

PRIORIZAÇÃO DE ÁREAS PARA PREVENÇÃO E COMBATE A INCÊNDIOS FLORESTAIS NA AMAZÔNIA BRASILEIRA

Ignácio Martins Pinho ¹, Alber Hamersson Sánchez Ipia ², Guilherme Mataveli ³ e Liana Oighenstein Anderson ⁴

¹ INPE, martins.pinhos@gmail.com, ² INPE, alber.ipia@inpe.br, ³ INPE, guilherme.mataveli@inpe.br e ⁴ Cemaden, lianaanderson@gmail.com

RESUMO

Os incêndios florestais representam um grande desafio para a conservação da Amazônia. O fogo na Amazônia é de origem humana e causa diversos impactos negativos em escalas locais, regionais e globais. Neste trabalho visamos desenvolver um método para priorização de áreas com previsão de ocorrência de incêndios florestais. A priorização foi feita a partir do desenvolvimento de um modelo de regressão *Random Forest* para previsão de ocorrência de incêndios florestais. Para o período de análise, a área determinada como prioritária continha 79,9% do total de área de floresta queimada na Amazônia brasileira. Os resultados ressaltam o potencial do uso de dados de sensoriamento remoto para gestão territorial da Amazônia.

Palavras-chave – Amazônia, incêndios florestais, *Random Forest*, modelagem preditiva, gestão territorial.

ABSTRACT

Forest fires are a major challenge for the conservation of the Amazon. Forest fires are human related and cause negative impacts on local, regional and global scales. We aimed to develop an area prioritization method with the prediction of forest fire occurrence. Prioritization was based on the development of a Random Forest regression model for predicting the occurrence of forest fires. For the period of analysis, the area determined as a priority contained 79.9% of the total area of burned forests in the Brazilian Amazon. The results highlight the potential of using remote sensing data for territorial management in the Amazon.

Key words – Amazonia, forest fire, *Random Forest*, predictive modeling, territorial management.

1. INTRODUÇÃO

A Amazônia é a maior floresta tropical do mundo, abriga a maior biodiversidade terrestre do planeta, estoca uma quantidade de carbono equivalente a 15 anos de emissões globais e tem um importante papel na regulação do clima e prestação de serviços ecossistêmicos [1]. Entretanto, o avanço de atividades antrópicas sobre a floresta e as mudanças climáticas decorrentes do aquecimento global ameaçam seu funcionamento. Neste cenário, os incêndios florestais representam um grande desafio para conservação do bioma [2]. Os incêndios florestais causam diversos impactos na dinâmica da floresta, como mortalidade vegetal e animal, redução da biomassa e da cobertura do dossel. A emissão

de carbono também representa um impacto de nível global. Os incêndios ainda causam impactos sociais e econômicos como mortalidade direta de pessoas, aumento de doenças respiratórias, insegurança hídrica, alimentar e energética e perda de lavouras agrícolas, de infraestrutura, de estoques de carbono e de serviços ecossistêmicos.

Na Amazônia as fontes de ignição que desencadeiam incêndios florestais são, majoritariamente, de origem humana. O fogo é utilizado para três fins principais: (1) conversão de floresta em áreas agropecuárias; (2) manutenção da área de pastagem para evitar o crescimento de vegetação secundária e (3) ferramenta de afronta a políticas públicas e gestão territorial [3]. Além das ignições humano, existem outros fatores que amplificam os incêndios: clima e paisagem [4]. Temperaturas altas e ausência de precipitação reduzem a umidade da vegetação, criando um cenário propício para ocorrência de incêndios, pois o combustível vegetal está seco. Em um cenário de aquecimento global com temperaturas mais altas e secas mais frequentes e longas, a floresta amazônica torna-se mais inflamável e, portanto, vulnerável ao fogo. Ainda, a estrutura da paisagem relaciona-se com a ocorrência de incêndios, pois as bordas das florestas, além de estarem mais próximas das fontes de ignição antrópicas, tem um microclima mais seco e quente, mais vento e menos biomassa. Por fim, a degradação florestal causada pelos próprios incêndios e pelo corte seletivo aumenta a quantidade de biomassa morta e reduz a cobertura do dossel, o que aumenta a incidência solar no interior da floresta, tornando-a mais quente e seca.

Para reduzir a ocorrência de incêndios florestais na Amazônia, é necessário que as políticas públicas de prevenção e combate tenham fundamentação científica nos padrões espaciais e temporais, processos, causas e consequências dos incêndios florestais, que são específicas para o bioma amazônico. Foram implementadas a Política de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal (PPCDAm) em 2023 e a Política Nacional do Manejo Integrado do Fogo em 2024. Embora a prevenção e o combate de incêndios florestais sejam contempladas dentro dessas políticas, pouco é estabelecido sobre o direcionamento de ações e definição de áreas prioritárias, tornando a gestão territorial menos eficiente. O objetivo deste trabalho é desenvolver um método que permita a delimitação de áreas prioritárias para estruturação de ações de resposta a incêndios na Amazônia com antecedência mensal. Para isso, foi desenvolvido um modelo para previsão de área de floresta queimada fundamentado em vetores de ocorrência de incêndios florestais na Amazônia.

2. MATERIAL E MÉTODOS

O desenvolvimento deste trabalho foi dividido em duas etapas. A primeira etapa é o levantamento e processamento de dados de produtos de sensoriamento remoto para estimar variáveis que representam os vetores de ocorrência de incêndios florestais na Amazônia. Para elaboração do trabalho, a área de estudo foi dividida em uma grade celular em que cada célula tem 0,5° de largura e foi definido período de análise entre 1 e 30 de novembro de 2023. Isso foi feito para compatibilizar os dados de previsão do tempo utilizados com o restante dos dados. A segunda etapa compreende o desenvolvimento de um modelo de regressão *Random Forest* para predição de área de floresta queimada no período de análise e posterior delimitação de áreas prioritárias. O modelo de priorização criado tem duas características importantes: (1) a previsão é sub-sazonal (período de um mês), permitindo um tempo hábil para direcionamento das ações de resposta; (2) a área classificada como prioritária compreende uma parcela significativa dos incêndios florestais mas não é extensa a ponto de impossibilitar as ações de resposta.

2.1. Área de Estudo

A área de estudo compreende todo o bioma amazônico brasileiro, totalizando uma área com mais de 4 milhões de km² que representa 60 % de toda a floresta amazônica. O bioma apresenta clima equatorial e é majoritariamente composto por formações florestais. Entretanto, dadas suas enormes dimensões espaciais, há muita heterogeneidade tanto no clima como na estrutura e composição da floresta.

2.2. Dados e Variáveis Utilizadas

Foram utilizados dados de diversos produtos de sensoriamento remoto para estimação de variáveis que representam os vetores de ocorrência de incêndios florestais na Amazônia. Foram estimados valores para cada célula da grade. Para a delimitação de o quê é floresta ou não, foi utilizada a máscara de cobertura florestal do PRODES de 2023. A seguir estão as descrições das variáveis utilizadas na construção dos modelos.

- **aq_deter**: Área de floresta queimada por célula durante o período de análise (1 e 30 de novembro de 2023). Dados de cicatrizes de incêndios florestais foram utilizados para calcular a área de floresta queimada. **DETER**.
- **aq_mes_ant**: Área de floresta queimada por célula nos 30 dias anteriores ao período de análise. **DETER**.
- **aq_2022**: Área de floresta queimada por célula no período entre 1 e 30 de novembro de 2022. **DETER**.
- **des_mes_ant**: Área dos alertas de desmatamento por célula nos 30 dias anteriores ao período de análise. **DETER**.
- **des_2022**: Área dos alertas de desmatamento por célula no período entre 1 e 30 de novembro de 2022. **DETER**.
- **f_mes_ant**: Quantidade de focos de calor em áreas não florestadas por célula nos 30 dias anteriores ao período de análise. **FIRMS/VIIRS**.

- **f_2022**: Quantidade de focos de calor em áreas não florestadas por célula no período entre 1 e 30 de novembro de 2022. **FIRMS/VIIRS**.
- **DB, NF, AMF, CF, HIDRO**: Densidade de borda de floresta, quantidade de fragmentos de floresta, área média dos fragmentos de floresta, porcentagem de cobertura florestal e porcentagem de cobertura de corpos d'água por célula. **PRODES**.
- **CP**: Porcentagem de cobertura de pastagens por célula. **MapBiomass**.
- **UC, TI, pp_car**: Porcentagem de área ocupada por unidades de conservação, territórios indígenas e propriedades privadas registradas no Cadastro Ambiental Rural (CAR) para cada célula. **Atlas Agropecuário - Imafloa**.
- **BM**: Biomassa média das áreas florestadas dentro de cada célula. **CCI - Biomass**.
- **DE**: Densidade de estradas para cada célula. **OpenStreetMap**.
- **prev_prec, prev_temp**: Previsão de precipitação média por dia e previsão de temperatura média por célula durante o período de análise. **BAM-1.2 - CPTEC/INPE**.
- **DHA**: Déficit hídrico acumulado até o período de análise [5]. **CHIRPS**.
- **anom_prec, anom_temp**: Ocorrência ou não de anomalia de precipitação e temperatura de superfície média no mês anterior ao período de análise para cada célula [6]. **CHIRPS e MOD21C3**.

2.3. Modelagem e Delimitação de Área Prioritárias

Com os dados descritos, foram criados modelos de regressão *Random Forest* para previsão da área de floresta queimada por célula durante o período de análise. O algoritmo não paramétrico de aprendizado de máquina *Random Forest* utiliza um conjunto de árvores de decisão para realizar predições a partir de um conjunto de dados. Este método foi utilizado pois o algoritmo lida bem com diferentes tipos de dados, tem fácil implementação e otimização e lida bem com *overfitting*. A implementação foi feita em R através do pacote *Ranger*.

Foi calculado o coeficiente de correlação de Pearson entre cada variável independente e a variável dependente (**aq_deter**). As variáveis foram então ranqueadas de acordo com o valor absoluto de seus respectivos coeficientes de correlação. Então, foi criada uma série de modelos, iniciando apenas com a variável melhor ranqueada e a cada passo adicionando mais uma variável. Isso foi feito até chegar em um modelo que utilizava todas as variáveis. Neste processo, os hiper-parâmetros do modelo foram otimizados através de uma busca em grade. A cada passo, foi calculada a acurácia do modelo (R²). Além disso, cada modelo foi utilizado para fazer predições de área de floresta queimada. A partir dessas predições, cada célula da grade foi classificada em classes de prioridade para prevenção e combate ao fogo. A classificação foi feita utilizando percentis. As classes de alta, média e baixa prioridade foram definidas pelos percentis maior que

90, entre 90 e 70 e menor que 70, respectivamente. Em seguida, foi feito o seguinte cálculo para avaliar a qualidade da classificação de cada modelo:

$$\% \text{ de área de floresta queimada por classe} = \frac{\text{Área de floresta queimada por classe}}{\text{Total de área de floresta queimada}}$$

Também foi realizada uma classificação de prioridade utilizando os dados de área de floresta queimada do DETER. Como esta classificação utilizou os dados reais, ela representa a melhor classificação a ser realizada. As classificações feitas a partir das predições dos modelos foram comparadas com esta classificação ideal para avaliar quão eficientes foram as predições.

3. RESULTADOS

Em novembro de 2023, a área de floresta queimada no bioma amazônico brasileiro totalizou 2123,68 km². Neste período, incêndios florestais ocorreram em 197 das 471 células que compõem a área de estudo. Houve uma concentração de incêndios nas regiões central e leste da Amazônia brasileira, embora também tenham ocorrido incêndios em outras regiões. Na Figura 1, está representada a distribuição espacial da área de floresta queimada no período.

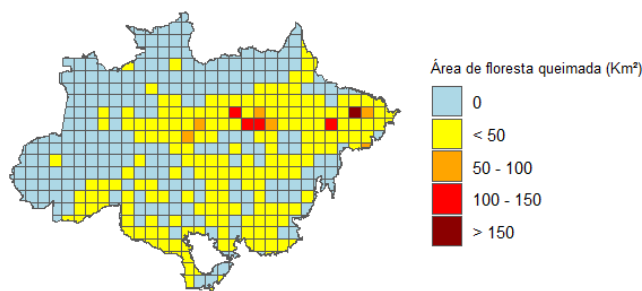


Figura 1: Área de floresta queimada (km²) por célula da grade entre 1 e 30 de novembro de 2023, segundo dados do DETER.

A Figura 2 apresenta o ranqueamento das variáveis em relação ao valor absoluto do coeficiente de correlação de Pearson. Seguindo a ordem deste ranqueamento, foram criados modelos de regressão.

O desempenho dos modelos, tanto em relação ao valor de R² como também em relação à porcentagem de área de floresta queimada nas células classificadas como alta prioridade, está representado na Figura 3. O valor máximo de R² (0,209) foi alcançado utilizando apenas as 3 primeiras variáveis do ranqueamento (des_mes_ant, f_mes_ant e prev_temp). Entretanto, esse modelo apresentou o terceiro pior desempenho em relação a porcentagem de área de floresta queimada nas células classificadas como alta prioridade (73,6%). Para essa métrica, o modelo com seis variáveis (des_mes_ant, f_mes_ant e prev_temp, f_2022, prev_prec e DB) apresentou o terceiro melhor desempenho (79,9%). A adição de mais variáveis não elevou essa porcentagem significativamente, apenas os modelos com 19 e 22 variáveis apresentaram uma porcentagem ligeiramente maior (80,1% e 80,2%, respectivamente). O modelo com apenas 6

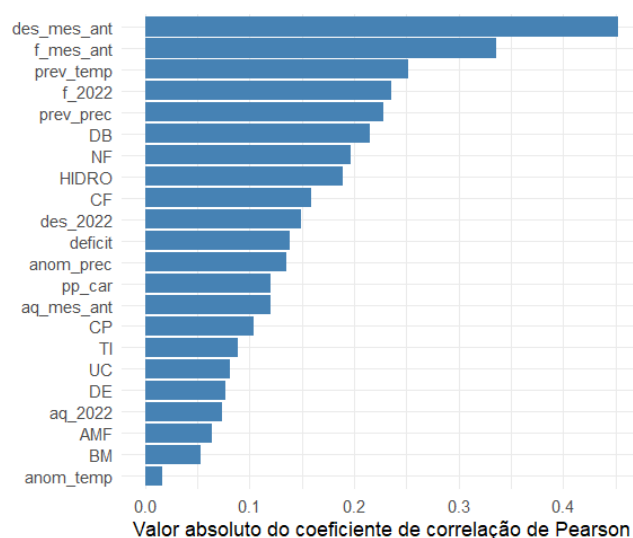


Figura 2: Ranqueamento das variáveis a partir do valor absoluto dos coeficientes de correlação entre as variáveis e a área de floresta queimada. As variáveis prev_prec, CF, DHA, anom_prec, TI, UC, AMF e BM apresentaram coeficiente com valor negativo.

variáveis foi considerado o melhor por dois fatores: (1) a diferença entre as porcentagens é muito pequena; (2) por ter uma quantidade muito menor de variáveis, o modelo é significativamente mais simples de ser explicado e portanto mais parcimonioso.

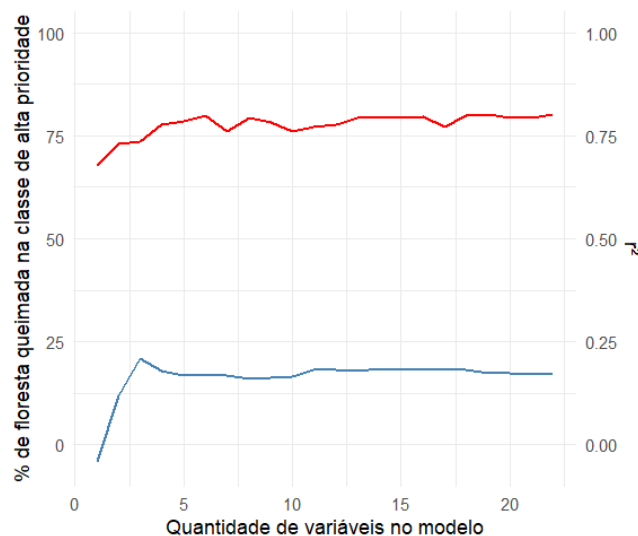


Figura 3: Desempenho dos modelos de acordo com a quantidade de variáveis. Em vermelho: porcentagem de área de floresta queimada que ocorreu em células classificadas com maior prioridade em relação ao total de área de floresta queimada. Em azul: R² de cada modelo.

A classificação de áreas prioritárias feita a partir do modelo selecionado está representada na Figura 4, assim como a classificação ideal feita a partir dos próprios dados de área de floresta queimada. A classificação feita a partir do modelo acertou a classificação de 32 células de alta prioridade e errou 15. Ambas as classificações concentraram as áreas de alta prioridade na região central e leste da Amazônia brasileira,

onde houve maior ocorrência de incêndios florestais. A classificação feita a partir do modelo não priorizou áreas no sul do bioma onde também houve quantidade significativa de incêndios. A Tabela 1 também apresenta uma comparação entre a classificação feita pelo modelo e a classificação feita com os dados de área queimada. Nela pode ser observado que a área de floresta queimada dentro da classe de alta prioridade determinada pelo modelo representa 89,6% da área de floresta queimada dentro da classe de alta prioridade determinada pela classificação ideal.

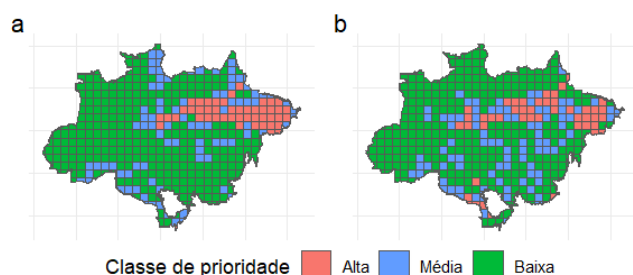


Figura 4: a: classificação de prioridade criada a partir das previsões do melhor modelo; b: classificação de prioridade criada a partir dos dados reais de área de floresta queimada do DETER.

Classificação a partir das previsões do modelo			Classificação a partir dos dados de área de floresta queimada do DETER		
Classe de prioridade	Área de floresta queimada (Km²)	Porcentagem da área de floresta queimada	Classe de prioridade	Área de floresta queimada (Km²)	Porcentagem da área de floresta queimada
Alta	1.696.23	79.87	Alta	1.893.27	89.15
Baixa	107.60	5.07	Baixa	18.59	0.88
Média	319.85	15.06	Média	211.82	9.97

Tabela 1: Comparação das classificações feitas a partir do modelo e a partir dos dados de área de floresta queimada do DETER.

4. DISCUSSÃO

A partir da comparação entre os modelos, percebe-se que o valor de R^2 traz pouca informação para avaliar qualidade do modelo. Como o objetivo não é fazer uma previsão exata da área de floresta queimada e sim fazer uma classificação de prioridade a partir destas previsões, ou seja, de suas espacializações, avaliar as próprias classificações é um melhor caminho para determinar a qualidade dos modelos. Portanto, foi selecionado como melhor o modelo composto por 6 variáveis. Estas variáveis condizem com os vetores de ocorrência de incêndios florestais na Amazônia. As fontes de ignição estão representadas pelas variáveis que quantificam o desmatamento e a utilização de fogo em áreas não florestadas. As condições meteorológicas estão representadas pela variáveis de previsão do tempo. Por fim, a estrutura da floresta está representada pela variável de densidade borda. A simplicidade, a concordância conceitual e o bom desempenho da classificação derivada deste modelo foram os motivos de sua seleção como modelo mais adequado.

A classificação feita a partir do modelo selecionado apresentou resultado satisfatório. Com 79,9% da área

de floresta queimada ocorrendo dentro da classe alta prioridade, o objetivo de delimitar áreas prioritárias pequenas que concentrassem parcela significativa da ocorrência de incêndios foi atingido. Deste modo, esse método de priorização de áreas para combate e prevenção de incêndios pode auxiliar o direcionamento de ações efetivas para reduzir e mitigar os impactos do fogo na Amazônia brasileira.

Apesar de apresentar bom potencial para delimitação de áreas prioritárias para combate e prevenção de incêndios, o método ainda tem limitações e precisa ser otimizado. Inicialmente, há limitações em relação aos próprios dados. Os produtos de sensoriamento remoto tendem a subestimar a ocorrência de fogo em florestas densas pois as copas das árvores dificultam a detecção de fogo no sub-bosque. Incertezas e limitações existem também em todos os outros dados utilizados. Portanto, a construção e comparação de modelos que utilizem outros dados como variável dependente deve ser feita. Por fim, é preciso testar a classificação do modelo a partir de conjuntos de dados de outras épocas do ano. Para que um modelo destes se torne operacional, é necessário que a classificação mantenha um padrão de desempenho ao longo do tempo. Esse pode ser o maior desafio, pois a importância relativa de cada variável não deve ser constante ao longo do tempo, visto que os vetores de ocorrência de incêndios florestais na Amazônia variam no tempo e espaço.

5. CONCLUSÕES

A ameaça do fogo, cada vez mais frequente e intensa, exige que o Brasil tome ações mais efetivas para reduzir a ocorrência de incêndios florestais na floresta amazônica. Isso é necessário para a manutenção do bioma, de suas interações ecológicas, serviços ecossistêmicos e para que o Brasil consiga cumprir suas metas de redução de emissão de carbono. Dadas as grandes dimensões da Amazônia, o uso de produtos de sensoriamento e aprendizado de máquina apresentam um grande potencial para auxiliar no planejamento da gestão territorial da Amazônia. Neste caso, o modelo criado possibilita auxiliar o direcionamento de ações de prevenção e combate de incêndios florestais.

6. REFERÊNCIAS

- [1] IPCC. Climate change 2021: The physical science basis (eds masson-delmotte, v. et al.). <https://www.ipcc.ch/report/ar6/wg1/FullReport>, Cambridge Univ. Press, 2021.
- [2] David M. Lapola et. al. The drivers and impacts of amazon forest degradation. *Science*, 379, 2023.
- [3] Jos Barlow et. al. Clarifying amazonia's burning crisis. *Global Change Biology*, 26, 2020.
- [4] Celso H.L. Silva-Junior et. al. Forest fragmentation and fires in the eastern brazilian amazon–maranhão state, brazil. *Fire*, 5, 2022.
- [5] Luiz Eduardo O.C. Aragão et. al. Spatial patterns and fire response of recent amazonian droughts. *Geophysical Research Letters*, 34, 2007.
- [6] Liana O. Anderson et. al. Remote sensing detection of droughts in amazonian forest canopies. *New Phytologist*, 187, 2010.