

# **A distribuição espacial das descargas elétricas na região central da Bacia Amazônica e o uso da árvore de decisão em dados de radar meteorológico para sua classificação**

**Adriano P. Almeida<sup>1</sup>, Alan J. P. Calheiros<sup>1</sup>, Rachel I. Albrecht<sup>2</sup>,  
Luiz A. T. Machado<sup>1</sup>, Rafael D. C. Santos<sup>1</sup>, Helvécio B. L. Neto<sup>1</sup>, Felipe C. Souza<sup>1</sup>**

<sup>1</sup>Instituto Nacional de Pesquisas Espaciais (INPE)  
CEP 12.227-010 – São José dos Campos – SP – Brasil

<sup>2</sup>Universidade de São Paulo (USP)  
CEP 05.508-090 – São Paulo – SP – Brasil

{adriano.almeida, alan.calheiros, luiz.machado, rafael.santos,  
helvecio.neto, felipe.carvalho}@inpe.br, rachel.albrecht@iag.usp.br

**Abstract.** *Severe storms have as their main characteristic the occurrence of lightning. Brazil is a country where a larger number of lightning are recorded during the year, it a relationship mainly with a lot of convection and its large territorial extension. Lightning can cause great risks and damage to society, being responsible for human fatalities, interruption of energy transmission and communication services, forest fires and losses in livestock herds in the rural area. Therefore, this work aims to present an analysis of the spatial distribution of lightning during fews months of 2014 in the central region of the Amazon Basin, and finally explore the main meteorological variables related to his intensification and cessation. The data come from the experiments of the GoAmazon project and the attributes selection process was made through the decision tree algorithm. The results showed that the decision tree technique was able to identify the patterns associated with the increase in the number of lightning, as well as in the classification of clouds with strong occurrences of lightning.*

**Resumo.** *As tempestades severas têm como uma de suas principais características as descargas elétricas atmosféricas. O Brasil é um dos países onde são registradas as maiores incidências de raios durante o ano, tendo relação sobretudo com sua grande extensão territorial e alta atividade convectiva. Os raios podem causar grandes riscos e prejuízos à sociedade, sendo responsáveis por fatalidades humanas, interrupção de serviços de transmissão de energia e comunicação, incêndios florestais e perdas em rebanhos de gados em zonas rurais. Diante disso, este trabalho tem como objetivo apresentar uma análise da distribuição espacial das descargas elétricas durante alguns meses de 2014 na região central da Bacia Amazônica, e por fim explorar os principais índices meteorológicos que estão associadas à sua intensificação ou não. Os dados são oriundos dos experimentos do projeto GoAmazon e a seleção dos atributos foi feita através do algoritmo árvore de decisão. Os resultados mostraram que com a técnica de árvore de decisão foi capaz de identificar os padrões associados ao aumento do número de descargas elétricas, assim como na classificação de nuvens com e sem raios.*

## 1. Introdução

As descargas elétricas atmosféricas têm sido alvo de estudos desde os tempos de Benjamin Franklin (1706-1790), quando o próprio cientista, em 1752, realizou um dos primeiros experimentos com este fenômeno da natureza. O experimento consistiu em soltar durante uma tempestade uma pipa manipulada por um fio de seda, o instrumento recebia cargas elétricas advindas de raios e descia até uma chave de metal onde o fio estava preso. Com isso, Franklin acabará de descobrir de forma concisa a natureza elétrica das descargas elétricas, dando origem ao primeiro tipo de para-raios [Krider 2006].

Os raios, como também são conhecidas as descargas atmosféricas, são utilizados como indicadores importantes para o monitoramento estratégico de tempestades, pois estão diretamente associados a tempo severo [dos Santos Filho et al. 2005]. As nuvens do tipo *cumulonimbus* são caracterizadas por possuírem um extenso perfil vertical, e o seu pleno desenvolvimento expressa o estágio maduro de uma tempestade severa [Rakov 2007]. Embora se desenvolvam tipicamente dentro das nuvens *cumulonimbus*, as descargas elétricas também podem se originar a partir de tempestades de neve, tempestades de areia ou erupções vulcânicas [Romero 2007].

Segundo [Souza 2017], os raios são eventos de natureza aleatória que advém das fortes interações entre os centros de cargas em sistemas convectivos. Essas interações chegam a desenvolver intensas atividades elétricas que podem chegar na ordem de milhares de ampères e ter uma grande extensão, sendo responsáveis pelos característicos efeitos sonoros e luminosos. As descargas elétricas podem interagir ou não com superfícies no solo. Na ocorrência de raios com a interação do solo, eles são denominados de nuvem-solo, já quando não há esse contato, eles são caracterizados como intra-nuvem (quando acontecem na região interna de uma única nuvem) ou nuvem-nuvem (quando o raio de uma nuvem interage com outras nuvens) [Rakov 2007]. A média global de ocorrência de descargas atmosféricas é superior a 100 por segundo, onde apenas 25% desse total chegam a atingir a superfície [Rakov 2007]. Embora os raios do tipo nuvem-solo sejam menos frequentes, são eles que oferecem maior perigo iminente para ser humano.

Além de afetarem diretamente a composição química da atmosfera e a qualidade do ar na alta troposfera (liberando altos índices de óxidos de nitrogênio e ozônio) [Weber et al. 2005], as descargas atmosféricas são responsáveis por uma série de impactos socioeconômicos. Os raios são os principais responsáveis por iniciarem incêndios florestais, interrupções de serviços de transmissão de energia e comunicação, e também por baixas em rebanhos de gado na zona em zona rural [Cardoso 2017]. As descargas atmosféricas também estão diretamente relacionadas ao alto índice de fatalidades contra seres humanos expostos à tempestades, conforme apresentado no trabalho de [Cardoso et al. 2014].

Por ser um país de grande extensão territorial, e está predominantemente localizado em zona tropical, o Brasil ocupa uma posição de destaque mundial no que diz respeito a grande incidência de raios, chegando a serem registrados aproximadamente 77,8 milhões por ano, de acordo com estimativa feita pelo Grupo de Eletricidade Atmosférica (ELAT)<sup>1</sup> do Instituto Nacional de Pesquisas Espaciais (INPE). As redes de detecção desempenham um papel fundamental no monitoramento de descargas elétricas no Brasil, re-

---

<sup>1</sup><http://www.inpe.br/webelat/homepage/menu/el.atm/perguntas.e.respostas.php>

presentando grande avanço nas pesquisas correlatas, principalmente na prevenção de fatalidades e auxílios na tomadas de decisão diante de tempos severos [Chinchay 2018]. Além das redes de detecção, os radares e satélites meteorológicos são de suma importância para provimento de dados que auxiliam pesquisas relacionadas às descargas atmosféricas [Calheiros et al. 2016].

Este trabalho tem como objetivo realizar a análise da distribuição espacial de descargas atmosféricas detectadas durante os experimentos da campanha GoAmazon [Machado et al. 2014] na região central da Bacia Amazônica, e identificar através da seleção de atributos do algoritmo de árvore de decisão os principais atributos meteorológicos que influenciaram na intensificação da ocorrência dos raios.

## **2. Contextualização da área de estudo**

Para este trabalho foram utilizados os dados do radar meteorológico SIPAM (Sistema de Proteção da Amazônia) [Saraiva et al. 2016] e da rede de detecção de raios LINET (sigla do inglês, *Lightning NETWORK*) [Betz et al. 2009]. Os dados são oriundos dos experimentos GoAmazon [Machado et al. 2018] que aconteceram no ano de 2014 na região central da Bacia Amazônica. O conjunto de dados, também é resultado do processamento feito no trabalho de [Pereira 2019]. A fim de se calcular alguns outros índices, foi utilizada a ferramenta proposta por [Queiroz 2009], que se trata de uma adaptação do algoritmo de rastreamento de nuvens ForTraCC (*Forecast and Tracking of Active Convective Cells*) [Vila et al. 2008].

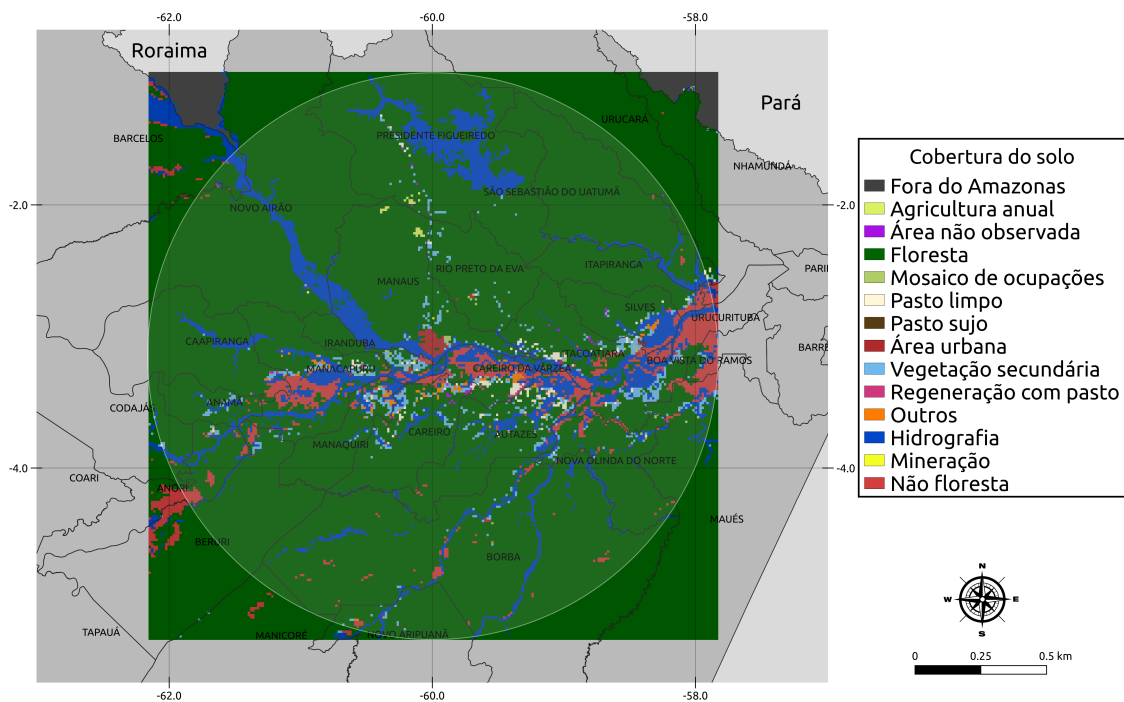
Os dados possuem resolução temporal de 12 minutos, e foram obtidos durante o período de 27 de agosto a 7 de outubro em uma área de 500 km<sup>2</sup>, cobrindo ao todo 20 municípios do estado do Amazonas. A floresta é o tipo de cobertura do solo predominante nessa região, e na área de estudo representa mais de 85%, como pode ser observado na Figura 1. Com isso, aumentam-se os riscos serem iniciados incêndios florestais em decorrência de raios nuvem-solo, principalmente a partir do mês de agosto, época em que se inicia o segundo Período de Operação Intensiva (IOP - sigla do inglês, *Intensive Operation Period*) [Marengo et al. 1994], que corresponde a transição entre as estações seca e chuvosa. É durante esse período que acontecem as chuvas mais intensas, desenvolvendo tempestades e conseqüentemente maior incidência de descargas elétricas [Machado et al. 2018].

## **3. Análise exploratória dos dados**

Nesta seção serão apresentadas algumas das análises que foram feitas a respeito da distribuição espacial das descargas elétricas atmosféricas e das principais variáveis medidas pelo radar SIPAM que estão associadas a estes eventos.

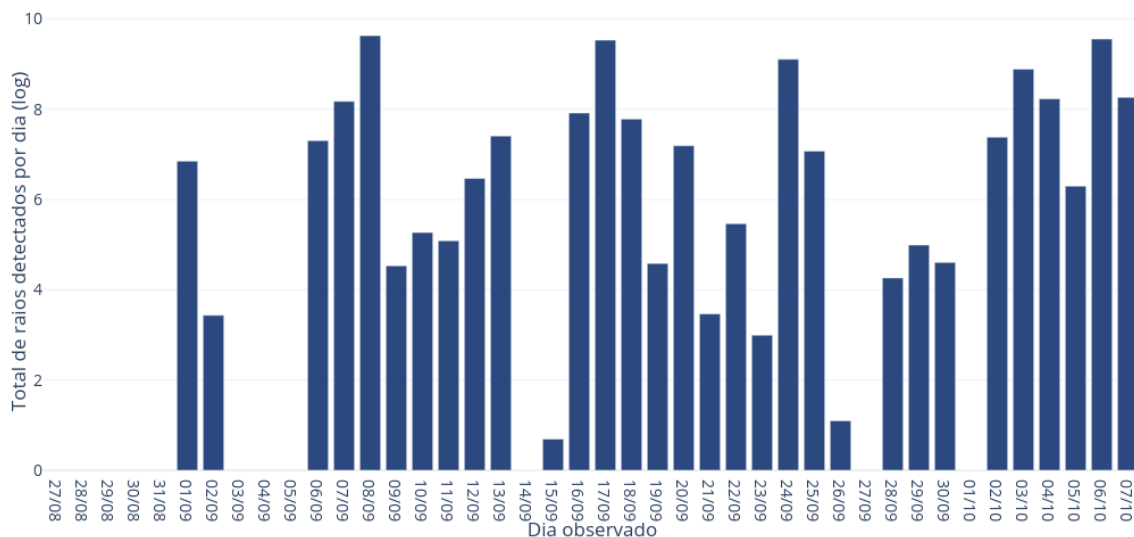
### **3.1. Densidade e distribuição de raios**

Do ponto de vista físico, os sistemas precipitantes são classificados em estratiformes ou convectivos [Damian 2011]. Essa classificação é baseada no trabalho de [Steiner et al. 1995], que fez uso principalmente dos índices de refletividade obtidos através dos ecos de radares meteorológicos. Enquanto nos sistemas estratiformes as chuvas acontecem de forma moderada e com distribuição uniforme, nos sistemas convectivos



**Figura 1. Área de estudo e sua cobertura do solo. Imagem disponível em: <https://bit.ly/3eCOdD1>**

elas acontecem de forma mais intensa e concentrada em regiões específicas, caracterizando uma condição de tempo severo. As descargas atmosféricas são fenômenos que acontecem principalmente nos sistemas convectivos.



**Figura 2. Total de raios detectados por dia pela rede LINET. Imagem disponível em: <https://bit.ly/3g1Vufl>**

Os dados utilizados neste trabalho são referentes ao segundo Período de Operação Intensiva (ou seja IOP 2), que como supracitado, corresponde ainda ao período mais seco

na região, sendo assim, as chuvas acontecem com pouca frequência e são oriundas principalmente de sistemas convectivos [Machado et al. 2018]. Esse comportamento pode ser observado no gráfico da Figura 2, onde durante todo o período de observação, somente em alguns dias houveram ocorrência de raios, e na maioria dos dias em que aconteceram, a atividade elétrica se manifestou de forma mais intensa.

Os valores do número de descargas elétricas observados na Figura 3, assim como em todo o trabalho, são providos de medidas realizadas pela LINET, que é uma rede que opera em baixa frequência, e faz a detecção de raios intra-nuvem e nuvem-solo [Betz et al. 2009]. Essa rede possui a bordo um sistema de GPS que possibilita também computar com precisão a localização onde ocorreram as descargas atmosféricas. Através desse recurso de geolocalização, é possível realizar análises mais aprofundadas do comportamento da atividade elétrica em certas regiões.

A Figura 3 foi gerada a partir da quantidade de raios detectados pela rede LINET em cada município da área de estudo. Careiro, Presidente Figueiredo e Manaus, foram os municípios em que foram detectadas durante o período as maiores ocorrências de raios com 12753, 10758 e 10305 detecções, respectivamente. Estudos como o que foi feito por [Pereira 2019], indicam que a atividade elétrica das nuvens tende a ser mais intensa

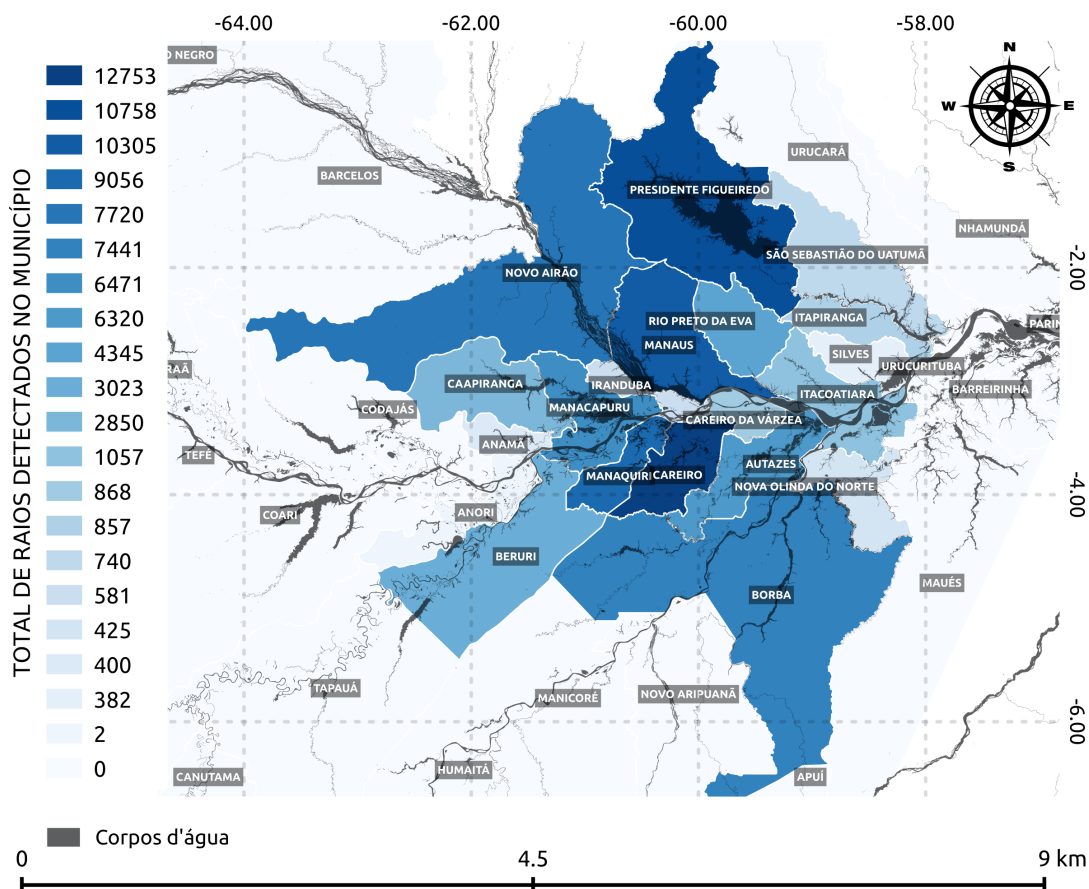
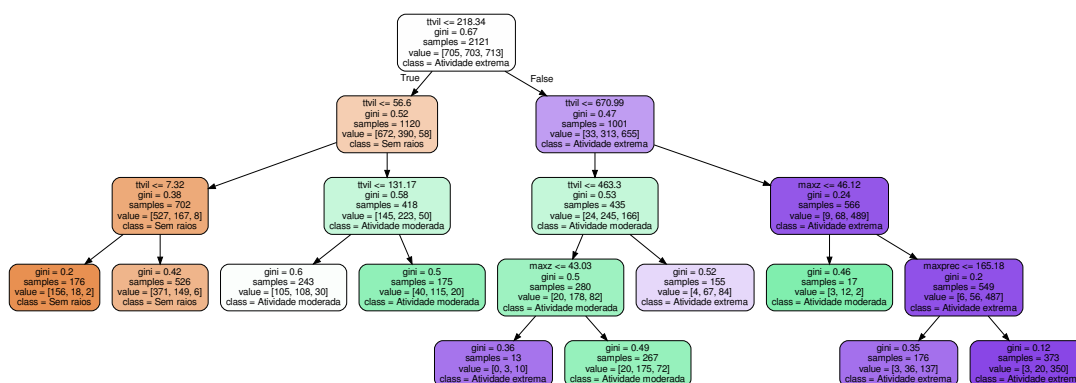


Figura 3. Densidade de descargas atmosféricas por município. Imagem disponível em: <https://bit.ly/3dEZhOy>

próximo aos rios, o que também pode ser observado na Figura 3, já que os municípios com maior incidência de descargas elétricas, ficam próximos a corpos d'água, como a comunidade de Careiro, que está situada às margens do rio Castanho.

### 3.2. Variáveis meteorológicas

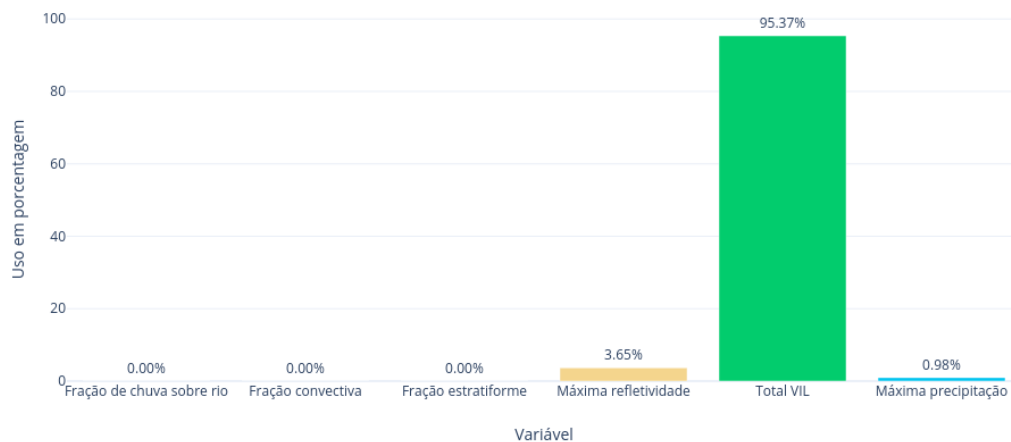
A refletividade ( $Z$ ) é a principal variável obtida através de radares meteorológicos. A partir desse índice que outros são gerados, como o conteúdo de água líquida integrado verticalmente (VIL - sigla do inglês, *Vertically Integrated Liquid*) [Greene and Clark 1972] e a taxa de precipitação ( $R$ ) [Marshall and Palmer 1948]. Essas foram as principais variáveis utilizadas no desenvolvimento deste trabalho. Sua escolha se deu através da seleção de atributos feita pelo algoritmo de aprendizado de máquina árvore de decisão [Safavian and Landgrebe 1991]. A árvore de decisão utilizada para classificar cada registro de acordo com intensidade de raios, definidas previamente como: Sem Raios (0 raios detectados), Atividade Moderada (1-10 raios) e Atividade mais intensa (>10 raios). A definição destes limiares foi feita de maneira subjetiva, análises estatísticas serão executadas no futuro para melhorar esta definição. A fim de aumentar a capacidade de generalização do modelo, a árvore de decisão foi configurada para ter no máximo 10 nós folhas. O esquema da árvore de decisão gerada após o treinamento pode ser visualizado na Figura 4.



**Figura 4. Esquema da árvore de decisão após o treinamento. Imagem disponível em: <https://bit.ly/3i4Yz05>**

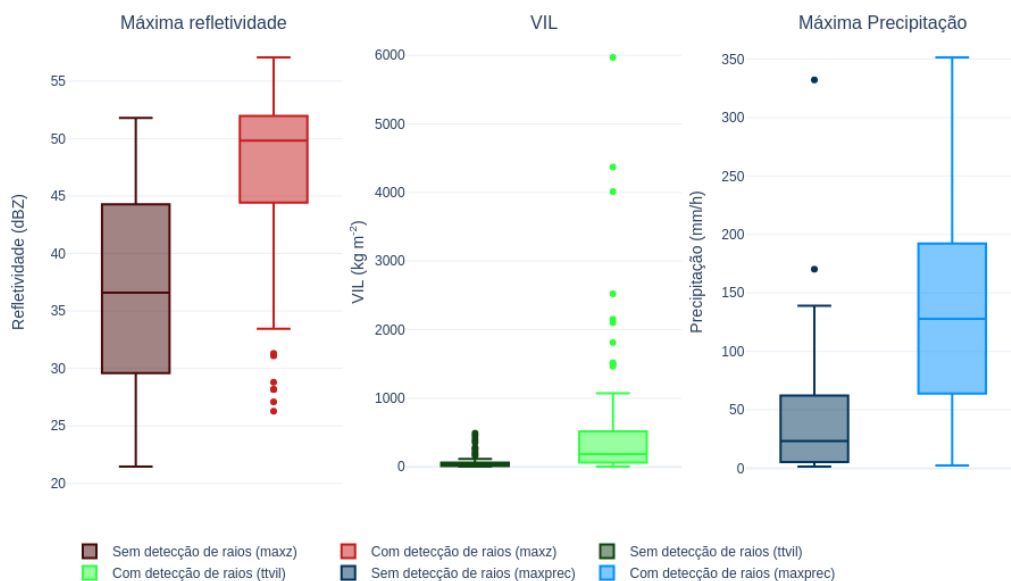
Os índices utilizados no modelo para a classificação da intensificação de raios foram: fração de chuva sobre os rios, fração convectiva, fração estratiforme, máxima refletividade, total VIL e máxima precipitação. Essas variáveis foram previamente obtidas a partir da base de dados preprocessada no trabalho de [Pereira 2019]. Ao final do treinamento, a árvore de decisão expôs os atributos que mais influenciaram na separação das classes, sendo eles: total VIL, máxima refletividade e máxima precipitação, com a porcentagem de importância de 95.37%, 3.65% e 0.98%, respectivamente, como mostrado no gráfico da Figura 5. Trabalhos futuros irão adicionar mais atributos a esta técnica, com o objetivo de melhorar ainda mais a precisão das classificações.

Baseado na seleção de atributos feita pela árvore de decisão, em conformidade com alguns trabalhos com propósitos semelhantes a esse [Pereira 2019] e [Chinchay 2018], as análises foram baseadas nas variáveis máxima refletividade, total



**Figura 5. Porcentagem da importância de cada atributo para a árvore de decisão. Imagem disponível em: <https://bit.ly/2A6n6kJ>**

Distribuição de alguns índices em subconjuntos com e sem detecção de raios (100 amostras de cada subconjunto)

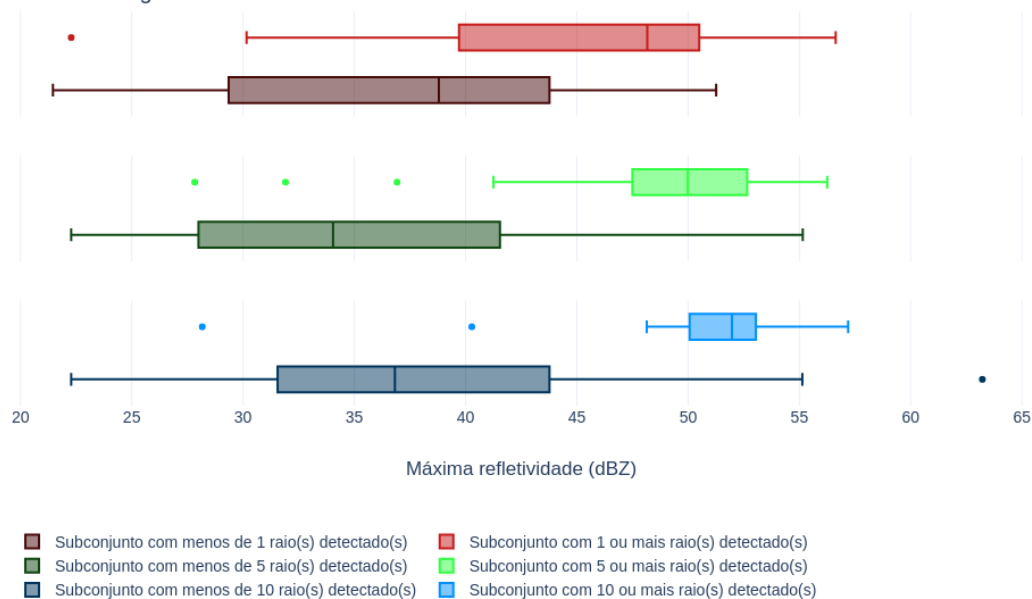


**Figura 6. Distribuição de alguns índices em subconjuntos com e sem detecção de raios. Imagem disponível em: <https://bit.ly/3gd6RSr>**

VIL e máxima precipitação. O comportamento dessas variáveis meteorológicas sofrem uma certa variação quando há ocorrência de descargas atmosféricas, como pode ser observado na Figura 6. Quando a atividade elétrica está mais intensa nos sistemas convectivos, as variáveis tendem a estarem também com valores mais altos, o que acaba caracterizando as condições de tempo severo.

O conjunto de dados possui uma baixa frequência de detecção de raios, ou seja, a

Comparação da distribuição do índice de **máxima refletividade** entre os subconjuntos com e sem descargas elétricas atmosféricas



**Figura 7. Máxima refletividade em registros com diferentes limiares de detecção de raios. Imagem disponível em: <https://bit.ly/2V1mAgY>**

quantidade de registros em que foram detectados raios é muito inferior que quantidade de registros onde não foram detectados. Isso pode provocar um certo ruído na distribuição dos valores. Com o propósito de suavizar este ruído, foram aplicados alguns limiares de detecções de raios, para que então, seja analisada a distribuição dos valores de máxima refletividade. Na Figura 7, pode ser observado que a medida que se aumenta o limite da quantidade de raios detectados, a variável máxima refletividade dos subconjuntos abaixo dos limiares ficam mais distantes dos subconjuntos acima definido pela árvore de decisão. Isso mostra que é possível quantificar faixas de ocorrências de descargas elétricas de modo mais discreto, o que favoreceria um sistema de previsão de raios baseado nestas variáveis, sendo de suma importância para a sociedade.

#### 4. Considerações finais

Este trabalho apresentou uma breve análise da distribuição espacial da ocorrência de descargas elétricas atmosféricas e dos principais atributos de dados de radar relacionados a este processo na região central da Bacia Amazônica. Notou-se que as variáveis utilizadas, que foram selecionadas a partir de critérios físicos, quando aplicadas técnicas de aprendizado de máquina, apresentam um comportamento que pode ser utilizado para classificar automaticamente faixas de ocorrência de descargas elétricas. Observou-se também a importância dos valores de refletividade, VIL e precipitação na separação dos dados com detecção de descargas elétricas em faixas mais discretas, ou seja, aumentando a ocorrência de raios, a discrepância entre os dados ficou cada vez mais evidente, o que mostra seu potencial para a aplicação em sistemas de previsão.

Como continuidade deste trabalho, pretende-se aproveitar a análise de atributos que foi realizada, e desenvolver um sistema de previsão a curto prazo de descargas elétricas utilizando algoritmos de aprendizado de máquina e redes neurais.



## 5. Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## Referências

- [Betz et al. 2009] Betz, H. D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W. P., Defer, E., Dziewit, Z., and Konarski, J. (2009). Linet—an international lightning detection network in europe. *Atmospheric Research*, 91(2-4):564–573.
- [Calheiros et al. 2016] Calheiros, A. J. P., Enoré, D. P., Mattos, E. V., da Costa, I. C., and Machado, L. A. T. (2016). *Sistema de Previsão Imediata: Descrição dos Produtos*. Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), Cachoeira Paulista-SP.
- [Cardoso et al. 2014] Cardoso, I., Pinto Jr, O., Pinto, I., and Holle, R. (2014). Lightning casualty demographics in brazil and their implications for safety rules. *Atmospheric Research*, 135:374–379.
- [Cardoso 2017] Cardoso, M. C. (2017). Determinação em tempo real dos riscos de desligamentos em linhas de transmissão devido a descargas atmosféricas. Dissertação (mestrado em engenharia elétrica), Pontifícia Universidade Católica - PUC, Pontifícia Universidade Católica, Rio de Janeiro-RJ.
- [Chinchay 2018] Chinchay, J. H. H. (2018). Uso de multi-canais do goes-16 para previsão imediata de densidade de descargas elétricas. Dissertação de mestrado, Instituto Nacional de Pesquisas Espaciais (INPE).
- [Damian 2011] Damian, E. A. (2011). Duas metodologias aplicadas à classificação de precipitação convectiva e estratiforme com radar meteorológico: Svm e k-means. Dissertação de mestrado, Universidade Federal do Paraná (UFPR), Curitiba-PR.
- [dos Santos Filho et al. 2005] dos Santos Filho, R. B., Gin, R. B., and Bianchi, R. A. (2005). Sistema de visão omnidirecional para o monitoramento de descargas atmosféricas.
- [Greene and Clark 1972] Greene, D. R. and Clark, R. A. (1972). Vertically integrated liquid water—a new analysis tool. *Monthly Weather Review*, 100(7):548–552.
- [Krider 2006] Krider, E. P. (2006). Benjamin franklin and lightning rods. *Physics Today*, 59(1):42.
- [Machado et al. 2014] Machado, L. A., Silva Dias, M. A., Morales, C., Fisch, G., Vila, D., Albrecht, R., Goodman, S. J., Calheiros, A. J., Biscaro, T., Kummerow, C., et al. (2014). The chuva project: How does convection vary across brazil? *Bulletin of the American Meteorological Society*, 95(9):1365–1380.
- [Machado et al. 2018] Machado, L. A. T., Calheiros, A. J. P., Biscaro, T., Giangrande, S., Silva Dias, M. A. F., Cecchini, M. A., Albrecht, R., Andreae, M. O., Araujo, W. F., Artaxo, P., Borrmann, S., Braga, R., Burleyson, C., Eichholz, C. W., Fan, J., Feng, Z., Fisch, G. F., Jensen, M. P., Martin, S. T., Pöschl, U., Pöhlker, C., Pöhlker, M. L., Ribaud, J.-f., Rosenfeld, D., Saraiva, J. M. B., Schumacher, C., Thalman, R., Walter, D., and Wendisch, M. (2018). Overview: Precipitation characteristics and sensitivities to environmental conditions during goamazon2014/5 and acridicon-chuva. *Atmospheric Chemistry And Physics*, 18:6461–6482.

- [Marengo et al. 1994] Marengo, J., Miller, J., Russell, G., Rosenzweig, C., and Abramopoulos, F. (1994). Calculations of river-runoff in the giss ggm: impact of a new land-surface parameterization and runoff routing model on the hydrology of the amazon river. *Climate Dynamics*, 10(6-7):349–361.
- [Marshall and Palmer 1948] Marshall, J. S. and Palmer, W. M. K. (1948). The distribution of raindrops with size. *Journal of meteorology*, 5(4):165–166.
- [Pereira 2019] Pereira, R. F. d. O. (2019). Propriedades físicas da precipitação e eletrificação dos sistemas precipitantes observados durante o projeto chuva-manaus e goamazon. Dissertação de mestrado, Universidade de São Paulo (USP).
- [Queiroz 2009] Queiroz, A. P. (2009). Monitoramento e previsão imediata de tempestades severas usando dados de radar. Dissertação (mestrado em meteorologia), Instituto Nacional de Pesquisas Espaciais - INPE, Instituto Nacional de Pesquisas Espaciais, São José dos Campos-SP.
- [Rakov 2007] Rakov, V. (2007). Lightning phenomenology and parameters important for lightning protection. *IX SIPDA, 26th-30th November*.
- [Romero 2007] Romero, F. (2007). *Avaliação do comportamento dos campos eletromagnéticos gerados por descargas atmosféricas nuvem-terra*. PhD thesis, Universidade de São Paulo.
- [Safavian and Landgrebe 1991] Safavian, S. R. and Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674.
- [Saraiva et al. 2016] Saraiva, I., Silva Dias, M., Morales, C., and Saraiva, J. (2016). Regional variability of rain clouds in the amazon basin as seen by a network of weather radars. *Journal of Applied Meteorology and Climatology*, 55(12):2657–2675.
- [Souza 2017] Souza, M. G. A. d. (2017). Uma análise teórica dos parâmetros típicos de descargas atmosféricas.
- [Steiner et al. 1995] Steiner, M., Houze, R. A., and Yuter, S. E. (1995). Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *Journal of Applied Meteorology*, 34(9):1978–2007.
- [Vila et al. 2008] Vila, D. A., Machado, L. A. T., Laurent, H., and Velasco, I. (2008). Forecast and tracking the evolution of cloud clusters (fortracc) using satellite infrared imagery: Methodology and validation. *Weather and Forecasting*, 23(2):233–245.
- [Weber et al. 2005] Weber, M., Evans, J., Wolfson, M., DeLaura, R., Moser, B., Martin, B., Welch, J., Andrews, J., and Bertsimas, D. (2005). Improving air traffic management during thunderstorms. In *24th digital avionics systems conference*, volume 1, pages 3–D. IEEE.