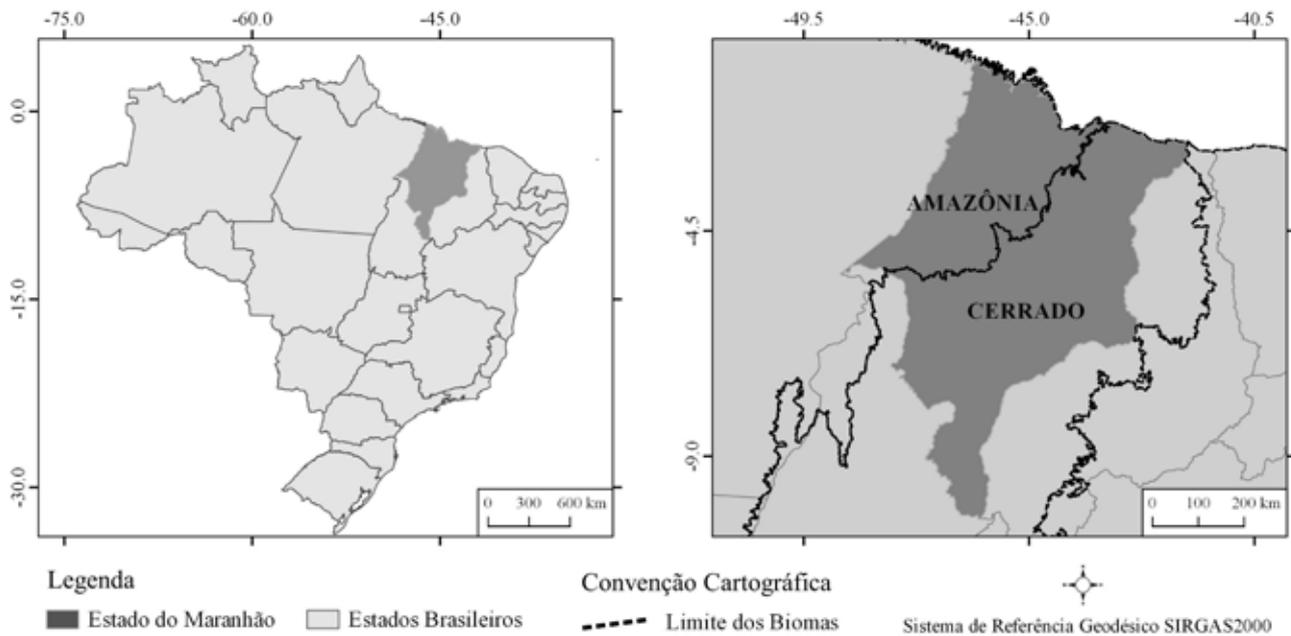


Figura 1 Mapa de Localização do Estado do Maranhão.

2.2 Dados

Satélite TERRA – Sensor MODIS: Produto MOD09A1 (Surface Reflectance8-Day L3 Global 500m).

O MOD09A1 fornece refletância de superfície através de 7 bandas espectrais, cobrindo as faixas do visível (VIS) até o infravermelho de ondas curtas (SWIR) com resolução espacial de 500 metros cobrindo todo o globo. No produto MOD09A1 cada pixel contém a melhor observação possível durante um período de 8 dias, conforme selecionado com base em alta cobertura de observação, baixo ângulo de visão, ausência de nuvens ou sombra das nuvens e carga de aerossol (MODIS, 2015). As imagens foram adquiridas para o mês de setembro (período de estiagem) de 2001 a 2019. Nesse período a probabilidade de ocorrência de focos de calor é maior em comparação aos outros meses.

Os dados hidrológicos foram adquiridos a partir de 5 estações meteorológicas do Instituto Nacional de Meteorologia – INMET, através da plataforma BDMEP (Banco de Dados Meteorológicos para Ensino e Pesquisa),

distribuídas no Estado do Maranhão (INMET, 2020). Seguem as estações segregadas quanto ao bioma: Alto Parnaíba (82970), São Luís (82280), Barra da Corda (82571), Imperatriz (82564) e Colinas (82676). Foram utilizados registros mensais de precipitação mensal e anual média, no período de 2001 a 2019. A partir do Banco de Dados de Queimadas (BD Queimadas) do Instituto Nacional de Pesquisas Espaciais, foram obtidos os dados pontuais de detecção de focos de calor compreendidos num raio de até 6 km (INPE, 2020).

2.3 Validação do Produto de Áreas Queimadas MCD64A1

De acordo com MODIS (2015), o produto MCD64A1 é o mais recente do conjunto MODIS Área Queimada (MODIS Burned Area). Este produto apresenta resolução espacial de 500 metros e contém informações de área queimada por pixel. O MCD64A1 é baseado em uma abordagem híbrida, que explora o potencial de incêndios ativos MODIS de 1 km e dados de entrada de

refletância de 500 m de superfície (MODIS, 2015). O algoritmo usa o Índice de Vegetação Sensível à Queima (VI) permite criar limites dinâmicos, que são aplicados aos dados compostos. O VI é derivado das bandas 5 (NIR) e 7 (SWIR) de refletâncias superficiais atmosféricas corrigidas por ondas curtas MODIS. Os mapas cumulativos de fogo ativo são usados para gerar funções regionais de densidade de probabilidade para a classificação de amostras de treinamento de áreas queimadas e não queimadas, que guiarão a determinação final de pixels queimados e não queimados (MODIS, 2015).

2.4 Classificação por *Random Forest*

A classificação por *Random Forest* (RF) é uma classificação de conjunto, que utiliza vários classificadores (Tian *et al.*, 2016). O método é baseado na combinação de preditores de árvores. Cada árvore depende dos valores de um vetor aleatório amostrado independentemente e com a mesma distribuição para todas as árvores da floresta (Ramo & Chuvieco, 2017). O RF produz várias árvores de decisão usando um subconjunto selecionado aleatoriamente de amostras e variáveis de treinamento. No algoritmo RF, o número de árvores de classificação precisa ser definido suficientemente alto para permitir a convergência do modelo (Guo *et al.*, 2016). O erro de generalização para florestas converge a um limite à medida que o número de árvores na floresta se torna grande: o erro de generalização de uma floresta de classificadores de árvores depende da força das árvores individuais na floresta e da correlação entre elas (Melchiori *et al.*, 2014).

O *Random Forest* gera a possibilidade de estimar a probabilidade da classe atribuída ao pixel, ou seja, a probabilidade de queimada mínima associada a cada pixel para classificá-lo como queimado (Ramo *et al.*, 2018).

Neste estudo, inicialmente, são pré-selecionados pixels aleatoriamente para a combinação de parâmetros contidos em cada nó para formação de uma árvore. Cada pixel é classificado de acordo com a classe que tem maior adesão de todos os elementos preditores das árvores no conjunto amostral. Para a classificação RF é necessário configurar os dados de entrada e assim fornecer o número de árvores, que representa o número de variáveis testadas em cada divisão do nó. Segundo Millard & Richardson (2015) o número de árvores é necessário para manter a precisão da classificação. A seleção de dados de treinamento e variáveis de entrada específicas (ou seja, canais de imagem) têm um grande impacto na precisão geral da classificação da imagem.

Neste contexto, foi testado para cada ano uma classificação considerando de 100 a 400 árvores. Um terço deste número foi usado para testar o erro associado às

previsões, também chamado de Erro Out of Bag (OOB). O parâmetro MTRY controla o número de variáveis disponíveis para divisão em cada nó em uma árvore em uma floresta aleatória (Cutler *et al.*, 2007). Para o parâmetro MTRY foi utilizado o valor padrão. Cutler *et al.* (2007) relataram que diferentes valores de MTRY não afetam as taxas de classificação no modelo e que outras métricas de desempenho se mantiveram estáveis sob diferentes valores de MTRY.

2.5 Procedimento Metodológico

As imagens MOD09A1 e MCD64A1 correspondentes ao mês de setembro (período de estiagem) de 2001 a 2019 foram obtidas gratuitamente através do servidor de Internet do Centro de Arquivo Ativo Distribuído por Processos Terrestres (LP DAAC) do U.S. GEOLOGICAL SURVEY (USGS, 2020). Todas as imagens foram reprojadas e processadas usando o software livre QGIS versão 2.14. Posteriormente, foi realizada uma composição das 5 bandas do produto MOD09A1 correspondente à faixa 0.45 – 2.15 μm para cada ano e a seleção de pixels de treinamento de AQ. Isto teve a finalidade de aplicar a classificação por *Random Forest*. A classificação consistiu na análise do Erro OOB por número de árvores, importância de cada banda e classificação seguindo o modelo binário de presença e não presença de AQ. As bandas 1 e 2 foram descartadas devido a sua diferença de 250 m de resolução espacial.

2.6 Avaliação da Classificação

A qualidade de mapas originados por sensoriamento remoto é geralmente avaliada com base numa comparação sistemática com outros mapas também derivados por sensoriamento remoto (Libonati *et al.*, 2015). A precisão dos resultados é geralmente caracterizada através da tabulação cruzada em relação aos conjuntos de dados de referência, contabilizando as coincidências e desacordos espaço-temporais: a abordagem é amplamente usada em projetos de mapeamento de área queimada (Tanase *et al.*, 2020).

O Erro Total (ET) reflete o acordo entre o produto da área queimada e o mapa de referência (ou seja, a precisão da classificação) e satisfaz o princípio da equivalência de eventos, uma vez que credita os pixels queimados e não queimados igualmente (Wilks, 2011). Os Erros de Omissão (EO) e de Comissão (EC), fornecem, respectivamente, informações sobre a confiabilidade e o poder de discriminação do produto de Área Queimada (AQ). Os erros EO e EC variam entre 0% a 100% e apresentam uma escala reversa, de modo que os valores menores são as melhores estimativas (Pereira *et al.*, 2017).

A precisão dos mapas de áreas queimadas obtidos pela classificação supervisionada *Random Forest* foi comparada com os perímetros de referência de incêndios extraídos de cada cena do produto MCD64A1. A análise dos mapas gerados foi baseada nas informações obtidas pela Matriz de Confusão, na qual foram calculadas diferentes medidas de verificação: Erro Total (ET), Erro de Omissão (EO), Erro de Comissão (EC) e Coeficiente de Kappa (Kp).

3 Resultados e Discussão

O modelo de classificação foi estabelecido usando diferentes números de árvores (100, 200, 300, 400) com o objetivo de identificar o número mais adequado em relação ao erro OOB como mostrado na Figura 2. Observa-se que as variações de erro OOB foram razoavelmente pequenas entre o número de árvores, porém as menores estimativas de erro foram encontradas no modelo baseado em 400 árvores. Dessa forma, optou-se em aplicar a classificação de imagens com 400 árvores como parâmetro de entrada. De acordo com Belgiu & Drăguț (2016), a classificação por RF com um número de árvores acima de 100 apresentam geralmente uma estabilidade, porém verificou-se em estudos anteriores que o número de árvores varia bastante dependendo de diversos fatores, como a quantidade de classes e tamanho amostral. O efeito do número de árvores é evidente na Figura 2A onde o Erro OOB diminui a partir do número padrão do parâmetro *MTRY* empregado.

A Figura 2B evidencia a importância das bandas do MOD09A1 no processo de classificação. A banda 5 do MOD09A1 mostrou ser a variável mais importante na previsão de pixels na classificação de (AQ). A alta importância da faixa de infravermelho médio da banda 5 (1,23-1,25 μm) provavelmente está associada às suas propriedades de sensibilidade à ausência de umidade da vegetação pós fogo. A banda 7 foi a segunda banda de maior importância e as outras bandas tiveram valores abaixo de 5%. Mellor *et al.* (2013) mostraram que as bandas do infravermelho do Landsat/TM apresentaram maior importância na classificação por RF para a separação espaço-espectral de florestas densas e desmatadas no sudoeste australiano. As bandas 5 e 7 coincidem com as faixas espectrais utilizadas no IV utilizado no produto MCD64A1 e com os principais índices de vegetação utilizados o mapeamento de queimadas florestais, como o NBR – Normalized Burn Ratio, BAI – Burn Area Index e NBRT-1 Normalized Burn Ratio Termal – 1. Os resultados das classificações binárias descritos acima foram utilizados como base para a geração dos mapas de áreas queimadas no estado do Maranhão nos anos de 2001 a 2019 (Figura 3).

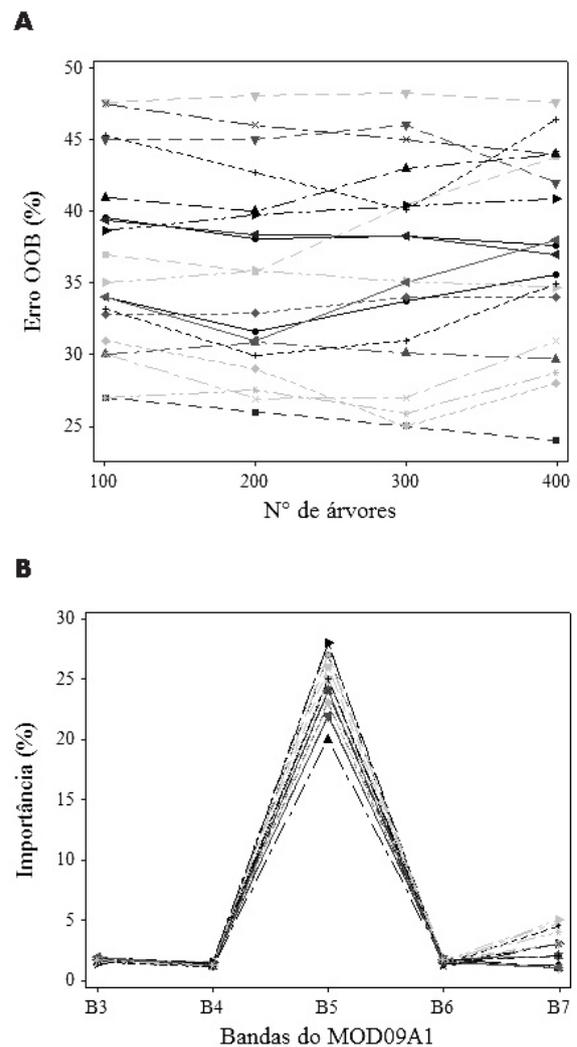


Figura 2 A. Relação entre os números de árvores com o Erro OOB (%). **B.** Importância de cada banda na classificação Random Forest.

Observa-se visualmente na Figura 3 que a região do Cerrado maranhense apresentou uma maior frequência de polígonos de AQ, com destaque para os anos 2007, 2010, 2012 e para o período de 2013 a 2016. Estes resultados corroboram com a dinâmica de desmatamento no Cerrado, que vem crescendo anualmente e grande parte deste desmatamento está relacionado com a atividade carvoeira (INPE, 2020).

Fatores climáticos influenciam significativamente os incêndios no Maranhão, principalmente, no setor sul do estado. Nesta região o microclima apresenta-se com baixos índices de precipitação e altas temperaturas do ar, que incorporado à vegetação secundária seca, propicia a ocorrência de queimadas naturais. Na parte da Amazônia

Maranhense, os mapas mostraram em sua grande maioria, baixa frequência de polígonos de AQ, embora os anos de 2007, 2010, 2012 e 2015 apresentaram feições de significativas de AQ. Masullo (2018) concluiu que a gradual conversão do estrato arbóreo (floresta) para arbustivo e herbáceo (vegetação aberta), deflagrada na última década, é o principal fator que ampliou a abrangência de áreas com maior risco de incêndios no Maranhão Amazônico. Na parte central do estado, correspondente aos municípios de Grajau, Barra da Corda, Mirador, Porto Franco e Itaipava do Grajaú, apresentaram alta frequência de áreas queimadas em todos os anos. Os referidos municípios lideram anualmente o ranking das cidades com maior número de focos de calor no mês de setembro (Sales *et al.*, 2019).

A Figura 4 apresenta a avaliação quantitativa dos resultados de AQ mapeado pelo RF e o Coeficiente de Determinação Linear (R^2): produto MCD64A1 nos anos de 2001 a 2019.

O gráfico da Figura 4 apresenta uma pequena variação de AQ nos anos de 2002 a 2006 correspondentes a valores estáveis, porém baixos de R^2 . Em 2007 houve um pico significativo de AQ e R^2 com valor de 22.541,60 km² e 0.35, respectivamente. Em 2008 e 2009 houve um declínio nos valores de áreas correspondentes à diminuição dos valores de R^2 . O ano de 2009 apresentou a menor variação de AQ e menor R^2 registrado da série. Em 2010 houve um aumento significativo de AQ em relação ao período homólogo, passando de 8.332,77 km² para 21.313,30 km².

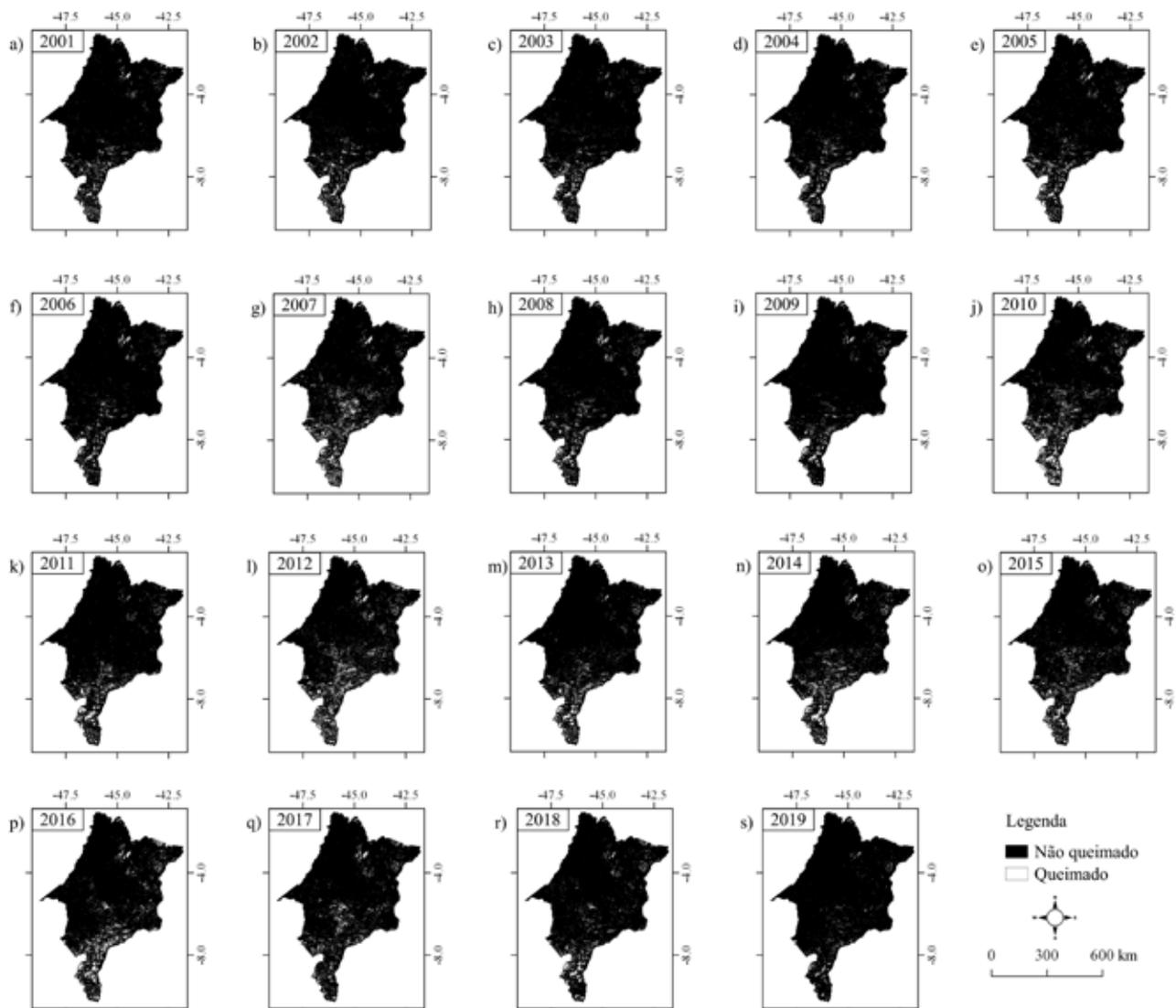


Figura 3 Mapas binários de áreas queimadas e não queimadas no estado do estado do Maranhão nos meses de setembro de 2001 a 2019.

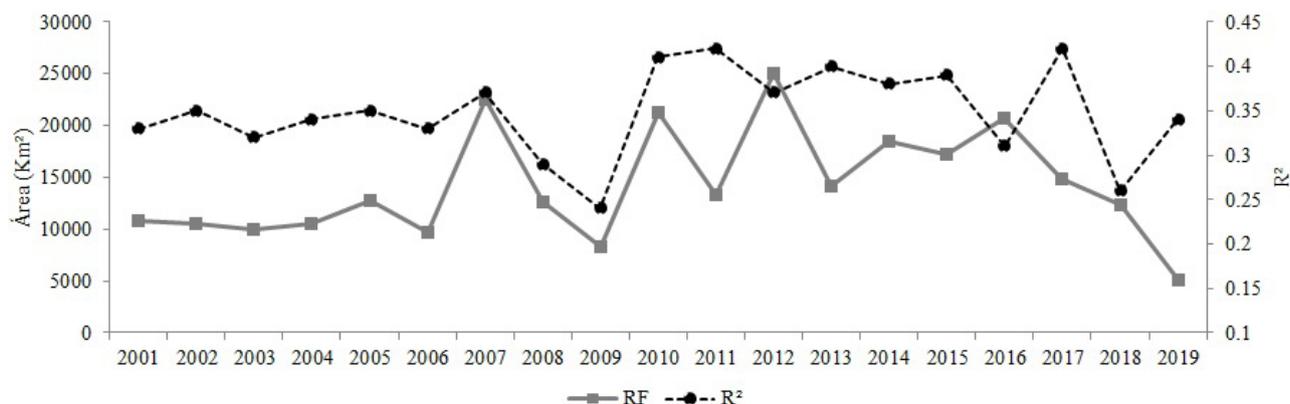


Figura 4 Gráficos de linhas sobre as Áreas Queimadas (AQ) Mapeadas pelo (RF) e o Coeficiente de Determinação (R^2) para os anos de 2001 a 2019.

O mesmo comportamento também foi visto para o R^2 , sendo este o segundo maior da série com o valor de 0.41.

De 2006 a 2013 houve variações importantes de AQ com tendência triangular e áreas variando entre 10.000 km² a 50.000 km². Em contrapartida, o R^2 apresentou valores estáveis com variações entorno de 0.4. De 2014 a 2016 não houve picos em AQ, entretanto para o R^2 a partir de 2015 houve 5 pontos de inflexão. Sendo em 2017 o maior valor da série (0.42). Ainda neste período, houve diminuição gradativa de AQ mapeadas pelo classificador. Chegando em 2019 ao menor valor registrado da série; 5.145,73 km². A Figura 5 apresenta a distribuição espacial dos Erros de Omissão (EO) e Comissão (EC) para os anos de 2001 a 2019, resultante da comparação entre a classificação RF e o produto MCD64A1.

A Figura 5 apresenta duas medidas de verificação derivadas das matrizes de confusão no período do estudo. Observa-se que de 2001 a 2011 houve variações de 55 a 70% de EC, ou seja, alta presença de pixels de falsos positivos. O ano de 2009 apresentou o maior valor de EC da série histórica, onde 4064 pixels foram classificados de forma equivocada. Estes pixels foram classificados como AQ. Quando, na verdade, são pixels de ANQ. Isto acarretou uma elevação da taxa de alarmes falsos.

Observa-se na Figura 5 que de 2011 a 2013 houve um declínio nos valores de EC. O ano de 2013 apresentou o menor valor da série, de 41%. Os EO de 2001 a 2007 não apresentaram picos significativos, com valores variando entre 25% a 40%. O período de 2008 a 2011 seguiu a tendência anterior e apresentou aumento gradativo. Ressalta-se, entretanto, que no ano de 2011 houve o maior EO da série (61.3%), correspondendo a um alto valor de EC. Acrescenta-se que em 2011 obteve-se a classificação com o

maior distanciamento do mapa de referência de toda a série histórica. Os anos de 2013 a 2019 apresentaram valores estáveis e semelhantes de EC e EO, com leve declínio em 2018 e 2019 e variações entre 40 e 50%. Mostrando assim, uma maior sobreposição média entre os pixels classificados como queimados pelo classificador e pelo MCD64A1. Corroborando com o início de estabilização e da redução dos valores de AQ mapeados apresentados na Figura 4. Os mapas classificados pelo algoritmo RF apresentaram no geral, altos valores de EO e EC. Os valores de EC foram sempre maiores que EO. Exceto no período de 2013 e 2019, quando os valores se aproximaram. Corroborando com as observações de Libonati *et al.* (2015).

Conforme Pulvirenti *et al.* (2020), o EO é causado, principalmente, por um mascaramento impreciso de nuvens, fumaça, sombras de nuvens e sombras de terreno (Pulvirenti *et al.*, 2020). Neste estudo, os erros também estão relacionados com a alta presença de ruídos ao longo de feições de áreas degradadas do MOD09A1. Apresentando pixels com tons escurecidos, causando confusões no classificador e conseqüentemente alta frequência de pixels de falsos alarmes. A Figura 6 apresenta o percentual do Erro Total (ET) e Coeficiente Kappa para os anos de 2001 a 2019.

Observa-se na Figura 6 baixos valores de ET, principalmente, nos anos de 2001 a 2009. Esses valores estão associados a boa classificação de ANQ, que elevou a frequência de pixels de verdadeiros negativos. Ocasionalmente assim um aumento da acurácia global dos mapas para esta classe. Ainda é possível observar bruscos picos nos anos de 2010 a 2012, corroborando com o aumento de AQ. Além dos altos valores de EC e OE, que caracterizam a baixa confiabilidade dos resultados mapeados associados a esses anos.

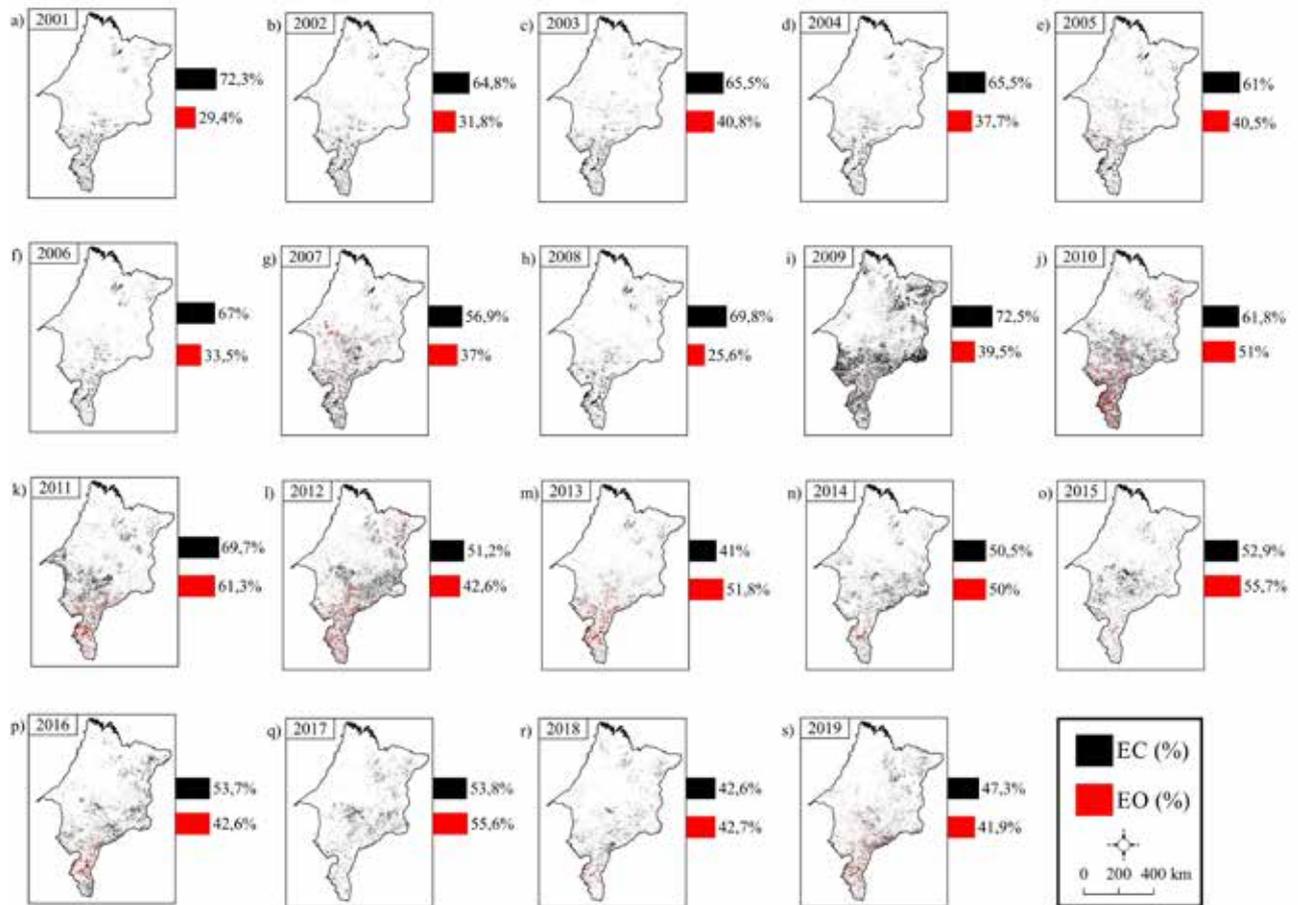


Figura 5 Erros de Omissão (EO) e Comissão (EC) para os anos de 2001 a 2019.

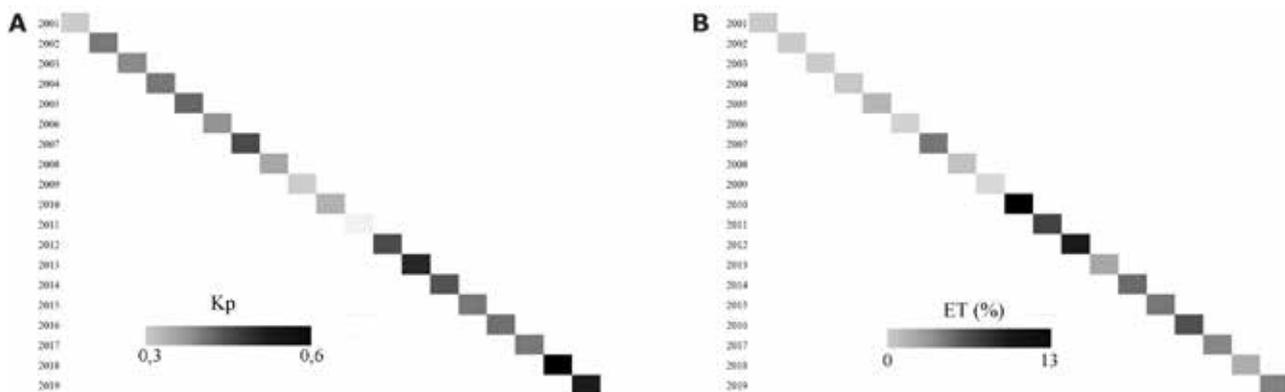


Figura 6 A Coeficiente Kappa (Kp) B Erro total (ET%) para os anos de 2001 a 2019.

O índice Kappa apresentou baixos valores concentrados abaixo de 0.6 em todos os anos. Os anos de 2010 e 2011 não apresentaram boa qualidade na classificação. Os anos de 2018 e 2019 apresentaram os maiores valores de índice Kappa, corroborando com a diminuição de AQ mapeada. O índice Kappa apresentou

valores estáveis entre 2001 á 2006 e de 2013 a 2017, que estão associadas a baixa variação de AQ mapeada nestes anos. Estes resultados retornam a discussão sobre a tendência de baixa sobreposição de AQ gerada pela classificação supervisionada e o produto MCD64A1.

O MCD64A1 é amplamente utilizado como produto de validação em abordagens de mapeamento de áreas queimadas (Libonati *et al.*, 2015; Chuvieco *et al.*, 2018; Hawbaker *et al.*, 2020). O algoritmo de geração do produto MCD64A1, como mencionado anteriormente, utiliza o índice de vegetação e pixels de fogo ativo gravado pelo sensor num intervalo de 8 dias. A classificação RF mapeou AQ de altas a baixas severidades, muitas vezes encontradas em áreas agrícolas com tamanho < 100 ha. Dessa forma, pixels de áreas atingidas pelo fogo de baixa gravidade, podem não ter sido incluídos no algoritmo de geração do MCD64A1. Assim, seus efeitos não acionaram o limiar de detecção, como discutido em (Ramo & Chuvieco, 2017). Isto ocasiona um grande número de pixels falsos positivos no resultado final do mapeamento. Justificando os altos valores de EC encontrados na Figura 6, que influenciou diretamente nos resultados de ET e Kp. Em contrapartida, o algoritmo pode até ter incluído pequenas áreas queimadas nos perímetros de referência, porém não são detectáveis numa precisão espacial de 500 m.

No geral, os erros encontrados também podem ser atribuídos à distribuição heterogênea de pequenos polígonos de (AQ) presentes nas áreas de estudo. Isto também foi relatado por Bastarrika *et al.* (2011). Estes autores mostraram, no algoritmo de classificação desenvolvido, que a área queimada compacta apresentava um melhor desempenho. Observaram também que a distribuição desigual da AQ pode estar relacionada a diferenças no tipo e arranjo territorial de uma área de transição entre biomas com vários tipos de cobertura do solo.

O enfoque dinâmico da relação das áreas queimadas mapeadas por RF e os parâmetros climáticos de focos de calor e precipitação são mostrados nas regressões cúbicas da Figura 7.

Observa-se na Figura 7A que boa parte do conjunto amostral está próximo à curva de regressão, embora os anos de 2007, 2010, 2012, 2014 e 2019 apresentaram grande distanciamento e um valor geral de R^2 baixo. Resultado semelhante foi publicado por Libonati *et al.* (2015). Os três anos com maiores registros anuais de AQ (2007, 2010 e 2012) são aqueles com as menores ocorrências de precipitação anual no Cerrado. De acordo com a Figura 7B, apenas os anos 2012 e 2016 apresentam maior dispersão do conjunto amostral, com moderado valor de R^2 . De uma maneira geral, as menores áreas (< 20000 km²) estão mais correlacionadas com a frequência de precipitação e focos de calor. Os resultados apresentaram-se coerentes com a dinâmica climática no estado do Maranhão. É importante comentar que apesar dos valores baixos de R^2 (< 0.5) e algumas disparidades pontuais, tomando uma visão macro, destaca-se a característica de proporcionalidade alta entre as AQ e os parâmetros climáticos. Uma precipitação abaixo de 50 mm e focos de calor acima de 1000 unidades são fatores suficientes e bastante favoráveis à ocorrência de queimadas naturais.

Acrescenta-se que além da contribuição de fatores climáticos na ocorrência de incêndios florestais, a ação antrópica também interfere na dinâmica do fogo nessa região e não pode ser desconsiderada.

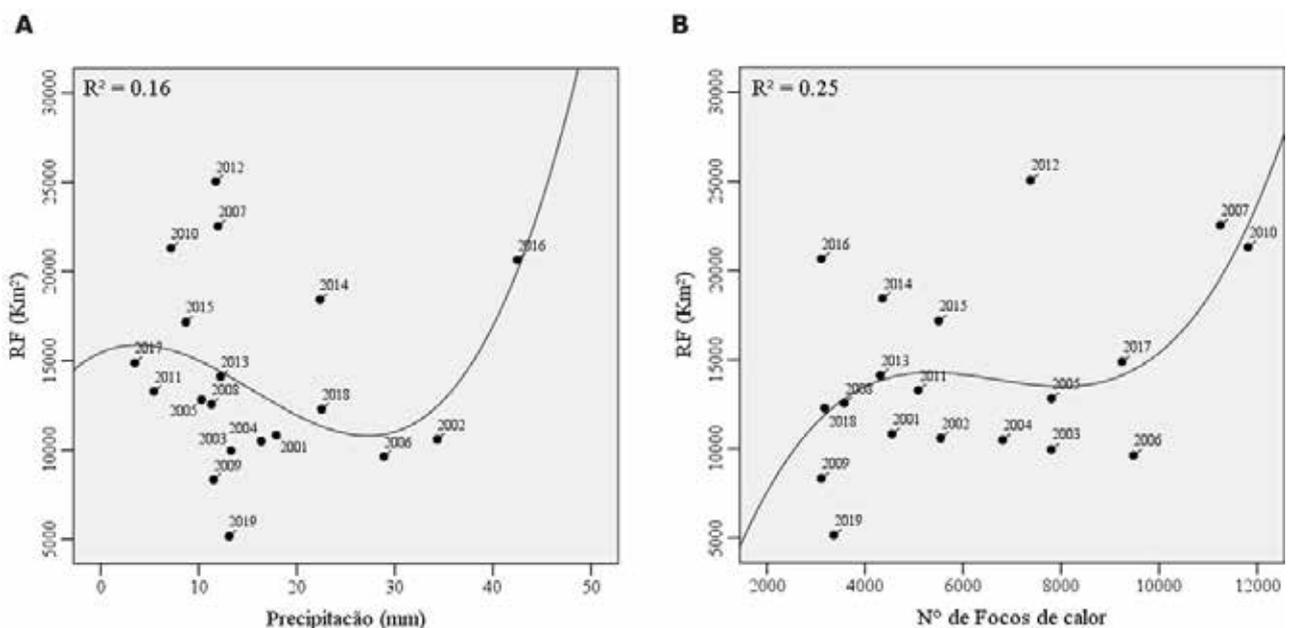


Figura 7 Regressão Cúbica entre as AQ mapeados pelo RF e A. Focos de calor e B. Precipitação Acumulada média para as 5 estações meteorológicas.

Observou neste estudo que séries temporais derivadas de imagens orbitais são materiais cruciais para a análise de terras, que sofrem perturbação na paisagem natural, sendo possível compreender anomalias ou tendências presentes na dinâmica espacial em escala regional e global. No entanto, imagens multiespectrais geralmente contêm pixels com problemas de mistura espectral, ou afetados por nuvens e as suas sombras, resultando em dados ruidosos, que muitas vezes são inadequados para a análise de áreas queimadas.

4 Conclusões

Os produtos do Satélite Terra/MODIS MOD09A1 e MCD64A1, integrados ao modelo de classificação Random Forest, permitiu a análise espaço-temporal de áreas de queimadas e pós-incêndio no estado do Maranhão.

O mosaico de imagens do produto MOD09A1 foi adequado para a identificação e classificação supervisionada de Áreas Queimadas, além da facilidade de pré-processamento e menor custo computacional, mesmo sendo aplicado numa região grande área. A disponibilização de 7 canais de reflectância da superfície deste produto, principalmente, na faixa espectral do infravermelho, foram importantes para a melhor identificação dos alvos de áreas atingidas pelo fogo em regiões de savana.

O uso do Modelo *Random Forest* para mapeamentos de áreas queimadas de baixa gravidade exerce boa funcionalidade, principalmente, quando se quer mapear áreas em que o produto Modis MCD64A1 ultrapassa a precisão espacial de 500 m. Localidades com classes de ocupação do solo opticamente menos complexas podem gerar resultados mais relevantes, pois aumentam a precisão do classificador em prever pixels espacialmente de feições queimadas. Apesar disso, foi observado neste trabalho, algumas limitações importantes ao usar o Random Forest por imagens de sensoriamento remoto. Na execução do *Random Forest* é extremamente importante que os dados de treinamento incluam amostras que cubram o máximo possível de toda a gama de áreas queimadas na área em estudo, pois a resposta da classificação por este algoritmo é a média de todas as árvores obtidas pelos dados de treinamento. Por exemplo, se os dados de treinamento para um modelo de classificação contiverem pixels com valores abaixo da média de cada árvore, e nenhum valor alto, não será possível classificar com precisão pixels de áreas queimadas com valores das árvores acima da média, quando o modelo for aplicado ao conjunto de dados completo. O efeito de classificar incorretamente feições de áreas queimadas contribui para aumentar a variabilidade intraclasse. Gerando assim uma diminuição da precisão dos mapas classificados a partir dessas informações.

Os resultados deste estudo mostraram que a abordagem apresentada é útil para determinar áreas

queimadas oriundas de dados de satélite de órbita heliocêntrica. Apresentando uma relação não linear entre a variabilidade espaço-temporal de incêndio e o clima em ecossistemas temperados.

A Metodologia estabelecida e validada neste estudo poderá ser estendida para outras regiões de climas semelhantes para detecção de outras paisagens queimadas, mesmo quando são necessárias medições espacialmente menos extensas.

5 Referências

- Bastarrika, A.; Chuvieco, E. & Martín, M.P. 2011. Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: balancing omission and commission errors. *Remote Sensing of Environment*, 115(4): 1003-1012. DOI. 10.1016/j.rse.2010.12.005.
- Belgiu, M. & Drăguń, L. 2016. Random forest in remote sensing: a review of applications and future directions. *Journal of Photogrammetry and Remote Sensing*, 114: 24-31. DOI.10.1016/j.isprsjprs.2016.01.011.
- Bezerra, D.S.; Dias, B.C.C.; Rodrigues, L.H.S.; Tomaz, R.B.; Santos, A.L. & Silva Junior, C.H.L. 2018. Análise dos focos de queimadas e seus impactos no Maranhão. *Revista Brasileira de Climatologia*, 22: 446-462.
- Breiman, L. 2001. Random forests. Kluwer Academic Publishers. Manufactured in The Netherland. *Machine Learning*, 45: 5-32.
- Calle, A. & Casanova, J.L. 2015. Forest Fires and Remote Sensing. *Earth Observation* 9: 177-202.
- Carmenta, A.R.; Parry, L.; Blackburn, A.; Vermeulen, S. & Barlow, J. 2011. Understanding human-fire interactions in tropical forest regions: A case for interdisciplinary research across the natural and social sciences. *Ecology and Society*, 16: 53-75.
- Chuvieco, E.; Giglio, L. & Justice, C. 2008. Global characterization of fire activity: toward defining fire regimes from earth observation data. *Global Change Biology*, 14(7): 1488-1502. DOI.10.1111/j.1365-2486.2008.01585.x.
- Chuvieco, E.; Lizundia-Loiola, J.; Pettinari, M.L.; Ramo, R.; Padilla, M.; Tansey, K.; Mouillot, F.; Laurent, P.; Storm, T. & Heil, A. 2018. Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth System Science Data*, 10(4): 2015-2031. DOI.10.5194/essd-10-2015-2018.
- Chuvieco, E.; Mouillot, F.; Werf, G.R.V.D.; Miguel, J.S.; Tanase, M.; Koutsias, N.; García, M.; Yebra, M.; Padilla, M. & Gitas, I. 2019. Historical background and current developments for mapping burned area from Satellite Earth Observation. *Remote Sensing of Environment*, 225: 45-64. DOI. 10.1016/j.rse.2019.02.013.
- Chu, T. & Guo, X. 2014. Review Remote Sensing Techniques in Monitoring Post-Fire Effects and Patterns of Forest Recovery in Boreal Forest Regions: A Review. *Remote Sensing*, 6: 470-520. DOI.10.3390/rs6010470.
- Collins, L.; Griffioen, P.; Newell, G. & Mellor, A. 2018. The utility of Random Forests for wildfire severity mapping. *Remote Sensing of Environment*, 216: 374 - 384.

- Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J. & Lawler, J.J. 2007. Random Forests for Classification in Ecology. *Ecology*, 88(11): 2783-2792. DOI.10.1890/07-0539.1.
- Giglio, L.; Van Der Werf, G.R.; Randerson, J.T.; Collatz, G.J. ; Kasibhatla, P.S. 2006. Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics*, 6: 957-974.
- Guo, F.; Wang, G.; Su, Z.; Liang, H.; Wang, W.; Lin, F. & Liu, A. 2016. What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. *International of Wildland Fire*, 25(5): 505-519.
- Hawbaker, T.J.; Vanderhoof, M.K.; Schmidt, G.L.; Beal, Y.; Picotte, J.J.; Takacs, J.D.; Falgout, J.T. & Dwyer, J.L. 2020. The Landsat Burned Area algorithm and products for the conterminous United States. *Remote Sensing of Environment*, 244: 111801. DOI. 10.1016/j.rse.2020.111801.
- INMET. 2020. INSTITUTO NACIONAL DE METEOROLOGIA - BDMEP - Banco de Dados Meteorológicos para Ensino e Pesquisa. Disponível em: <http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep>. Acesso em: 05 de março de 2020.
- INPE. 2020. INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite. São José dos Campos. 2020. Disponível em: <<http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>>. Acesso em: 10 de março. 2020.
- Klink, C.A. & Machado, R.B. 2005. Conservation of the Brazilian Cerrado. *Conservation Biology*, 19(3): 707-713. DOI. 10.1111/j.1523-1739.2005.00702.
- Libonati, R.; Dacamara, C.; Setzer, A.; Morelli, F. & Melchiori, A. 2015. An Algorithm for Burned Area Detection in the Brazilian Cerrado Using 4 μm MODIS Imagery. *Remote Sensing*, 7(11): 15782-15803. DOI.10.3390/rs71115782.
- Lizundia-Loiola, J.; Pettinari, M.L. & Chuvieco, E. 2020. Temporal Anomalies in Burned Area Trends: satellite estimations of the amazonian 2019 fire crisis. *Remote Sensing*, 12(1): 151-158. DOI. 10.3390/rs12010151.
- Maranhão. 2011. Plano de ação para prevenção e controle do desmatamento e das queimadas no estado do Maranhão. Decreto nº 27.317, de 14 de abril de 2011. Disponível em: <http://www.oads.org.br/leis/2503.pdf>. Acesso em: 10 de março de 2020.
- Masullo, G.A.Y. 2018. Análise preditiva de ocorrências de incêndios no bioma amazônico do Maranhão. *GeoTextos*, 14(2): 185-211.
- Melchiori, E.A.; Setzer, A.W.; Morelli, F.; Libonati, R.; Candido, P. & Jesus, S.A. Landsat-Tm/Oli algorithm for burned areas in the Brazilian Cerrado—Preliminary results. In: Proceedings of VII International Conference on Forest Fire Research, Coimbra, Portugal, 17–20 November 2014.
- Mellor, A.; Haywood, A.; Stone, C. & Jones, S. 2013. The Performance of Random Forests in an Operational Setting for Large Area Sclerophyll Forest Classification. *Remote Sensing*, 5(6): 2838-2856. DOI.10.3390/rs5062838.
- Millard, K. & Richardson, M. 2015. On the Importance of Training Data Sample Selection in Random Forest Image Classification: a case study in peatland ecosystem mapping. *Remote Sensing*, 7(7): 8489-8515. DOI. 10.3390/rs70708489.
- MMA. 2018. MINISTÉRIO DO MEIO AMBIENTE. Análises no Cerrado. Disponível em: <<http://combateadesmatamento.mma.gov.br/analises-no-cerrado>>. Acesso em: 1 jan. 2020.
- MODIS. 2015. Moderate Resolution Imaging Spectroradiometer Surface Reflectance - User's Guide. NASA (USA). Disponível em: <https://modis-land.gsfc.nasa.gov/pdf/MOD09_UserGuide_v1.4.pdf>. Acesso em: 6 mar. 2020.
- Oertel, C.; Matschullat, J.; Zurbach, K.; Zimmermann, F. & Erasmí, S. 2016. Greenhouse gas emissions from soils—A review. *Erde*, 76: 327-352
- Pereira, A.; Pereira, J.; Libonati, R.; Oom, D.; Setzer, A.; Morelli, F.; Machado-silva, F. & Carvalho, L. 2017. Burned Area Mapping in the Brazilian Savanna Using a One-Class Support Vector Machine Trained by Active Fires. *Remote Sensing*, 9(11): 1161. DOI.10.3390/rs9111161.
- Pereira, A.A.; Teixeira, R.F.; Libonati, R.; Melchiori, A.E. & Carvalho, T.M.L. 2016. Avaliação de índices espectrais para identificação de áreas queimadas no cerrado utilizando dados LANDSAT TM. *Revista Brasileira de Cartografia*, 8(68): 1665-1680.
- Pulvirenti, L.; Squicciarino, G.; Fiori, E.; Fiorucci, P.; Ferraris, L.; Negro, D.; Gollini, A.; Severino, M. & Puca, S. 2020. An Automatic Processing Chain for Near Real-Time Mapping of Burned Forest Areas Using Sentinel-2 Data. *Remote Sensing*, 12(4): 1-27. DOI. 10.3390/rs12040674.
- Ramo, R. & Chuvieco, E. 2017. Developing a Random Forest Algorithm for MODIS Global Burned Area Classification. *Remote Sensing*, 9(11): 1193. DOI. 10.3390/rs9111193.
- Ramo, R.; García, M.; Rodríguez, D. & Chuvieco, E. 2018. A data mining approach for global burned area mapping. *International Journal of Applied Earth Observation And Geoinformation*, 73: 39-51. DOI.10.1016/j.jag.2018.05.027.
- Sales, G.M.; Pereira, M.C.; Thalês, R.; Pocard-chapuis, R. & Almeida, S.A. 2019. Emprego dos focos de calor na avaliação das áreas queimadas e em incêndios florestais em Paragominas, Pará, Brasil. *Boletim do Museu Paraense Emílio Goeldi*, 14(1): 55-77. Disponível em: <<https://boletimcn.museu-goeldi.br/bcnaturais/article/view/140>>. Acesso em: 22 mai. 2020.
- Tanase, M.A.; Belenguer-Plomer, M.A.; Roteta, E.; Bastarrika, A.; Wheeler, J.; Fernández-Carrillo, Á.; Tansey, K.; Wiedemann, W.; Navratil, P. & Lohberge, R.S. 2020. Burned Area Detection and Mapping: intercomparison of sentinel-1 and sentinel-2 based algorithms over tropical Africa. *Remote Sensing*, 12(2): 334. 10.3390/rs12020334.
- Tian, S.; Zhang, X.; Tian, J. & Sun, Q. 2016. Random Forest Classification of Wetland Landcovers from Multi-Sensor Data in the Arid Region of Xinjiang, China. *Remote Sensing*, 8(11): 954. DOI. 10.3390/rs8110954.

USGS. 2020. United States Geological Survey. Disponível em: <<https://lpdaac.usgs.gov/products/mcd64a1v006/>>. Acesso em: 20 de maio de 2020.

Wilks, D.S. Statistical Methods in the Atmospheric Sciences. Cambridge: Academic Press, 2011. 704 p.

Yao, J.; Raffuse, S.M; Brauer, M.; Williamson, J.; Bowman, D.M.J.S.; Johnston, H. & Henderson, S.B. 2018. Predicting the minimum height of forest fire smoke within the atmosphere using machine learning and data from the CALIPSO satellite. *Remote Sensing of Environment*, 206(1): 98-106.

Recebido em: 06/07/2020

Aprovado em: 28/12/2020

Como citar:

Pacheco, A.P. & Silva-Jr., J.A. 2021. Análise Espaço-Temporal de Áreas de Queimadas no Estado do Maranhão a partir de Imagens MODIS e Classificação Random Forest. *Anuário do Instituto de Geociências*, 44: 36119. DOI 1982-3908_2021_44_36119