

MODELAGEM DE OCORRÊNCIA DE INCÊNDIOS FLORESTAIS RELACIONADOS A FATORES ANTRÓPICOS NA ÁREA DE PROTEÇÃO AMBIENTAL COCHÁ E GIBÃO

Livia Caroline César Dias¹, Luiz Eduardo Moschini², Yosio Edemir Shimabukuro³

¹ Universidade Federal de São Carlos, Rodovia Washington Luís, km 235 - SP-310 São Carlos – SP, diasclivia@gmail.com;

² Universidade Federal de São Carlos, Rodovia Washington Luís, km 235 - SP-310 São Carlos – SP, lemoschini@ufscar.br e

³ Instituto Nacional de Pesquisas Espaciais - INPE, Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil, yosio@dsr.inpe.br

RESUMO

Prever a ocorrência de incêndios florestais é uma etapa crucial para o planejamento da prevenção e do combate destes incêndios. Assim o objetivo deste trabalho foi desenvolver um modelo de suscetibilidade a incêndios florestais considerando a relação de variáveis topográficas, vegetacionais e antrópicas com a ocorrência dos incêndios florestais. Para isto, uma regressão logística foi aplicada na Área de Proteção Ambiental Cochá e Gibão. Utilizei focos de incêndios disponibilizados pelo INPE para gerar a variável dependente do modelo, e oito variáveis explicativas descritas como topográficas, vegetacionais e antrópicas. A escolha do melhor modelo foi baseada no Critério de Informação de Akaike. Os resultados mostraram uma AUC, que é a taxa de acerto do modelo, de 80,9%.

Palavras-chave — Regressão logística binária, AIC, incêndios florestais, sensoriamento remoto.

ABSTRACT

Predicting the occurrence of wildfires is a crucial step in planning for fire prevention and fighting. Thus, the objective of this work was to develop a susceptibility model to forest fires considering the relationship of topographic, vegetation, anthropic variables with the occurrence of forest fires. In this work, we the logistic regression that was applied in the Cocha and Gibão Environmental Protection Area. We use fire spots provided by INPE to generate a model-dependent variable, and eight explanatory variables such as topographic, vegetational and anthropic. The choice of the best model was based on the Akaike Information Criterion. The results showed an AUC, which is the model's hit rate, of 80.9%.

Key words — Binary logistic regression, AIC, forest fires, remote sensing.

1. INTRODUÇÃO

Extensas temporadas de incêndios têm resultado em impactos negativos em ecossistemas de diferentes biomas do Brasil, como Pantanal, Amazônia e o Cerrado. O fogo ocorre naturalmente em vários ecossistemas ao redor do mundo,

desde as florestas boreais até as savanas tropicais [1]. O Cerrado é um ecossistema dependente do fogo e sua vegetação apresenta inúmeras adaptações morfológicas e fisiológicas resultado de anos de evolução [2].

As políticas de supressão de fogo, principalmente em Unidades de Conservação, pode resultar em um alto acúmulo de combustíveis terrestres inflamáveis, ocasionando no aumento da incidência de incêndios florestais [3]. Somado a isso, e principal fator, o avanço acelerado da fronteira agrícola em áreas de cerrado e a crescente influência antrópica neste bioma tem aumentado a ocorrência de incêndios florestais, e a recorrência deste evento prejudica os benefícios do fogo e danifica o ecossistema mais rápido do que ele pode se recuperar [4].

Atualmente, os incêndios florestais tem demonstrado uma relação altamente dependente com as atividades humanas, resultando em impactos significativos nas esferas social, econômica e ecológica [5]. No Cerrado, as principais fontes de ignição estão associadas diretamente às ações humanas, seja utilizando o fogo para desmatar, limpar o terreno ou renovar o pasto, e estas atividades ocorrem principalmente durante a estação seca [1]. Quanto mais ao norte do Brasil, mais as áreas de Cerrado cruzam com a expansão da fronteira agrícola e áreas de vegetação nativa são perdidas devido à pressão antrópica.

O sensoriamento remoto pode ser uma ferramenta valiosa na aquisição de dados de incêndios pois pode auxiliar na construção de modelos focados em causas específicas e deve ser considerado uma meta global de gestão de incêndios, com potencial para mitigação de risco mais eficiente associado a investigações e coleta de campo. Neste contexto, o objetivo deste trabalho foi desenvolver um modelo de regressão logística a fim de estimar possíveis áreas de risco de incêndios florestais na Área de Proteção Ambiental Cochá e Gibão considerando fatores de origem antrópica e dados topográficos.

2. MATERIAL E MÉTODOS

2.1 Área de estudo

O presente estudo desenvolveu-se na Área de Proteção Ambiental Estadual Cochá e Gibão (APA Cochá e Gibão) que tem como principal objetivo proteger as formações de Cerrado, sendo uma área de extrema importância para a conservação. A APA Cochá e Gibão foi criada pelo Decreto nº 43.911 de 05 de novembro de 2004, na região dos

municípios de Januária, Bonito de Minas e Cônego Marinho no norte de Minas Gerais com uma área de 2.853,26 km² [7] (Figura 1). O domínio vegetacional predominante na APA Cochá e Gibão é o Cerrado.

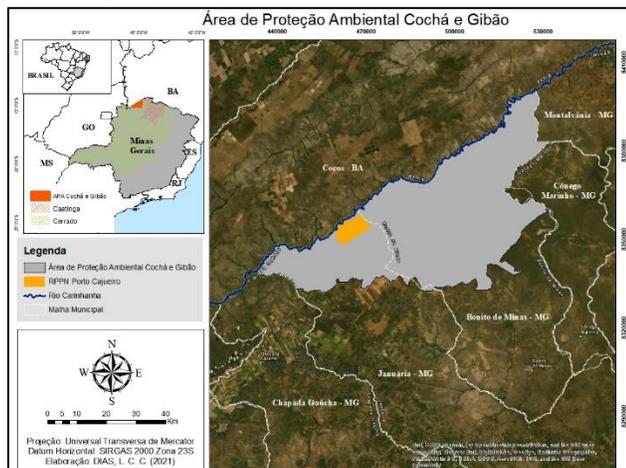


Figura 1: Localização da APA Cochá e Gibão no Estado de Minas Gerais

2.2 Modelo de Regressão Logística

Para a elaboração do modelo de suscetibilidade a ocorrência de incêndios florestais proposto, inicialmente foi necessário realizar a coleta, organização e estruturação de todo o banco de dados nos padrões necessários para que a metodologia escolhida fosse aplicada.

Os dados brutos obtidos em diferentes plataformas foram manipulados no software ArcGis® 10.7.1 para os padrões de entrada da metodologia utilizada, como podemos observar na tabela (Tabela 1).

Material	Ano	Resolução	Fonte	Formato
Focos de Incêndio	1999-2019	1km a 4km**	INPE	Point shapefile
NDVI (MOD13Q1)	2005-2019 (mensal)	30m	NASA	Raster
Altitude	2008	30m	TOPODATA	Raster
Declividade	2008	30m	TOPODATA	Raster
Distância de Hidrografia	2010	30m	A autora	Raster
Distância de Rodovias	2017	30m	A autora	Raster
Densidade Populacional	2010	30m	A autora	Raster
Distância de Habitações	2010	30m	A autora	Raster
Uso do solo coleção 5.0	2005, 2009, 2013, 2017	30m	MapBiomas	Raster

Tabela 1: Materiais utilizados na elaboração do modelo de suscetibilidade a ocorrência de incêndios florestais

A abordagem de *machine learning* escolhida para este trabalho foi a regressão logística binária. Existem vários trabalhos em localidades por todo mundo, em que a regressão logística é utilizada.

A regressão logística é muito utilizada pois permite associações a partir de conjuntos de variáveis que podem ser contínuas (numéricas), discretas (categóricas) ou uma mistura de qualquer um desses tipos, além possibilitar o conhecimento sobre as relações e os pontos fortes entre as variáveis [8].

O modelo de regressão logística representa a relação entre um conjunto de variáveis explicativas (representadas pela matriz de dados X) e a variável resposta/dependente (representada pelo vetor de resposta Y).

Variável Resposta

Para criar a nossa variável dependente, nós utilizamos focos de incêndios do período de 2005 a 2019. Estes focos estão disponíveis no Banco de Dados de Queimadas do INPE, e são representados por pixels cujo tamanho varia de acordo com o satélite utilizado. Para o modelo de regressão logística optamos por utilizar as detecções de focos de todos os satélites disponíveis e não apenas os “satélites de referência”. Quando o objetivo do estudo considera relevante toda e qualquer detecção de fogo, o INPE recomenda a utilização de todos os satélites.

No modelo de regressão logística binária a variável resposta é representada por dados categóricos 1 = ocorrência do evento e 0 = não ocorrência do evento. Neste caso, os focos de incêndios do INPE representam a variável resposta de categoria 1 (ocorrência de incêndios), para a categoria 0 (não ocorrência de incêndios) foi gerado, de forma aleatória, um número igual ao número de focos do INPE em regiões que não haviam ocorrências de incêndios. Dessa forma, para a APA Águas Vertentes obtivemos um total de 14.500 pontos (7.250 focos do INPE e 7.250 gerados aleatoriamente), ou seja, 50% da variável resposta são 1's e 50% são 0's, no intervalo de 15 anos (2005 a 2015).

Variáveis Explicativas

Um total de oito variáveis explicativas foram sugeridas para o modelo baseado no conhecimento prévio da área de estudo e em sua aplicação em estudos relacionados à previsão do risco de incêndios.

Foram escolhidas quatro variáveis consideradas características do ambiente, sendo elas: Altitude, declividade, NDVI (mensal) e distância de hidrografias. E quatro variáveis consideradas antrópicas, que foram: Distância de rodovias, distância de habitações, uso e ocupação do solo (para os anos 2005, 2009, 2013 e 2017) e densidade populacional.

Neste estudo realizamos uma regressão logística binária, pois a variável dependente ou variável resposta tem distribuição de Bernoulli (binomial). Portanto estamos trabalhando com a ocorrência (1) ou não ocorrência (0) de incêndios florestais, sendo a variável dependente binária com duas categorias (1 e 0 ou “fogo” e “não fogo”), enquanto que as variáveis explicativas independentes podem ser tanto numéricas quanto categóricas, a fim de explicar a sua relação com a ocorrência de incêndios florestais.

A seleção das variáveis e consequentemente a escolha do melhor modelo foi baseada no Critério de Informação de Akaike (AIC). Quanto menor o valor do AIC, melhor é o

modelo. Todas as combinações de conjuntos de variáveis foram testadas, mas o critério que utilizamos para decidir qual o modelo melhor para determinada APA foi o menor valor de AIC.

A estimação dos coeficientes da regressão logística é efetuada pelo uso da máxima verossimilhança, que busca encontrar as estimativas mais prováveis dos coeficientes e maximizar a probabilidade de que o evento ocorra. A qualidade do ajuste do modelo pode ser avaliada por diversas métricas, neste trabalho escolhemos o pseudo-R² (Tjur), a Raiz do Erro Quadrático Médio (RMSE – *Root Mean Squared Error*) e a Área sobre a Curva ROC (AUC – *Area Under the ROC Curve*).

3. RESULTADOS

Considerando o período de 2005 a 2019, de acordo com os dados do INPE, foram registrados um total de 7250 focos de incêndios para a APA Cochá e Gibão. A maior parte destes focos ocorreu nos meses de setembro e outubro, que são meses com baixa precipitação na região e por isso acredita-se que a ignição destes focos esteja relacionada com ações antrópicas e não a ocorrências naturais como por exemplo, relâmpagos.

Podemos observar na tabela (Tabela 2) a habilidade preditiva do modelo.

AUC	Sensibili- dade	Especifici- dade	R ² (Tjur)	RMSE
80,9%	73,8%	81,2%	0,416	0,377996

Tabela 2: Índices e métricas para analisar a performance dos modelos de predição de ocorrência de incêndios florestais: Área Abaixo da Curva ROC (AUC), Sensibilidade, Especificidade, pseudo-R² (tjur) e Raiz do Erro Quadrático Médio (RMSE)

Todas as variáveis explicativas utilizadas foram relevantes para o modelo e assim nenhuma foi excluída do mesmo. De modo geral, os valores de AUC são interpretados como: 0,5-0,6 (péssimo); 0,6-0,7 (ruim); 0,7-0,8 (pobre); 0,8-0,9 (bom) e >0,9 (excelente), o que representa a eficácia do modelo em porcentagem, ou seja, uma AUC de 0,7 reflete uma chance de classificação correta de 70% do evento [9]. O modelo da APA Cochá e Gibão apresentou um AUC de 80,9%, que pode ser considerado um modelo bom.

A sensibilidade (taxa de verdadeiros-positivos) representa a capacidade do modelo em avaliar o evento $\hat{Y}=1$ (estimado) dado que ele é evento real $Y=1$, ou seja, a capacidade do modelo em acertar os eventos em que houve a ocorrência de fogo. A especificidade (taxa de falsos-positivos) representa o poder de predição do modelo em avaliar o “não evento” $\hat{Y}=0$ (estimado) dado que ele é evento real $Y=0$, ou seja, a capacidade do modelo acertar quando não houve a ocorrência de fogo.

O R² de Tjur para todos os 0s observados na tabela de dados, calcula o valor médio previsto. Da mesma forma, para todos os 1s observados na tabela de dados, calcula esse valor médio previsto. O R² de Tjur é a distância (valor absoluto da

diferença) entre as duas médias. Assim, um valor de R² de Tjur se aproximando de 1 indica que há uma separação clara entre os valores previstos para os 0s e 1s.

O RMSE representa o erro do modelo, quanto mais próximo de 0 menor é o erro e melhor é o modelo.

Tanto a sensibilidade quanto a especificidade apresentaram taxas acima de 70% para o Modelo da APA Cochá e Gibão, o que significa que o modelo teve uma boa taxa de acerto ao prever incêndios, mas errou algumas vezes quando estes realmente não aconteceram. O modelo da APA Cochá e Gibão não se mostrou muito bem ajustado apenas com essas variáveis com um R² = 0,416, mas um erro baixo com um RMSE = 0,377996.

O Modelo da APA Cochá e Gibão possui um bom poder de predição. Os focos de incêndios dos anos de treinamento (2005-2015) estão distribuídos por todo o território da APA, e por isso podemos notar um mapa de probabilidades que apresenta várias regiões com uma chance razoável de que ocorra incêndios florestais (Figura 2).

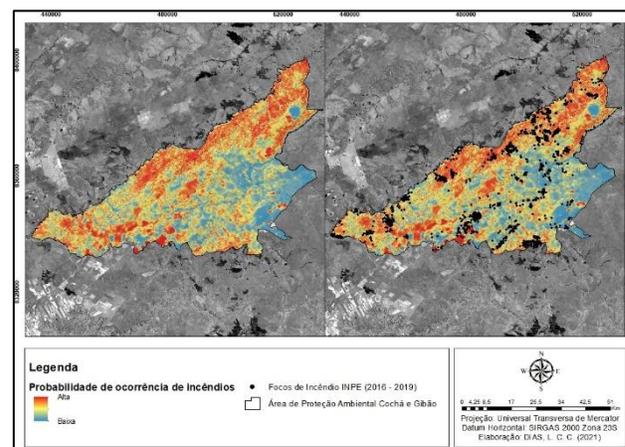


Figura 2: Mapa de probabilidade de incêndios florestais para a APA Cochá e Gibão

A área com mais alto risco de incêndios encontra-se na porção mais baixa da APA Cochá e Gibão e condiz com a região da Comunidade do Gibão e do Sumidouro, e como o modelo considerou a distância de habitações como uma variável relevante, podemos considerar esta relação direta com uma maior probabilidade de incêndios florestais nesta região. Na região mais a oeste da APA, há culturas de grandes produtores, diferente das pequenas agriculturas rurais da Comunidade do Gibão nestas áreas criam gado e a limpeza de pastos e lavouras para novas plantações utilizando fogo é muito comum.

4. DISCUSSÃO

O fogo é de fato um elemento natural no Cerrado, e sua ocorrência é importante para manter processos ecológicos neste bioma [4]. Já é de conhecimento comum que a supressão do fogo, fato comum em Unidades de Conservação de Proteção Integral, resulta no acúmulo de biomassa combustível, que quando acometida pelo fogo pode levar a

incêndios florestais de grandes proporções e difíceis de controlar [10].

Entretanto, as Áreas de Proteção Ambiental que são Unidades de Conservação de Uso Sustentável, apresentam outro tipo de manejo e diretrizes de conservação e os incêndios que ali ocorrem nem sempre estão relacionados com a supressão do fogo, mas sim causas antrópicas. Dessa forma, é necessário propor um plano de conservação que vise a participação das comunidades rurais que vivem daqueles recursos e que podem auxiliar na sua conservação, além de focar em ações que visem a colaboração dos latifundiários da região que vivem em conflito direto com as ações de manejo sustentável.

5. CONCLUSÕES

Notamos com este estudo que ao modelar incêndios florestais em Unidades de Conservação, inúmeros fatores devem ser levados em consideração muito além da paisagem. Porém estudos que preveem a ocorrência de incêndios no Cerrado, principalmente em Unidades de Conservação, estão crescendo devido ao cenário atual de incêndios florestais. Assim como em outras áreas do mundo, a Regressão Logística apresentou resultados satisfatórios para a modelagem de incêndios florestais.

A utilização de um conjunto diverso de variáveis explicativas proporcionou um resultado muito satisfatório para este trabalho. A metodologia utilizada pode ser replicada em outras áreas, inclusive as variáveis sugeridas neste trabalho são recomendadas. Entretanto, assim como neste estudo, para se atingir os melhores resultados desta metodologia, diferentes combinações devem ser testadas.

Para ações futuras seria interessante utilizar outros métodos de *machine learning* a critério de comparação entre modelos. No Brasil este tipo de metodologia tem se ampliado bastante e tem se mostrado um campo de pesquisa muito promissor.

Por fim, podemos afirmar que com os resultados obtidos, o mapa de ocorrência espacial e as estimativas correspondentes de novos incêndios obtidos poderão ser ferramentas valiosas para a prevenção e planejamento de tomada de decisões.

8. REFERÊNCIAS

- [1] U. OLIVEIRA, F. SOARES-FILHO, W. L. S. COSTA, L. GOMES, M. BUSTAMANTE and H. MIRANDA. Modeling fuel loads dynamics and fire spread probability in the Brazilian Cerrado. *Forest Ecology and Management*, v. 482, pp. 11889, 2021.
- [2] M. F. SIMON, R. GREYER, L. P. QUEIROZ, C. SKEMA, R. T. PENNINGTON and C. E. HUGHES C. E. Recent assembly of the cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Science*, v. 106, pp.20359-20364, 2009..
- [3] L. C. MOURA, A. O. SCARIOT, I. B. SCHMIDT, R. BEATTY and J. RUSSELL-SMITH. The legacy of colonial fire management policies on traditional livelihoods and ecological sustainability in savannas: Impacts, consequences, new directions. *Journal of Environmental Management*, v. 232, pp. 600-606, 2019.
- [4] V. R. PIVELLO. The use of fire in the cerrado and Amazonian rainforests of Brazil: past and present. *Fire Ecology*, v. 7, n. 1, pp. 24-39, 2011.
- [5] D. A. DRISCOLL, D. B. LINDENMAYER, A. F. BENNETT, M. BODE, R. A. BRADSTOCK, G. J. CARY, M. F. CLARKE, N. DEXTER, R. FENSHAM, G. FRIEND, M. GILL, S. JAMES, G.

KAY, D. A. KEITH, C. MACGREGOR, J. RUSSELL-SMITH, D. SALT, J. E. M. WATSON, R. J. WILLIAMS and A. YORK. Fire management for biodiversity conservation: key research questions and our capacity to answer them. *Biological Conservation*, v. 143, n. 9, pp. 1928-1939, 2010.

[6] G. F. ROCHA, L. G. FERREIRA, N. C. FERREIRA, and M. E. FERREIRA. Detecção de desmatamento no bioma Cerrado entre 2002 e 2009: Padrões, Tendências e Impactos. *Revista Brasileira de Cartografia*, v. 63, n. 3, pp. 341-349, 2011.

[7] FUNATURA. Fundação Pró Natureza. *Plano de Desenvolvimento Territorial de Base Conservacionista do Mosaico Sertão Veredas – Peruaçu*. Brasília, 2008. 60 p.

[8] K. GHOLAMNIA, T. G. NACHAPPA, O. GHORBANZADEH, and T. BLASCHKE. Comparisons of Diverse Machine Learning Approaches for Wildfire Susceptibility Mapping. *Symmetry*, v. 12, p. 604, 2020.

[9] T. C. F. POLO and H. A. MIOT. Aplicações da curva ROC em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, v. 19, 2020.

[10] A. FIDELIS, S. T. ALVARADO, A. C. S. BARRADAS and V. R. PIVELLO. The Year 2017: Megafires and Management in the Cerrado. *Fire*, v. 1, p. 49, 2018.