



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/04.29.03.12-TDI

**USO DE REDES NEURAIIS PARA PREVISÃO DE  
DESCARGAS ELÉTRICAS NUVEM-SOLO A PARTIR  
DE DADOS DE RADAR: UM ESTUDO DE CASO NA  
AMAZÔNIA**

Adriano Pereira Almeida

Dissertação de Mestrado do Curso  
de Pós-Graduação em Computação  
Aplicada, orientada pelo Dr. Alan  
James Peixoto Calheiros, aprovada  
em 28 de maio de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44JMNDL>>

INPE  
São José dos Campos  
2021

**PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE  
Coordenação de Ensino, Pesquisa e Extensão (COEPE)  
Divisão de Biblioteca (DIBIB)  
CEP 12.227-010  
São José dos Campos - SP - Brasil  
Tel.:(012) 3208-6923/7348  
E-mail: pubtc@inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):**

**Presidente:**

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra (CGCT)

**Membros:**

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)  
Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e Ciência Espaciais (CGCE)  
Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e Pesquisas Aplicadas (CGIP)  
Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

**BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon  
Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

**REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)  
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

**EDITORAÇÃO ELETRÔNICA:**

Ivone Martins - Divisão de Biblioteca (DIBIB)  
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2021/04.29.03.12-TDI

**USO DE REDES NEURAIIS PARA PREVISÃO DE  
DESCARGAS ELÉTRICAS NUVEM-SOLO A PARTIR  
DE DADOS DE RADAR: UM ESTUDO DE CASO NA  
AMAZÔNIA**

Adriano Pereira Almeida

Dissertação de Mestrado do Curso  
de Pós-Graduação em Computação  
Aplicada, orientada pelo Dr. Alan  
James Peixoto Calheiros, aprovada  
em 28 de maio de 2021.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/44JMNDL>>

INPE  
São José dos Campos  
2021

Dados Internacionais de Catalogação na Publicação (CIP)

---

Almeida, Adriano Pereira.

Al64u      Uso de redes neurais para previsão de descargas elétricas  
nuvem-solo a partir de dados de radar: um estudo de caso na  
Amazônia / Adriano Pereira Almeida. – São José dos Campos :  
INPE, 2021.

xxiv + 183 p. ; (sid.inpe.br/mtc-m21c/2021/04.29.03.12-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto  
Nacional de Pesquisas Espaciais, São José dos Campos, 2021.

Orientador : Dr. Alan James Peixoto Calheiros.

1. Tempestades severas. 2. Raios. 3. RADAR. 4. Análise  
exploratória de dados. 5. Inteligência artificial. I.Título.

CDU 004.032.26:551.594.22

---



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



## INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

### DEFESA FINAL DE DISSERTAÇÃO DE ADRIANO PEREIRA ALMEIDA BANCA Nº 151/2021, REG 35022/2019

No dia 28 de maio de 2021, às 08h30min, por Vídeo Conferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Mestre em Computação Aplicada. O trabalho precisa da incorporação das correções sugeridas pela Banca Examinadora e revisão final pelo(s) orientador(es).

**Novo título: “Uso de Redes Neurais para Previsão de Descargas Elétricas Nuvem-Solo a partir de Dados de Radar: Um Estudo de Caso na Amazônia”**

Eu, Rafael Duarte Coelho dos Santos, Presidente da Banca Examinadora, assino esta ATA, em nome de todos os membros, com o consentimento dos mesmos.

#### Membros da Banca

Dr. Rafael Duarte Coelho dos Santos - Presidente - INPE  
Dr. Alan James Peixoto Calheiros - Orientador - INPE  
Dr. Elcio Hideiti Shiguemori - Membro Interno - IEAv  
Dr. Carlos Augusto Morales Rodriguez - Membro Externo - USP



Documento assinado eletronicamente por **Rafael Duarte Coelho dos Santos, Tecnologista**, em 31/05/2021, às 13:57 (horário oficial de Brasília), com fundamento no art. 6º do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site <http://sei.mctic.gov.br/verifica.html>, informando o código verificador **7500496** e o código CRC **377ED6BD**.



*“Existe uma coisa que uma longa existência me ensinou: toda a nossa ciência, comparada à realidade, é primitiva e inocente; e, portanto, é o que temos de mais valioso”.*

A. EINSTEIN





*À minha mãe Ana Maria Pereira de Carvalho.*



## AGRADECIMENTOS

Agradeço profundamente à minha família por todo o amor e carinho que sempre manifestaram por mim e pelos sábios conselhos e apoio que me dão à cada decisão importante da minha vida. Em especial à minha amada mãe, que é a maior fonte de inspiração para mim, me mostrando de diversas formas o verdadeiro significado do amor.

Ao meu orientador Dr. Alan James Peixoto Calheiros, pela dedicada orientação e todos os ensinamentos que transcendem o escopo desta dissertação. Agradeço também pela confiança depositada em mim e paciência ao longo dessa jornada.

Aos docentes do programa de pós graduação da CAP do INPE, por compartilharem um pouco de seus conhecimentos e experiências. Em especial ao Dr. Rafael Santos, pela dedicação e comprometimento à frente da coordenação do programa.

Aos membros da banca examinadora pela disposição em analisar esta dissertação.

Aos meus amigos de longa data Maiely, Kássio, Pedro e Ruth, pelo incentivo e momentos divertidos, mesmo à distância.

Aos amigos que tive o privilégio de conhecer no INPE, pelo convívio, momentos de estudo, descontração e, principalmente, pela troca de experiências. Em especial, Amita, Guilherme, Felipe, Arturo, Helvécio, Isela e David.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Instituto Nacional de Pesquisas Espaciais (INPE) pelo financiamento econômico e oferecerem condições para o desenvolvimento deste trabalho.

Finalmente, agradeço a Deus e a todos aqueles que direto ou indiretamente contribuíram, de alguma maneira, na realização deste trabalho.



## RESUMO

A região central da Bacia Amazônica é caracterizada por ter uma vegetação densa e condições atmosféricas favoráveis para o desenvolvimento de tempestades severas e com alta incidência de raios. Diante disso, foi desenvolvido neste trabalho um estudo de algumas parametrizações através de técnicas de aprendizado de máquina para prover subsídios para o desenvolvimento de sistemas de previsão de descargas elétricas a curto prazo, levando em consideração as condições locais dessa região. Essas parametrizações partiram de conclusões obtidas na análise exploratória dos dados de radar meteorológico e da seleção de variáveis por meio de classificações com a árvore de decisão. Além disso, foram criados modelos baseados na rede neural *perceptron* multicamadas (MLP) com diversas configurações de entrada e variações dos seus principais hiper-parâmetros, que tiveram como intuito avaliar a sensibilidade das variáveis em diferentes intervalos de tempos anteriores à previsão. As previsões foram feitas para 12 minutos à frente, utilizando os valores das variáveis em instantes anteriores à previsão:  $t-1$  (12 minutos);  $t-2$  (24 minutos); e  $t-3$  (36 minutos). Embora a maioria das variáveis meteorológicas utilizadas possuam dependências temporais mais restritas, os modelos que utilizaram os três instantes de tempo, tiveram melhor desempenho. Junto com a degradação dos dados de entrada e informações sobre as características dos perfis de refletividade radar foi possível definir critérios que podem ajudar na melhorias das previsões de raios via redes neurais.

Palavras-chave: Tempestades severas. Raios. RADAR. Análise exploratória de dados. Inteligência artificial.



# USE OF NEURAL NETWORKS TO PREDICT CLOUD-TO-GROUND LIGHTNING FROM RADAR DATA: A CASE STUDY IN THE AMAZON

## ABSTRACT

Lightning associated with severe storms can cause several socio-economic problems. Due to its high convective activity, predominantly equatorial location, and large territorial extension, Brazil is one of the countries where the highest incidence of lightning strikes occurs globally. Studies show that the most significant cause of human fatalities related to storms in Brazil is associated with lightning and that a large part of them appear in rural areas. The material damages caused by these events are also very recurrent. Therefore, the interest in creating forecasting systems for this natural event is of paramount importance. However, several challenges are associated with the creation of lightning forecasting systems. Sometimes these challenges are related to the characteristics of the event itself, such as its high variability and its relationship with complex physical processes that are difficult to model computationally. The local factors that can influence the development of electrically active storms should also be considered. The central region of the Amazon Basin is characterized by dense vegetation and favorable atmospheric conditions for the development of severe storms and a high incidence of lightning. Therefore, a study of some parameterizations through machine learning techniques was developed in this work to provide subsidies for the development of lightning forecasting lightning in the short-term systems, considering the local conditions of the central region of the Amazon Basin. These parameterizations started from conclusions obtained in the exploratory analysis of meteorological radar data and the selection of variables through classifications with the decision tree. In addition, models based on the neural network multilayer layer perceptron (MLP) with different input configurations were created to evaluate the sensitivity of the variables at different time intervals before the forecast. The forecasts were made for 12 minutes ahead, using the previous times values of the variables:  $t-1$  (12 minutes);  $t-2$  (24 minutes); and  $t-3$  (36 minutes). Although most meteorological variables used have more restricted temporal dependencies, the models that used three previous time instants performed better. Along with the degradation of the input data and information on the characteristics of the radar reflectivity profiles, it was possible to define criteria that can help to improve the lightning forecasts via neural networks.

Keywords: Severe storms. Lightnings. RADAR. Exploratory data analysis. Artificial intelligence.





## LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Esquema do ciclo de vida da célula de tempestade. . . . .	6
2.2 Ramificação das principais redes neurais artificiais. . . . .	12
3.1 Classes de coberturas do solo e orografia da área de estudo. . . . .	24
3.2 Área delimitada para extração das variáveis. . . . .	26
3.3 Taxa de chuva horária acumulada nas estações chuvosa (a) e seca (b). Os valores de taxa de chuva computados foram aqueles superiores a 0,5 mm/h e inferiores a 160 mm/h (machado2018overview). . . . .	27
3.4 Etapas para a conversão de <i>strokes</i> em <i>flashes</i> . . . . .	31
3.5 Distribuição da multiplicidade de <i>strokes</i> por <i>flash</i> nas estações (a) chu- vosa e (b) secas. . . . .	31
3.6 Esquema da R-Tree. . . . .	32
3.7 Esquema do processo de degradação da matriz. . . . .	33
3.8 Esquema do funcionamento da árvore de decisão. . . . .	39
3.9 Variação da entropia e índice <i>Gini</i> em relação a probabilidade de um elemento pertencer à uma classe do total de duas. . . . .	40
3.10 Esquema do processo de previsão das descargas elétricas. . . . .	42
3.11 Esquema da arquitetura da rede <i>perceptron</i> multicamadas. . . . .	43
3.12 Esquema do neurônio artificial. . . . .	44
3.13 Gráfico das principais funções de ativações. . . . .	44
3.14 Esquema do <i>k-fold cross validation</i> . . . . .	48
3.15 Esquema da tabela de contingência. . . . .	48
4.1 Mapas de densidades dos <i>flashes</i> nas estações chuvosa, seca e suas tran- sições. . . . .	51
4.2 Densidade dos <i>strokes</i> por estações nos municípios da área de estudo. . . . .	52
4.3 Estatísticas dos <i>strokes</i> por tipo de solo durante todas as estações. . . . .	55
4.4 Estatísticas dos <i>strokes</i> por orografia durante todas as estações. . . . .	55
4.5 Média e acumulado horário de <i>flashes</i> em todas as estações. . . . .	57
4.6 Histogramas das frequências relativas das refletividades do produto CAPPI em 2,5 km com diferentes incidências de <i>flashes</i> em todas esta- ções. . . . .	59
4.7 Histogramas das frequências relativas das refletividades do produto CAPPI em 5 km com diferentes incidências de <i>flashes</i> em todas esta- ções. . . . .	60

4.8	Histogramas das frequências relativas das refletividades do produto CAPPI em 7 km com diferentes incidências de <i>flashes</i> em todas estações. . . . .	61
4.9	Histogramas das frequências relativas das Máximas refletividade (2-7km) com diferentes incidências de <i>flashes</i> em todas estações. . . . .	62
4.10	Histogramas das frequências relativas das alturas de Waldvogel (em metros) com diferentes incidências de <i>flashes</i> em todas estações. . . . .	63
4.11	Histogramas das frequências relativas das alturas do 35 dBZ ( <i>echo top</i> em metros) com diferentes incidências de <i>flashes</i> em todas estações. . . . .	64
4.12	Boxplots dos valores da taxa de precipitação com diferentes incidências de <i>flashes</i> em todas estações. . . . .	65
4.13	Boxplots dos valores do VIL com diferentes incidências de <i>flashes</i> em todas estações. . . . .	66
4.14	Boxplots dos valores da Densidade do VIL com diferentes incidências de <i>flashes</i> em todas estações. . . . .	67
4.15	Relações assíncronas das variáveis em diferentes tempos anteriores à observação durante as estações chuvosa e seca para atividade e inatividade de raios. $t$ (tempo atual), $t-1$ (12 minutos), $t-2$ (24 minutos) e $t-3$ (36 minutos). . . . .	69
4.16	Matriz de refletividade degradada com diferentes <i>kernels</i> para o dia 11 de março de 2014 às 16:35. . . . .	71
4.17	Matriz de <i>strokes</i> degradada com diferentes <i>kernels</i> e a soma como função de agregação para o dia 11 de março de 2014 às 16:35. . . . .	72
4.18	Índice de co-localização por variável em todas as estações com degradações em diferentes tamanhos de <i>kernels</i> . . . . .	74
4.19	Média dos Índices de co-localização por estação para todas variáveis com degradações em diferentes tamanhos de <i>kernels</i> . . . . .	76
4.20	Índice de correlação por variável em todas as estações com degradações em diferentes tamanhos de <i>kernels</i> . . . . .	77
4.21	Grau de importância atribuído às variáveis pela árvore de decisão em $t-1$ , $t-2$ e $t-3$ . . . . .	80
4.22	Árvore de decisão gerada pela classificação da atividade de raios para o tempo $t$ . . . . .	81
4.23	Árvore de decisão gerada pela classificação da atividade de raios para o tempo $t-1$ . . . . .	82
4.24	Árvore de decisão gerada pela classificação da atividade de raios para o tempo $t-2$ . . . . .	83

4.25	Árvore de decisão gerada pela classificação da atividade de raios para o tempo $t-3$ . . . . .	84
4.26	Esquema da preparação dos dados de entrada da RNA. Em que $V_i$ são as variáveis utilizadas em cada instante de tempo $T_j$ . . . . .	89
4.27	Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade $>50\%$ . . . . .	94
4.28	<i>Flashes</i> observados e previstos com as diferentes configurações de RNA. . . . .	97
4.29	Mosaico de previsão com as configurações criadas a partir da inclusão dos dados de cobertura do solo e hora. . . . .	107
4.30	<i>Flashes</i> observados e previstos com as configurações MLP05 e MLP09 com dados de cobertura do solo e hora. . . . .	108
4.31	Mosaico de previsão com as configurações criadas a partir da entrada de MLP05 com os dados degradados. . . . .	110
B.1	Histogramas da estação chuvosa. . . . .	162
B.2	Histogramas da transição entre a estação chuvosa e seca. . . . .	163
B.3	Histogramas da estação seca. . . . .	164
B.4	Histogramas da transição entre a estação seca e chuvosa. . . . .	165
C.1	Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade $>70\%$ . . . . .	180
C.2	Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade $>80\%$ . . . . .	181
C.3	Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade $>90\%$ . . . . .	182
C.4	Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade $>98\%$ . . . . .	183



## LISTA DE TABELAS

	<u>Pág.</u>
3.1 Áreas da Bacia Amazônica por países. . . . .	23
3.2 Variáveis calculadas a partir dos dados do radar meteorológico. . . . .	28
3.3 Hiper-parâmetros definidos para as configurações das MLPs. . . . .	45
3.4 Hiper-parâmetros definidos para as configurações das MLPs. . . . .	47
4.1 <i>Strokes</i> e <i>flashes</i> nos municípios durante as estações. . . . .	53
4.2 Índices de co-localização e correlação para diferentes <i>kernels</i> para o dia 11 de março de 2014 às 16:35. . . . .	73
4.3 <i>Skills</i> dos testes realizados com as diferentes configurações da RNA ba- seadas na EDA. . . . .	96
4.4 As 3 combinações com as melhores <i>skills</i> segundo teste de sensibilidade das variáveis em três instantes de tempos. . . . .	104
4.5 <i>Skills</i> dos testes com a inserção de dados de cobertura do solo e hora. . .	105
4.6 Comparação entre as <i>Skills</i> dos testes realizados com os modelos utili- zando todas as variáveis sem e com as degradações nos diferentes <i>kernels</i> do caso de estudo. . . . .	109
A.1 Estatísticas síncronas das variáveis e teste de hipótese. . . . .	142
B.1 Estatísticas dos os índices de co-localização das variáveis em diferentes com degradações em diferentes <i>kernels</i> . . . . .	166



## LISTA DE ABREVIATURAS E SIGLAS

ABI	– do inglês Advanced Baseline Imager
CAPPI	– do inglês Constant Altitude Plan Position Indicator
CG	– do inglês Cloud-Ground
DT	– do inglês Decision Tree
DVIL	– Vertically Integrated Liquid Density
ED	– Erro da densidade
FAR	– Falso Alarme
ForTraCC	– Forecasting and Tracking the Evolution of Cloud Clusters
GLM	– do inglês Geostationary Lightning Mapper
GOES	– do inglês Geostationary Operational Environmental Satellite
HFM	– Hopfield Model
IA	– Inteligência Artificial
IC	– do inglês Intra-Cloud
ICOL	– Índice de co-localização
ICOR	– Índice de correlação
INPE	– Instituto Nacional de Pesquisas Espaciais
IOP	– do inglês Intensive Operation Period
LF	– do inglês Low Frequency
LINET	– Lightning NETWORK
ML	– do inglês Machine Learning
MLP	– Multilayer Perceptron
NWP	– do inglês Numerical Weather Prediction
POD	– Probabilidade de Detecção
R	– Taxa de Precipitação
RBFN	– do inglês Radial Basis Function Network
RF	– do inglês Random Forests
RNA	– Redes Neurais Artificiais
RNN	– do inglês Recurrent Neural Network
SCM	– Sistemas Convectivos de Mesoescala
SIPAM	– Sistema de Proteção da Amazônia
SRRF	– Spatiotemporal Relational Random Forest
STARNET	– Sferics Timing Ranging Network
SVM	– Support Vector Machine
TITAN	– Thunderstorm Identification, Tracking, Analysis, and Nowcasting
VDRAS	– do inglês The Variational Doppler Radar Analysis System
VIL	– do inglês Vertically Integrated Liquid
VLF	– do inglês Very Low Frequency
Z	– Refletividade do radar
dBZ	– Decibels de Z





## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
1.1 Objetivos . . . . .	4
<b>2 REFERENCIAL TEÓRICO</b> . . . . .	<b>5</b>
2.1 Visão geral sobre tempestades . . . . .	5
2.1.1 Processo de eletrificação das nuvens . . . . .	7
2.1.2 Descargas elétricas atmosféricas . . . . .	8
2.2 Sistemas para previsão e monitoramento de tempestades . . . . .	10
2.2.1 Sistemas tradicionais . . . . .	10
2.2.2 Sistemas baseados em inteligência artificial . . . . .	11
2.2.2.1 Redes neurais artificiais . . . . .	12
2.2.2.2 Árvore de decisão e similares . . . . .	14
2.2.2.3 Outras técnicas . . . . .	15
2.3 Principais variáveis meteorológicas para previsão . . . . .	16
2.4 Influência da superfície na formação de tempestades . . . . .	21
<b>3 MATERIAIS E MÉTODOS</b> . . . . .	<b>23</b>
3.1 Área de estudo . . . . .	23
3.2 Dados . . . . .	25
3.2.1 Dados de radar . . . . .	25
3.2.2 Dados dos sensores de VLF . . . . .	29
3.2.2.1 Agrupamento dos <i>strokes</i> em <i>flashes</i> . . . . .	30
3.2.2.2 Indexação espacial para o mapeamento do raios . . . . .	32
3.2.3 Mapeamento baseado na degradação dos dados . . . . .	33
3.2.4 Índices de co-localização e correlação . . . . .	34
3.2.5 Classes de cobertura de solo e orografia . . . . .	35
3.3 Metodologia . . . . .	35
3.3.1 Análise espacial das descargas elétricas . . . . .	36
3.3.2 Classificação dicotômica dos raios . . . . .	37
3.3.3 Seleção das variáveis meteorológicas . . . . .	37
3.3.4 Método de classificação . . . . .	38
3.3.4.1 Árvores de decisão . . . . .	38
3.3.5 Previsão de raios a curto prazo . . . . .	41

3.3.5.1	<i>Perceptron</i> multicamadas . . . . .	43
3.3.6	Tecnologias utilizadas . . . . .	46
3.3.7	Métricas de avaliação dos modelos . . . . .	47
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>51</b>
4.1	Análises da distribuição espacial e sazonal dos raios . . . . .	51
4.1.1	Variação espacial dos raios em relação à cobertura de solo . . . . .	54
4.1.2	Análise do ciclo diurno . . . . .	56
4.2	Análise exploratória de dados . . . . .	58
4.2.1	Análises das relações síncronas . . . . .	58
4.2.2	Análises das relações assíncronas . . . . .	67
4.3	Análise dos dados degradados . . . . .	70
4.4	Classificação da atividade de raios por Árvore de Decisão . . . . .	79
4.5	Previsão de raios com <i>perceptron</i> multicamadas . . . . .	88
4.5.0.1	Testes de sensibilidade para os dados de entrada . . . . .	89
<b>5</b>	<b>CONCLUSÃO . . . . .</b>	<b>113</b>
5.1	Trabalhos futuros . . . . .	117
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>121</b>
	<b>APÊNDICE A - ESTATÍSTICAS SÍNCRONAS DAS VARIÁVEIS</b>	<b>141</b>
	<b>APÊNDICE B - ESTATÍSTICAS ÍNDICE DE CO- LOCALIZAÇÃO . . . . .</b>	<b>161</b>
	<b>APÊNDICE C - MOSAICOS COM AS PREVISÕES DE RAIOS DIFERENTES LIMIARES DE PROBABILIDADE . . . . .</b>	<b>179</b>

# 1 INTRODUÇÃO

As tempestades severas são responsáveis por causarem diversos problemas socioeconômicos. Esse tipo de fenômeno natural tem as descargas elétricas atmosféricas como uma de suas principais características. O Brasil é um dos países onde são registradas as maiores incidências de raios durante o ano, 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). A grande incidência de raios no país tem relação sobretudo com sua grande extensão territorial e alta atividade convectiva, devido sua localização predominantemente em região tropical, tornando-o propício para ser um dos principais pontos onde acontecem esses eventos (ALBRECHT et al., 2016).

Os raios podem causar grandes riscos e prejuízos à sociedade. As fatalidades humanas em decorrência dos raios são uma das principais preocupações quando se trata das condições de severidade. Com base nos dados registrados durante o período de 2000 à 2009 pela Secretaria Nacional de Defesa Civil e Ministério da Saúde no Brasil, Cardoso et al. (2014) realizaram um estudo acerca das fatalidades causadas por raios no país, identificando uma média de 132 mortes por ano. A maioria das fatalidades humanas causadas por raios estão relacionadas às atividades rurais, onde as pessoas ficam mais expostas às tempestades (CARDOSO et al., 2014).

Além do risco direto iminente à vida humana, os raios também são responsáveis por uma outra série de problemáticas. Alguns setores podem sofrer grandes prejuízos em consequência da ocorrência de descargas elétricas, dentre eles o da aviação, serviços de transmissão de energia e comunicação e agropecuária (CARDOSO, 2017).

As tempestades severas com alta atividade elétrica no Brasil podem ser monitoradas principalmente por meio das redes de detecção de raios, radares e satélites meteorológicos. Os instrumentos de observação utilizados para o monitoramento de raios, complementados por dados de radares, possuem alta resolução espacial e temporal, podendo auxiliar em estudos de aspectos importantes relacionados aos processos físicos das tempestades elétricas (HUAMÁN, 2018). No entanto, embora alguns desses instrumentos de superfície cubram todo o território brasileiro (e.g. STARNET), ainda há limitações de cobertura contínua em algumas regiões do Brasil (HUAMÁN, 2018). Essas limitações estão relacionadas principalmente com a falta de cobertura dos radares meteorológicos.

---

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

Os satélites também são instrumentos que contribuem com grande parte da demanda de dados meteorológicos no Brasil. A nova geração de satélites *Geostationary Operational Environmental Satellite* (GOES-16), lançado pela *National Oceanic and Atmospheric Administration* (NOAA) em 2016, possibilitou avanços significativos com o fornecimento de dados para todo o Hemisfério Ocidental com resoluções temporais e espaciais maiores que seus antecessores (NOAA, 2016). Os principais instrumentos a bordo deste satélite são o *Advanced Baseline Imager* (ABI), que fornece imagens em 16 bandas espectrais da superfície e atmosfera com resolução espacial de até 500 metros e o *Geostationary Lightning Mapper* (GLM), detector óptico que opera no infravermelho-próximo detectando eventos de alta velocidade e natureza transiente na atmosfera que indicam a presença de descargas elétricas (SCHMIT et al., 2005; GOODMAN et al., 2012; GOODMAN et al., 2013).

Os raios são eventos de natureza aleatória (LOBODA, 2009; RAKOV, 2009), e que diante de uma tempestade não podem ser ignorados, restando a árdua tarefa de mitigar o máximo possível os danos que eles podem causar. Os sistemas para previsão de tempestades e raios podem desempenhar um papel importante nesse cenário, ajudando as autoridades competentes e defesa civil na tomada de decisões estratégicas e alguns setores para o planejamento de suas atividades. O Brasil possui características e padrões meteorológicos bastante variados, sendo assim, sistemas para previsão desenvolvidos para a região podem não funcionar muito bem para outras (SILVA et al., 2019). A antecipação de informações confiáveis a respeito da evolução de eventos que podem causar danos, como as tempestades severas, é de suma importância para a fundamentação de tomadas de decisão das entidades competentes e alertar a população sobre os possíveis riscos. A antecipação dessas informações a curtíssimo prazo é denominada *nowcasting*, escala temporal utilizada para previsões meteorológicas em até 6 horas (BROWNING, 1980; CALHEIROS et al., 2016).

Durante os anos de 2014 e 2015 aconteceram os experimentos da campanha GoAmazon na região central da Bacia Amazônica (MACHADO et al., 2018). Durante os experimentos foram utilizados diversos instrumentos com o intuito de estudar as características meteorológicas do ponto de vista natural e sob a influência humana na região nos períodos chuvosos, seco e de transição (MARTIN et al., 2015). Os dados obtidos pela campanha GoAmazon serviram como base para diversos estudos para um melhor entendimento dos regimes precipitantes e sua influência na formação e intensificação das atividades elétricas (PEREIRA, 2019).

O uso de técnicas de inteligência artificial na previsão de tempestades severas tem mostrado resultados promissores. Além das tarefas de previsões de tempestades, as abordagens baseadas em inteligência artificial são amplamente utilizadas em diversas outras aplicações associadas a eventos meteorológicos: previsão de chuva (ALLEN; MARSHALL, 1994; DESHPANDE, 2012; ESTEVES et al., 2019; VELASCO et al., 2019), previsão de descargas elétricas (FRANKEL; DRAPER, 1992; MINIERE et al., 1996; EMAMGHOREISHI et al., 2001; ABDULLAH et al., 2018; RAMZI et al., 2018), previsão de tempo severo (MARZBAN; WITT, 2001; CHAUDHURI, 2010; SILVA et al., 2017; CHAKRABARTY et al., 2018), dentre outras. A partir de dados obtidos através de instrumentos de observação, os algoritmos de aprendizado de máquina, que entre eles se destacam as redes neurais artificiais, podem encontrar padrões, que com a interpretação dos especialistas, podem prover novas informações que são importantes para a melhoria de sistemas de previsão. Que por sua vez podem ser tão robustas quanto os métodos tradicionais de previsão baseado em modelos matemáticos e físicos (COLLINS; TISSOT, 2016).

Redes neurais são técnicas muitas vezes consideradas "caixas pretas". Contudo, em ciência exatas esta é uma abordagem que apresenta uma certo grau de resistência pela comunidade, que visa entender os processos físicos e assim modelá-los. Neste estudo, a ideia de realizar testes de sensibilidade é justamente para entender os impactos de diferentes configurações nos resultados da rede neurais escolhida. Logo, espera-se que os modelos aqui criados não sejam os melhores a serem implementados com o propósito de realizar previsões de raios satisfatórias. Este estudo visa apenas definir os desafios e estratégias para as próximas gerações de modelos de previsão de raios usando dados de radares meteorológicos.

Diante do exposto, este trabalho visou caracterizar a distribuição dos raios nuvem-solo na região da Bacia Amazônica e apresentar subsídios para o desenvolvimento de sistemas de previsão a curto prazo utilizando técnicas de aprendizado de máquina. Visto que existem diversos aspectos que devem ser considerados do ponto de vista de ciências de dados e meteorologia, que são um desafio para a previsão de raios do tipo nuvem-solo. Uma vez que esses sistemas não são tão frequentes quanto os outros tipos e ainda precisam de maior aprofundamento sobre sua gênese. Ainda, os dados utilizados neste estudo como preditores, ou seja, os dados de radar, são frequentemente afetados por problemas sistemáticos e má representatividade de algumas de suas medidas. A junção de ambos os dados, sensores de raios e radar, apresentam certas limitações que precisam ser contabilizadas e ajustadas para sua melhor co-localização. Em suma, não é uma tarefa trivial. Contudo, este trabalho

focou na melhor descrição dos aspectos técnicos e científicos necessários para o uso de aprendizado de máquina na previsão de raios nuvem-solo, considerando suas limitações.

## 1.1 Objetivos

O *nowcasting* tem como característica principal levar em consideração as condições ambientais locais e assimilar a maior quantidade de informações no menor tempo possível. Neste sentido, para prever condições de ocorrência ou cessação de raio (neste estudo o tipo nuvem-solo) sobre a região central da Bacia Amazônica, o uso de inteligência artificial se mostra uma ferramenta promissora. Principalmente, devido a sua capacidade de generalização e rapidez de decisão. Logo, o objetivo geral deste trabalho foi analisar os impactos de diferentes abordagens na previsão de raios nuvem-solo via redes neurais artificiais. Para tanto foi necessário antes entender as características dos padrões espaço-temporal dos raios sobre a região de estudo e sua relação com as possíveis variáveis preditoras. Assim, elucidando questões importantes no desenvolvimento de um sistema de previsão de descargas elétricas. Com base no objetivo geral, os principais objetivos específicos foram:

- Realizar um estudo aprofundado da literatura a respeito das principais metodologias empregadas no desenvolvimentos de sistemas de previsão de tempestades a partir de algoritmos de inteligência artificial;
- Identificar na literatura as principais técnicas para seleção de variáveis meteorológicas para previsão de raios;
- Explorar a variabilidade dos raios com relação a aspectos intrínsecos da região, como características de superfície, ciclo diurno e estações;
- Analisar o impacto das relações síncronas e assíncronas entre as variáveis meteorológicas e a atividade/inatividade de raios;
- Avaliar o impacto de diferentes configurações de entrada e hiperparâmetros dos modelos de aprendizado de máquina para a previsão da atividade de raios a curto prazo.

## 2 REFERENCIAL TEÓRICO

Ao decorrer deste capítulo serão apresentados alguns dos conceitos fundamentais para o pleno entendimento deste trabalho. Na Seção 2.1 será apresentada uma visão geral acerca das tempestades, bem como o processo de eletrificação de suas células e condições favoráveis para a ocorrência de descargas elétricas atmosféricas. Na Seção 2.2 serão apresentadas as principais abordagens utilizadas no desenvolvimento de sistemas de previsão de fenômenos meteorológicos, com ênfase na previsão de tempestades severas e descargas elétricas em escala de *nowcasting*. Na Seção 2.3 serão expostas as principais variáveis meteorológicas utilizadas para a previsão de tempestades severas. E por fim, na Seção 2.4 será apresentada a influência da superfície na formação de tempestades severas.

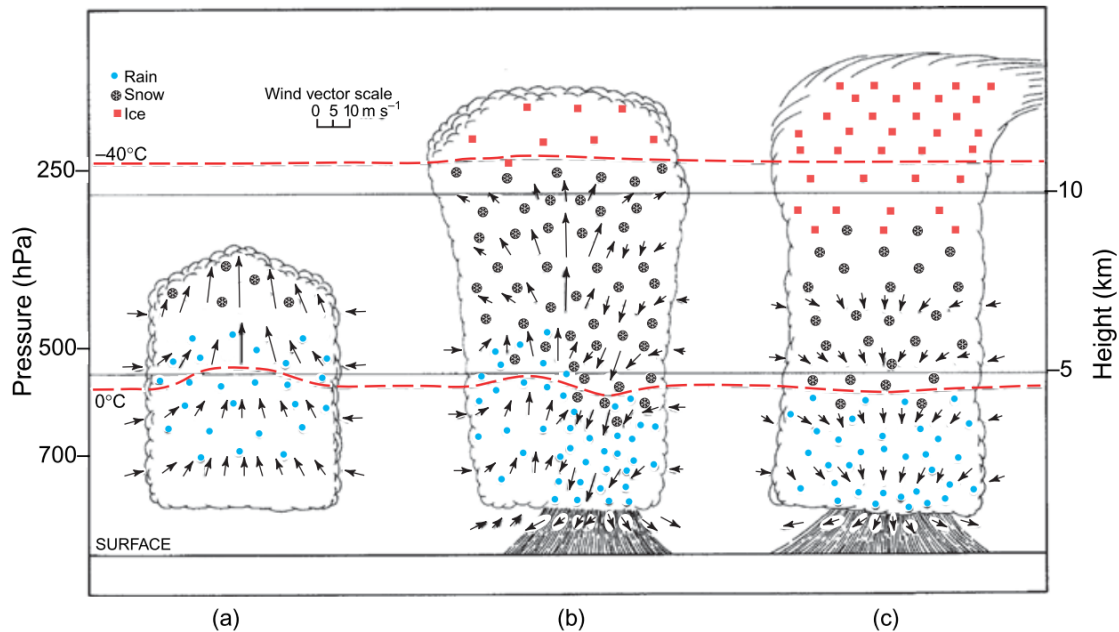
### 2.1 Visão geral sobre tempestades

De acordo com o glossário da Sociedade Americana de Meteorologia, as tempestades são caracterizadas por quaisquer perturbações ou instabilidade na atmosfera que podem acarretar na ocorrência de raios, granizos e fortes chuvas e ventos (AMS, 2012). As tempestades severas, também conhecidas como células convectivas, são caracterizadas por possuírem um ciclo de vida curto, alcance pequeno e forte poder destrutivo (WANG et al., 2020).

Segundo Ravi et al. (1999), todo o ciclo de vida de uma tempestade isolada dura em média entre 30 minutos e uma hora, salvo em casos excepcionais. Geralmente o raio de ação dos radares meteorológicos possuem cobertura suficiente para identificar e monitorar todo o ciclo de vida das células de tempestades (WANG et al., 2020), pois sua extensão horizontal pode variar entre 8 e 16 km quando encontra-se plenamente desenvolvida (YU et al., 2006 apud WANG et al., 2020). O poder destrutivo das tempestades pode causar diversos problemas sócio-econômicos oriundo dos seguintes eventos: enchentes (proveniente das fortes chuvas), fortes ventos, granizo e descargas elétricas atmosféricas (MOLLER et al., 1994).

As tempestades convectivas são desenvolvidas principalmente em nuvens do tipo *cumulonimbus*, a qual possui grande extensão vertical que se desenvolve a partir das fortes correntes de ar ascendentes, componente vertical de transferência de calor por convecção e outros processos termodinâmicos (COLLINS; TISSOT, 2016). Geralmente, as células de tempestades possuem um ciclo de vida bem definido (BYERS; BRAHAM, 1948; WALLACE; HOBBS, 2006; WANG et al., 2020).

Figura 2.1 - Esquema do ciclo de vida da célula de tempestade.



Esquema dos estágios durante o ciclo de vida de uma tempestade, onde (a) estágio *cumulus*, (b) estágio maduro e (c) estágio de dissipação. A escala horizontal da figura está reduzida em cerca de 30% em relação à escala vertical.

Fonte: Wallace e Hobbs (2006).

Segundo Byers e Braham (1948), o ciclo de vida das tempestades é dividido em três estágios: *cumulus*, maduro e dissipação, conforme mostrado na Figura 2.1. O estágio *cumulus* (Figura 2.1a) marca o início do desenvolvimento das células de tempestades. Esse estágio é caracterizado principalmente pelas correntes ascendentes, que são oriundas do ar quente e úmido vindo de níveis inferiores, fortalecendo assim a sua intensificação e crescimento vertical (BYERS; BRAHAM, 1948). O estágio maduro (Figura 2.1b) é marcado pelo início da precipitação e geração do fluxo de ar descendente que acontece em paralelo com as fortes correntes de ar ascendentes (WANG et al., 2020). No estágio de dissipação (Figura 2.1c) o topo da nuvem se aproxima da tropopausa e começa a se espalhar horizontalmente (WALLACE; HOBBS, 2006). De acordo com Wallace e Hobbs (2006), à medida que a precipitação se desenvolve nessa última fase, o fluxo de ar descendente gradativamente se torna mais intenso que as correntes ascendentes, impedindo que as gotículas de nuvem continuem crescendo por consequência da falta da corrente de ar supersaturada que não é mais gerada devido ao processo de dissipação que já contempla toda a nuvem.



Na subseção a seguir serão apresentados alguns dos processos que levam a eletrificação de nuvens e ao possível surgimento de descargas elétricas nas tempestades.

### 2.1.1 Processo de eletrificação das nuvens

A eletrificação das nuvens é um processo característico no desenvolvimento das tempestades. Esse processo acontece em consequência da separação das cargas elétricas resultantes dos movimentos verticais do ar e das partículas que compõem as nuvens (LHERMITTE; WILLIAMS, 1983). Ainda segundo os autores, as teoria da convecção e precipitação são as principais hipóteses que buscam descrever o processo de criação e desenvolvimento da estrutura elétrica das nuvens.

Uma das hipóteses para o processo de eletrificação das nuvens se baseia no transporte de cargas por convecção, acumulando-as em regiões específicas das nuvens (LHERMITTE; WILLIAMS, 1983). Esse processo faz parte das teorias propostas por Grenet (1947) e Vonnegut (1955), em que os hidrometeoros capturam os íons livres presentes na atmosfera. Os íons positivos são conduzidos pelas correntes ascendentes até a parte superior da nuvem, que por interação eletrostática acabam conduzindo os íons negativos, aprisionando-se em estreitas camadas das nuvens, denominadas de camadas de blindagem (GRENET, 1947; VONNEGUT, 1955; FAZOLO, 2019; PEREIRA, 2019). As eventuais correntes descendentes de ar que se originam durante o desenvolvimento da nuvem, começam a coexistir com as correntes ascendentes, fazendo com que a camada de blindagem se desloque para a base da nuvem, originando novos centros de cargas, que realimentam o processo (VONNEGUT, 1955; FAZOLO, 2019; PEREIRA, 2019). Embora essa hipótese seja eficiente para explicar o carregamento elétrico das nuvens, principalmente em níveis mais altos da atmosfera, o processo de eletrificação baseado na precipitação é a principal hipótese que descreve esse procedimento.

A teoria da precipitação é a mais utilizada para descrever o processo de eletrificação das nuvens, e baseia-se principalmente na concentração de cargas negativas e positivas acumuladas em níveis diferentes da atmosfera por influência dos movimentos verticais dos hidrometeoros, que acontecem em decorrência das correntes de ar ascendentes e força da gravidade (LHERMITTE; WILLIAMS, 1983). A eletrificação das nuvens baseia-se principalmente na teoria da precipitação, e pode ser explicada pelos processos denominados de não-indutivo e indutivo.

O processo não-indutivo não depende de um campo elétrico e é caracterizado pela separação de cargas através das colisões entre as partículas de gelo em acreção (*rimers*) e outras partículas de gelo menores, como os cristais de gelo (REYNOLDS et al., 1957; CARANTI, 1980). A colisão faz com que as partículas de gelo em acreção tornem-se carregadas negativamente e concentrem-se no meio da nuvem, enquanto as partículas mais leves adquirem sinais opostos e são conduzidos através das correntes ascendentes para os níveis mais altos do sistema (REYNOLDS et al., 1957; CARANTI, 1980; PEREIRA; ALBRECHT, 2016). O processo indutivo é baseado nos mecanismos de carregamentos microfísicos propostos por Sartor (1963) e Latham (1981). Diferente do processo não-indutivo, o processo indutivo depende da presença de um campo elétrico. Este campo elétrico induz a polarização dos hidrometeoros que estão em movimentos verticais no interior das nuvens, concentrando as cargas negativas na parte superior e as cargas positivas na parte inferior da partícula. Ao colidirem, as partículas transferem cargas entre si, deixando as maiores carregadas negativamente no centro da nuvem ou na base da nuvem e as menores ficam carregadas positivamente e sobem com as correntes ascendentes (SARTOR, 1963; LATHAM, 1981; FAZOLO, 2019; PEREIRA, 2019).

O processo de eletrificação acontece, geralmente, em nuvens convectivas a partir de uma extensão vertical apropriada, indicando um certo grau de desenvolvimento da tempestade. Tal processo, é o principal fator para o favorecimento da ocorrência de descargas elétricas, que acontecem quando é estabelecido um campo elétrico em um intervalo que separam as cargas elétricas de sinais opostos e sua rigidez dielétrica é quebrada. A subseção seguinte discute sobre a definição de uma descarga elétrica.

### **2.1.2 Descargas elétricas atmosféricas**

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 consistia 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 acaba descobrindo de forma concisa a natureza elétrica das descargas, dando origem ao primeiro tipo de pára-raios (KRIDER, 2006).

Segundo uma definição dada em Rakov (2016, p. 1), o raio é uma descarga elétrica transitória de alta corrente no ar composta por gás ionizado, cuja a temperatura pode chegar até 30000 K, cinco vezes a temperatura da superfície do sol. O rompimento da resistência do ar pelo canal ionizado por onde o raio se desenvolve, que acontece entre dois centros de cargas de polaridade oposta, faz com que seja gerado um aquecimento intenso acompanhado de efeitos luminosos (conhecidos como relâmpagos) e sonoros (conhecidos como trovões) (VISACRO, 2005, p. 40). Embora os raios se desenvolvam tipicamente a partir de tempestades convectivas, outros processos também podem provocar esse evento como as tempestades de neves (MARKET; BECKER, 2009), tempestades de areia (SISSAKIAN et al., 2013) e a fumaça das erupções vulcânicas (BENNETT et al., 2010). Este trabalho está delimitado dentro do escopo dos raios presentes nas tempestades convectivas.

De acordo com as definições e terminologias definidas nos estudos de Schonland (1956) e Uman e McLain (1969), as descargas elétricas mais comuns podem ser classificadas em nuvem-solo (CG - sigla do inglês, *cloud-to-ground*) e intra-nuvem (IC - sigla do inglês, *intra-cloud*). O raios CG são aqueles que interagem entre as nuvens e a superfície da terra ou com objetos sobre ela (e.g. construções, árvores, etc), podendo estes serem positivos (CG+), quando cargas de prótons trilham em direção ao solo, ou negativos (CG-), quando partículas de elétrons vão de encontro ao solo a partir das nuvens (SHAFER et al., 2000; LYNN et al., 2012). Os raios IC são considerados as descargas elétricas que não interagem com a superfície da terra, podendo acontecer tanto dentro no interior de uma nuvem, como também interagir com outras nuvem ou com o ar (WILLIAMS, 2001, p. 550-552).

O interesse em compreender a frequência de raios sobre a Terra tem sido objeto de estudo desde muito tempo, sendo o trabalho de Brooks (1925) um dos precursores. Conforme estimativa feita por Rakov (2016, p. 1), a taxa global de raios varia entre algumas dezenas até 100 por segundo. Segundo extrapolação feita a partir do trabalho de Stano (2012), no qual estuda as tempestades do verão de 2010 na Flórida, estima-se que a taxa global de raios intra-nuvem é muito superior ao tipo nuvem-solo, variando de 75% para até 99% das ocorrências, dependendo da tempestade.

As descargas atmosféricas são utilizadas como importantes indicadores para o monitoramento estratégico de tempestades, pois estão diretamente associados a condições de tempo severo, como as fortes correntes de ar ascendente (DEIERLING; PETERSEN, 2008). Trabalhos como o de MacGorman e Nielsen (1991), Williams et al. (1999), Gatlin e Goodman (2008) mostram que existe uma forte correlação entre a ocor-

rência de descargas elétricas e condições de tempo severo. Sendo assim, a maioria dos sistemas de previsão de descargas atmosféricas, segundo Wu et al. (2018), são construído sob sistemas de previsão imediata de tempestades, também conhecidos como sistemas de *nowcasting*, conceito proposto por Browning (1980), e susodito.

## 2.2 Sistemas para previsão e monitoramento de tempestades

O *nowcasting*, conceito introduzido por Browning (1980), é a previsão do tempo em um período curto de até 6 horas, com a utilização de diversas técnicas e fontes de dados. Segundo Browning, as técnicas de *nowcasting* proporcionam uma alta taxa de acerto nos instantes imediatamente após o atual, e decaem à medida que a previsão se estende para tempos muito à frente, uma vez que a maior parte das técnicas parte de metodologias de extrapolação espacial e temporal.

Neste sentido, esta seção foi dividida em sistemas tradicionais e aqueles que usam técnicas de IA.

### 2.2.1 Sistemas tradicionais

Uma das primeiras abordagens com o intuito de realizar a previsão do tempo foi proposta por Richardson (1922), e tinha como objetivo prever para um determinado local a pressão atmosférica com a antecedência de 6 horas. Devido às limitações computacionais da época, a previsão não foi tão eficiente, porém abriu-se as possibilidades para que a ideia fosse progredida até chegarmos nos complexos sistemas de previsão do tempo que temos nos dias atuais (FREITAS et al., 2017; NCAR, 2019). Com os avanços computacionais que ocorreram na década de 1950, os mapas com previsão do tempo começaram a serem produzidos de forma rotineira, tendo máquinas com dedicação exclusiva para tal aplicação (MOURA, 1996). Embora as abordagens de previsão do tempo atuais tenham tido uma grande evolução, ainda continua sendo um processo contínuo, com muitos dados multidimensionais dinâmicos e caóticos, fazendo com que essa tarefa ainda seja um grande desafio (FATHI et al., 2021).

A maioria dos sistemas de previsão atualmente são baseados em modelos matemáticos conhecidos como modelos de previsão numérica de tempo (NWP - sigla do inglês, *numerical weather prediction*), introduzidos por Richardson (1922) e que usam o processo de assimilação de dados para melhorar os resultados das previsões (FREITAS et al., 2017; NCAR, 2019). O processo de assimilação de dados usa o NWP como problema de valor inicial e as condições atuais da atmosfera para prever a sua evolução no futuro com a maior precisão a possível (NAVON, 2009, p. 22). Além de

previsões a curto prazo, os modelos de previsão baseados em NWP, também, podem realizar previsões para escalas climáticas a longo prazo.

Ao longo dos anos diversos algoritmos foram desenvolvidos a partir de técnicas de extrapolação para a previsão de tempestades. Entre eles podemos citar o *The Forecasting and Tracking the Evolution of Cloud Clusters* (ForTraCC) (VILA et al., 2008) que tem como propósito fazer o rastreo e previsão das estruturas morfológicas das nuvens associadas com sistemas convectivos de mesoescala (SCM) a partir de imagens do satélite meteorológico GOES. Com base na utilização conjunta das ferramentas ForTraCC e do Hydro-Estimator (VICENTE et al., 1998), do qual trata-se um estimador de precipitação global a partir de informações de satélite e modelos meteorológicos, Calheiros (2008) desenvolveu o HydroTrack, uma ferramenta na qual realiza a previsão a curto prazo da precipitação nas regiões sul e sudeste do Brasil com base em análises estatísticas do ciclo de vida das estruturas de precipitação em SCM. Também com base em técnicas estatísticas e de extrapolação, Dixon e Wiener (1993) propuseram o *Thunderstorm Identification, Tracking, Analysis, and Nowcasting* (TITAN), algoritmo que realiza a identificação, monitoramento e previsão de tempestades em tempo real a partir de dados de radares meteorológicos.

Embora as abordagens pontuadas até agora sejam tradicionalmente mais utilizadas em sistemas de previsão, os algoritmos baseados em IA e aprendizado de máquina (ML - sigla do inglês, *machine learning*) tem mostrado uma grande eficiência nos últimos anos. A robustez desses modelos fazem com que tarefas complexas de previsão sejam executadas em um tempo curto e de forma robusta. Isso acontece pois, uma vez que os modelos de ML estão treinados (tarefa que pode exigir um tempo e custo computacional maior), a previsão acontece através da sua ativação com os novos dados de entrada.

As seções seguintes descrevem algumas técnicas aplicadas à meteorologia que utilizam inteligência artificial.

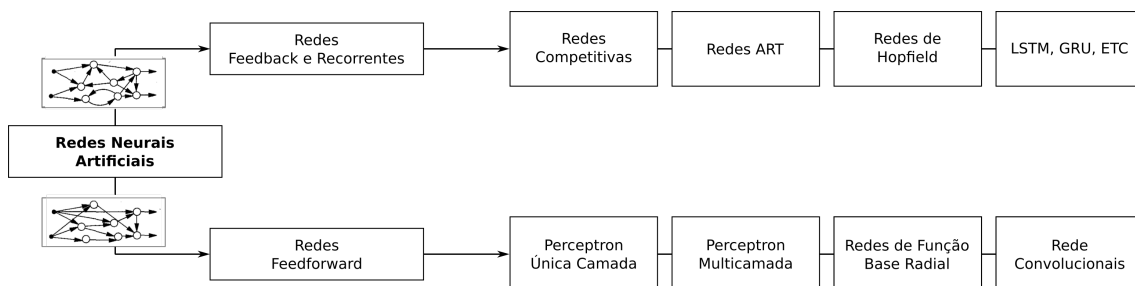
### 2.2.2 Sistemas baseados em inteligência artificial

Em 1990, durante a *16th conference on Severe Local Storms* que aconteceu no Canadá, foi apresentado por Moninger (1990) uma visão geral do cenário da IA aplicada na previsão de tempestades, o que acabou encorajando o desenvolvimento de novas pesquisas usando essa abordagem, principalmente empregando o uso das redes neurais artificiais (MCCANN, 1992).

### 2.2.2.1 Redes neurais artificiais

As redes neurais artificiais (RNA) são sistemas motivados pela capacidade que o cérebro tem em realizar uma computação de dados massivamente paralela e distribuída, permitindo a realização de tarefas complexas tal como o reconhecimento de padrões (HAYKIN, 1994, p. 23). Elas são representadas através de grafos direcionais ponderados, no qual simulam o processo de sinapse da malha neural biológica, onde os neurônios são representados pelos vértices da estrutura e a transferência da informação é feita com base nos pesos atribuídos às arestas após aplicar uma função matemática de ativação ou não do neurônio. A informação que chega ao final do grafo é comparada com a informação real, e então tem os seus pesos atualizados com base na propagação do erro (HAYKIN, 1994; HASSOUN, 1995; NUSSBAUM; CHAFFIN, 1996). Baseado no padrão de conexão do grafo (arquitetura), as RNA podem ser classificadas em dois grupos (JAIN et al., 1996): *feedforward* e *feedback*. A Figura 2.2 mostra os dois principais grupos de conexão da RNA, bem como algumas de suas respectivas redes. De forma geral, esses dois grupos se distinguem pela forma que é feito o fluxo de informações entre os neurônios. Enquanto que, nas redes do tipo *feedforward* os neurônios transmitem as informações somente para as camadas subsequentes, as redes *feedback* podem conter *loops* em seus neurônios, ou seja, as informações podem ser ativadas e utilizadas pelos mesmos neurônios. Além disso, as redes *feedforward* possuem uma topologia direcionada, já as redes *feedback* podem ter topologias cíclicas, em que todos os neurônios são inter-conectados (JAIN et al., 1996).

Figura 2.2 - Ramificação das principais redes neurais artificiais.



Fonte: Adaptado Jain et al. (1996).

Uma das principais características das RNA é a sua não linearidade intrínseca, o que pode proporcionar uma vantagem em relação aos métodos lineares na descrição de eventos meteorológicos complexos (MANZATO, 2007). Um dos modelos de RNA mais utilizado é *multilayer perceptron* (MLP). Com o intuito de prever tempestades a curto prazo, McCann (1992) propôs a implementação de duas MLPs para prever os valores de alguns índices de instabilidade atmosférica em regiões diferentes. As RNAs foram treinadas com configurações distintas, variando seu número de neurônios nas camadas ocultas e sua função de ativação, tendo como entrada a sensibilidade aos padrões que os especialistas em previsões reconheciam como fatores que indicavam a ocorrência de fortes tempestades, como os valores dos índices de levantamento e convergência da umidade na superfície.

As redes MLP pertencem ao grupo das arquiteturas do tipo *feedforward*, ou seja, as informações processadas pelos neurônios são transmitidas somente para as camadas subsequentes. O outro grupo possui arquiteturas que incluem também *loops* os quais permitem armazenar as informações processadas pelo próprio neurônio. Nesse grupo estão as redes neurais recorrentes (RNN, sigla do inglês, *recurrent neural network*). O método de funcionamento das RNN tradicionais usam suas conexões de *feedback* para armazenar as representações de eventos de entradas recentes, agindo como uma memória de curto prazo. A partir desse modelo surgiram outras variantes de redes recorrentes como a Long Short-Term Memory (LSTM) (HOCHREITER; SCHMIDHUBER, 1997) e Gated Recurrent Unit (GRU) (CHUNG et al., 2014).

As RNN são comumente utilizadas para previsão e imputação de dados em séries temporais, e devido à sua memória as informações do passado persistem durante algumas interações do processo de treinamento, influenciando diretamente na previsão. As redes LSTM podem aprender a unir intervalos de tempo acima de 1000 etapas, permitindo que sequências com eventuais ruídos não afetem o desempenho das previsões (HOCHREITER; SCHMIDHUBER, 1997). As RNN vem sendo amplamente utilizadas em tarefas de previsões meteorológicas, por exemplo, no trabalho de Stensö (2018) foi utilizada a LSTM para prever trajetória de tempestades tropicais a partir de dados de raios. No trabalho de Qing e Niu (2018) foi proposta uma abordagem para prever a curto prazo a irradiância solar a partir de variáveis meteorológicas.

Devido a sua não linearidade e às vezes a presença de uma quantidade muito grande de neurônios e camadas ocultas para o processamento de uma significativa quantidade de dados, as denominadas redes neurais profundas podem prover resultados

satisfatórios, porém, de forma às vezes incompreensíveis, comportamento conhecido como “*black box*” (HUSSAIN, 2019). Diante disso, algumas outras abordagens podem ser utilizadas com o intuito de acompanhar a execução do algoritmo e a interação dos dados com maior clareza, como os algoritmos de aprendizado de máquina supervisionado apresentado na próxima subseção.

### 2.2.2.2 Árvore de decisão e similares

As árvores de decisão (DT - sigla do inglês, *decision tree*) tiveram sua origem a partir da teoria de aprendizado de máquina no trabalhos introduzidos por Quinlan (1986), sendo uma ferramenta extremamente eficiente para solução de problemas de classificação e regressão. As DTs são baseadas em um esquema de decisão por etapas ou de forma hierárquica, diferente de outras abordagens de classificação, que utilizam faixas de valores para executar a classificação em uma única etapa de decisão (XU et al., 2005). Essa abordagem tem sido frequentemente utilizada na previsão de eventos meteorológicos por ser simples de entender o seu funcionamento e interpretar seus resultados (PETRE, 2009).

Pouco tempo após sua introdução, as DTs começaram a ser utilizadas em previsões meteorológicas. Por exemplo, no trabalho de Colquhoun (1987) foram utilizados classificadores baseados em árvores de decisões para realizar previsões de tempestades severas e tornados, utilizando regras baseadas em limiares de algumas variáveis termodinâmicas de instabilidade atmosférica. Dubrovskỳ (1995) também utilizou as DTs para obter a probabilidade da ocorrência de tempestades em curto espaço de tempo. Já no trabalho de (GUILLOT et al., 2008) as árvores de decisão foram utilizadas para classificar tempestades, quanto à sua severidade e seu potencial risco de desenvolver tornados. Essa técnica também foi utilizada para o desenvolvimento de um importante módulo de *nowcasting* de raios do sistema de identificação, rastreamento e previsão de tempestades proposto no trabalho de Meng et al. (2019).

A floresta aleatória (RF - sigla do inglês, *random forest*) é um algoritmo de ML para classificação e regressão baseados em uma síntese dos resultados da combinação de árvores de decisão (BREIMAN, 2001). A construção do conjunto de preditores depende de uma amostra aleatória retirada de um vetor de aprendizagem, no qual é utilizado para a criação dos nós de cada árvore de decisão (BREIMAN, 2001; GENUER et al., 2010). Segundo Breiman (2001), o erro de generalização, que depende da força individual das árvores e das correlações entre si, converge para o mínimo quando o número de preditores do conjunto se torna grande.



O algoritmo RF foi utilizado no trabalho de Ruiz e Villa (2008) para previsão de tempestades, onde apresentou melhor desempenho em relação à abordagem utilizando regressão logística, tendo uma maior taxa de acerto e uma menor quantidade de falsos alarmes. O RF também foi utilizado no trabalho de Bates et al. (2018) para previsão de descargas elétricas na Austrália. Já no trabalho de Hill et al. (2020) o RF foi utilizado para prever eventos meteorológicos severos com antecedência de 1 a 3 dias no Estados Unidos da América (EUA).

A Spatiotemporal Relational Random Forest (SRRF) é uma técnica proposta por Supinie et al. (2009) que inclui recursos espaço-temporais no algoritmo de RF. Essa técnica foi inicialmente validada com dados simulados e dados reais de turbulência induzidos por convecção. Tal técnica foi aplicada no trabalho de McGovern et al. (2010), e teve três propósitos principais. Em um primeiro momento, o objetivo foi prever turbulências induzidas por tempestades que podem afetar aeronaves. O segundo domínio da aplicação da técnica teve como o objetivo prever tornados e ter um melhor entendimento a respeito do seu ciclo de vida. E por fim, foi possível realizar um estudo em escala climatológica a respeito da progressão das secas durante um período de 134 anos e a influência dos tornados e turbulências. Além de redes neurais e árvores de decisão, existem outras técnicas que já foram usadas em meteorologia e que serão citadas na subseção seguinte.

### 2.2.2.3 Outras técnicas

O *support vector machine* (SVM) (CORTES; VAPNIK, 1995) é um algoritmo de aprendizado de máquina amplamente utilizado em tarefas de classificação e também regressão. Essa técnica foi utilizada por Radhika e Shashi (2009) para previsão da temperatura atmosférica, onde apresentou melhor performance em comparação a MLP. O SVM foi utilizado no trabalho de Han et al. (2017) como complemento de um sistema de previsão espaço-temporal de tempestades convectivas a partir dos valores de algumas variáveis meteorológicas, que foram obtidas a partir do produto de assimilação de dados de radar denominado *The variational Doppler radar analysis system* (VDRAS). O sistema de *nowcasting* de Han et al. (2017) foi capaz de prever, com precisão superior a 60%, o local e intensidade de tempestades convectivas com antecedência de 30 minutos após serem detectados pelo radar os primeiros ecos de  $\geq 30$  dBZ. O SVM também foi utilizado para prever o acumulado de precipitação e potenciais tempestades em curto intervalo de tempo em diversas localidades (FAN et al., 2009; QIU et al., 2010; WANG et al., 2013; MAI et al., 2019).

Muitas abordagens, em busca de um melhor resultado na previsão, tentam combinar técnicas (RAVI et al., 1999; TAYLOR; BUIZZA, 2002; MAQSOOD et al., 2004). No trabalho proposto por Ravi et al. (1999), foram combinadas as técnicas de método gráfico e múltiplas regressões para prever tempestades na cidade de Delhi (capital da Índia) durante alguns períodos entre os anos de 1985 e 1994 a partir dos dados de índices de estabilidade, ventos e temperatura. Nesta mesma linha, Maqsood et al. (2004) criaram um modelo combinando as redes MLP, RNN, redes de função de base radial (RBFN - sigla do inglês, *radial basis function network*), modelo de Hopfield (HFM - sigla do inglês, *Hopfield model*) e algumas técnicas de regressão, para prever a temperatura, velocidade do vento e umidade durante as quatro estações dos anos de 2000 e 2001 na província de Saskatchewan, no Canadá. Wang (2011) combinou uma variação da árvore de decisão e as SVM para construir um modelo de previsão da capacidade aeroportuária em relação às condições meteorológicas de alguns aeroportos dos EUA.

As técnicas e recursos computacionais cada vez mais robustos, munidos da vasta gama de recursos disponíveis para a aquisição de dados meteorológicos, faz com que sejam construídos sistemas de previsão cada vez mais eficientes. Porém, como parte da construção desses sistemas, é de suma importância a escolha das variáveis que melhor representam o evento que se deseja observar. A literatura possui uma sólida base acerca das principais variáveis meteorológicas relacionadas com condições de tempo severo, as quais as descargas elétricas atmosféricas estão associadas, essas serão abordados na próxima seção.

### **2.3 Principais variáveis meteorológicas para previsão**

Este trabalho, como pode ser observado, tem como foco apresentar subsídios para o desenvolvimento de um sistema de previsão da ocorrência ou não dos raios CG sobre uma região em específico, neste caso, a região amazônica. As descargas elétricas atmosféricas são fenômenos caóticos, sinuosos e instáveis, o que acaba tornando a sua previsão um grande desafio (WILLIAMS, 2008). As informações acerca da ocorrência de raios é de suma importância para várias tomadas de decisões. Por exemplo, a Estação da Força Aérea de Cabo Canaveral (EUA), utilizam em larga escala as informações de sistemas previsores de descargas elétricas para o planejamento de suas missões (LAMBERT et al., 2005), no futuro o INPE pode auxiliar a Força Aérea Brasileira em missões similares já existentes no Centro de Lançamento de Alcântara.

De acordo com Huamán (2018), o monitoramento de descargas elétricas no Brasil é feito principalmente através das redes de detecção, informações de radares e satélites meteorológicos. As redes de detecção de raios são instrumentos importantes para o estudo desse fenômeno, e no Brasil algumas delas são aplicadas tanto em pesquisa como operacionalmente, e entre elas estão por exemplo: BrasilDat (NACCARATO; PINTO, 2012), STARNET (MORALES et al., 2014), LINET (BETZ et al., 2009) e GLD360 (DEMETRIADES et al., 2010). Posteriormente, no Capítulo 3, serão descritas as principais características da STARNET, rede de detecção que originaram os dados de descargas elétricas utilizado neste trabalho.

Atualmente, além dos sensores de superfície, uma das principais fontes de dados de descargas elétricas para o Brasil são oriundas da nova geração de satélites geoestacionários GOES-16. O GOES-16 foi lançado em novembro de 2016 em uma missão<sup>1</sup> liderada pela *National Oceanic and Atmospheric Administration* (NOAA) e *National Aeronautics and Space Administration* (NASA), fornecendo os primeiros dados em 2017. Goodman et al. (2012) destaca que as tecnologias a bordo do novo GOES apoia “a detecção expandida de fenômenos ambientais, resultando em previsões e avisos mais oportunos e precisos”, em relação às gerações anteriores do satélite. Dos instrumentos a bordo do satélite destacam-se o *Advanced Baseline Imager* (ABI) (SCHMIT et al., 2018) e *Geostationary Lightning Mapper* (GLM) (GOODMAN et al., 2013), ambos com alta resolução temporal e espacial. Este último, o GLM, fornece informações sobre a ocorrência de descargas elétricas sobre todo o território brasileiro de forma contínua e em tempo quase real.

Como visto na seção anterior, diversas técnicas precisam dessa informação sobre raios para definir modelos matemático de modo a realizar a previsão do mesmo e validar seus resultados. Prever raios utilizando somente dados de raios, é extremamente difícil e requer modelos que representem a alta variabilidade espaço-temporal deste fenômeno (FAZOLO, 2019). A partir desses dados de descargas elétricas é possível definir outras variáveis que são correlacionadas no espaço e no tempo com outros parâmetros e variáveis meteorológicas, e estes podem ser usados para criar índices que auxiliam na previsão de raios. A seguir vamos listar algumas variáveis que estão relacionadas com tempestades e que podem ser usadas na previsão de raios.

Uma das principais fontes de dados em meteorologia para definição das tempestades são os radares meteorológicos. Esses instrumentos desempenham um papel importante nas tarefas de identificação, monitoramento e previsão de tempestades, devido

---

<sup>1</sup><https://www.goes-r.gov/mission/mission.html>

à grande quantidade de variáveis altamente correlacionadas com a microfísica das nuvens que podem ser obtidas a partir de seus dados (KLEINA et al., 2016). De acordo com Voormansik et al. (2017) a eletrificação das tempestades e outras características de severidade estão fortemente associadas com o desenvolvimento das nuvens e suas células convectivas. As interações ativas das ondas eletromagnéticas emitidas nos escaneamentos dos radares meteorológicos podem traçar os perfis verticais das nuvens, sendo possível acompanhar o seu desenvolvimento quase que instantaneamente (ZIPSER; LUTZ, 1994; STEINER et al., 1995).

O princípio do funcionamento dos radares meteorológicos consiste em emitir uma onda eletromagnética que se propaga pelo ar na velocidade da luz e ao colidir com algum alvo (e.g. hidrometeoros suspensos nas nuvens) é retro-espalhada uma fração da energia enviada, ao qual é denominada como refletividade ( $Z$ ) (MORELL, 2018). A refletividade é definida pela Equação 2.1, onde  $N_v(D)dD$  representa o número médio das gotas de chuva (ou qualquer outro hidrometeoro) com diâmetros esféricos equivalentes entre  $D + dD$  mm presentes por unidade de volume de ar. Os valores de refletividade é medido em  $\text{mm}^6/\text{m}^3$ , e quando normalizados por faixa podem variar em diversas ordens de magnitude, então com o intuito de simplificar os valores para o uso operacional,  $Z$  é exibido em em escala logarítmica (dBZ) (ROGERS; YAU, 1989, p. 190). A transformação de  $Z$  para o fator de refletividade em decibéis é dada pela Equação 2.2.

$$Z = \int_0^{\infty} D^6 N_v(D) dD \quad (2.1)$$

$$dBZ = 10 \log_{10} Z \quad (2.2)$$

Os radares meteorológicos realizam varreduras completas (em  $360^\circ$ ) estratificadas em diferentes altitudes. Esses dados volumétricos são transformados em um produto representado por uma grade cartesiana tridimensional, denominado *Constant Altitude Plan Position Indicator* (CAPPI). Tal distribuição permite o levantamento de todo o perfil vertical de refletividade do radar dentro da área de cobertura do sensor (ATLAS, 2015). O perfil vertical de refletividade do radar (VPRR - sigla do inglês, *vertical profile of radar reflectivity*) permite o acompanhamento da evolução dos núcleos convectivos durante todo o seu ciclo de vida, uma vez que, a variação temporal do topo da nuvem ( $dh/dt$ ) obtida a partir dele está associada aos movimentos verticais, onde os valores negativos e positivos indicam movimentos descendentes

e ascendentes, respectivamente (MACHADO; LAURENT, 2004). O VPRR também é utilizado como um importante indicador do aumento do processo de eletrificação das nuvens e intensificação das células de tempestade (ZIPSER; LUTZ, 1994). Com o auxílio de dados de satélite, Liu et al. (2012) exploraram a relação entre as descargas elétricas e a variação do VPRR sobre diferentes superfícies, onde foi possível identificar limiares de refletividade em diferentes níveis associados ao total de raios registrados.

A partir da distribuição dos dados de refletividade do radar é possível definir outras variáveis associadas a condições de tempo severo. A máxima refletividade vertical indica o maior valor de refletividade (em dBZ) nos níveis escaneados pelo radar, e é acompanhada de outra variável informando a altura (em metros) em que foi obtido esse valor. A diferença de altura entre o nível de 45 dBZ e o nível de congelamento (0°C) caracterizam células de tempestade com potencial para produzir granizo (QUEIROZ, 2009), essa variável é denominada altura de Waldvogel (WALDVOGEL et al., 1979). A altura máxima do eco de 35 dBZ também é uma informação importante para o acompanhamento da estrutura convectiva vertical das tempestades, uma vez que a essa região pode concentrar uma grande quantidade de cargas negativas, estando associada à presença de raios (TORACINTA et al., 1996).

A precipitação é uma variável que pode ser calculada a partir dos dados de radar e que possuem forte relação com a ocorrência de descargas elétricas (SEITY et al., 2001). A precipitação acontece a partir de processos microfísicos nas nuvens, em que hidrometeoros formados a partir do vapor de água condensado na atmosfera cedem para a atração gravitacional e caem na forma de chuva, granizo ou neve (SUI et al., 2007; ATKINSON, 2019). Vários trabalhos estudaram as relações entre os processos de eletrificação das nuvens, associados aos raios e a precipitação em sistemas convectivos, principalmente aquelas com presença significativa na composição das nuvens (LHERMITTE; WILLIAMS, 1983; GOODMAN et al., 1988; ZIPSER; LUTZ, 1994).

A precipitação ( $R$ ) (medida em  $mmh^1$ ) pode ser derivada a partir do valor de refletividade por uma relação conhecida como  $ZR$  (MARSHALL; PALMER, 1948). Essa relação ( $ZR$ ) pode ser obtida através da lei de potência expressa pela Equação 2.3, onde  $a$  e  $b$  são coeficientes que dependem de alguns fatores locais, sazonais e dos tamanhos das gotas que variam em função do tipo de chuva. Algumas complicações podem surgir quando existe erros nas medidas de refletividade, segundo (HUNTER, 1996 apud CALHEIROS et al., 2016) “uma diferença de 1 dBZ na calibração dos instrumentos pode causar um erro de 17% na taxa de precipitação e a existência de

incertezas entre o valor da refletividade e sua relação com as condições de superfície (relação  $ZR$ )”, além da variação da distribuição de tamanho de gotas. Os valores de dBZ são comumente convertidos para taxa de precipitação em  $mmh^1$  usando a Equação 2.4, proposta por Marshall e Palmer (1948).

$$Z = a.R^b \quad (2.3)$$

$$R = \left( \frac{10^{(dBZ/10)}}{200} \right)^{\frac{5}{8}} \quad (2.4)$$

O conteúdo de água integrado verticalmente (VIL - sigla do inglês, *vertically integrated liquid water*) (GREENE; CLARK, 1972), é outra variável importante que está associada a ocorrência de raios devido a sua relação com a microfísica de nuvens (WATSON et al., 1995). Assim como  $R$ , o VIL é uma variável derivada da refletividade do radar. Essa variável transforma a refletividade, através de uma função não linear, em conteúdo de água líquida ou gelo que está presente em uma camada vertical atmosférica, em que sua massa em volume unitário é dada em  $kg/m^{-2}$ . O VIL é definido pela Equação 2.5, onde  $Z$  é a refletividade do radar (em  $mm^6m^3$ ),  $a$  e  $b$  são os limites inferior e superior da camada atmosférica e  $h$  é a altura do feixe do radar (em metros). A altura do feixe do radar é expressa pela Equação 2.6, onde  $\theta$  é o ângulo de elevação (em radianos),  $r$  é a linha de alcance do radar (em metros) e  $a$  é uma constante representando o raio da terra.

$$VIL = \sum_{i=a}^b 3,14 \times 10^{-6} \left[ \left( \frac{Z_i + Z_{i+1}}{2} \right) \right]^{\frac{4}{7}} \Delta h \quad (2.5)$$

$$h = \frac{r^2 \times \cos^2 \theta}{\frac{8}{3} \times a} + r \times \sin \theta \quad (2.6)$$

Com o intuito de identificar algumas características de tempo severo a partir dos hidrometeoros, Amburn e Wolf (1997) propuseram uma nova métrica que analisasse a densidade do VIL dentro da nuvens, o DVIL (*Vertically Integrated Liquid Density*). Segundo Lahiff (2005), partindo da hipótese que tempestades intensas com altos valores de VIL nem sempre produziam granizos grandes, enquanto que tempestades com valores de VIL poderiam produzir, Amburn e Wolf (1997) propuseram dividir o valor do VIL de uma tempestade pelo eco do topo do radar, para assim

ser normalizado, produzindo um valor comum para tempestades com granizos grandes, independente das características da massa de ar. A variável DVIL (em  $g/m^{-3}$ ) pode ser expressa pela Equação 2.7, onde  $ET$  é o topo de uma área de precipitação indicado pelo radar.

$$DVIL = \frac{VIL}{ET} \times 1000 \quad (2.7)$$

Notem que todas as variáveis listadas nesta seção estão associadas a uma alta resolução espacial e temporal, geralmente entregue por sensores remotos, como radares meteorológicos. Isso se dá pelo fato que o *nowcasting* de tempestade depende de um fluxo de dados com baixa latência, de modo a detectar variações rápidas na estrutura dinâmica e microfísica das tempestades. Outras informações sobre a estrutura termodinâmica do ambiente circunvizinho a tempestades seriam importantes, contudo, para este estudo, devido a sua resolução, não serão aplicadas.

Outros fatores podem influenciar no processo de formação e eletrificação das tempestades, entre eles estão o tipo de solo e orografia do local. Na próxima seção será abordado como essas características podem se relacionar na intensificação da ocorrência de raios.

## 2.4 Influência da superfície na formação de tempestades

As interações entre a atmosfera e as superfícies desempenham um importante papel no desenvolvimento de tempestades, sendo amplamente utilizadas como parâmetros para modelos de previsão (PIELKE, 2001; WANG et al., 2013). Pielke (2001) analisou a relação entre a umidade em diferentes superfícies e os fluxos de calor no desenvolvimento das nuvens *cumulonimbus*. As perturbações de temperatura têm alta influência nos processos não lineares de eletrificação das nuvens (WILLIAMS, 2005). Pielke (2001) observou que as temperaturas de uma região tendem a ser influenciadas pela mudança na vegetação.

Na região da Bacia Amazônica, os fluxos de calor sofrem grande influência do gradiente de temperatura da superfície provocada pelas brisas fluviais, que podem evoluir com as correntes ascendentes favorecendo a formação de nuvens (DIAS et al., 2004; PEREIRA, 2019). Durante a estação chuvosa dessa região Machado et al. (2018) não observaram dependência entre o tipo de superfície e a taxa de precipitação, entretanto, na estação seca foram registrados maiores taxas de chuva em áreas urbanas e taxas médias menores em regiões desmatadas. Os autores também notaram que

as taxas de precipitação durante o período chuvoso não tiveram grandes variações para diferentes topografias, já durante a estação seca, esses valores acompanharam o aumento da elevação do relevo. Tais resultados estavam provavelmente associados a tipo de precipitação que ocorre nas diferentes estações, como a convectiva durante o período mais seco do ano.

As mudanças no uso do solo também podem influenciar na polaridade da atividade elétrica das tempestades (ALBRECHT et al., 2011; PEREIRA, 2019). Pereira (2019) observou que tempestades com maior incidência de raios CG+ se formam sobre regiões de pastagem durante o final da estação seca no sudeste da Amazônia, enquanto que na estação chuvosa não há predominância de polaridades dos raios. Na estação chuvosa também foi observado a influência da poluição no desenvolvimento de tempestades, uma vez que em ambientes mais poluídos tendem a acontecer mais tempestades que ambientes limpos (PEREIRA, 2019; MARTIN et al., 2017).



### 3 MATERIAIS E MÉTODOS

Este Capítulo tem como finalidade descrever a área de estudo, dados e a metodologia que foi aplicada neste trabalho. Na Seção 3.1 será apresentada uma visão geral sobre a área de estudo, dados e instrumentos utilizados em sua coleta. Na Seção 3.3 serão apresentadas as metodologias que foram aplicadas no trabalho referentes à análise espacial das descargas elétricas, abordagem para definição de classes, seleção das variáveis meteorológicas e os métodos de classificação e previsão. Ainda nessa seção, serão expostas as tecnologias utilizadas para o desenvolvimento do trabalho e também as abordagens de avaliação dos modelos de aprendizado de máquina.

#### 3.1 Área de estudo

A Bacia Amazônica possui cerca de 6,8 milhões de km<sup>2</sup>, sendo a maior bacia hidrográfica do mundo (BARTHEM et al., 2005). A Bacia Amazônica está localizada predominante no território brasileiro, no entanto, ela está presente também em outros seis países da América do Sul, conforme mostrado na Tabela 3.1. De acordo com Mosmann et al. (1986), a Bacia Amazônica é subdividida em três regiões, tendo como referência o Rio Amazonas, são elas: Baixo, Central e Alto Amazonas. A delimitação correspondente à área de estudo deste trabalho está localizada na região Central da Bacia Amazônica.

Tabela 3.1 - Áreas da Bacia Amazônica por países.

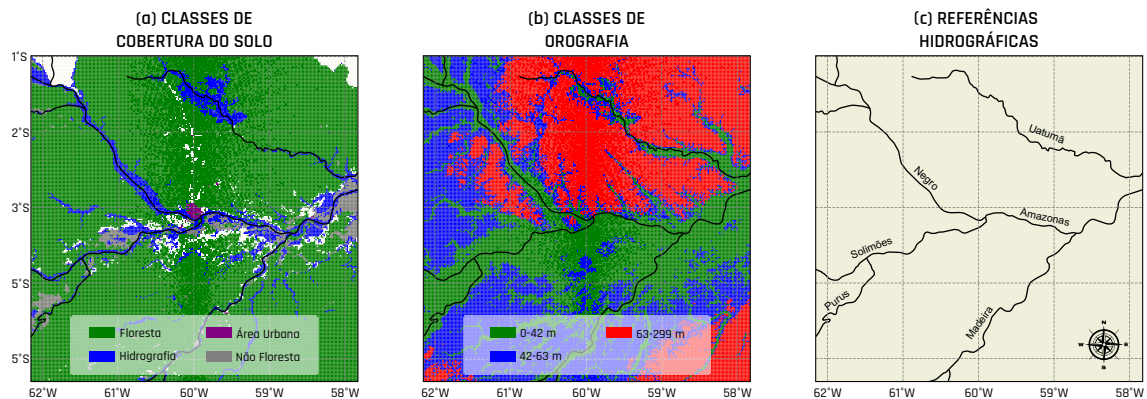
PAÍS	ÁREA EM KM <sup>2</sup>	%
Brasil	4746479	69,1
Peru	783066	11,4
Bolívia	734983	10,7
Colômbia	405271	5,9
Equador	137380	2,0
Guayana	6869	0,1
Venezuela	54952	0,8
<b>TOTAL</b>	<b>6869000</b>	<b>100.0</b>

Fonte: Adaptado de Barthem et al. (2005).

A área de estudo está situada em uma região que abrange a capital do Amazonas, Manaus, e outros 19 municípios aos arredores, totalizando uma área de aproximadamente 480 km<sup>2</sup>. A floresta é o tipo de cobertura do solo predominante nessa região, e

na área de estudo representa mais de 85%. A classe que representa os corpos d'águas é o segundo tipo de cobertura de solo mais presente, essa representatividade se dá principalmente pelo Rio Amazonas e a confluência dos rios Negro e Solimões e alguns de seus principais afluentes, como o Rio Madeira. A Figura 3.1(a) mostra o mapa de cobertura do solo com as quatro classes mais comuns na área de estudo, enquanto que na Figura 3.1(b) é mostrado o mapa com as classes de orografias criadas a partir do balanceamento feito pelos percentis de 33,33 e 66,66 em relação às alturas do terreno. Já a 3.1(c) mostra as referências hidrográficas.

Figura 3.1 - Classes de coberturas do solo e orografia da área de estudo.



Fonte: Adaptado de Machado et al. (2018).

A área de estudo possui duas principais características sazonais ao longo do ano, sendo elas a estação chuvosa e a estação seca. A estação chuvosa, quando ocorreu o IOP1 (IOP1 - sigla do inglês, *intensive operation period*) do experimento GoAmazon, é caracterizada principalmente pela ocorrência de sistemas estratiformes, ou seja, chuvas que se estendem por um longo período e acontece de forma distribuída e uniforme em uma grande área. Enquanto que, na estação seca (quando ocorreu o segundo IOP; IOP2) se destacam as chuvas convectivas, que acontecem de forma mais intensa e concentrada em regiões específicas e normalmente são acompanhadas por uma alta incidência de raios (MACHADO et al., 2018). Os meses que correspondem a estação chuvosa são janeiro, fevereiro e março. Já a estação seca acontece nos meses de agosto, setembro e outubro, sendo o intervalo entre esses períodos as estações de transição (MACHADO et al., 2018).

## 3.2 Dados

Os dados utilizados neste trabalho são oriundos de alguns dos principais instrumentos utilizados durante a campanha *Green Ocean Amazon* (GoAmazon) (ARM, 2018) que aconteceu nos anos de 2014 e 2015. Os dados da campanha GoAmazon serviram como base para diversos outros estudos com o intuito de investigar a relação entre as particularidades da região e algumas características de tempo severo. Martin et al. (2015) buscaram compreender e quantificar as interações do ponto de vista natural e através da intervenção humana entre a vegetação, química da atmosfera, precipitação, dentre outros. Pereira (2019) identificou a influência de algumas características particulares dessa região com a intensificação da atividade elétrica durante os períodos chuvosos, seco e transição. Sendo assim, a aplicação desse estudo sobre essa área poderá ser uma contribuição importante para a previsão de descargas elétricas durante os períodos chuvosos e secos na região central da Bacia Amazônica.

Nos trabalhos desenvolvidos a partir do GoAmazon, foi observado que nesta região as tempestades apresentam características distintas que são influenciadas por fatores locais. Alguns desses fatores serão levados em consideração como possíveis atributos no treinamento dos modelos de ML, como por exemplo: cobertura de solo, relevo e sazonalidade. Outros fatores de maior importância, como distribuição de aerossóis e algumas propriedades microfísicas associadas a cada tempestades (exceto aquelas estimadas via radar e satélites) são difíceis de se definir a uma taxa de assimilação necessária para o *nowcasting* e por isso não serão utilizadas inicialmente. Tendo em vista isso, neste trabalho foram utilizados dados provenientes de um dos radares meteorológico da rede SIPAM (Sistema de Proteção da Amazônia) (SARAIVA et al., 2016), a rede de detecção de raios STARNET (*Sferics Timing Ranging Network*) (MORALES et al., 2014), classes de cobertura e uso do solo foram obtidas a partir dos mapas do produto *European Space Agency Climate Change Initiative Land Cover* ESA (2017) e orografia calculada pelo modelo de elevação digital a partir dos dados do *Shuttle Radar Topographic Mission* (SRTM) (JARVIS et al., 2008).

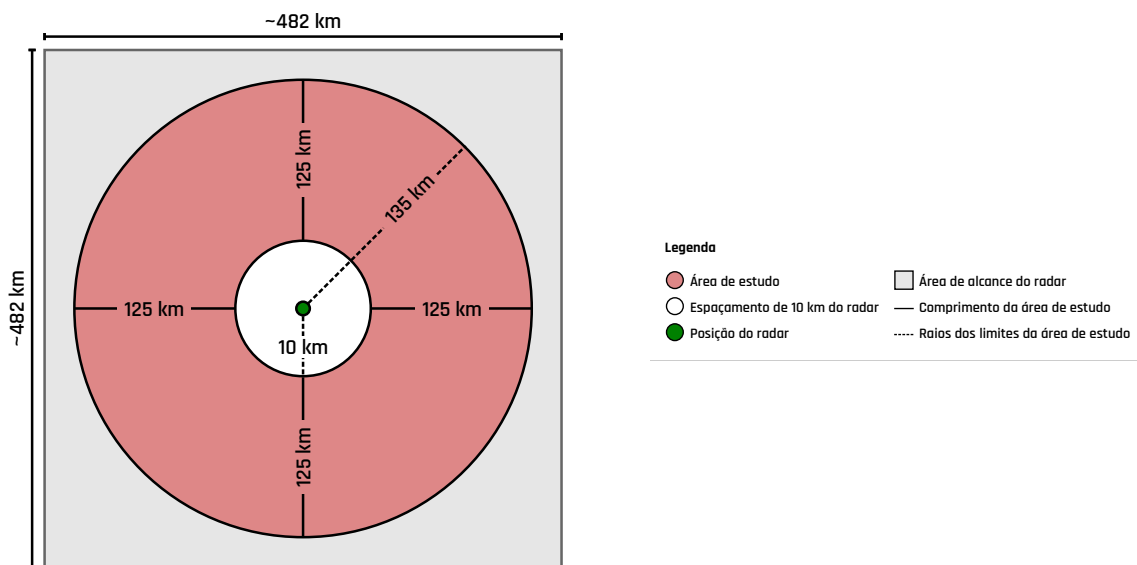
### 3.2.1 Dados de radar

Os radares meteorológicos doppler banda S da rede do SIPAM possuem raio de 240 km de alcance em 10 municípios da Amazônia Legal: Boa Vista (RR), Cruzeiro do Sul (AC), Macapá (AP), Manaus (AM), Porto Velho (RO), Santarém (PA), São Gabriel da Cachoeira (AM), São Luís (MA), Tabatinga (AM) e Tefé (AM). Os dados utilizados neste trabalho pertencem ao radar de banda S (frequência 2 - 4 GHz) localizado em Manaus (AM). Esse radar possui um feixe com 2° de abertura,

realizando varreduras volumétricas em 17 graus de elevação (de 0,9° para 19,5°) a cada 12 minutos. Os produtos CAPPI (ARM, 2018) são constituídos por 40 níveis de 241 linhas e 241 colunas, formando uma grade cartesiana tridimensional com a resolução espacial horizontal de 2 km e vertical de 0,5 km (SARAIVA et al., 2016; MACHADO et al., 2018).

Embora o alcance do radar possua originalmente um raio de 240 km, em partes deste trabalho (i.e. na análise exploratória, classificação e regressão) foi aplicada uma delimitação dessa área definida entre a distância de 10 a 135 km da posição do radar, como ilustrado na Figura 3.2. Essa delimitação foi feita de acordo com o trabalho de Machado et al. (2018) e tem como propósito reduzir alguns problemas relacionados ao escaneamento do radar como o cone cego, altura e largura do feixe. A Figura 3.3 mostra o acumulado da taxa de chuva horária para cada estação. Como pode ser observado nessa figura e conforme descrito descrito por Machado et al. (2018), a área a uma distância entre aproximadamente 10 km 135 km do radar possui uma melhor representatividade dos dados de chuva.

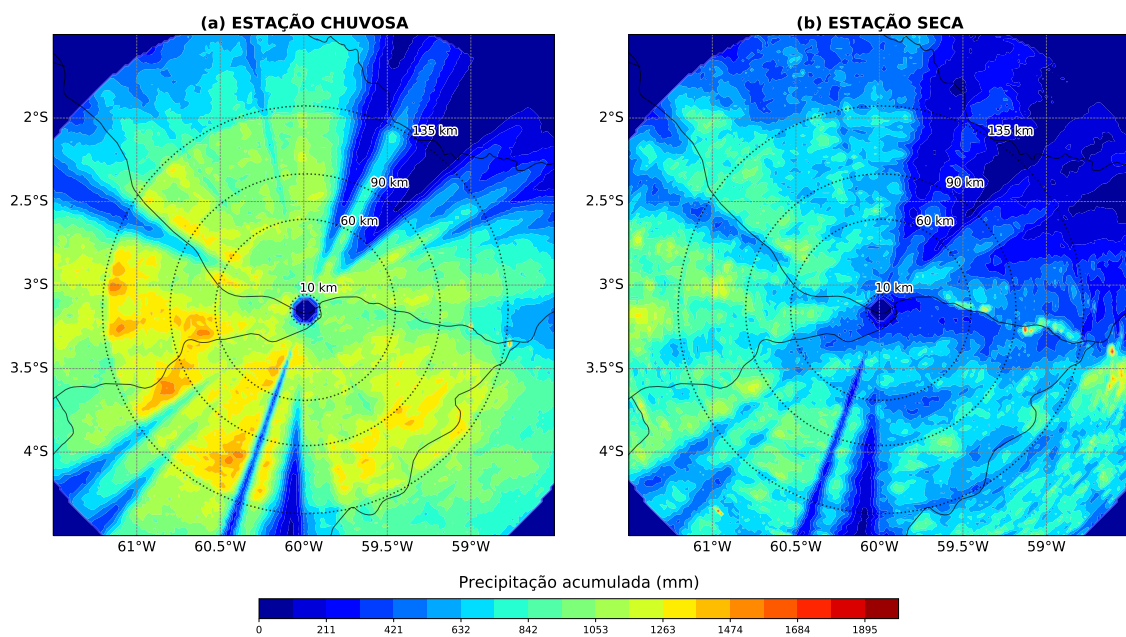
Figura 3.2 - Área delimitada para extração das variáveis.



Fonte: Produção Autor.

Os dados de radar utilizados neste trabalho são disponibilizados originalmente em decibéis de refletividade (dBZ). A partir desses dados foram extraídas outras variáveis meteorológicas importantes para realizar análises em cada estação. Essas variáveis estão descritas na Tabela 3.2. Os dados de refletividade foram extraídos a partir do CAPPI para as alturas de 2 a 7 km de modo evitar problemas de "cone cego", *cluter* de superfície e a má representatividade dos feixes do radar fora dessas alturas (RINEHART, 1991). Esses problemas, assim como a análise da altura do feixe em relação à sua distância para esse radar foram explorados no trabalho de Biscaro (2019).

Figura 3.3 - Taxa de chuva horária acumulada nas estações chuvosa (a) e seca (b). Os valores de taxa de chuva computados foram aqueles superiores a 0,5 mm/h e inferiores a 160 mm/h (machado2018overview).



Fonte: Produção Autor.

Contudo, algumas variáveis como máxima refletividade em cada pixel foi obtida a partir de toda a extensão vertical do escaneamento do radar, ou seja, para os 40 produtos CAPPI presentes nas altura de 0,5 a 20 km. Este último se deu devido ao fato que alguns núcleos mais intensos poderiam estar em alturas mais altas e não poderiam deixar de ser contabilizados. A taxa de chuva foi obtida a partir do produto

CAPPI de 2,5 km de altura, utilizando a relação ZR com os valores dos parâmetros  $a$  e  $b$  definidos de acordo com a documentação dos dados de RADAR disponíveis no repositório do GoAMAZON/ARM: 174,8 e 1,56, respectivamente (ARM, 2018). O conteúdo de água líquida integrada verticalmente (VIL, em  $kg/m^2$ ) foi calculado de acordo com a Equação 2.5, onde os limites da camada atmosférica  $a$  e  $b$  são os mesmos utilizados para os valores de refletividade, ou seja, entre as alturas de 2 a 7 km. As variáveis da altura máxima de 35 dBZ, altura de Waldvogel e variação do topo da nuvem também foram calculadas entre as alturas de 2 a 7 km. Maiores detalhes sobre estas variáveis podem ser encontradas na Seção 2.3.

Tabela 3.2 - Variáveis calculadas a partir dos dados do radar meteorológico.

VARIÁVEL	UN. MEDIDA	DESCRIÇÃO
Refletividade	dBZ	A refletividade radar dos produtos CAPPI entre 2 e 7 km
Máxima refletividade	dBZ	A máxima refletividade foi calculada entre os produtos CAPPI 2 e 7 km.
Taxa de chuva	mm/h <sup>-1</sup>	A taxa de chuva foi computada a partir do produto CAPPI 2,5 km, utilizando a relação: $Z = 174,8R^{1,56}$ .
VIL	kg/m <sup>2</sup>	O VIL foi computado a partir do produtos CAPPI entre as alturas de 2 e 7 km.
VILD	g/m <sup>-3</sup>	O VILD foi computado a partir do VIL, entre 2 e 7 km.
Altura máxima de 35 dBZ	km	A altura máxima de 35 dBZ foi computada entre níveis de 2 e 7 km.
Altura de Waldvogel	km	A altura Waldvogel foi computada entre níveis de 0,5 e 20 km.
Variação do topo da nuvem	dh/dt	A variação do topo da chuva foi calculada a partir de duas imagens consecutivas do radar no produtos CAPPI entre as alturas de 0,5 e 20 km.

### 3.2.2 Dados dos sensores de VLF

A ocorrência ou propagação das descargas elétricas atmosféricas são responsáveis por emitirem radiação eletromagnética que variam desde frequências muito baixas (VLF - sigla do inglês, *very low frequency*) até raios gama (MORALES, 2001; RAKOV, 2016). O ruído natural que acontece na faixa de 3-30 kHz consiste em impulsos eletromagnéticos causados pela atividade elétrica da atmosfera da terra, esses impulsos podem ser detectados por alguns sensores e são denominados *sferics*, estando geralmente associados aos raios (LEWIS et al., 1960). A STARNET é uma rede de detecções de raios que possui sensores operando em VLF, capaz de medir continuamente o campo elétrico vertical na faixa de 7 kHz a 15 kHz, podendo detectar até 130 *sferics* por segundo (MORALES et al., 2014).

A STARNET possui antenas receptoras sincronizadas temporalmente por GPS em diversas regiões brasileiras, países da América Latina e África, sendo mantida desde 2006 pelo Laboratório de Sensoriamento Remoto de Tempestades (STORM-T) da Universidade de São Paulo (USP) e desde 2011 detecta raios sobre toda a América do Sul, auxiliando na elaboração de estudos e fornecendo informações de alta resolução temporal e espacial para as mais diversas aplicações associadas às descargas elétricas (MORALES, 2018). Com a cobertura de até 7000 km, a STARNET pode detectar essencialmente raios do tipo CG, e sua precisão estimada é entre 1,9 e 5,5 km dentro do território brasileiro e de 5 a 10 km nas extremidades da América do Sul 2 a 2 km com eficiência de detecção entre 32,2 e 39,2% durante o dia e até 89% no período noturno.

As formas de onda capturadas por diferentes sensores da STARNET, pertencentes a um mesmo raio, são comparadas através de pares distintos a fim de calcular a correlação temporal cruzada e aquelas superiores a 0,95 são consideradas como um raio válido e, diante disso, é aplicado o método para estimar a sua localização (LEE, 1986; MORALES, 2018). O método utilizado pela STARNET para identificar as posições dos raios é baseada na diferença do tempo de chegada (ATD - sigla do inglês *arrival time difference*), metodologia proposta por Lee (1986). Esse método utiliza como referência espacial as coordenadas geográficas onde os sensores da rede estão instalados, e a partir da diferença do atraso de um sinal válido recebido por dois sensores são obtidas funções hiperbólicas, em que o ponto de intersecção entre elas define a posição em que o raio aconteceu (LEE, 1986; MORALES, 2018).

Os *sferics* computados pela STARNET são identificados neste trabalho como *strokes*. Uma única descarga elétrica pode emitir múltiplos pulsos elétricos (LEWIS et al., 1960), podendo ser computado pela STARNET como *strokes* diferentes, sendo assim, com o intuito de representar melhor as descargas elétricas distintas, foram criados a partir desses dados os *flashes*. Os *flashes* foram criados a partir de um agrupamento espaço-temporal dos *strokes*, seguindo os critérios considerados por Cummins et al. (1998), Murphy e Nag (2015) e Lopes (2019). Na criação dos *flashes*, os agrupamentos temporais dos *strokes* foram realizados com o intervalo de até 500 ms e a partir desses grupos realizado o agrupamento espacial com o raio de 2,5 km a partir dos *strokes* agrupados temporalmente.

### 3.2.2.1 Agrupamento dos *strokes* em *flashes*

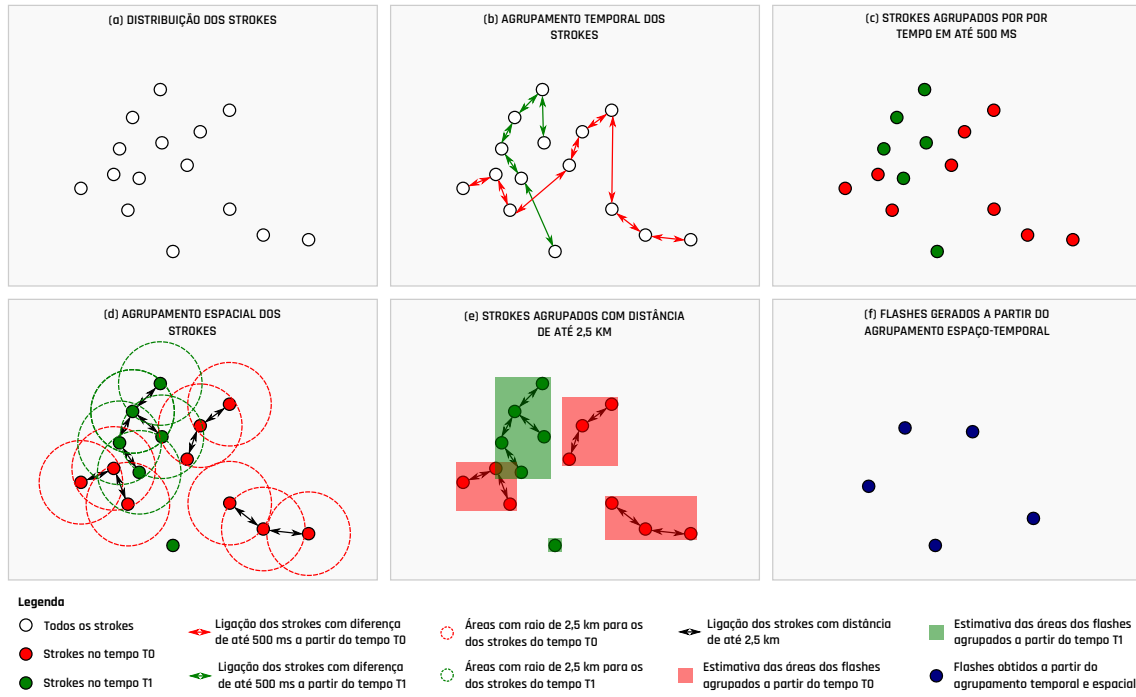
O Algoritmo de agrupamento utilizado na criação dos *flashes* foi o *Density-Based Spatial Clustering of Application with Noise* (DBSCAN) (ESTER et al., 1996). O DBSCAN é um algoritmo de aprendizado de máquina não supervisionado que tem como intuito criar agrupamentos a partir de um conjunto de dados de  $n$  dimensões. Esse algoritmo consiste em realizar agrupamentos a partir da densidade de pontos em uma vizinhança, que é definida por uma distância máxima entre os pontos. O algoritmo inicia selecionando aleatoriamente um ponto qualquer, calcula sua vizinhança a partir de um raio predefinido a partir dele e repete esse processo de forma recursiva com o ponto mais próximo. Os principais parâmetros desse algoritmo são o  $\epsilon$  e  $m$ , indicando, respectivamente, o raio máximo da vizinhança e a quantidade mínima de pontos para formar um grupo.

Os *flashes* foram criados de acordo as etapas descritas no exemplo da Figura 3.4. A Figura 3.4(a) representa um exemplo de um conjunto de *strokes* computados pela STARNET. A Figura 3.4(b) mostra o agrupamento temporal unidimensional dos *strokes* e a Figura 3.4(c) os dois grupos que foram obtidos a partir deste processamento. A Figura 3.4(d) mostra o processo de agrupamento espacial, onde as circunferências possuem raio de 2,5 km. A Figura 3.4(e) mostra a delimitação espacial do segundo agrupamento e por fim, a Figura 3.4(f) mostra os pontos que representam os *flashes* provenientes do centroide do retângulo envolvente ilustrado na etapa anterior. Já a Figura 3.5 mostra as distribuições da multiplicidade de *strokes* por *flash* para as estações chuvosa e seca, ou seja, a distribuição da quantidade de *strokes* que foram agrupados para gerar um *flash*. A multiplicidade mostra uma visão geral da conversão de como os dados foram agrupados, e que pode ser utilizada para nortear a escolha dos valores dos parâmetros de agrupamento espacial e



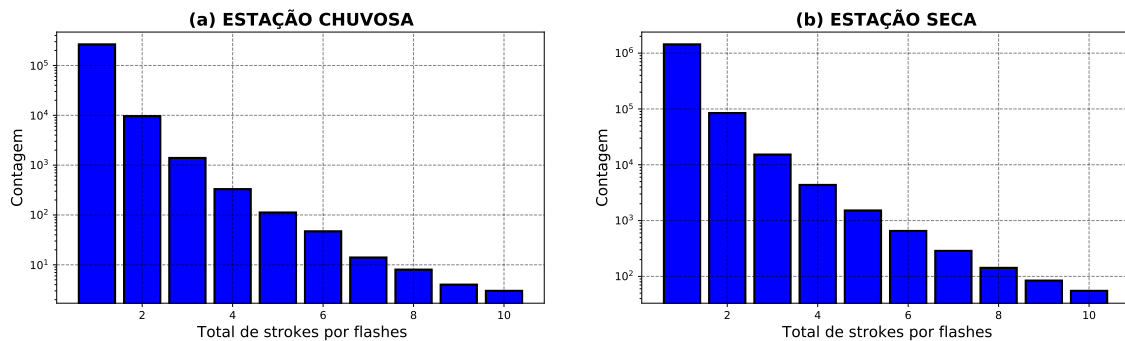
temporal do algoritmo (CUMMINS et al., 1998), descritos na Figura 3.4.

Figura 3.4 - Etapas para a conversão de *strokes* em *flashes*.



Fonte: Produção Autor.

Figura 3.5 - Distribuição da multiplicidade de *strokes* por *flash* nas estações (a) chuvosa e (b) secas.



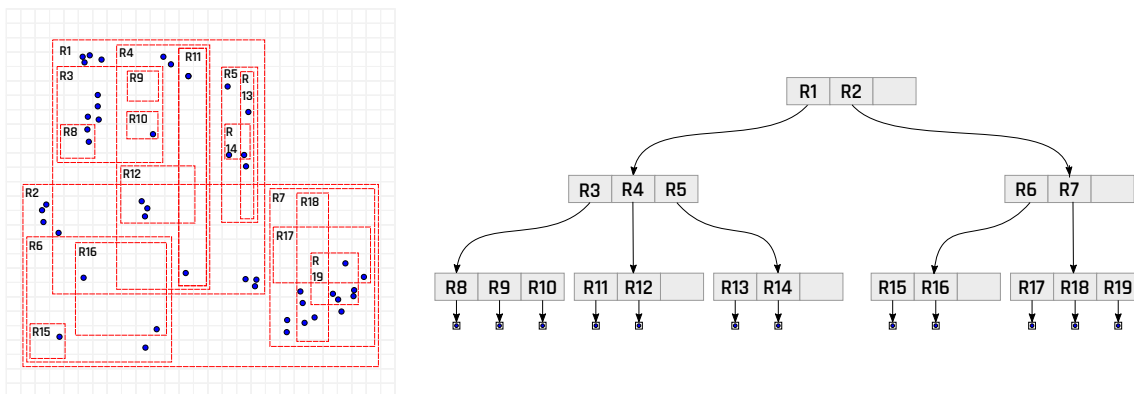
Fonte: Produção Autor.

### 3.2.2.2 Indexação espacial para o mapeamento do raios

Como mencionado anteriormente, os dados do radar estão dispostos em uma grade horizontal  $241 \times 241$  pixels. Devido a esse formato matricial dos dados, as análises de múltiplas variáveis para uma mesma área são feitas a partir de suas referências espaciais representadas pelos índices da grade. Entretanto, originalmente os dados do sensor VLF são disponibilizados no formato vetorial, sendo necessário fazer a indexação dos *strokes* e *flashes* em uma matriz com a mesma dimensão das outras variáveis.

Devido à grande quantidade de dados, o custo computacional necessário para sua conversão de vetorial para matricial de forma convencional seria muito alto, sendo assim, buscou-se a utilização de índices espaciais para realizar essa tarefa. Os índices espaciais são estruturas de dados que permitem o acesso rápido a dados espaciais, e são comumente utilizados em bancos de dados geográficos na realização de consultas e acesso aos dados de forma eficiente (MANOLOPOULOS et al., 2009). O índice espacial utilizado neste trabalho foi a R-Tree, metodologia desenvolvida por Guttman (1984).

Figura 3.6 - Esquema da R-Tree.



Fonte: Adaptado de Guttman (1984).

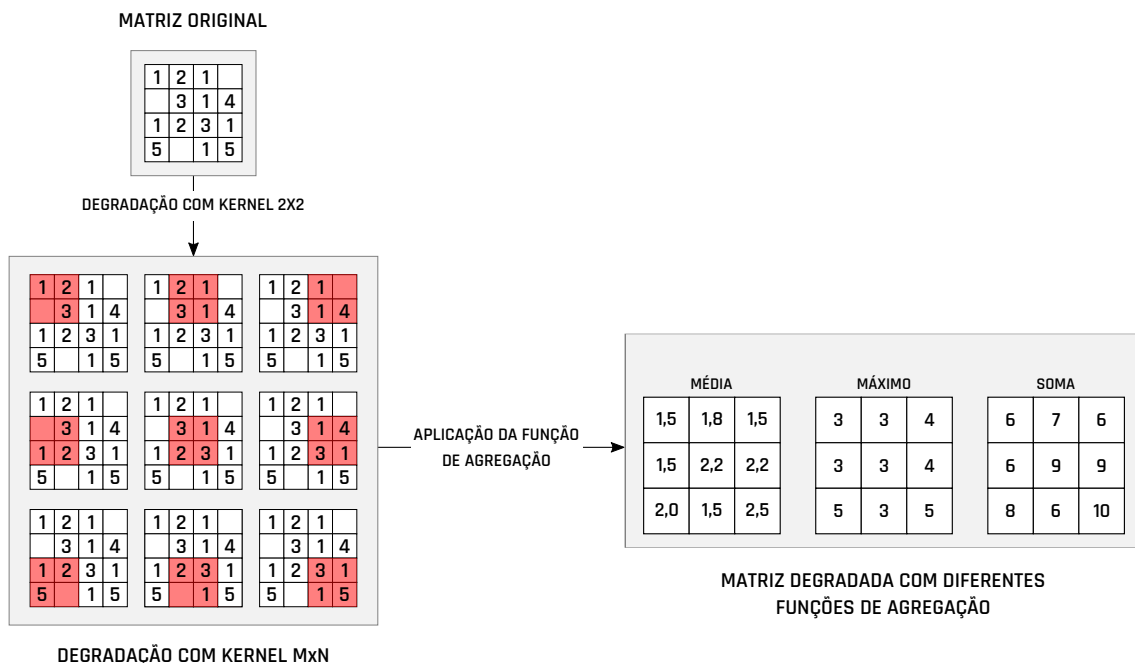
A R-Tree é uma estrutura de dados em árvore que possui altura balanceada e tem como premissa agrupar objetos próximos através de mínimos retângulos envolventes. Os mínimos retângulos envolventes são representados pelos nós da árvore, de tal forma que o nó raiz seja o mais externo abrangendo mais objetos e nós folhas o mais interno delimitando um único objeto. A Figura 3.6 mostra o esquema do

funcionamento da R-Tree em um exemplo de mapeamento dos raios na matriz. No lado esquerdo da figura, os pontos azuis representam os raios e as caixas tracejadas em vermelho representam os mínimos retângulos envolventes, neste exemplo pode ser observado que os pixels que possuem ausência de raios, não são considerados pelo algoritmo. Já a estrutura à direita é a representação parcial em forma de árvore do enquadramento gerado pela R-Tree.

### 3.2.3 Mapeamento baseado na degradação dos dados

Os dados das observações do radar e dos sensores VLF foram co-localizados em uma grade regular, de tal forma que um mesmo pixel possa apresentar informações diferentes, sejam das variáveis obtidas a partir da refletividade ou dos produtos provenientes da STARNET. Por se tratar de fontes diferentes, a co-localização pode, em alguns casos, perder a sua referência espacial, pois além de serem de formatos espaciais diferentes (matriciais e vetoriais), ainda tem o erro associado à acurácia dos sensores. Sabendo disso, foram aplicadas algumas degradações nas matrizes, para que os ruídos fossem suavizados com a menor perda de informações possível.

Figura 3.7 - Esquema do processo de degradação da matriz.



Fonte: Produção Autor.

A estratégia de degradação utilizada é baseada nas camadas de *pooling* presentes na arquitetura das redes neurais convolucionais (CNN) (LECUN et al., 1998). Nas CNN, as camadas de *pooling* são complementares ao processo de convolução da rede e são responsáveis pela extração dos mapas de características dos dados através do deslocamento de um *kernel* associado à uma função de agregação. Sendo assim, a abordagem utilizada consiste em aplicar *kernels* de diferentes tamanhos e funções de agregação nas matrizes originais, conforme mostrado no esquema da Figura 3.7. Neste trabalho, os *kernels* são abstrações que agrupam em diferentes tamanhos as células das matrizes, e então aplicam alguma função de agregação, que Figura 3.7 são representados pela área em vermelho. Para todos os dados foram geradas matrizes utilizando os *kernels* de tamanhos  $2 \times 2$ ,  $3 \times 3$ ,  $4 \times 4$  e  $5 \times 5$ , conseqüentemente, tendo um agrupamento espacial das informações de no máximo  $100 \text{ km}^2$ . O *kernel* é deslocado por colunas e linhas, respectivamente, saltando um único índice. Para os dados do radar foi aplicada a função de agregação média, então para cada iteração da degradação é obtida a média dos valores das matrizes correspondente à área do *kernel* e adicionada na matriz de saída. Para as matrizes correspondente aos raios, além da média, também foram aplicadas as funções de agregação de máximo valor e soma.

### 3.2.4 Índices de co-localização e correlação

O mapeamento espacial dos dados de diferentes sensores deve estar coerente com as condições reais, para que sejam mitigados o máximo possível os ruídos na etapa de classificação e previsão. Então, para avaliar a qualidade dos mapeamentos foram aplicados os índices de co-localização (ICOL) e correlação (ICOR), atributos que serão de suma importância para a seleção de atributos utilizados durante o treinamento dos modelos de ML.

O valor do índice de co-localização é baseado nos pixels coincidentes em que foram detectados raios e qualquer outra variável do radar. O índice de co-localização é dado pela Equação 3.1, em que  $A$  é o conjunto das coordenadas dos pixels onde foram detectados raios e  $B$  é o conjunto das coordenadas dos pixels onde foram detectados valores para a variável do radar. Através do índice de co-localização é possível definir quais variáveis estão mais presentes na relação espacial entre o mapeamento das observações nos diferentes sensores, de tal forma que o valor 1 representa máxima relação e 0 nenhuma.

$$ICOL = \frac{|A \cap B|}{|A|} \quad (3.1)$$

Já o índice de correlação é medido pelo coeficiente de Pearson, dado pela Equação 3.2, em que  $\bar{x}$  e  $\bar{y}$  são as médias aritméticas de cada uma das variáveis analisadas e  $n$  é a quantidade de observações. O índice de correlação foi utilizado para avaliar se há relação entre os raios e as variáveis do radar, e caso haja, se essa relação é inversa ou não. O que pode ajudar a identificar possíveis relações de intensificação ou decaimento.

$$ICOR = r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (3.2)$$

### 3.2.5 Classes de cobertura de solo e orografia

Além das variáveis meteorológicas, também foram utilizados dados associados à superfície, dados de cobertura de solo e orografia. O dado de cobertura do solo é uma matriz bidimensional com rótulos para 16 classes de coberturas e uso do solo que incluem hidrografia, florestas, áreas desmatadas, áreas de mineração, agricultura, entre outras. As classes de cobertura e uso do solo são provenientes do mapa criado por Almeida et al. (2016) a partir de imagens dos sensor MODIS do satélite Landsat-5/TM.

Os dados com os efeitos orográficos foram obtidos a partir do modelo de elevação digital da missão SRTM e originalmente possuem resolução espacial de 90 metros no Equador. Neste trabalho, assim como as demais variáveis, os dados de orografia e cobertura do solo, foram interpolados e também estão dispostos em uma matriz bidimensional de  $241 \times 241$ , onde cada pixel possui a estimativa da altura (em metros) para aquela área e tipo de uso predominante, respectivamente.

## 3.3 Metodologia

O primeiro passo desta pesquisa foi realizar uma análise exploratória dos dados. Foram levados em consideração a distribuição espacial e temporal (sazonalidade e ciclo diurno) de todas as variáveis envolvidas, os efeitos de superfície, como a orografia local, a vegetação e a rede fluviométrica da região. Para complementar essa análise, foi aplicada a técnica de árvore de decisão, a fim de identificar as variáveis mais importantes na separação dos dados com inatividade e atividade de raios. Os

resultados dessas análises serviram como subsídios para a definição das configurações das entradas dos modelos de previsão baseados em redes neurais artificiais. Uma vez criado o modelo de previsão, validações foram realizadas para determinar o grau de confiabilidade do modelo a partir de seu desempenho. Nas seções seguintes serão apresentadas as etapas de execução dessas metodologias.

### 3.3.1 Análise espacial das descargas elétricas

Como susodito, um dos primeiros passos desta pesquisa é a análise da distribuição espacial das descargas elétricas e as outras variáveis envolvidas nesta pesquisa sobre a região de estudo. O intuito desta análise exploratória dos dados é verificar a variabilidade espacial e temporal das descargas elétricas com relação as diferentes informações disponibilizadas pelo radar meteorológico e dados de superfície. Como citado anteriormente, os períodos chuvoso, seco e suas transições na Amazônia apresentam características distintas de regime de chuva (MACHADO et al., 2018). Tais variações são em função não só da própria sazonalidade dos eventos de tempestades, mas também devido aos efeitos locais da própria floresta, rios e da cidade de Manaus na microfísica e termodinâmica dessas nuvens e ambiente circundante. Além disso, identificar e analisar a espacialidade destes eventos em cada município, tipo de superfície e orografia foram de suma importância para descrever como os efeitos locais são assimilados pelos algoritmos de aprendizado de máquina. Estas informações também foram importantes para definir as variações sazonais da ocorrência de descargas elétricas e definir também os limiares que utilizados para determinar a classificação da densidade de raios, como será discutido posteriormente.

Inicialmente, foram criados gráficos e mapas temáticos com a densidade de raios em cada ponto de grade em função de diferentes *shapefiles*, como os de municípios, tipo de vegetação e orografia para verificar sua distribuição. Além disso, foram analisadas também as variáveis do radar que serão utilizadas como preditoras de eventos de raios de modo a determinar sua relação. Além da análise visual dos histogramas e *boxplots*, foram também consideradas algumas informações estatísticas, como média, mediana, desvio padrão, variância, assimetria e curtose, algumas delas apenas para consulta no apêndice.

Esperou-se que, após a análise exploratória, o uso de técnicas de aprendizado de máquina (seções posteriores), junto as essas variáveis escolhidas, nos forneçam uma base estatística para definir alguns processos físicos que estejam associados a esta associação de variáveis, de modo a extrairmos ajustes locais e novas relações físico-empíricas que represente a natureza do sistemas geradores de descargas elétricas.

### 3.3.2 Classificação dicotômica dos raios

Uma vez realizada a análise exploratória dos dados, foi possível definir algumas variáveis que apresentam certa discriminação sobre a ocorrência ou não de descargas elétricas nuvem solo. A classificação da ocorrência de raios será simples e dependerá da relação entre as variáveis do radar e a própria descarga. Apenas duas classes foram criadas inicialmente:

- Classe 1 (Sem atividade) - Essa classe representa os registros em que não foram detectadas descargas elétricas em seu intervalo de tempo. Essa classe é importante porque não só o aumento de raios é considerado na previsão, mas também a cessação deles.
- Classe 2 (Com atividade) - Essa classe representa os registros em que ocorreram pelo menos 1 raio, indicando a atividade de descargas elétricas no pixel.

Devido à natureza dos dados, existe de ante mão um grande desbalanceamento entre classes (cerca de 80%), haja vista que a quantidade de registros sem detecção de raios é muito superior que a quantidade de registros com detecção. Além disso, o número de eventos nuvem-solo são ainda menos frequentes se comparados ao tipo intra-nuvem, o que muitas vezes impede sua ampla classificação em diferentes classes, o que ajudaria na definição da intensidade de raios em diferentes faixas de intensidade. As classes acima descritas serão o alvo dos modelos de classificação e previsão aqui desenvolvidos. Contudo, antes de entrar nas questões associadas as técnicas de aprendizado de máquina, é preciso apresentar quais variáveis serão inicialmente utilizadas para determinar a classificação dessas classes de raios.

### 3.3.3 Seleção das variáveis meteorológicas

Para que seja possível avaliar a criação de um sistemas de previsão de raios é necessário selecionar as variáveis inferidas a partir dos dados de radares meteorológicos que podem ser associadas às tempestades severas e indiquem a tendência de ocorrência das descargas elétricas. A Tabela 3.2 apresenta todas as variáveis que foram utilizadas ao longo deste estudo, e supracitadas no capítulo anterior. A seleção dessas variáveis foi feita através de uma das abordagens de aprendizado de máquina e da análise exploratória dos dados. A árvore de decisão foi um dos métodos de aprendizado de máquina aplicado para a seleção de variáveis, como veremos em maiores detalhes na seção seguinte.

A seleção das variáveis na árvore de decisão é feita por consequência da construção dos nós que compõe sua estrutura. Esses nós são criados a partir da definição automática de limiares nos valores das variáveis que proporcionam o maior ganho de informação com base em alguma função. Esse processo é feito de forma recursiva até atingir alguma condição de parada definida previamente ou até que não seja mais possível fazer a separação dos dados (QUINLAN, 1986). Nessa abordagem é atribuído um grau de importância para aquelas variáveis que proporcionam os melhores ganhos de informação, ao fazer a separação dos dados na construção dos nós.

Um dos métodos de seleção dos dados de entrada das configurações de RNAs testadas, foi feito de acordo com os graus de importância atribuídos às variáveis pelas árvores de decisão construídas para classificação de inatividade e atividade de raios. Outros critérios utilizados para a seleção de variáveis foram com base na análise exploratória dos dados. Através das análises de gráficos e da extração de informações estatísticas, foram identificadas as variáveis de radar que possuem melhor relação com os dados de raios. Todos esses aspectos foram considerados na definição de algumas configurações dos modelos de previsão testados. Além disso, foram realizados também testes de sensibilidades com todas as variáveis, a fim de encontrar as melhores combinações que melhorassem o desempenho dos modelos baseados em RNAs. Ou seja, as configurações com as diversas combinações dos dados de entrada foram executadas, de modo a encontrar a melhor combinação baseada nos índices de validação listados em seção posterior. O método de classificação será melhor descrito na seção que segue.

### **3.3.4 Método de classificação**

Como dito anteriormente, o método de classificação teve como finalidade definir automaticamente a ocorrência de descargas elétricas com base nos valores das variáveis meteorológicas obtidas a partir dos dados de radar. Essa classificação foi aplicada para contribuir a seleção de variáveis, para isso foi utilizado o algoritmo de aprendizado de máquina árvore de decisão. Sua escolha se deu pela fácil interpretação de seus resultados, e um modo mais intuitivo de entender a física envolvida nas escolhas automáticas através das suas representações.

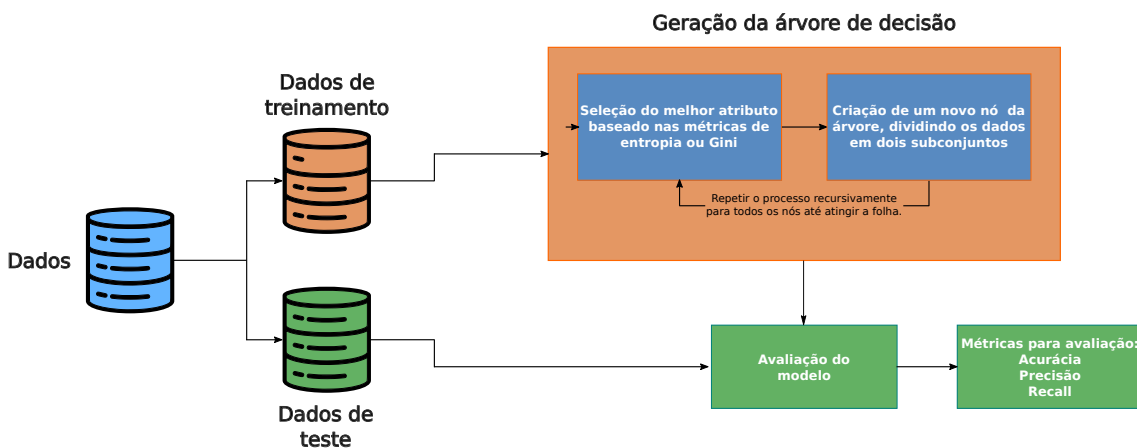
#### **3.3.4.1 Árvores de decisão**

As árvores de decisão são algoritmos de aprendizado de máquina supervisionado, ou seja, necessitam que seu treinamento seja feito com dados rotulados. Os primei-



ros conceitos acerca desta abordagem, foram propostos principalmente por [Quinlan \(1986\)](#). Embora possam ser utilizadas também para tarefas de regressão, as árvores de decisão são comumente utilizadas para classificação, sendo essa, a aplicação desse algoritmo neste trabalho. A Figura 3.8 mostra o esquema do funcionamento da árvore de decisão para a tarefa de classificação.

Figura 3.8 - Esquema do funcionamento da árvore de decisão.



Fonte: Produção Autor.

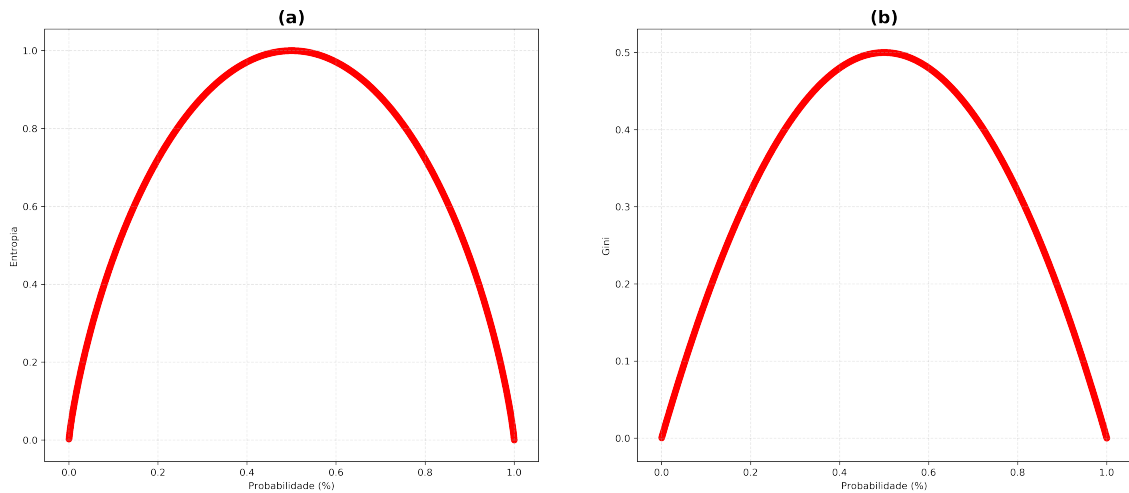
O funcionamento base deste algoritmo para a tarefa de classificação, consiste em realizar cortes ortogonais em um plano cartesiano através de uma estrutura de dados modelada como árvore. Os cortes ortogonais definidos pelo o algoritmo, indicam a criação de um novo nó na árvore de decisão e, conseqüentemente, a relevância das variáveis para o modelo. Nos casos em que a árvore é utilizada para classificação, os critérios mais utilizados para realizar as partições são baseados na entropia e índice Gini ([ONODA; EBECKEN, 2001](#)).

A entropia baseia-se em uma medida utilizada na teoria da informação que tem como finalidade calcular o ganho de informação na separação de dados heterogêneos ([SILVA, 2005](#)). Sendo assim, esse critério mede a falta de homogeneidade dos dados em relação à sua classe. O melhor ganho de informação para esse critério acontece quando o valor de entropia se aproxima de zero, o que acontece quando a probabilidade de um elemento escolhido aleatoriamente pertencer à uma classe for próximo de 0% ou 100%, conforme mostrado na Figura 3.9(a). A entropia é definida pela Equação 3.3, onde  $c$  é o total de classes e  $p_j$  é a probabilidade do elemento  $p$

pertencer à classe  $j$ .

$$Entropia = \sum_j^c p_j \log_2 p_j \quad (3.3)$$

Figura 3.9 - Variação da entropia e índice *Gini* em relação a probabilidade de um elemento pertencer à uma classe do total de duas.



Fonte: Produção Autor.

O ganho de informação baseado no índice Gini indica o grau de pureza das classes em relação à frequência em que um elemento escolhido aleatoriamente é classificado incorretamente (ONODA; EBECKEN, 2001). O índice Gini é definido pela Equação 3.4, onde  $c$  é o total de classes e  $p_j$  é a probabilidade do elemento  $p$  pertencer à classe  $j$ . Nessa abordagem, quanto mais próximo de zero o valor estiver, melhor será o ganho de informação, enquanto que, quando a probabilidade de um elemento pertencer a uma classe é balanceada com as demais, pior será o valor do índice, ou seja, menos pura será a separação dos dados, conforme mostrado na Figura 3.9(b).

$$Gini = 1 - \sum_j^c p_j^2 \quad (3.4)$$

Segundo Silva (2005), quando as árvores de classificação com partições binárias utilizam o critério do índice Gini, os registros da classe mais frequente tende-se a se

isolar em um ramo da árvore, enquanto que, quando se utiliza o critério de entropia para esse cenário, o número de registros são balanceados em cada ramo.

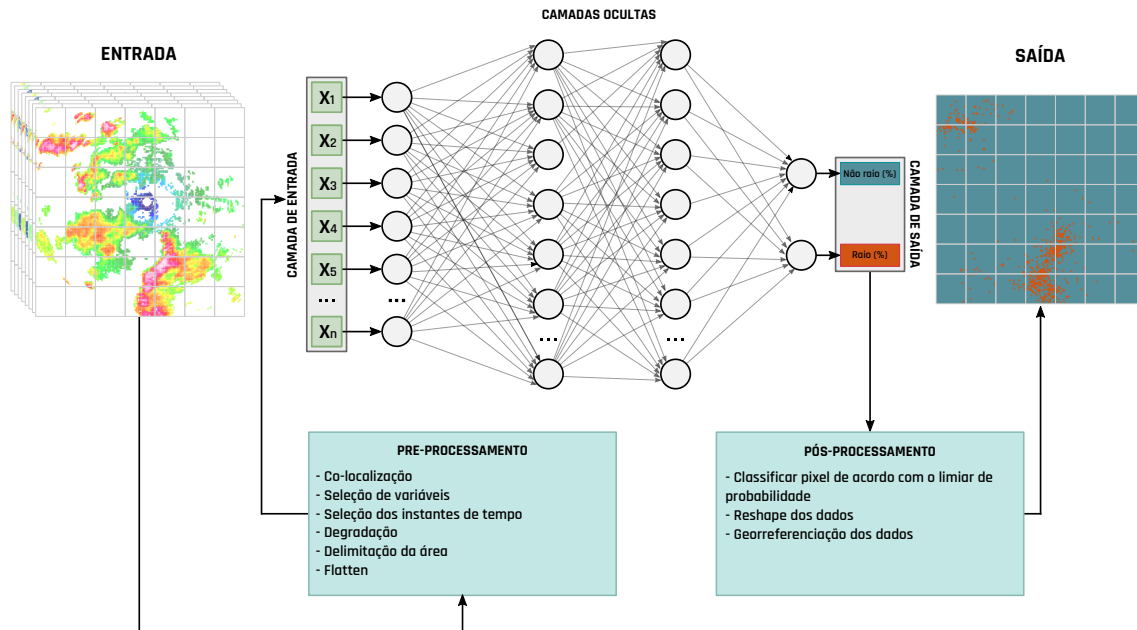
Uma vez classificadas as atividades de raios e definidas as variáveis que podem ser utilizadas para determinar o surgimento/manutenção e a cessação de raios (conforme serão apresentados no próximo capítulo), o próximo passo foi a previsão.

### 3.3.5 Previsão de raios a curto prazo

As previsões das descargas elétricas foram feitas com a utilização de redes neurais artificiais. A escolha das RNAs se dá pela sua capacidade de generalização na criação de relações complexas entre variáveis preditoras e variáveis alvos. Como a variabilidade dos raios é alta, tanto espacial como temporal, as técnicas clássicas muitas vezes não conseguem ter representação bem definida de eventos menos frequentes, como o raios nuvem-solo. Diante disso, as redes neurais se tornam uma alternativa viável para a realização de tais tarefas.

A rede neural utilizada neste trabalho é a *perceptron* multicamadas (MLP) (GARDNER; DORLING, 1998). A escolha dessa arquitetura se deu devido à sua alta capacidade de generalização na criação de relações complexas em tempo hábil. As diversas configurações de entrada testadas neste trabalho variam de acordo com a quantidade de variáveis e instantes de tempos que serão fornecidos. Os testes feitos com as diferentes configurações de entrada foram padronizados, de tal forma que fossem executados com os mesmos dados e arquitetura, mudando apenas a camada de entrada da MLP após o ajuste dos hiper-parâmetros. As previsões obtidas a partir desses modelos se deram pelos valores probabilísticos fornecidos pela camada de saída. As variáveis utilizadas para treinar a rede estão descritas na Tabela 3.2.

Figura 3.10 - Esquema do processo de previsão das descargas elétricas.



Fonte: Produção Autor.

A Figura 3.10 mostra o esquema do processo para realizar as previsões com as diferentes configurações de entrada da MLP. Uma estrutura de dados foi criada para agrupar os conjuntos de treinamento, validação e teste. Cada estrutura possui todas as informações nos diferentes instantes de tempo. Todas as configurações passam por um etapa de pré-processamento antes da execução da rede. Esse pré-processamento tem como intuito preparar os dados de acordo com as parametrizações da camada de entrada do modelo. Inicialmente é feita a co-localização dos dados, em que cada índice da estrutura de dados fique associado com as diferentes informações de seus respectivos pixels. Em seguida é feita a seleção dos índices da estrutura de dados em que estão alocadas os valores das variáveis e instantes de tempo referentes às parametrizações de entrada da rede. Caso se aplique, é feita a degradação (um teste específico) dos dados de acordo com as configurações do *kernel* indicado, e então feita a extração dos pixels da área delimitada. Por fim, é feita a transformação dos dados para unidimensional e apresentados à camada de entrada da rede. Para a conclusão da previsão, a saída da rede é submetida a um pós-processamento. A classificação dicotômica da atividade de raios é feita de acordo com a superação do limiar de probabilidade definido na parametrização da configuração. Em seguida os

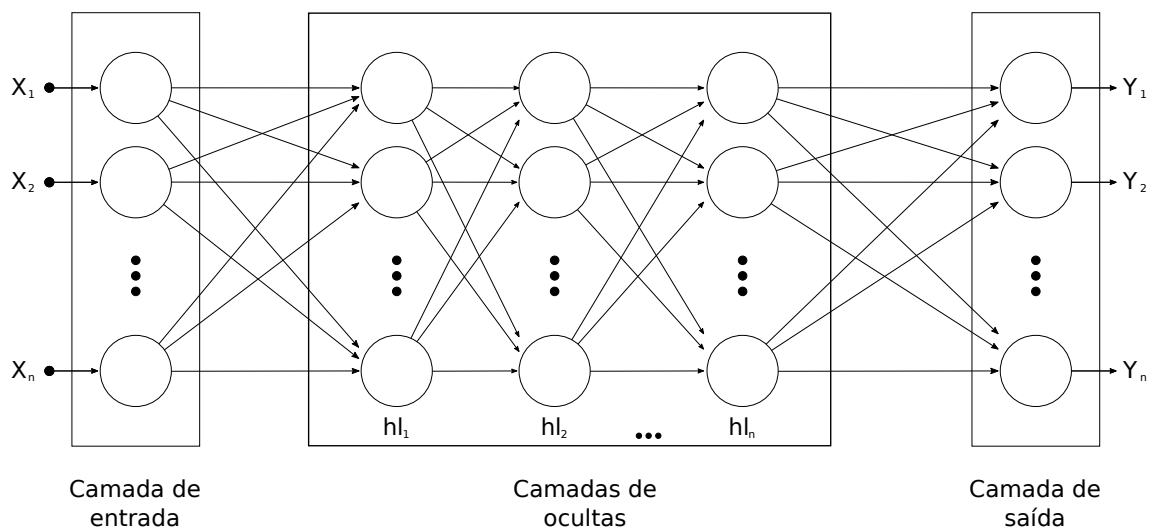
dados são revertidos estruturas bi-dimensionais georreferenciadas.

A seguir o funcionamento das redes *perceptron* multicamadas serão descritas com mais detalhes, assim como as principais configurações de hiper-parâmetros utilizadas neste trabalho.

### 3.3.5.1 *Perceptron* multicamadas

As redes neurais utilizadas neste trabalho terão como finalidade avaliar parametrizações associadas às variáveis meteorológicas para previsão de raios a curto prazo. As redes utilizadas são baseadas na arquitetura *perceptron* multicamadas. Como já introduzido em seções anteriores, a MLP é uma rede neural artificial que pode ser utilizada para tarefas de classificação e regressão. Essa rede pertence à classe de redes neurais artificiais do tipo *feedforward*, ou seja, as conexões entre os seus nós não formam ciclos (JAIN et al., 1996). Essa rede surgiu com a proposta de superar a limitação de seu antecessor, o *perceptron* de única camada, que era capaz de resolver somente problemas linearmente separáveis. A MLP é constituída por uma camada de entrada, uma camada de saída e entre essas duas um número arbitrário de camadas ocultas, conforme mostrada na Figura 3.11.

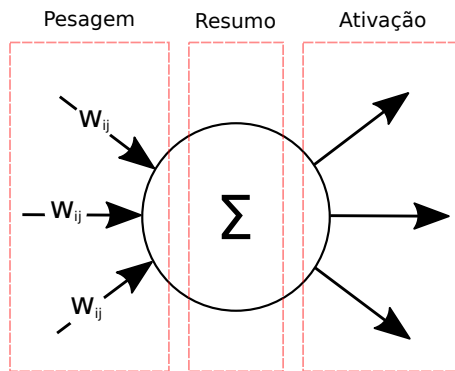
Figura 3.11 - Esquema da arquitetura da rede *perceptron* multicamadas.



Fonte: Produção Autor.

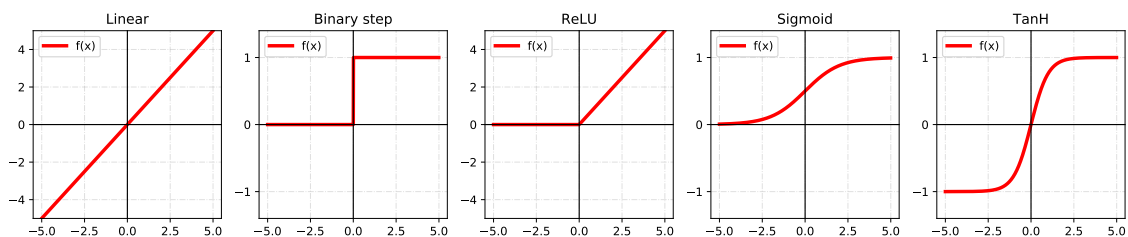
A camada de entrada recebe os atributos previsores ( $X_i$ ) e passa para as camadas ocultas. As camadas ocultas são compostas por neurônios artificiais (Figura 3.12), que realizam um processo dividido em três etapas: pesagem, resumo e ativação (OBERNDORFER, 2017). Na etapa de pesagem, cada sinal é multiplicado pelo seu respectivo peso da aresta. Na etapa de resumo, é realizado o somatório dos pesos e acrescentado um bias. Por fim, o valor obtido é submetido à uma função de ativação, e o resultado é transferido para os neurônios da próxima camada. Os gráficos das funções de ativação mais utilizadas estão apresentados na Figura 3.13.

Figura 3.12 - Esquema do neurônio artificial.



Fonte: Produção Autor.

Figura 3.13 - Gráfico das principais funções de ativações.



Fonte: Adaptado de Oberndorfer (2017).

O treinamento da MLP é feito de forma supervisionada. Sendo assim, as informações apresentadas pelo modelo na camada de saída são comparadas com o valor real a

cada iteração. Essas iterações são chamadas de épocas, em que pode ser executadas com lotes de dados. Para cada época é calculado o erro entre o valor real e o valor apresentado pelo modelo. Esse erro é calculado com base em alguma função que depende do tipo de tarefa que está sendo executada (e.g. regressão ou classificação). Para isso existem funções que calculam o erro que são mais apropriadas para cada caso. Tipicamente o erro calculado por essa função é utilizado para a atualização dos pesos das arestas em um processo denominado *backpropagation* (WERBOS, 1974; RUMELHART et al., 1986). Embora hajam outras formas de atualizações dos pesos das MLPs, o *backpropagation* é uma das técnicas mais difundida (JAIN et al., 1996).

Embora hajam grandes esforços para auxiliar na definição de hiper-parâmetros mais apropriados para a criação de modelos de aprendizado de máquina (BENGIO, 2000; SMITH, 2018; RENDA et al., 2020), a sua escolha às vezes está diretamente associada com o tipo de problema e dados que estão sendo utilizados. Diante disso, a configuração dos hiper-parâmetros foram definidas partir de testes de sensibilidade. Com relação aos parâmetros da arquitetura da rede, foram testadas diferentes funções de ativação, parâmetros de otimização, número de camadas ocultas e neurônios. Para o treinamento os parâmetros testados foram a função de erro, número de épocas e tamanho do lote. Os parâmetros definidos para as MLPs estão apresentados na Tabela 3.3. As combinações foram escolhidas de acordo com a avaliação das métricas de validação descritas na Subseção 3.3.7.

Tabela 3.3 - Hiper-parâmetros definidos para as configurações das MLPs.

<b>HIPER-PARÂMETRO</b>	<b>ESCOLHIDO</b>
Nº camadas ocultas	2
Nº neurônios na 1ª camada oculta	22
Nº neurônios na 2ª camada oculta	44
Função de ativação na 1ª camada oculta	<i>Sigmoid</i>
Função de ativação na 2ª camada oculta	<i>ReLU</i>
Função de ativação na camada de saída	<i>Softmax</i>
Otimizador	adam
Nº de épocas	150
Tamanho do lote	32
Função de erro	<i>Focal Loss</i>

Os testes de sensibilidade partiram da execução com o número de camadas ocultas variando entre 1 e 10. A quantidade de neurônios nessas camadas, inicialmente foi variada entre 1 e 100, com base nas melhores combinações, foram testadas algumas configurações de forma subjetiva, resultando nos valores apresentados na Tabela 3.3. As funções de ativação testadas nas camadas ocultas foi a sigmoide, tangente hiperbólica e *ReLU*. Na camada de saída foram testadas a sigmoide, tangente hiperbólica e *softmax*. Os otimizadores testados foram o gradiente descendente estocástico e Adam. O número de épocas foi definido de com a convergência do modelo monitorada por funções auxiliares (*callbacks*). O tamanho do lote foi escolhido entre as opções 8, 16, 32, 64, 128, 256 e 512. E por fim, as funções de erro testadas foram *binary-cross entropy* e *focal loss*. A *focal loss* é uma função concebida com o intuito de reduzir a especialização de modelos de classificação em dados desbalanceados (LIN et al., 2018).

### 3.3.6 Tecnologias utilizadas

Todas as ferramentas utilizadas no desenvolvimento deste trabalho são *OpenSource*. O trabalho foi desenvolvido majoritariamente utilizando a linguagem de programação Python (versão 3.7.3). A produção dos mapas temáticos foi feita com o *software* QGIS (versão 3.4.6). A leitura e processamento dos dados foi feita através das bibliotecas Pandas<sup>1</sup> (versão 0.24.2) e NumPy<sup>2</sup> (versão 1.16.4). O gráficos foram construídos com base na bibliotecas Matplotlib<sup>3</sup> (versão 3.1.0) e Plotly<sup>4</sup> (versão 4.1.0). Para a utilização dos modelos de *machine learning* e redes neurais foram utilizadas as bibliotecas Scikit-learning<sup>5</sup> (versão 0.21.2) e Keras<sup>6</sup> (versão 2.0.8). A execução do código para realizar os cálculos das variáveis meteorológicas e algumas etapas de pré-processamento foram realizadas com alguns *scripts* auxiliares em *Shell Script*. Esses *scripts* contém o *pipeline* de execução dos processos em *cluster*, gerenciado pelo *Slurm Workload Manager* <sup>7</sup>.

Os códigos-fonte desenvolvidos neste trabalho, em sua maioria, foram executados de forma paralela em *cluster*. As especificações do *cluster* e do computador utilizado para a execução dos códigos estão descritas na Tabela 3.4.

---

<sup>1</sup><https://pandas.pydata.org/>

<sup>2</sup><https://numpy.org/>

<sup>3</sup><https://matplotlib.org/>

<sup>4</sup><https://plotly.com/python/>

<sup>5</sup><https://scikit-learn.org/>

<sup>6</sup><https://keras.io/>

<sup>7</sup><https://slurm.schedmd.com/>



Tabela 3.4 - Hiper-parâmetros definidos para as configurações das MLPs.

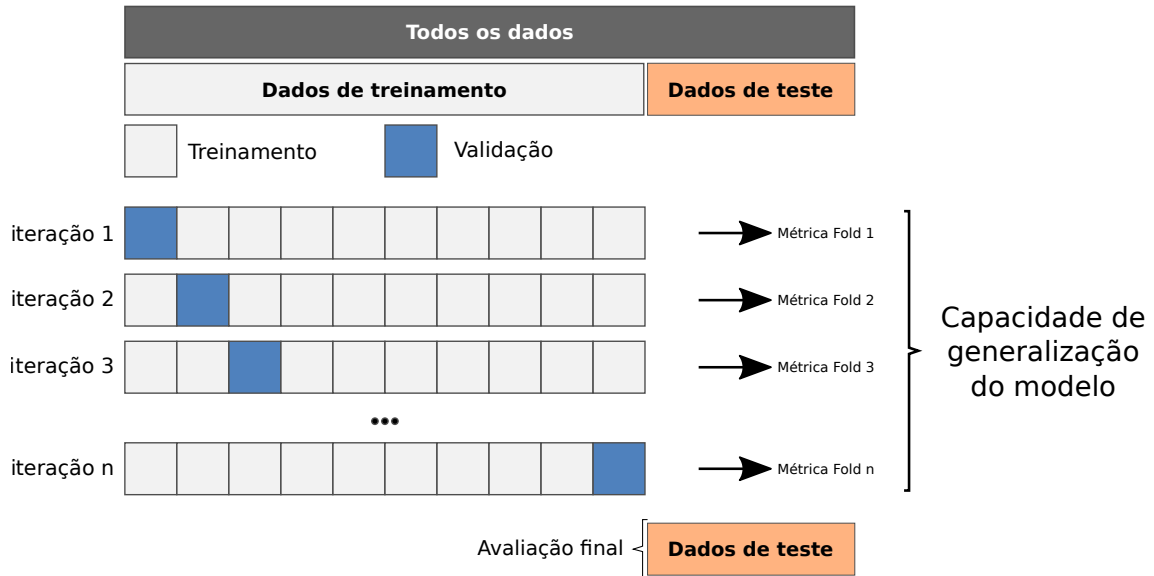
PROCESSADOR	MEMÓRIA RAM	PARTE DO CLUSTER	Nº NÓS
Intel(R) Xeon(R) CPU E5-2620 0 @ 2.00GHz (24 cores)	64GB	sim	8
Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz (32 cores)	128GB	sim	8
Intel(R) Xeon(R) CPU E3-1230 V2 @ 3.30GHz (32 cores)	32GB	sim	1
Intel(R) Core i7-8700T CPU @ 2.40GHz (12 cores)	32GB	não	-

### 3.3.7 Métricas de avaliação dos modelos

A primeira etapa do processo de validação dos modelos foi feita através da validação cruzada. A validação cruzada é uma técnica que busca avaliar a capacidade de generalização de um modelo a partir de um conjunto de dados. A técnica utilizada nesse processo será o *k-fold cross validation*, introduzido por Kohavi et al. (1995). Essa técnica consiste particionar o conjunto de treinamento em estruturas denominadas *folds*. Para cada *fold* os dados são divididos em conjunto de treinamento e validação, conforme mostrado na Figura 3.14. O principal parâmetro a ser definido nesse tipo de validação é quantidade de *folds* ( $k$ ). Segundo Ashfaque (2019), 10 é um valor ideal para definir a quantidade de *folds*, permitindo uma observação melhor do viés e variância do modelo, sendo esse o valor definido no processo de validação cruzada deste trabalho.

A segunda etapa de validação dos modelos foi feita com parcelas de dados separados previamente. Os dados foram divididos em três conjuntos: treinamento (75%), validação (15%) e teste (10%). O conjunto de treinamento foi utilizado para treinar os modelos. O segundo conjunto foi utilizado para validar os modelos durante seu treinamento, a fim de acompanhar sua convergência ao longo das épocas. E o terceiro conjunto é utilizado apenas para checar as performances dos modelos treinados. As métricas de avaliação dos modelos (*skills*) são obtidas a partir da tabela de contingência (WILKS, 2011), descritas a seguir.

Figura 3.14 - Esquema do *k-fold cross validation*.



Fonte: Adaptado de Ashfaque (2019).

Figura 3.15 - Esquema da tabela de contingência.

		EVENTO OBSERVADO	
		SIM	NÃO
EVENTO PREVISTO	SIM	Verdadeiro Positivo a	Falso Positivo b
	NÃO	Falso Negativo c	Verdadeiro Negativo d

Fonte: Adaptado de Wilks (2011).

A Figura 3.15 mostra o esquema da tabela de contingência. O valor de  $a$  é quantidade de eventos observados que foram previstos pelo modelo. O valor de  $b$  é quantidade de eventos que não foram observados, mas que o modelo previu. O valor de  $c$  trata-se da quantidade de observações que o modelo não previu. E por fim,  $d$  é a quantidade de registro que o modelo não previu e que também não foram observados. Através desses valores é possível obter métricas para avaliar as *skills* dos modelos de previsão. As principais métricas extraídas da tabela de contingência são a Probabilidade de detecção (POD) e o Falso Alarme (FAR).

O POD avalia a qualidade do modelo em relação à previsão de quando acontece o evento que está sendo observado. Neste caso, o total de raios que ocorreram e que o modelo previu. O valor de POD, descrito pela Equação 3.5, pode variar entre 0 (pior) e 1 (melhor).

$$POD = \frac{a}{a + c} \quad (3.5)$$

O FAR avalia a sensibilidade do modelo em prever eventos que não aconteceram. Neste trabalho, essa métrica computa o total de raios previstos pelas diferentes configurações de RNA, mas que não foram observados pela STARNET. O valor de FAR pode variar entre 0 (melhor) e 1 (pior), e é definido pela Equação 3.6.

$$FAR = \frac{b}{a + b} \quad (3.6)$$

A qualidade das configurações de RNAs testadas dependem do balanceamento entre POD e FAR. Então para avaliar os modelos segundo esse critério, foi considerado a diferença  $POD - FAR$ . Para essa *skill*, os valores mais próximos de 1 tem o balanceamento mais vantajoso.

Além dessas métricas, também foi considerado o Erro da Densidade (ED), descrito pela Equação 3.7, para a avaliação das configurações de alguns modelos. Trata-se da diferença entre o total de raios previsto e observados. Essa *skill* busca identificar a superestimação (valores positivos) ou subestimação (valores negativos) do modelo. O valor de ED pode variar entre  $-\infty$  e  $+\infty$ , em que o valor ideal é 0, indicando que a quantidade de raios previstas, foi a mesma observada.

$$ED = (a + b) - (a + c) \quad (3.7)$$

Cabe ressaltar que, análises sobre a influência da vegetação, ciclo diurno, degradação e combinações de entrada da RNA serão avaliadas. O intuito é definir os impactos dos dados de entrada na criação da RNA para a previsão de ocorrência de raios na Bacia Amazônica Central.

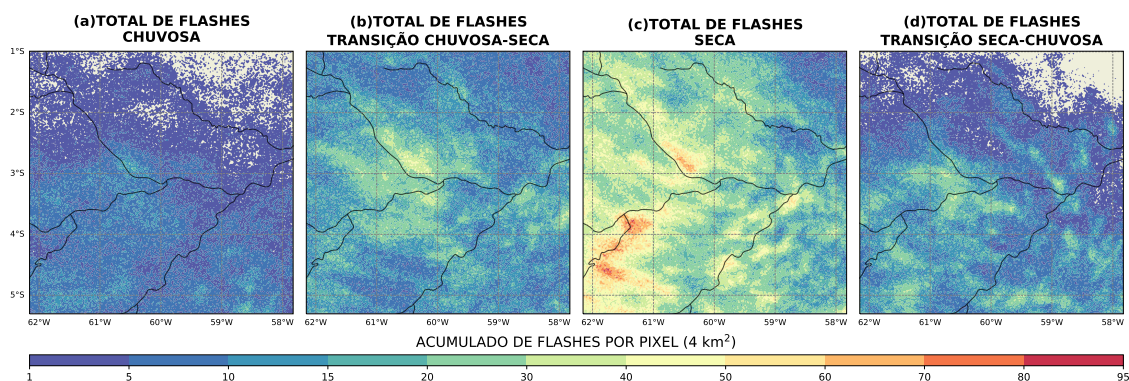
## 4 RESULTADOS

Neste Capítulo serão mostrados os resultados encontrados ao longo desta pesquisa. O primeiro passo foi definir o comportamento espacial e temporal da variável a qual espera-se prever, as descargas elétricas. Seguindo os conceitos de ciências dos dados, o segundo passo realizar uma análise exploratória dos dados, cujo objetivo apresentar as relações síncronas entre os possíveis preditores e o que será previsto. Em seguida, a aplicação de técnicas de aprendizado de máquina foi realizada para entender os aspectos que influenciam na previsão das descargas elétricas na região da bacia amazônica central.

### 4.1 Análises da distribuição espacial e sazonal dos raios

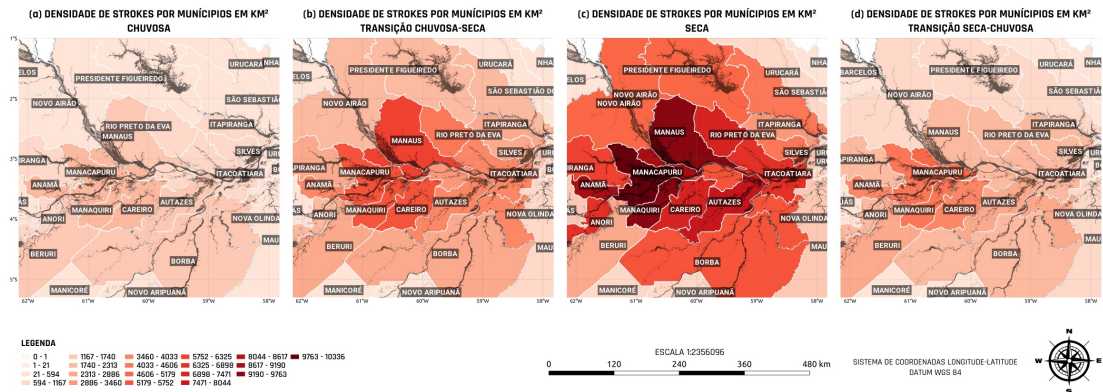
A fim de identificar padrões espaciais e sazonais das descargas elétricas, os resultados iniciais partiram de análises mais gerais dos dados. Foram analisadas as densidades de descargas elétricas durante as estações chuvosa, seca e suas transições. Essas densidades foram computadas a partir das áreas dos pixels da grade criada e dos municípios. Foram realizadas também análises em busca de padrões das descargas elétricas em relação à cobertura de solo e topografia.

Figura 4.1 - Mapas de densidades dos *flashes* nas estações chuvosa, seca e suas transições.



Fonte: Produção Autor.

Figura 4.2 - Densidade dos *strokes* por estações nos municípios da área de estudo.



Fonte: Produção Autor.

O mapa apresentado na Figura 4.1 mostra o acumulado de *flashes* detectados por pixels na área de estudo durante as estações para os anos de 2014 e 2015, juntos. Na primeira sub-figura nota-se que para estação chuvosa observa-se uma distribuição uniforme e intensidade baixa das descargas elétricas. Embora a uniformidade espacial seja uma característica predominante no mapa da estação chuvosa, algumas áreas possuem leves padrões de intensificação, essas áreas estão principalmente às margens do rio Solimões, sobre as áreas de sua confluência com o rio Negro para o rio Amazonas e em regiões próximas às margens do rio Madeira. Tal resultado mostra efeitos possíveis de brisas fluviais no desenvolvimento de sistemas eletricamente ativos. Algumas regiões ao norte da área de estudo apresentam baixa densidade de ocorrência de raios em relação às demais áreas, principalmente no nordeste, região esta que tem uma topografia mais elevada (Figura 3.1). Os efeitos topográficos no desenvolvimento de tempestades nesta região em específico não parece ser eficiente, provavelmente porque não são áreas tão elevadas, diferente dos trabalhos de (LUMB, 1970; KATONA et al., 2016) que analisaram regiões montanhosas. A segunda sub-figura refere-se ao mapa de densidade de raios durante a estação seca. Nota-se na Figura que este período apresenta maior incidências de raios, provavelmente devido a maior ocorrência de tempestades em decorrência de uma condição termodinâmica mais propícia ((MACHADO et al., 2018)). As áreas realçadas no mapa da estação chuvosa ficam ainda mais evidentes durante a estação seca, expandindo sua cobertura ao das margens dos rios e com intensificações em outras regiões sobre o rio Purus. Fica evidente na figura que a maior parte da ocorrência de raios CG são próximas

a rios, o que sugere efeitos de brisas fluviais, evapotranspiração da floresta, contrastes térmicos e/ou de umidade sobre o desenvolvimento das tempestades sobre esta região (GRECO et al., 1992; OLIVEIRA; FITZJARRALD, 1993; FISCH et al., 1998). As outras figuras mostram o que seriam as estações de transição, e como tal mostram as condições de aumento na atividade de raios CG da chuvosa para seca e de cessação da seca para a chuvosa.

Tabela 4.1 - *Strokes e flashes* nos municípios durante as estações.

<b>ESTAÇÃO</b>	<b>MUNICÍPIO</b>	<b>ÁREA</b>	<b>STROKES</b>	<b>FLASHES</b>
<b>ESTAÇÃO CHUVOSA</b>	MANAQUIRI	4,01 km <sup>2</sup>	2230/km <sup>2</sup>	2120/km <sup>2</sup>
	ANAMÁ	2,48 km <sup>2</sup>	2151/km <sup>2</sup>	2054/km <sup>2</sup>
	CAREIRO	6,16 km <sup>2</sup>	2038/km <sup>2</sup>	1939/km <sup>2</sup>
	IRANDUBA	2,23 km <sup>2</sup>	2010/km <sup>2</sup>	1927/km <sup>2</sup>
	MANACAPURU	7,41 km <sup>2</sup>	1954/km <sup>2</sup>	1881/km <sup>2</sup>
<b>CHUVOSA - SECA</b>	IRANDUBA	2,23 km <sup>2</sup>	6047/km <sup>2</sup>	5659/km <sup>2</sup>
	MANAUS	11,51 km <sup>2</sup>	5796/km <sup>2</sup>	5423/km <sup>2</sup>
	MANAQUIRI	4,01 km <sup>2</sup>	5597/km <sup>2</sup>	5187/km <sup>2</sup>
	MANACAPURU	7,41 km <sup>2</sup>	5554/km <sup>2</sup>	5211/km <sup>2</sup>
	CAREIRO	6,16 km <sup>2</sup>	5440/km <sup>2</sup>	4942/km <sup>2</sup>
<b>ESTAÇÃO SECA</b>	BERURI	17,48 km <sup>2</sup>	10336/km <sup>2</sup>	9601/km <sup>2</sup>
	MANACAPURU	7,41 km <sup>2</sup>	9961/km <sup>2</sup>	9229/km <sup>2</sup>
	MANAQUIRI	4,01 km <sup>2</sup>	9651/km <sup>2</sup>	8935/km <sup>2</sup>
	MANAUS	11,51 km <sup>2</sup>	9466/km <sup>2</sup>	8776/km <sup>2</sup>
	ANAMÁ	2,48 km <sup>2</sup>	9137/km <sup>2</sup>	8591/km <sup>2</sup>
<b>SECA - CHUVOSA</b>	IRANDUBA	2,23 km <sup>2</sup>	4866/km <sup>2</sup>	4441/km <sup>2</sup>
	MANACAPURU	7,41 km <sup>2</sup>	4540/km <sup>2</sup>	4152/km <sup>2</sup>
	MANAQUIRI	4,01 km <sup>2</sup>	4170/km <sup>2</sup>	3835/km <sup>2</sup>
	CAREIRO	6,16 km <sup>2</sup>	3901/km <sup>2</sup>	3581/km <sup>2</sup>
	ANAMÁ	2,48 km <sup>2</sup>	3878/km <sup>2</sup>	3589/km <sup>2</sup>

Faz parte das atividades do INPE traduzir as informações para os tomadores de decisão (INPE, 2021). Muitas prefeituras as vezes precisam informar a sua população sobre a quantidade de raios que caem em seu território. A Figura a seguir é apenas ilustrativa e traz as informações mostradas anteriormente, mas a nível de município. Nota-se na Figura que a incidência de raios (representada pelos *strokes*) na estação chuvosa é mais fraca e uniforme, como dito anteriormente, tendo uma atividade maior nos municípios próximos da confluência dos rios Negro e Solimões. Na estação seca a intensificação aumenta e atinge mais municípios próximos a outras áreas ao longo das margens dos rios Purus, Amazonas e Madeira. Na transição entre as estações chuvosa e seca os municípios entre os rios Solimões e Madeira tem uma

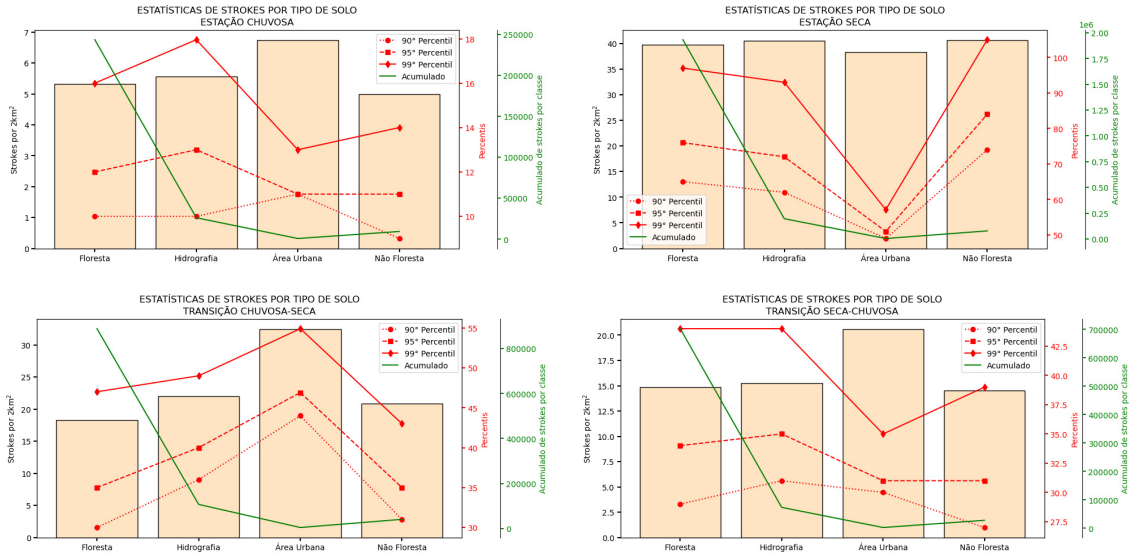
taxa elevada de ocorrências, assim como em áreas coincidentes com a estação seca, no entanto há uma taxa significativa na cidade de Manaus, o que pode ter relação com os efeitos dos aerossóis de poluentes e das queimas de biomassas que se iniciam nesse período (MACHADO et al., 2018). Por fim, na transição entre as estações seca e chuvosa, as taxas mais elevadas se concentram nos municípios próximos ao rio Solimões, enquanto que nos demais municípios a cessação das incidências convergem para os padrões apresentados na estação chuvosa. A Tabela 4.1 mostra os 5 municípios com mais *strokes* em cada estação. Alguns municípios aparecem no topo da lista em todas as estações ou em mais de uma, o que corresponde às regiões mais ativas durante todo o ano, como citado anteriormente.

#### 4.1.1 Variação espacial dos raios em relação à cobertura de solo

Com o intuito de analisar a variação espacial dos raios em relação aos principais tipos de cobertura de solo da área de estudo, foram calculadas algumas estatísticas durante todas as estações, conforme descritas na Figura 4.3. As estatísticas obtidas para cada período foram: a densidade de *strokes* por pixel (primeiro eixo vertical) que aqui normaliza a quantidade de raios pela área; os percentis 90, 95 e 99 (segundo eixo vertical) que indicam a intensidade dos eventos mais severos; e o acumulado (terceiro eixo vertical) para cada uma das classes de cobertura do solo que foram analisadas que é dependente da área da classe. Em números absolutos, os acumulados ficaram de acordo com a representatividade das classes de solo, sendo assim, a floresta, que é o tipo de solo predominante teve um número elevado em todas as estações, enquanto que a área urbana possui o menor valor, sendo ela o tipo menos frequente na classificação da área de estudo. Entretanto, quando a ocorrência de raios é comparada proporcionalmente com a quantidade de pixels que representam as classes, ou seja a densidade, a área urbana, embora menor, se sobressai bastante, tendo uma densidade maior, com exceção na estação seca, cuja a variação é pequena entre as classes. Os percentis mostram comportamentos distintos, eventos mais intensos aparentemente ocorrem com maior frequência sobre: os rios na estação chuvosa; áreas urbanas na transição chuvosa-seca; não-floresta na seca; e novamente sobre os rios na estação seca-chuvosa. De um modo geral, as estações chuvosa e seca-chuvosa são similares, dada a devida proporcionalidade. Enquanto que as outras mais secas diferem em diferentes aspectos. Tal resultado pode ser um reflexo da maior atividade termodinâmica nesse período, assim como os possíveis efeitos de aerossóis de queimadas de biomassas no desenvolvimento das tempestades, como investigado nos trabalhos de Camponogara et al. (2014), Machado et al. (2018), Pereira (2019).

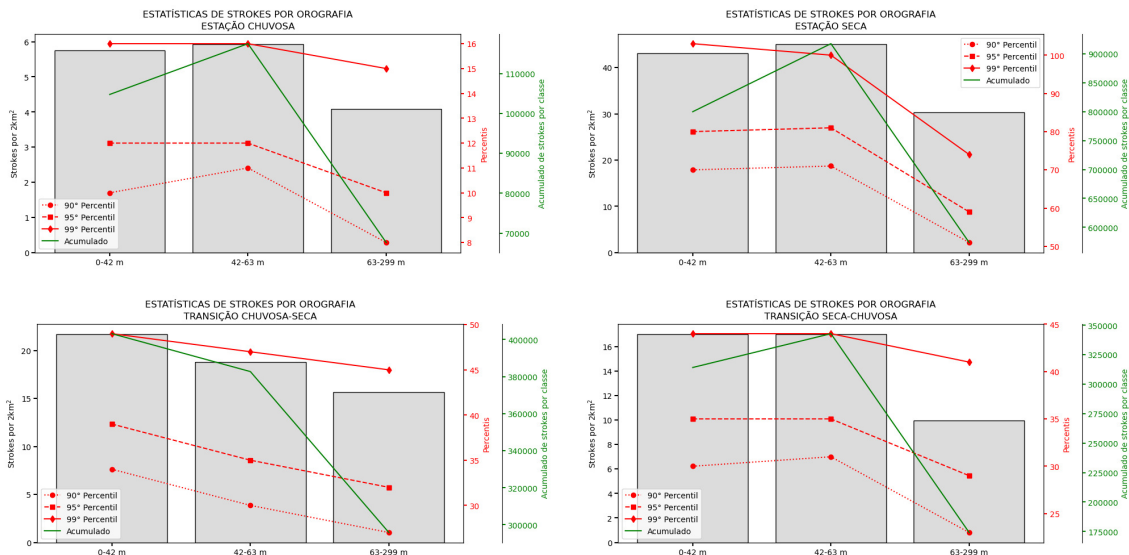


Figura 4.3 - Estatísticas dos *strokes* por tipo de solo durante todas as estações.



Fonte: Produção Autor.

Figura 4.4 - Estatísticas dos *strokes* por orografia durante todas as estações.



Fonte: Produção Autor.

Assim como para as classes de cobertura de solo, foram analisadas as estatísticas dos *strokes* para as orografias também, conforme mostradas na Figura 4.4. As classes aqui foram definidas pelos percentis de 33 e 66%, dividindo em três classes: 0-42m; <42-63m; >63-99. Dada as devidas proporcionalidades para quantidade de *strokes*, nota-se nas Figuras que o comportamento é bem similar entre as variáveis estatísticas analisadas para todas as classes e estações, exceto para o acumulado (que depende da área). Onde existe uma maior densidade de *strokes*, assim como os eventos mais severos (percentis), em áreas mais baixas sobre a região. Contudo, salienta-se que as estações chuvosa e seca apresentam valores um pouco maiores nas elevações intermediárias.

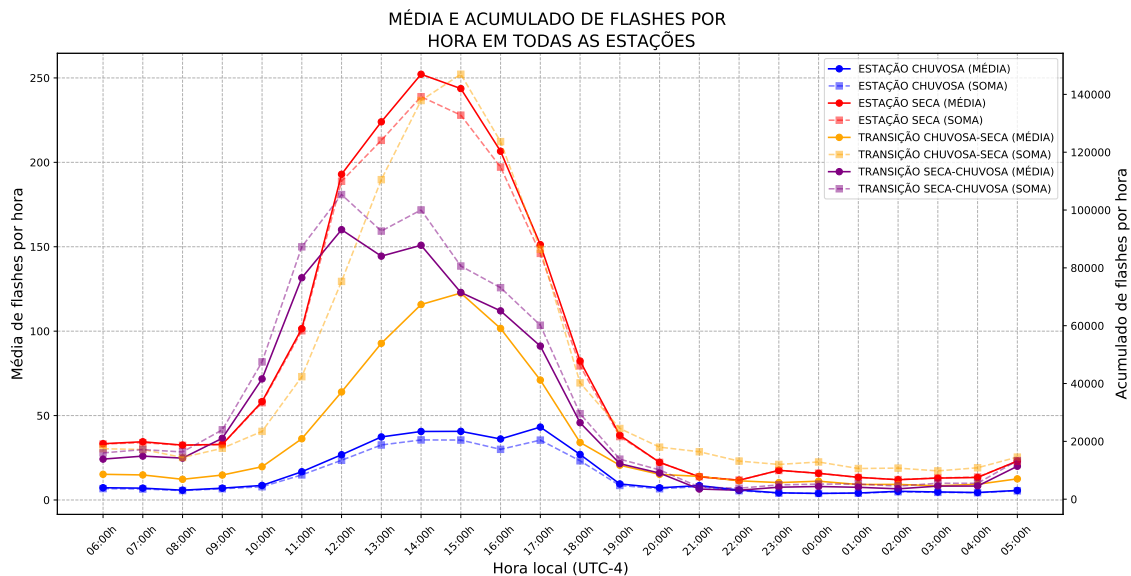
Em suma, os resultados aqui encontrados mostram que existe uma variação espacial relativamente alta entre as estações que dependem também dos aspectos de superfície. Os rios aparecem como áreas importantes no desenvolvimento de raios sobre a região, uma possível relação com os padrões locais diários das brisas fluviais apontados por [Martin et al. \(2015\)](#), que desempenham um papel importante no processo de transporte das massas de ar. Ainda, a alta da incidência de raios no período correspondente à estação seca pode estar associada com as emissões antrópicas de aerossóis, sobretudo, relacionado à cultura de queimadas que se intensifica nesse período e que favorece o aumento dos processos de oxidação, levando ao aumento das concentrações dos núcleos de condensação das nuvens ([CECCHINI et al., 2017](#)). Como apontado por [Cecchini et al. \(2017\)](#), as nuvens desenvolvidas em condições similares de aerossóis tendem a ter distribuição do tamanho de gotas semelhante em todo o perfil vertical da nuvem, então, a injeção de aerossóis provenientes da atividade humana nesse período pode favorecer o desenvolvimento de nuvens profundas com alta presença de gelo, o que talvez pudesse ser um dos motivos da maior ocorrência de raios. Contudo, estudos mais profundos são necessários para definir a real importância dos aerossóis na eletrificação atmosférica dessa região. Além disso, regiões de planície mostram-se mais favoráveis a formação de tempestades eletricamente ativas. Neste sentido, fica claro que um sistema de previsão dessa variável, raios CG, devem levar em consideração as características de superfície e outros aspectos locais em suas entradas.

#### **4.1.2 Análise do ciclo diurno**

Como observado nas seções anteriores, existe uma grande variação espacial e sazonal dos dados de descargas elétricas nuvem-solo da rede STARNET na região da Bacia Amazônica. É sabido que os processos convectivos na região amazônica oferecem

forte ciclo diurno (MACHADO et al., 2018). Neste sentido, de modo a contabilizar esta influência sobre a área de estudo a Figura 4.5 mostra as médias e acumulados de raios CG ao longo do dia. Nota-se na figura que existem comportamentos distintos entre as estações, um possível reflexo da atividade de tempestades nesta região. As estações secas (linhas vermelhas) e transição chuvosa-seca (linhas amarelas) foram aquelas que apresentaram uma maior quantidade de raios. A atividade dos raios tem sua ascensão ainda no período da manhã por volta das 9 h no horário local (UTC-4) e começa a decair por volta das 14 h, padrão semelhante ao observado por Angelis et al. (2004). Observa-se que, a estação seca é aquela com maior incidência de raios, com o pico máximo entre as 12 h e 14 h. A estação de transição seca-chuvosa, mostra uma pequena antecipação no pico das ocorrências, que acontece entre 10 h e 12 h. Na estação chuvosa, a incidência de raios é mais acentuada e se prolonga por grande parte do dia, entre as 10 h e 19 h. Logo, conclui-se que, para a criação de um sistema de previsão de raios é necessário levar em consideração também o ciclo diurno destes fenômenos. Não só pela natureza dos mesmos, mas pela eficácia da rede de detecção que é sensível ao período do dia que está sendo observado.

Figura 4.5 - Média e acumulado horário de *flashes* em todas as estações.



Fonte: Produção Autor.

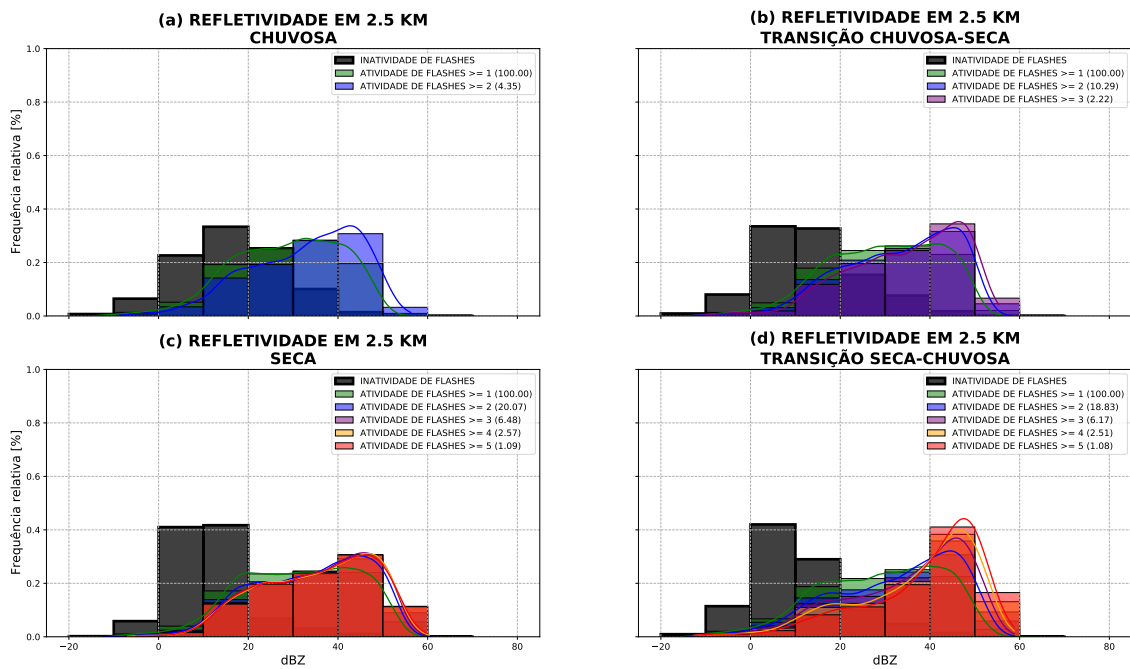
## 4.2 Análise exploratória de dados

Com base nos conceitos de ciência dos dados, nesta seção foram analisadas as variáveis do radar meteorológico em banda-S do SIPAM definidas pela refletividade radar e sua relação com os raios, neste caso o *flash*. A ideia desta seção é entender as relações diretas que existem entre as variáveis e os raios. Estas relações irão permitir verificar os aspectos físicos associados e entender os resultados encontrados pelas RNAs e DTs construídas ao longo desta pesquisa.

### 4.2.1 Análises das relações síncronas

Como apontado no trabalho de Almeida et al. (2020), as distribuições dos dados com diferentes atividades de raios podem servir como importantes indicativos para identificar tendências. Diante disso, foram criados histogramas com as frequências relativas dos valores de cada variável, utilizando os diferentes limiares de incidências de raios CG em todas as estações. A Figura 4.6 mostra para todas as estações os histogramas das frequências relativas de refletividade radar na altura de 2,5 km com diferentes limiares de atividade de raios. Os limiares de incidências de raios foram divididos em: inatividade elétrica (*flashes* igual a 0) e atividade elétrica (*flashes* acima de 1, 2, 3, 4 e 5). Além disso, o histograma é exibido somente quando a quantidade de pontos que representam a classe é superior a 1% dos dados, isso é feito com o intuito de evitar ruídos e provocar interpretações equivocadas a respeito dos dados com pouca representatividade. Através dessa figura é possível observar que o pico máximo das frequências da classe de inatividade de raios acontece na faixa de 10-20 dBZ, decrescendo significativamente nas faixas seguintes. É a partir dessa faixa que as frequências das classes de atividades de raios começam a ficar mais evidentes, diferenciando-se ao longo das estações.

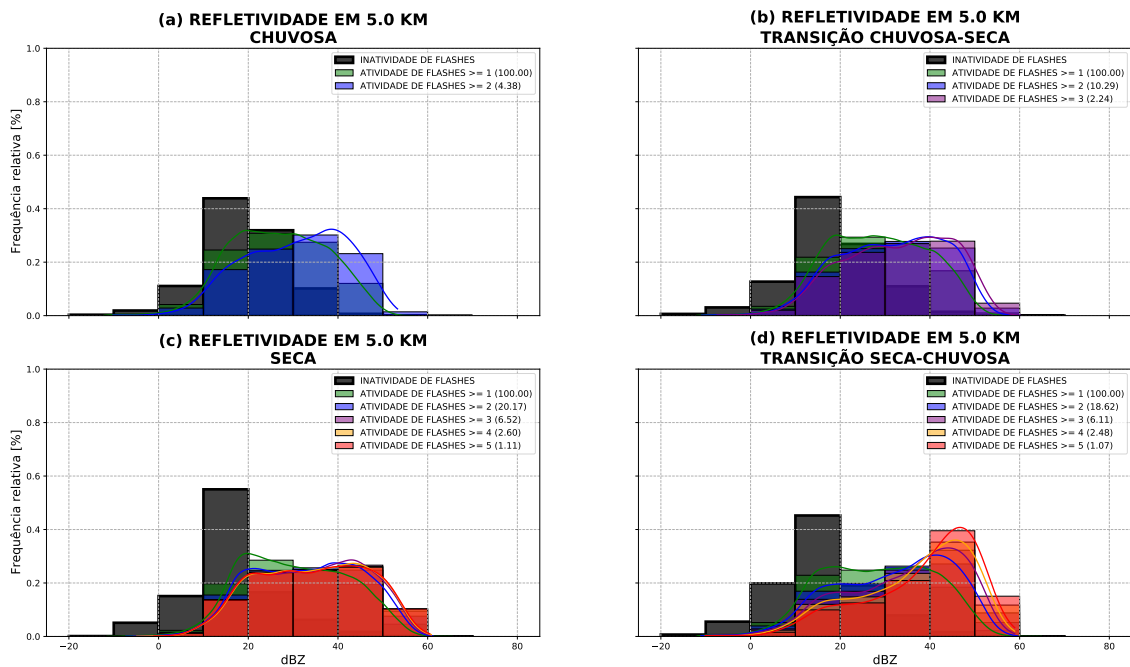
Figura 4.6 - Histogramas das frequências relativas das refletividades do produto CAPPI em 2,5 km com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

Na estação chuvosa a sobreposição das frequências nas diferentes classes de atividade é maior que nas demais, isso se dá devido ao comportamento uniforme das chuvas e baixa incidência de raios característico desse período, como observado na seção anterior. Na estação de transição chuvosa-seca há uma intensificação das relação mostrando uma separação entre classes. Comportamento este que continua sendo observado na estação seca, como sabido, período em que a incidência de raios por pixel é maior. Notou-se na Figura referente à estação seca que a sobreposição das frequências é bem menor, um indicativo de maior separação entre as classes, em que a inatividade de raios se concentra principalmente na faixa de 0-20 dBZ, tendo o pico máximo de frequência em relação às demais estações. Os valores de refletividade ligados a raios CG se concentram em valores maiores com um padrão bimodal. Ressalta-se que há uma maior acentuação para valores mais altos quando existem mais raios CG. Na transição entre as estações seca e chuvosa esta acentuação fica ainda mais clara nos histogramas, o que mostra que nessa estação um sistema de classificação e previsão podem ficar ainda mais eficientes.

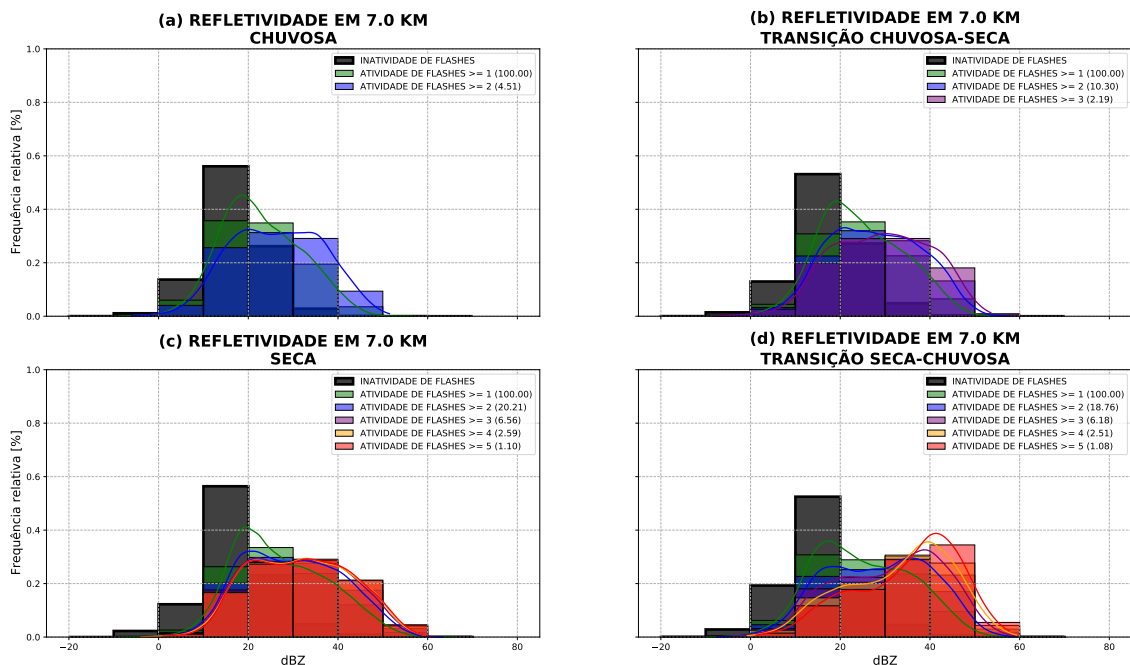
Figura 4.7 - Histogramas das frequências relativas das refletividades do produto CAPPI em 5 km com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

Analisar as características dos perfis de refletividade é de suma importância para definir condições propícias para a ocorrência de descargas elétricas (DONALDSON, 1961; ZIPSER; LUTZ, 1994; LIU et al., 2012). Neste sentido, as próximas figuras mostram os histogramas apresentados anteriormente, contudo, para diferentes níveis da atmosfera: 5km (Figura 4.7) e 7km (Figura 4.8). Nota-se nas figuras que, a medida que se elava a altura, os valores de refletividade vão sendo atenuados, ou seja, o sinal encontrado para diferenciar raios CG de não-raios se torna cada vez menos expressivo. Este comportamento pode ser observado para todas as estações. Contudo, esta mudança não é abrupta, logo, devem existir eventos mais intensos que ainda mantém o sinal em todos os níveis. Além disso, os picos de não-atividade de raios ficam ainda mais concentrados em refletividades mais baixas na maior altura. Contudo, não é possível mensurar que essa informação seja ou não válida para um sistema de previsão. Uma vez que, variações nos topos das tempestades associadas a fortes correntes ascendentes podem ser responsáveis por injetar mais hidrometeoros em camadas mais altas e favorecer a ocorrência de descargas elétricas (RYZHKOV; ZRNIC, 2019), pode existir uma relação assíncrona entre elas, como veremos posteriormente.

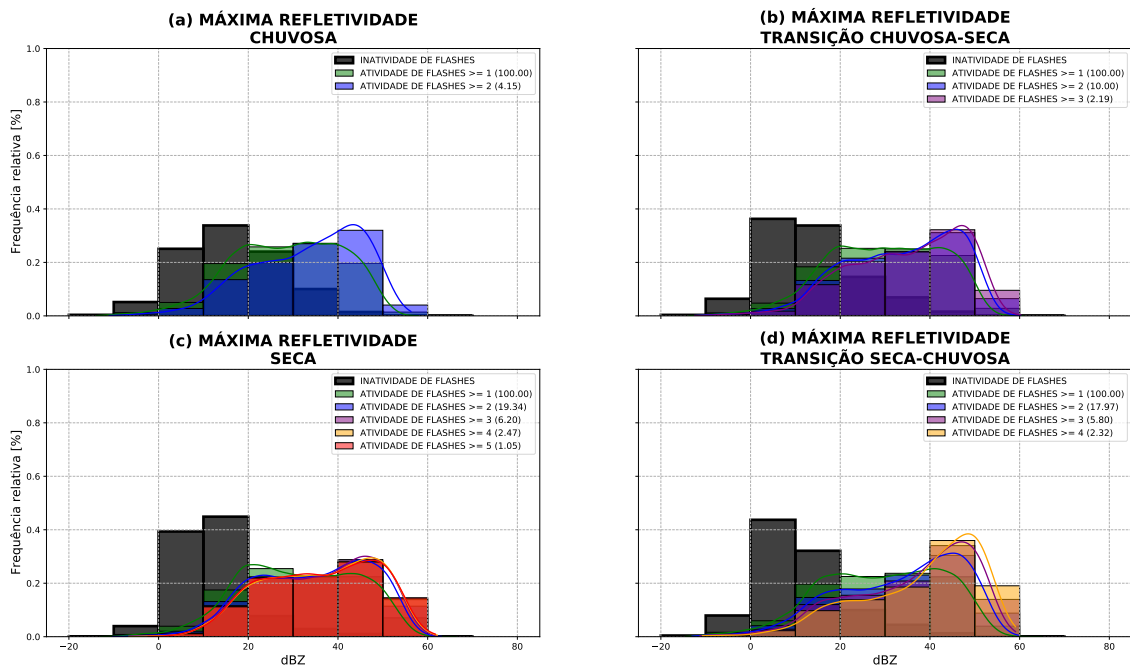
Figura 4.8 - Histogramas das frequências relativas das refletividades do produto CAPPI em 7 km com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

A maioria das variáveis estudadas neste trabalho foram baseadas nos perfis de refletividade radar. A Figura 4.9 mostra os mesmos tipos de histogramas apresentados anteriormente, contudo, agora para a máxima refletividade observada no perfil para cada pixel. Nota-se na figura que as distribuições são semelhantes aqueles de refletividades em níveis mais baixos. Este comportamento é esperado, uma vez que sistemas mais convectivos apresentam valores mais intensos abaixo do nível de congelamento ( 4,5 km na Amazônia) (YUTER; HOUZE, 1995). Contudo, cabe ressaltar que os valores médios e desvios-padrão ficam ainda mais realçados, dando a esta variável uma capacidade maior de classificação entre atividade e não-atividade.

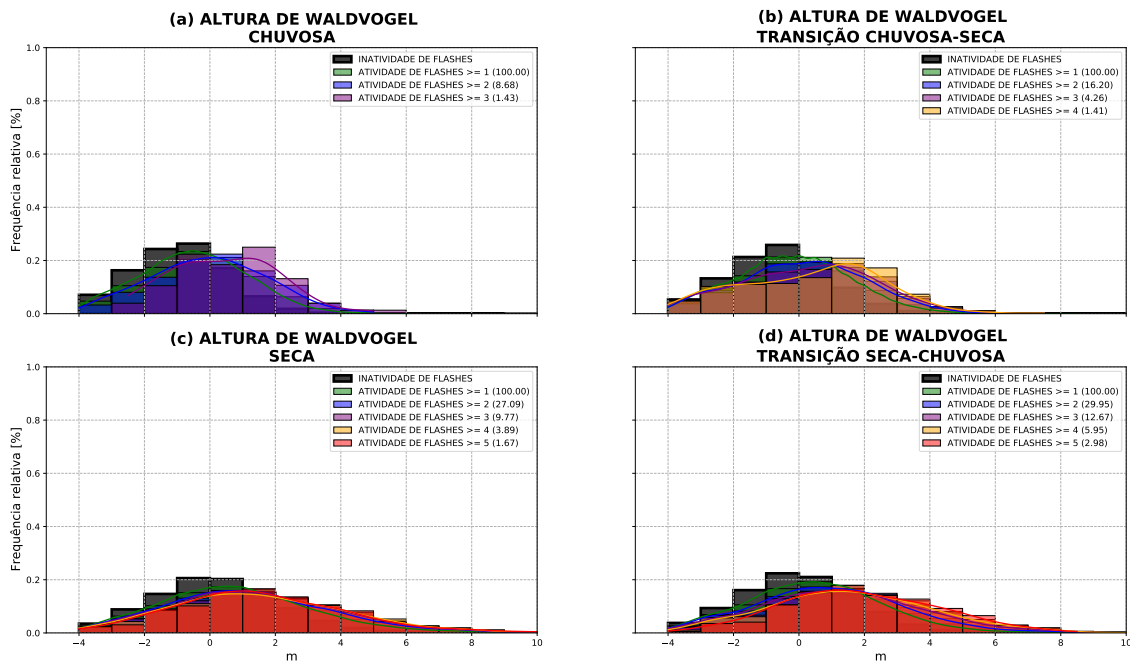
Figura 4.9 - Histogramas das frequências relativas das Máximas refletividade (2-7km) com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.



Figura 4.10 - Histogramas das frequências relativas das alturas de Waldvogel (em metros) com diferentes incidências de *flashes* em todas estações.

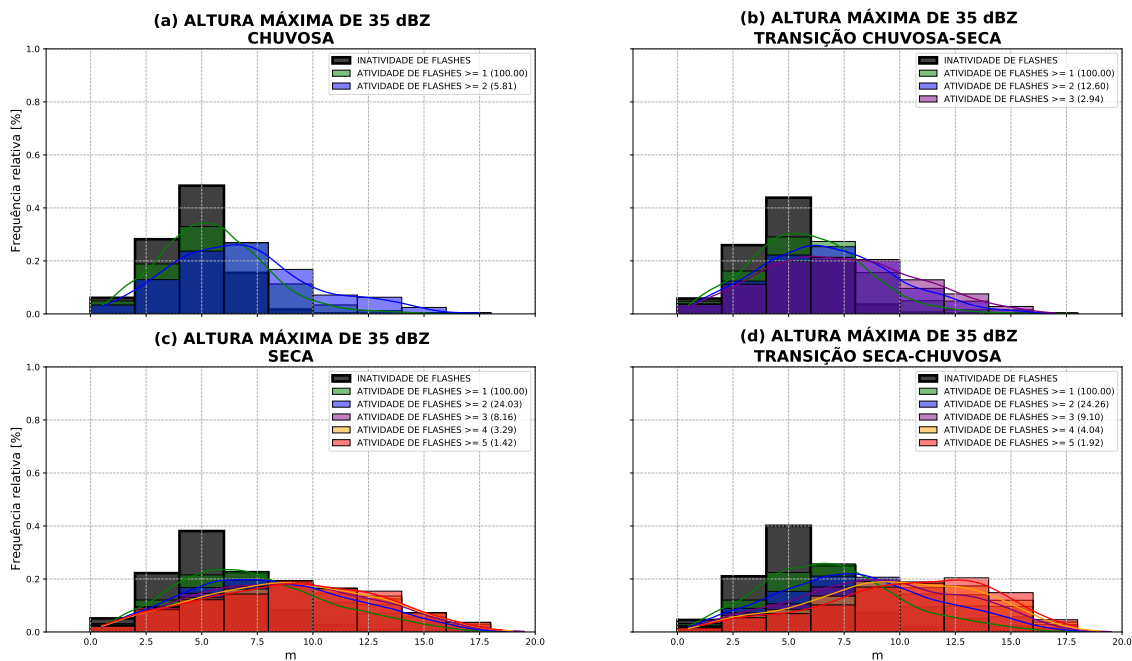


Fonte: Produção Autor.

Outras variáveis podem realçar características da microfísica da coluna atmosférica, levando em consideração outras informações do ambiente ou do próprio perfil da precipitação. Um exemplo é a Altura de Waldvogel (Figura 4.10), que geralmente é utilizado como um indicador da presença de forte corrente ascendente. Onde uma camada entre o nível de congelamento (aqui fixado em 4,5km) e a altura do *echo top* de 45 dBZ poderia ser realçada pela presença de gotas de água superesfriadas e/ou gelo injetados por corrente ascendente pronunciada, muito usada na detecção de granizo em tempestades (WALDVOGEL et al., 1979). Ao analisar a Figura 4.10 observa-se que existem grandes áreas de intersecções entre a classe de inatividade e as classes de atividade em todas as estações. Contudo, a distribuição apresenta uma forma mais plana, suavizada. Logo, quando verificado os valores médios apresentados na Tabela A.1 do Apêndice A, observa-se que existe uma distribuição maior de valores mais altos na condição de atividade, principalmente em estação com maior atividade convectiva. O que não é o caso da estação chuvosa, que pouco é abstraído de suas relações.

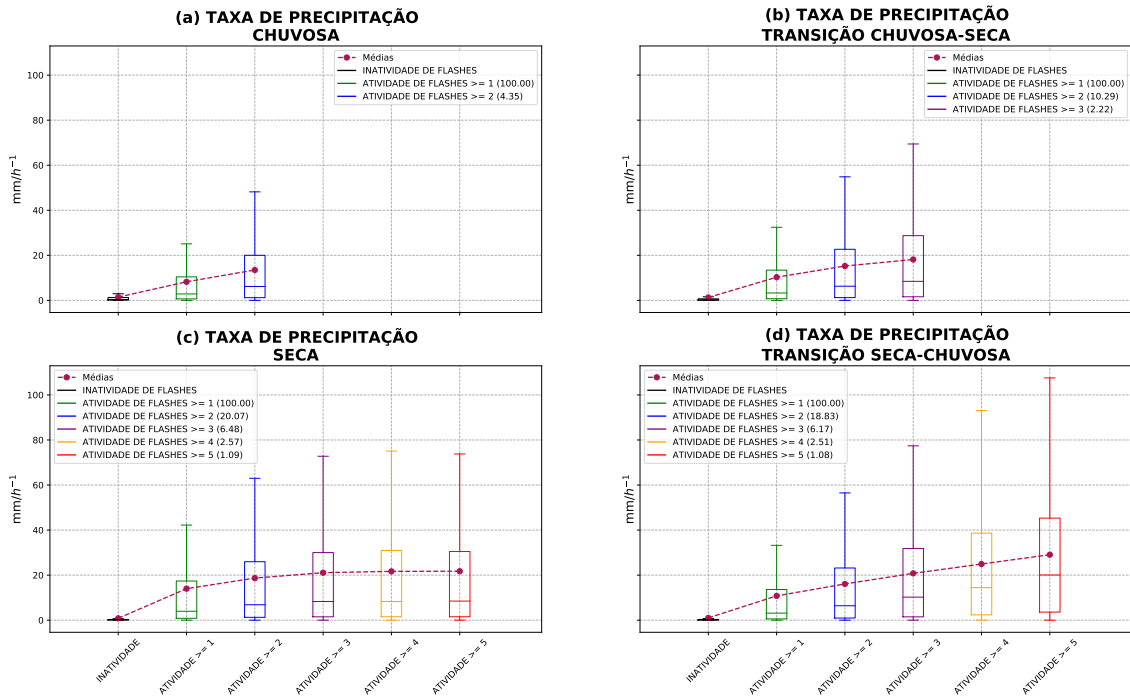
Muitos estudos mostram que a ocorrência de raios CG está relacionada a um rápido aumento na altura de certos valores de refletividade radar, chamados de *echo tops* (STOLZENBURG, 01 Aug. 1994; VOORMANSIK et al., 2017). Assim como feito para a altura de Waldvogel, os gráficos da Figura 4.11 mostram os histogramas para o *echo top* de 35 dBZ. Nota-se na figura que as distribuições são similares aquelas observadas para Waldvogel. Todavia, observa-se que as distribuições para valores maiores de raios apresentaram-se enviesadas para valores mais altos de altura. Tal resultado pode estar associado a uma relação maior com tempestades mais bem desenvolvidas verticalmente e com maior densidade de raios CG. Isto acontece principalmente nos períodos de maior atividade elétrica, como a estação seca e seca-chuvosa. Logo, espera-se que esta variável seja importantes em algoritmos de regressão, que quantifiquem a densidade de raios. Entretanto, não fica claro se teriam boa classificação (i.e. triagem) para a ocorrência de raios. Maiores detalhes sobre estas questões serão dados em seções posteriores.

Figura 4.11 - Histogramas das frequências relativas das alturas do 35 dBZ (*echo top* em metros) com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

Figura 4.12 - Boxplots dos valores da taxa de precipitação com diferentes incidências de *flashes* em todas estações.

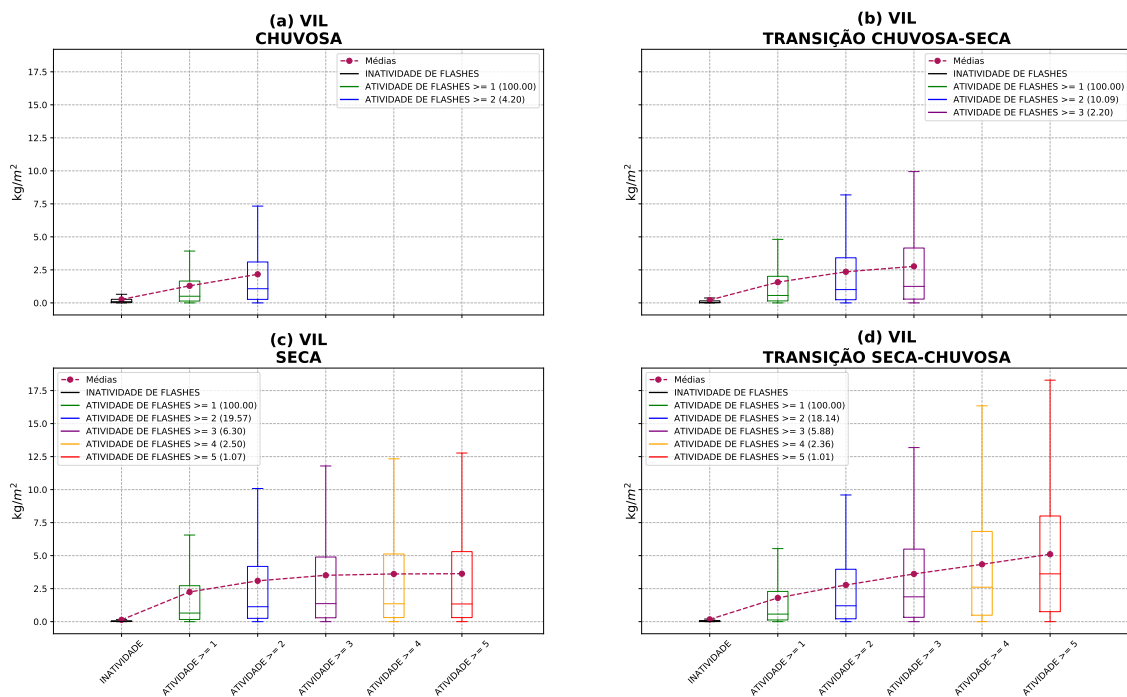


Fonte: Produção Autor.

Algumas Variáveis analisadas neste estudo não apresentaram muitas informações que possam ser retiradas de análises visuais dos histogramas, principalmente porque existem muitos valores próximos a zero. Contudo, é sabido que elas podem ajudar no monitoramento e previsão de eventos com atividade elétrica. Neste sentido, para melhor apresentar os resultados, foram analisados os gráficos do tipo *boxplot*. Na Figura 4.12 observa-se que as variações das taxas de precipitação (Equação 2.3) aumentam quase que linearmente a medida que mais raios CG são registrados. No entanto, há uma variação muito pequena das medianas, que pode ser um indicativo de melhor classificação de ocorrência de raios, mas nada indica sob uma possível estimativa de valores a partir de métodos de regressão.

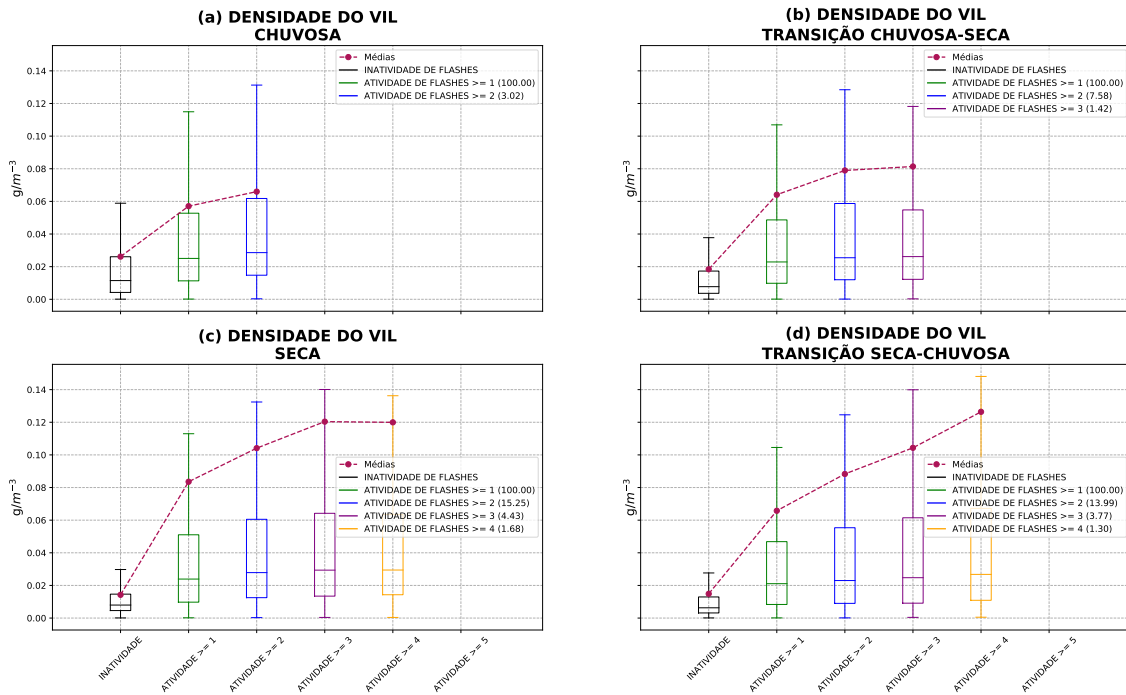
O mesmo comportamento foi observado para o VIL (Figura 4.13) e a densidade de VIL (Figura 4.14). Na literatura o VIL é um variável correlacionada a eventos severos, como a presença de granizo e descargas elétricas (GREENE; CLARK, 1972). As médias (linhas na cor bordô) mostram que entre as variáveis mostradas, a densidade de VIL foi aquela que apresenta a maiores diferenças entre atividade e inatividade. O que pode tornar esta variável importante em um processo de classificação de ocorrência de raios. A densidade do VIL é uma importante variável para a rápida identificação de tempestades com alta refletividade em relação à sua altura e características que geralmente indicam núcleos de granizo. A medida que o valor da densidade do VIL aumenta, esses núcleos tendem a serem mais intensos e profundos, resultando em granizo de diâmetros maiores (AMBURN; WOLF, 1997), que podem favorecer a alta incidência de raios.

Figura 4.13 - Boxplots dos valores do VIL com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

Figura 4.14 - Boxplots dos valores da Densidade do VIL com diferentes incidências de *flashes* em todas estações.



Fonte: Produção Autor.

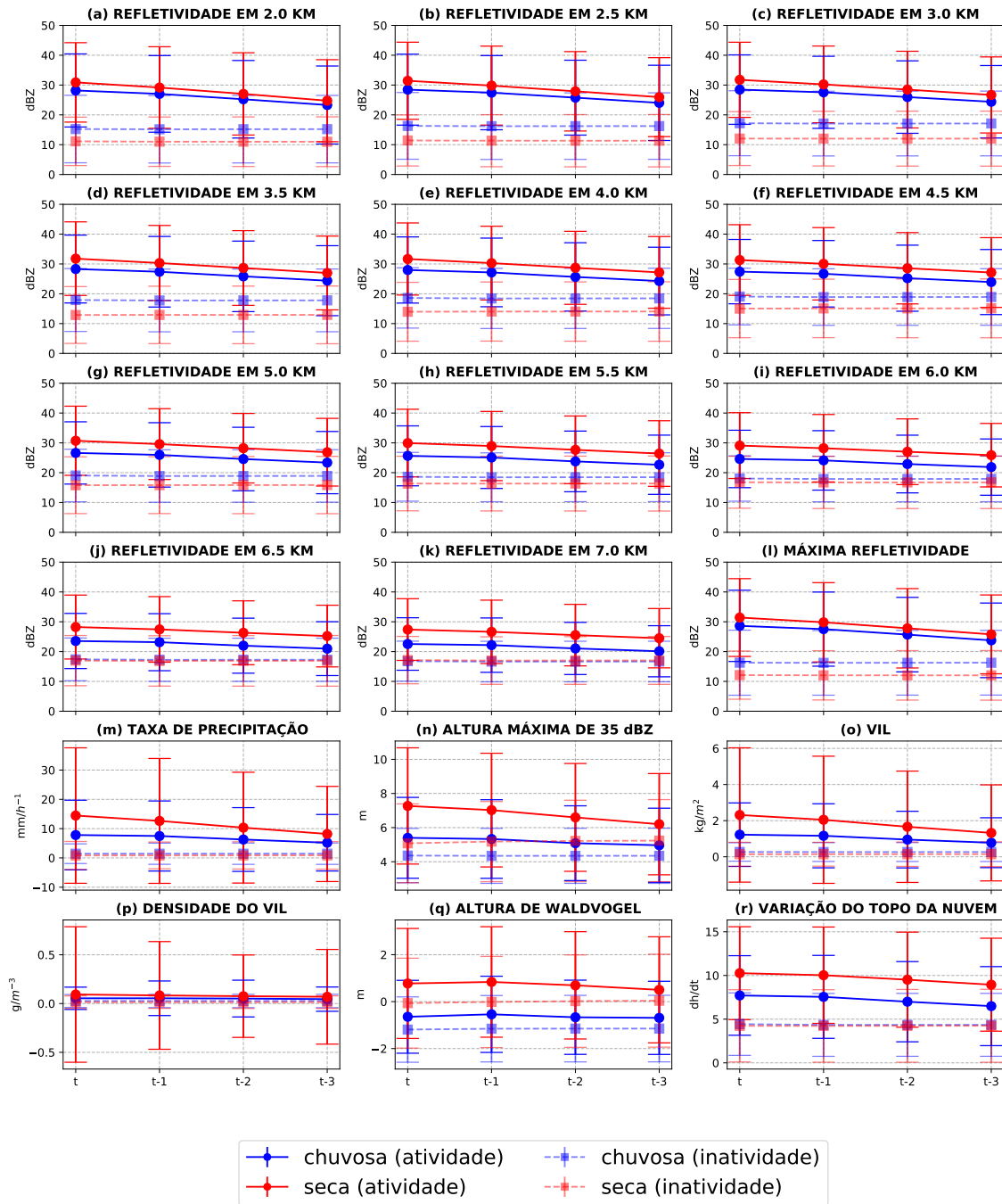
#### 4.2.2 Análises das relações assíncronas

As descargas elétricas são fenômenos atmosféricos que dependem de processos microfísicos e dinâmicos que ocorrem dentro da nuvem, como mostrado em seções anteriores. Estes processos podem apresentar impactos direto nas medidas do radar. Contudo, pode existir uma relação assíncrona entre o momento de ocorrência do raio e as tendências de variação observadas por variáveis meteorológicas (WILLIAMS et al., 1999; VINCENT et al., 2003). Neste sentido, foi necessário realizar uma análise em medidas anteriores do radar e compará-las aos *flashes* em um determinado instante.

A Figura 4.15 mostra os valores médios e seus respectivos desvios-padrão (barra vertical) para todas as variáveis do radar analisadas neste trabalho para os instantes  $t$  (quando ocorreram os *flashes*),  $t-1$ ,  $t-2$  e  $t-3$ , para inatividade (linha tracejada) e atividade (linha contínua) de raios durante a estação seca (vermelho) e chuvosa (azul). Essas informações foram obtidas a partir do instante de tempo atual ( $t$ ) e para os três intervalos de tempo anteriores na resolução temporal dos dados do radar, ou seja, 12 minutos ( $t-1$ ), 24 minutos ( $t-2$ ) e 36 minutos ( $t-3$ ). Nota-se na figura que durante na estação seca, todos os valores médios e de desvio das variáveis radar associadas a raios CG foram superiores a aqueles observados durante a estação chuvosa. O que mostra que durante a estação seca as relações entre raios e as variáveis são mais acentuadas, uma possível consequência da maior quantidade de raios que acontece durante esse período.

Uma análise importante que deve ser endereçada a este gráfico é justamente a comparação entre inatividade e atividade. Quanto maior for a diferença, maior deve ser a probabilidade de discriminação pelo método de aprendizado de máquina. Logo, como pode ser verificado na figura, os dados de refletividade em diferentes níveis apresentam um maior distanciamento entre as médias. Vale ressaltar que, a medida que aumenta-se a altura do CAPPI essa diferença é atenuada. Tanto que as diferenças entre inatividades nas estações são ínfimas. As outras variáveis que dependem do perfil de refletividade mostram comportamentos distintos. A precipitação, o VIL e a máxima refletividade no perfil apresentam comportamento semelhantes à refletividade. Sendo que a máxima refletividade foi aquela que apresentou a maior diferença na média (todos os valores podem ser consultados na Tabela A.1 do Apêndice A). Contudo, as variáveis associadas às alturas, como altura de Waldvogel, altura do *echo top* de 35dBZ e DHDt mostram-se similares. Contudo, o DHDt é aquele que apresenta as maiores diferenças. Aquela que apresenta um resultado discrepante de todas as outras variáveis é a densidade de VIL. Nela as médias são similares, contudo, o desvio padrão é muito maior para eventos de atividade do que para inatividade. Tal resultado pode ser um indicativo de maior separação de classes, uma vez que a média mais o desvio-padrão da atividade ultrapassaria a mesma operação para inatividade. Foi aplicado o teste de hipótese *t-student* para verificar se as diferenças entre as médias são estatisticamente significativas em um intervalo de confiança de 95%. Os resultados dos testes estatísticos corroboram com a análise visual que foi descrita, e os valores, assim como outras informações estatísticas, estão descritos na Tabela A.1 do Apêndice A.

Figura 4.15 - Relações assíncronas das variáveis em diferentes tempos anteriores à observação durante as estações chuvosa e seca para atividade e inatividade de raios. t (tempo atual), t-1 (12 minutos), t-2 (24 minutos) e t-3 (36 minutos).



Fonte: Produção Autor.

As variações entre os diferentes tempos podem apresentar características importantes a um sistema de previsão. A tendência na evolução temporal da variável antes da ocorrência dos raios pode ser definido como um possível parâmetro preditor. Analisando a figura notamos que existe uma tendência linear de crescimento nos valores médios para quase todas as variáveis. Entretanto, é importante notar que para a classe de inatividade esse comportamento não se aplica, tendo uma variação quase inexistente nas médias das variáveis entre a ocorrência dos raios e os tempos anteriores à ela. A medida que se afasta do momento de ocorrência da descarga elétrica, as diferenças entre a atividade e a inatividade se tornam menores. Este tipo de comportamento sugere que para algumas variáveis, seu uso em tempos passados não seja uma informação proveitosa, como por exemplo, a precipitação em  $t - 3$ . Através da distribuição desses dados fica evidente que pouca informação relevante para um sistema de previsão pode ser extraída de dados mais antigos. Contudo, essas relações não são lineares, pois o desvio padrão varia para alguns tempos, o que mostra que a distribuição não é semelhante. Apenas analisando o ciclo de vida dos sistemas de tempestades e sua evolução seria possível determinar com maior precisão as contribuições. De modo a avaliar os modelos de previsão que foram criados, foram necessários realizar testes de sensibilidade para definir empiricamente os processos de causa efeito da variável no raio, ou vice-versa, conforme serão apresentados em seções posteriores.

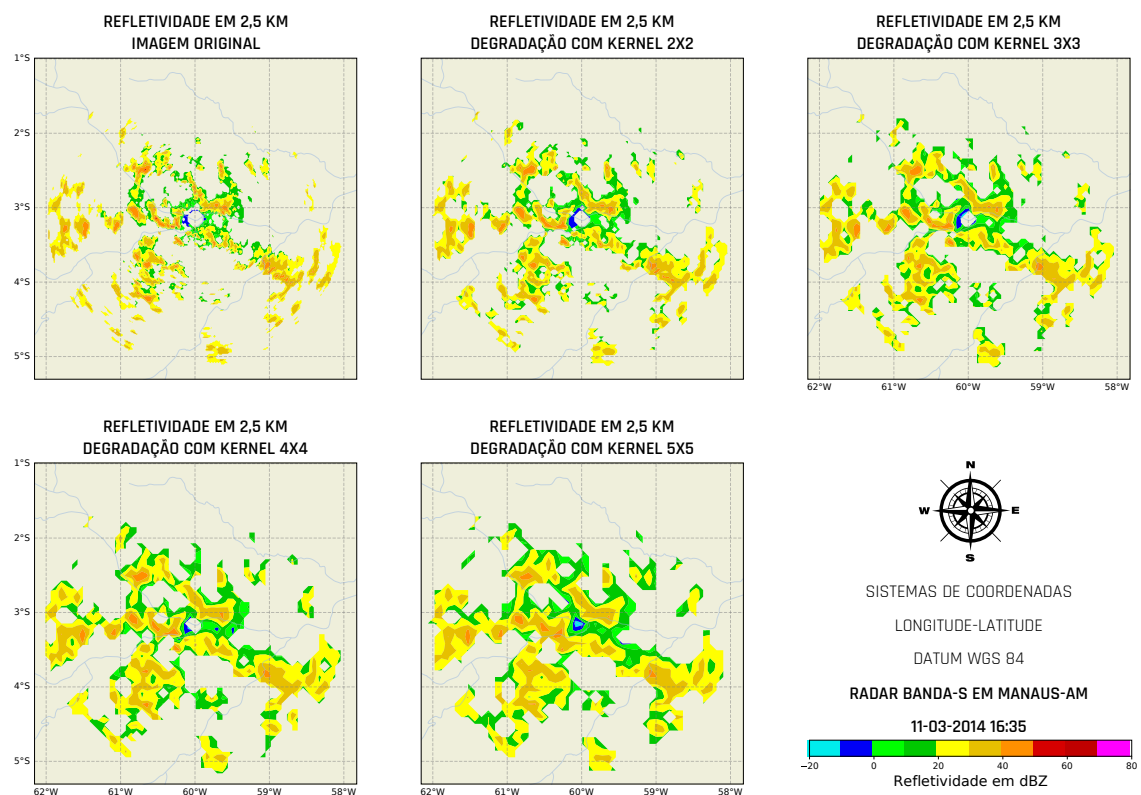
### 4.3 Análise dos dados degradados

Os dados provenientes de ambos instrumentos (i.e. radar e STARNET) podem estar sujeitos à imprecisões entre os pixels de uma vizinhança, o que conseqüentemente pode interferir nos resultados de suas co-localizações. A fim de mitigar esses problemas foram aplicadas as degradações descritas no capítulo anterior. A degradação ajuda a conter o erro de co-localização dos dados, resultando em uma melhor capacidade de associação entre os dados. Em contrapartida, esse processo acaba impactando na resolução espacial dos dados, o que pode influenciar no erro de localização do evento a ser previsto, já que a nova área a ser analisada será uma simplificação generalizada a partir do conjunto de outras. Entretanto, tal abordagem permite extrair informação de escalas maiores, importantes a natureza do sistema de tempestades.



A Figura 4.16 mostra o resultado da degradação de uma matriz de refletividade do produto CAPPI em 2,5 km com *kernels* de diferentes tamanhos. A imagem mostra os campos de refletividade radar na área de estudo no dia 11 de março de 2014 às 16:35. O exemplo corresponde a uma imagem da estação chuvosa onde diversos núcleos de precipitação desorganizados, alguns com refletividades associadas com taxas de chuvas que superam os 40 dBZ em várias partes da região.

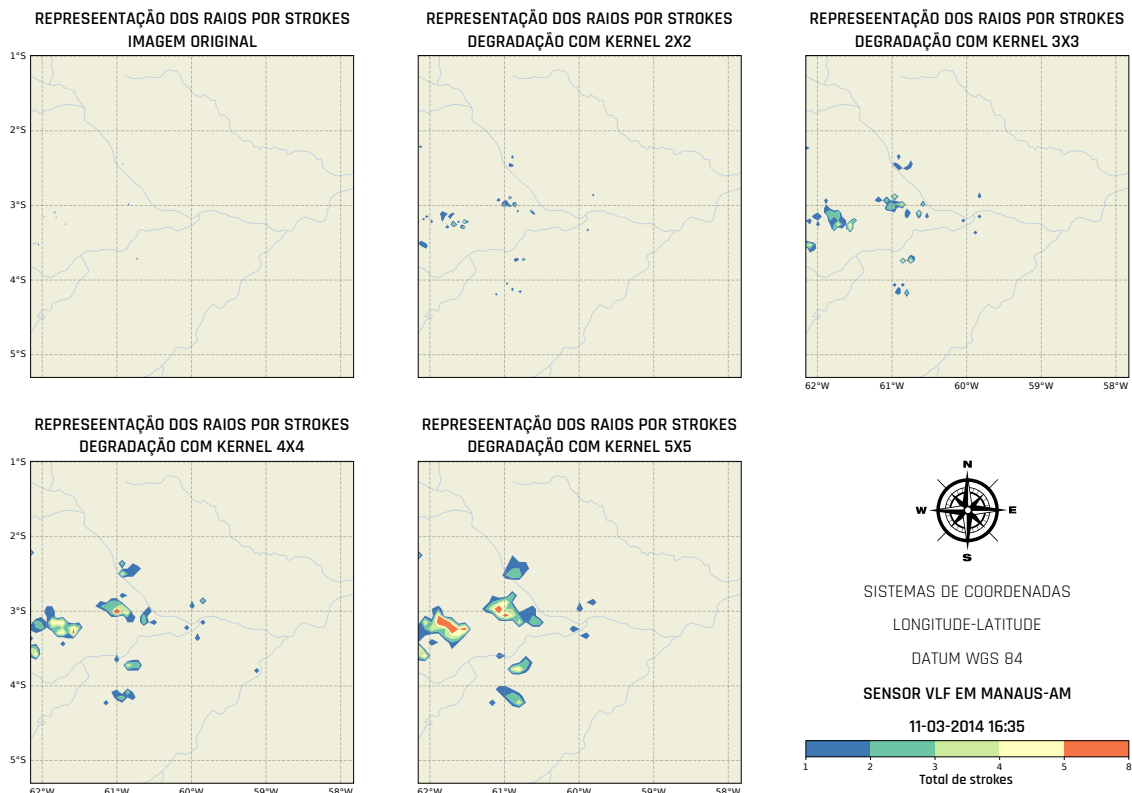
Figura 4.16 - Matriz de refletividade degradada com diferentes *kernels* para o dia 11 de março de 2014 às 16:35.



Fonte: Produção Autor.

Devido a desorganização dos sistemas ali mostrados na Figura 4.16, a ocorrência de descargas elétricas tendem a acontecer de forma mais dispersa, fazendo com que a densidade de raios por pixels na matriz seja ruidosa no dado original. Tal distribuição dificulta a co-localização entre os dados. Ao degradar também as matrizes de flashes, as regiões com maior ocorrência de dados ficam mais evidentes, facilitando a co-localização dos dados, mas aumentando o erro na localização. Todavia, a ideia aqui é mostrar uma possível melhor relação entre os dados que representam relações físicas mais evidentes, isto é, os valores mais intensos de refletividade com as maiores taxas de raios.

Figura 4.17 - Matriz de *strokes* degradada com diferentes *kernels* e a soma como função de agregação para o dia 11 de março de 2014 às 16:35.



Fonte: Produção Autor.

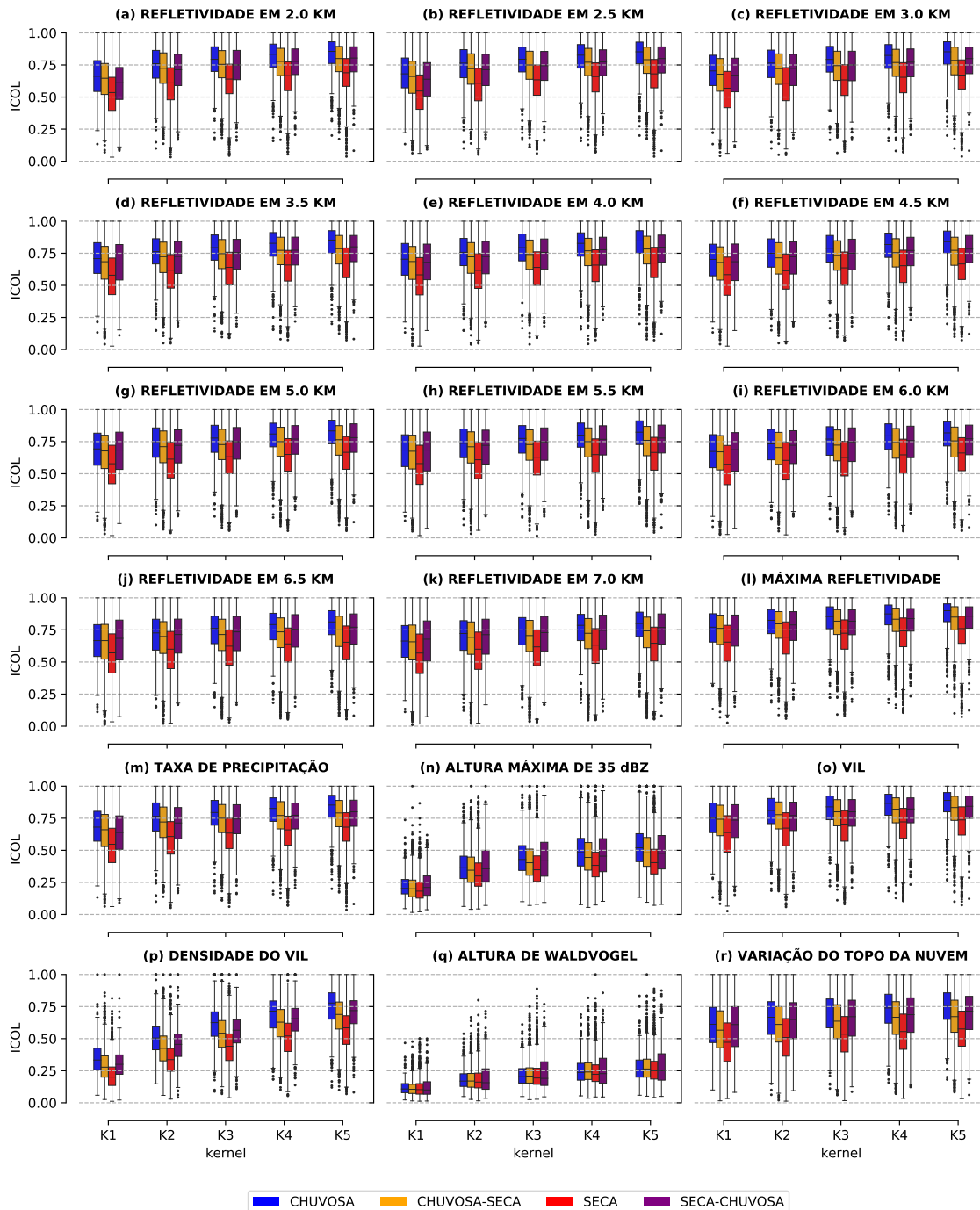
A Figura 4.17 mostra a imagem original de raios e suas derivações degradadas na mesma faixa de tempo e espaço da imagem do radar mostrada anteriormente. Nota-se que na imagem original a visualização das áreas onde ocorreram raios é pouco perceptível, mas quando degradada, as regiões mais ativas começam a ficar mais visíveis. E como apontado, ao aumentar muito o tamanho dos *kernels* de degradação, a perda de informações também começa a acontecer, como pode ser verificada pela fusão de diversas regiões ativas *kernels* maiores. Essa perda de informação também pode ser observada na Tabela 4.2, em que mostra os índices de co-localização e correlação para a imagem analisada.

Na Tabela 4.2 pode ser notado que, o índice de co-localização tem uma melhora muito significativa ao realizar a degradação da imagem. A correlação entre as variáveis do radar e os dados de raios também sofre mudanças com as modificações no tamanho do *kernel* utilizado na degradação. Na tabela também pode ser observado que a correlação obtida a partir dos dados degradados com diferentes funções de agregação também variam, em que para esse caso específico o ganho é mais relevante ao utilizar a soma como função de agregação. Entretanto os valores de correlação ainda são pouco representativos, havendo ganho somente no índice de co-localização. Esses ganhos no índice de co-localização podem impactar positivamente nas *skills* dos modelos de previsão como veremos na última seção deste trabalho. Entretanto, quanto maiores forem os tamanhos dos *kernels*, menos acurados os modelos serão, em relação à posição de onde aconteceu o evento, devido à perda de resolução decorrente da degradação.

Tabela 4.2 - Índices de co-localização e correlação para diferentes *kernels* para o dia 11 de março de 2014 às 16:35.

INTEGRIDADE DA IMAGEM	ICOL	ICOR		
		MÁXIMO	MÉDIA	SOMA
Imagem original (sem degradação)	0.57931	-0.07608	-0.07608	-0.07608
Imagem degradada com <i>kernel</i> 2×2	0.66165	-0.09758	-0.07198	0.06198
Imagem degradada com <i>kernel</i> 3×3	0.67257	-0.06722	-0.01314	0.09469
Imagem degradada com <i>kernel</i> 4×4	0.77320	-0.09025	-0.10825	0.15850
Imagem degradada com <i>kernel</i> 5×5	0.83333	-0.05409	-0.07680	0.13265

Figura 4.18 - Índice de co-localização por variável em todas as estações com degradações em diferentes tamanhos de *kernels*.

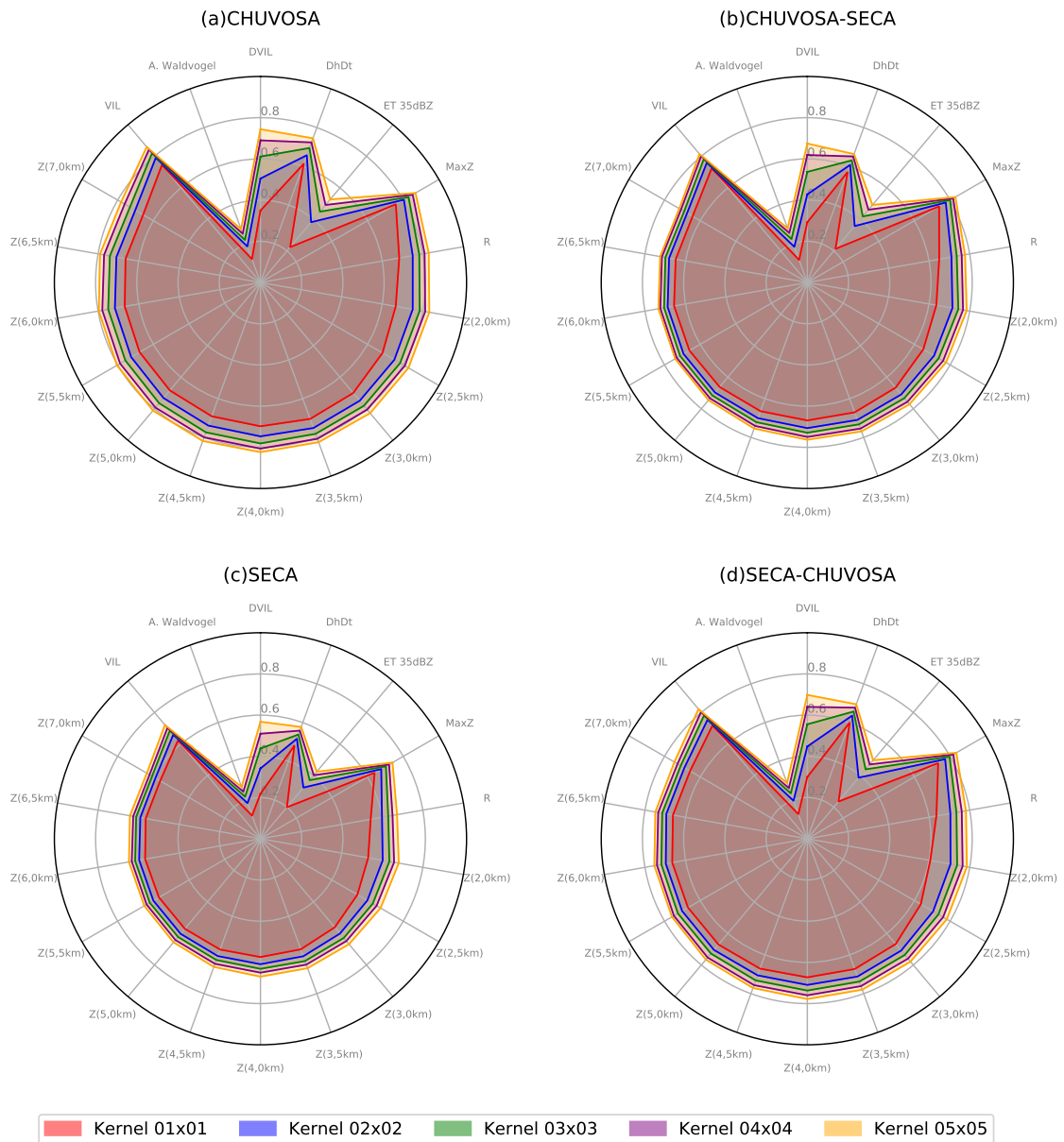


Fonte: Produção Autor.

A viabilidade do uso dos dados degradados nos modelos de previsão foi avaliada através dos índices de co-localização e correlação. Através das análises feitas a partir do índice de co-localização, é possível identificar padrões distintos com relação às variáveis, *kernels* e estações. A Figura 4.18 mostra os valores para o índice de co-localização para todas as variáveis e estações nos gráficos do tipo *boxplots*. Para todas as variáveis, o índice de co-localização melhora ao aumentar o tamanho do *kernel* de degradação. Entretanto, algumas variáveis há melhora mais significativa que outras. Na Figura 4.18 é possível notar que para as variáveis com informações do perfil vertical (e.g altura máxima de 35dBZ, densidade do VIL e altura de Waldvogel), essa melhora é mais expressiva que nas variáveis com informações de apenas um nível, como as refletividade em uma altura específica. É possível notar também que há uma discrepância nos valores de co-localização das diferentes estações. A diferença mais explícita é entre a estação chuvosa (em azul) e seca (em vermelho). Na estação chuvosa, independente da degradação, as variáveis estão melhor co-localizadas com os raios do que na estação seca. Isso pode estar associado, principalmente ao fato que a precipitação acontece de forma mais distribuída, o que pode favorecer a melhor co-localização dos dados nesta estação. Já na estação seca, a alta incidência de raios e o seu desenvolvimento em sistemas mais intensos e isolados, contribui para uma menor taxa de co-localização com variáveis. Essa tendência na variação no erro de co-localização entre as estações pode ser observada também nos valores referente a transição chuvosa-seca (em laranja) e seca-chuvosa (em roxo).

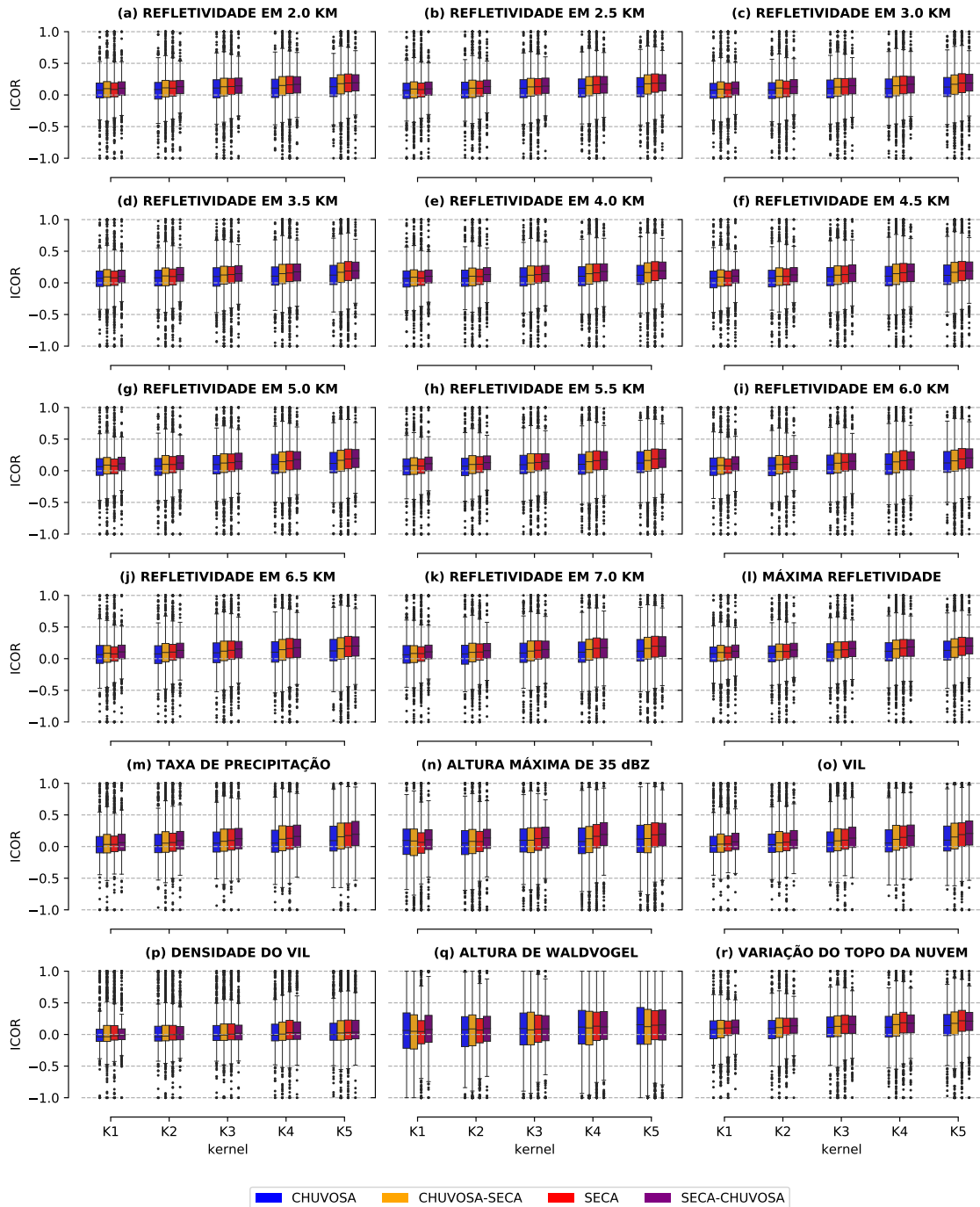
Como complemento desta análise, a Figura 4.19 apresenta as comparações entre as médias dos índices de co-localização para todas as variáveis nas estações e suas transições com todos o *kernels* aplicados. Nota-se pela figura que as variáveis que dependiam do desenvolvimento da camada mais alta da nuvem (e.g. *echo top* de 35dBZ, Altura de Waldvogel, DVIL, Dhdt) foram as com o menor índice médio de co-localização. Tal resultado reflete ao fato que estas variáveis ocupam uma menor quantidade de pixels da matriz de observação, uma vez que precisam de uma coluna de chuva mais bem desenvolvida. Logo, cobrindo uma área menor para a co-localização. Além disso, as Figuras B.1, B.2, B.3 e B.4 do Apêndice B contém os histogramas das frequências relativa das médias e desvios-padrão do índice de co-localização de cada variável em suas respectivas estações, chuvosa, chuvosa-seca, seca e seca-chuvosa. E na Tabela B.1 do mesmo Apêndice, além destas, estão algumas informações estatísticas mais detalhadas dos valores do índice de co-localização para consulta.

Figura 4.19 - Média dos Índices de co-localização por estação para todas variáveis com degradações em diferentes tamanhos de *kernels*.



Fonte: Produção Autor.

Figura 4.20 - Índice de correlação por variável em todas as estações com degradações em diferentes tamanhos de *kernel*s.



Fonte: Produção Autor.

O impacto da degradação também foi avaliado através do índice de correlação entre as variáveis meteorológicas e os dados da STARNET. Na Figura 4.20 são exibidos os *boxplots* com as distribuições dos valores do índice de correlação para todas as variáveis em cada estação com degradações com os diferentes *kernels*. Alguns padrões identificados na análise anterior não foram percebidos nos valores dos índices de correlação. Para esse índice, Todas as estações, e principalmente a seca, mostraram as distribuições dos valores de correlação levemente superiores à estação chuvosa. Este resultado reflete ao fato que, por ter baixa atividade convectiva neste período do ano, apesar de ter alta co-localização, a correlação é esperada ser mais baixa. De um modo geral, observa-se que as medianas se apresentam acima do valor zero e a medida que aumenta-se a degradação para cada *kernels* esses valores também aumentam. Isso mostra que existe correlação positiva na maioria dos casos, contudo, existe uma alta variabilidade dos valores observados principalmente pelos *outliers*. Tal comportamento pode estar associado a sub-notificação associada aos raios que se encontraram fora da região chuvosa devido a problemas de co-localização. Estes problemas se dão devido ao processo de medição de ambos os equipamentos, onde os raios se posicionam a frente da chuva devido ao rápido deslocamento dos sistemas, que muitas vezes não são identificados pelo radar na construção do CAPPI. Vale ressaltar que, assim como para os valores do índice de co-localização, não foram aplicados testes de hipótese para verificar se há diferença estatisticamente significantes nos valores do índice de correlação.

A função de agregação no índice de co-localização é indiferente, uma vez que, esse índice leva em consideração a atividade do pixel, independente do total de raios. Já a correlação pode variar de acordo com a função de agregação utilizada, conforme pôde ser visto na Tabela 4.2. Entretanto, as funções de agregação baseadas na soma e máximo, não foram avaliadas neste trabalho. Pois as configurações de RNAs testadas em seções posteriores são para previsões dicotômicas da inatividade/atividade, sendo assim, a função de agregação utilizada não teria efeito. Por isso, foi utilizada apenas uma, a média. Contudo, para trabalhos futuros, serão analisadas também outros processos de degradação como a média simples dentro de cada caixa, sem a janela deslizante, o vizinho mais próximo e outros métodos de interpolação comuns na meteorologia. Uma vez que este é uma das técnicas aplicadas na computação em redes neurais convolucionais. Além disso, essas degradações serão também aplicadas ao dado de entrada, ao invés de se utilizar apenas a média, realçando ainda mais alguns aspectos físicos observados nos dados de radar.

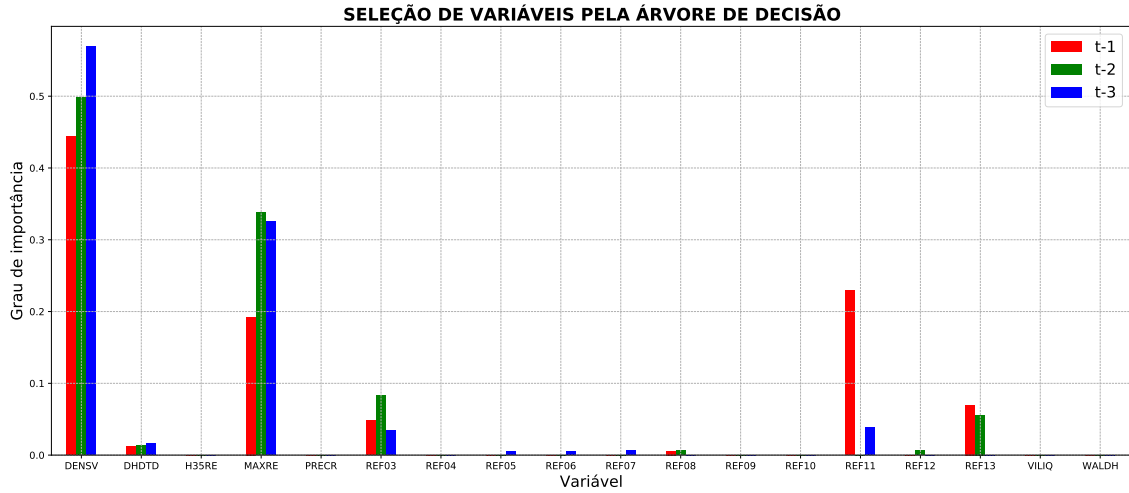


#### 4.4 Classificação da atividade de raios por Árvore de Decisão

Para aproveitar o máximo da robustez das técnicas de aprendizado de máquina é importante não só a configuração dos hiper-parâmetros dos modelos, mas também a seleção e pré-processamento dos dados. Uma parte expressiva do entendimento dos dados foi compreendida na etapa da sua análise exploratória, apresentada na Seção 4.2. No entanto, para refinar ainda mais a quantidade de testes a ser feito, a partir de uma perspectiva mais objetiva, foi realizada a classificação da atividade de raios em diferentes tempos com a árvore de decisão. A partir dessas classificações foram selecionadas as variáveis com grau de importância maior que zero nos tempos anteriores à previsão. Cada teste consistiu em fornecer como entrada as variáveis provenientes dos dados de radar que foram analisadas nas seções anteriores e descritos no Capítulo 2, totalizando 18 variáveis. Como saída, foi fornecido um rótulo aos modelos que indica a ocorrência de raios de forma dicotômica provenientes dos dados da STARNET. Esses testes se deram de forma individual para três tempos:  $t-1$  (12 minutos anterior à previsão),  $t-2$  (24 minutos anterior à previsão) e  $t-3$  (36 minutos anterior à previsão). Cabe ressaltar que a partir deste ponto todas as análises foram baseadas em dados da estação seca. Isso se dá pelo fato que foi nesta estação onde foram observados a maior quantidade de descargas elétricas na região de estudo. Além disso, por questões de maior rapidez nas análises de testes com a MLP, apenas dados associados ao IOP2 do projeto GoAmazon foram utilizados.

Essa abordagem foi utilizada para complementar as interpretações obtidas a partir da análise exploratória dos dados, realçando assim, as variáveis com mais relevância na classificação e/ou previsão da ocorrência de raios a partir de aprendizado de máquina. A Figura 4.21 mostra o grau de importância atribuído pela árvore de decisão para cada variável nos tempos  $t-1$ ,  $t-2$  e  $t-3$ . A densidade do VIL foi a variável com o melhor aproveitamento na classificação para os três tempos, seguido pela máxima refletividade, refletividade em 2 km e variável inferida que indica a variação do topo da nuvem. Essas variáveis, com exceção da refletividade em 2 km, possuem em comum características associadas ao perfil vertical, sendo um indicativo de que a previsão para tempos muito anteriores à previsão tem uma dependência assíncrona do desenvolvimento vertical da tempestade. Embora contra-intuitivo, o resultado apontado na figura sugere que mais variáveis foram utilizadas no tempo mais distante à previsão ( $t-3$ ). Entretanto, para algumas dessas variáveis o grau de importância foi muito baixo, como é o caso das variáveis de refletividade em 3km, 3,5km e 4km.

Figura 4.21 - Grau de importância atribuído às variáveis pela árvore de decisão em  $t-1$ ,  $t-2$  e  $t-3$ .

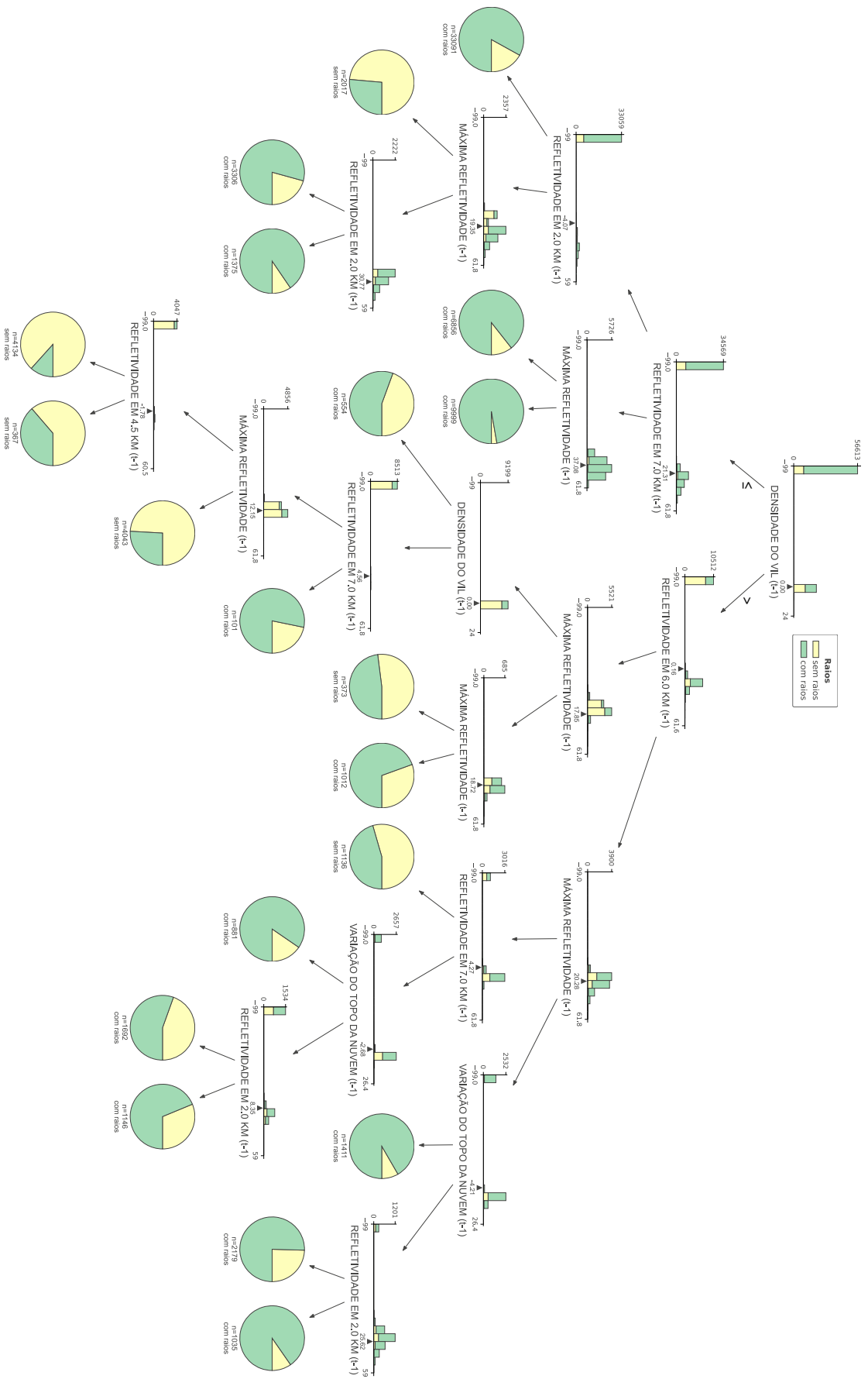


Fonte: Produção Autor.

A classificação da ocorrência de raios feita pela árvore de decisão não tem como propósito gerar os modelos finais de previsão, mas sim utilizar a sua capacidade de seleção de atributos. A partir de seus resultados será possível identificar variáveis mais relevantes em diferentes tempos para que possam ser utilizadas como entrada nas redes neurais, conforme será abordado na próxima seção. Além disso, a análise visual das representações das árvores de decisão são utilizadas para complementar a análise exploratória dos dados, uma vez que, através dela os possíveis limiares que separam melhor as classes de atividades e inatividade são realçados, o que é mais difícil de conseguir a partir de um modelo mais complexo, como as RNAs. Cabe ressaltar que, esses resultados podem também serem usados em outros estudos, como por exemplo, o rastreamento de tempestades eletricamente ativas a partir de processos de limiarização e clusterização, como aqueles aplicados por Vila et al. (2008). Uma vez que as variáveis aqui encontradas podem realçar as áreas em dados de radar que estão associadas com maior ocorrência de raios CG.

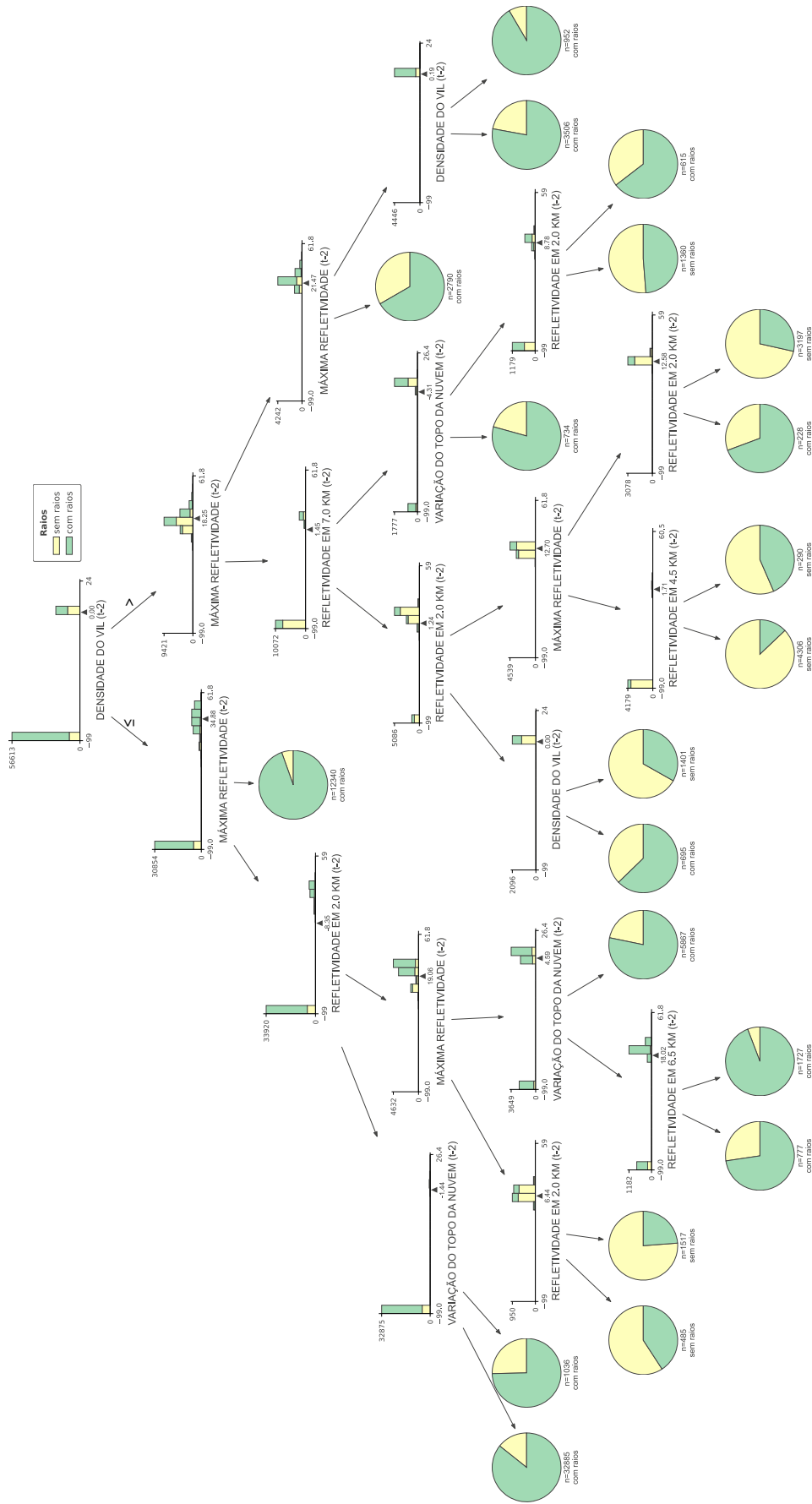


Figura 4.23 - Árvore de decisão gerada pela classificação da atividade de raios para o tempo *t-1*.



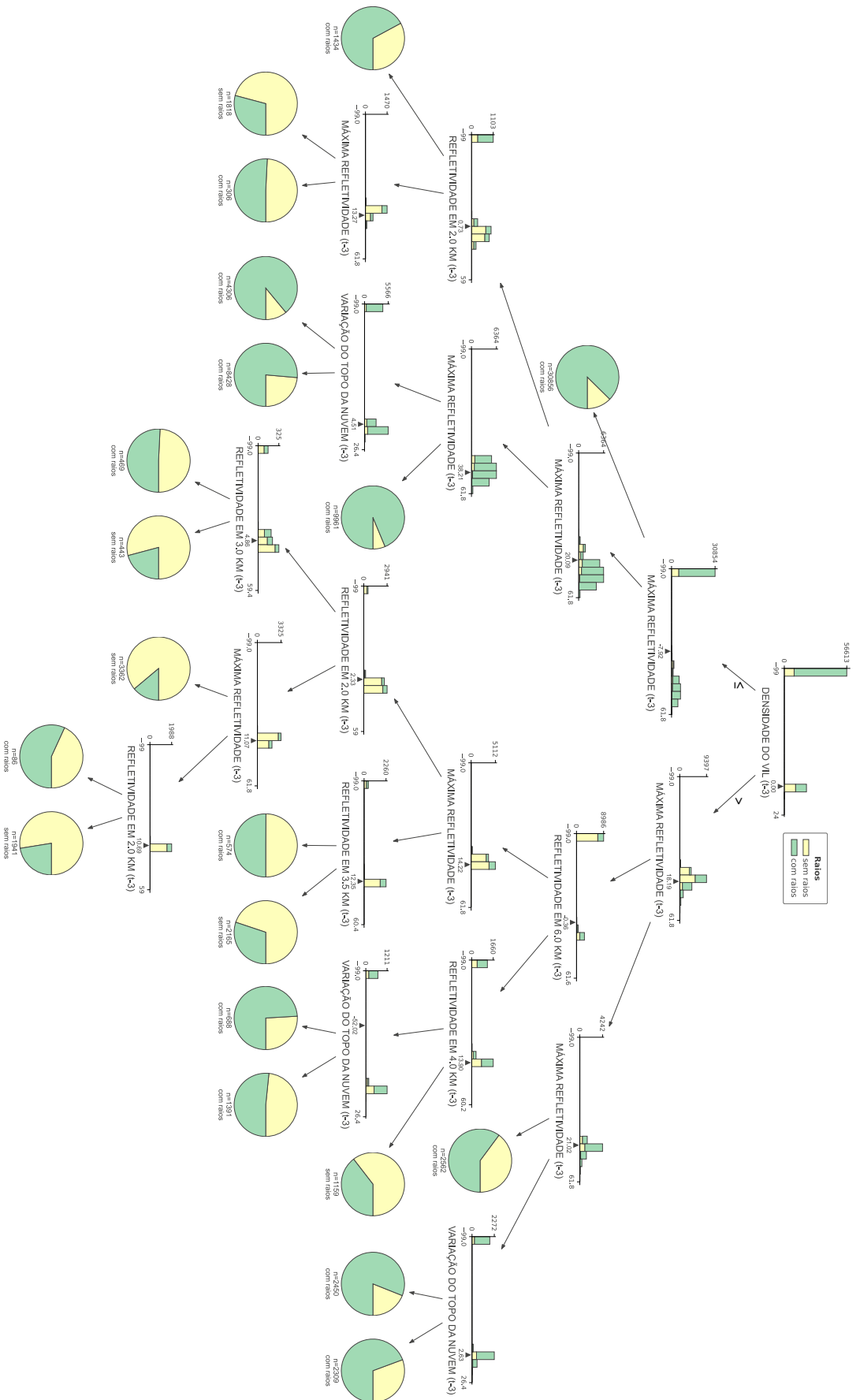
Fonte: Produção Autor.

Figura 4.24 - Árvore de decisão gerada pela classificação da atividade de raios para o tempo *t-2*.



Fonte: Produção Autor.

Figura 4.25 - Árvore de decisão gerada pela classificação da atividade de raios para o tempo *t*-3.



Fonte: Produção Autor.

As Figuras 4.22, 4.23, 4.24 e 4.25 mostram as árvores de decisão resultantes dos ajustes aos dados referentes ao IOP2 para os tempos  $t$ ,  $t-1$ ,  $t-2$  e  $t-3$ , respectivamente. O esquema da árvore de decisão é composto principalmente pelas representações dos dados através de histogramas e gráficos de pizzas, em que a cor amarela representa os dados sem a ocorrência de raios e em verde os dados com a ocorrência de raios. Os nós das árvores são representados pelos histogramas e os limiares criados pela DT são indicados pelos marcadores triangulares em suas bases. A cada nova partição dos dados para criação de novos nós, são geradas sub-árvores à esquerda e à direita, indicando, respectivamente, os dados menores e maiores que os limiares do nó atual. Por fim, os nós folhas são representados pelos gráficos de pizza, que mostram as proporções de dados de cada classe e os rótulos com suas respectivas classificações.

Nota-se nas figuras que o primeiro nó raiz da árvore corresponde à separação feita na densidade do VIL. Embora essa variável possua baixa representatividade numérica em relação às demais, por ter poucos pixels com essa informação, ela é uma das que melhor se destacou na separação com valores indefinidos, no primeiro ajuste aos dados. Note que geralmente a separação para valores acima dos limiares tendem a procurar áreas com raios, um indicativo que as variáveis com tendência de crescimento são melhores na definição inicial de ocorrência de raios. Em algumas novas partições o resultado não muda, o que significa que o nó anterior já seria suficientes para definir a ocorrência ou não de raios. Em outros casos a divisão não fica clara, com os gráficos de pizza mostrando uma certa semelhança para raio e não-raio, um possível problema de distribuições dos dados que pode estar associado a uma má co-localização dos dados.

A primeira figura desta análise sobre Árvores de Decisão (Figura 4.22) remete mais a uma questão de classificação e não previsão, pois verifica-se uma análise síncrona. Ao analisar seus resultados nota-se que alguns processos físicos podem remeter a processos dinâmicos e microfísicos das nuvens observadas. Observa-se que os valores de refletividade em diferentes níveis são selecionados como aqueles que melhor definem a ocorrência de raios. Isto remete ao fato que diferentes processos dentro da nuvem podem modificar as características físicas, principalmente aquelas associadas as quantidade de hidrometeoros dentro da nuvem. Como exemplo, a presença de fortes correntes ascendente, que transportam gotículas de nuvens superesfriadas e partículas de gelo para níveis superiores (HOUZE, 2014). O desenvolvimento de áreas mais altas, como a região de mistura da nuvem e/ou a parte glaciada podem ajudar no desenvolvimento de descargas elétrica dependendo do local de desenvolvimento da tempestade (DWYER; UMAN, 2014). Uma das sub-árvores é bem instrutiva, como

a segunda da esquerda para a direita. Onde a presença de raio é dada quando ocorre valores de refletividade em 7km acima de 26,31 dBZ e em baixos níveis (2km) acima de 31,7 dBZ. Tais valores podem ser aplicados em sistemas de rastreios para melhor identificar áreas em dados de radar com capacidade de gerar descargas elétricas. Contudo, este não faz parte do escopo desta pesquisa, mas podem ser considerado em questões futuras.

A Figura 4.23 mostra a Árvore de Decisão para os dados de radar em  $t-1$  associados aos raios em  $t$ . Alguns aspectos físicos podem ser associados aos critérios definidos pela DT, podendo ser observados a partir da análise de sua representação. Os dados da sub-árvore à esquerda do nó raiz já não tem mais co-localização com a variável com maior relevância (densidade do VIL), segundo os critérios da DT, sendo assim é escolhida a refletividade em níveis mais altos (i.e. 7km), onde o valor limítrofe foi de 21,31 dBZ. Já o nó seguinte está associados à refletividade máxima no perfil, onde um valor relativamente alto de 37,08 dBZ foi escolhido como limiar. Esse comportamento poder realçar a presença das forte correntes ascendentes, que através da sua intensificação, pode produzir maior saturação, levando a produção de mais gotículas de água e/ou gelo (PRUPPACHER; KLETT, 2010). Nesse processo, as partículas mais leves geradas pelas correntes ascendentes são transportadas para os níveis superiores, enquanto que, as mais pesadas se sustentam na base da nuvem, criando camadas mais intensas nos níveis inferiores, um típico perfil de nuvens convectivas (LHERMITTE; WILLIAMS, 1983; YUTER; HOUZE, 1995). Nota-se a sub-árvore mais a esquerda que foram necessárias mais interações (partições) para definir se ocorre ou não raios. Neste caso o nó folha que definiu “com raio” identificou que quando os níveis inferiores apresentam certa intensidade (limiar maior que 30,77 dBZ em 2km). Tal resultado reflete um possível maior desenvolvimento das camadas quentes das nuvens. As sub-árvores mais a direita tiveram um comportamento similar, contudo focando mais em níveis médios e variações na altura do topo da nuvem (DHDT). Sendo este último responsável por duas folhas importantes na definição de ocorrência de raios.

Para a Figura 4.24 mostra a árvore de decisão para o tempo  $t-2$ . Fica evidente na figura que a variável mais importante (após a separação do DVIL) é a máxima refletividade no perfil (em dBZ), como pode ser verificado também na Figura 4.21. Em seguida a refletividade em níveis inferiores é mais relevante. Tal resultado mostra que a dinâmica já não fica tão evidente. Os sinais ali apresentados mostram que apenas pixeis com valores relativamente altos podem ainda conter algum sinal da ocorrência de raios. Um exemplo claro é a sub-árvores do segundo nó (folha a direita),



onde apenas as refletividades máximas foram suficientes para definir a ocorrência de descargas. Cabe ressaltar que isso pode ter uma baixa precisão em sua classificação.

A medida se distancia do tempo de ocorrência dos raios, as relações se tornam ainda mais difíceis de ter algum tipo de relação física direta. Nota-se que a sub-árvore mais a esquerda, cujo o nó folha diz que tem raio em valores muito baixo de refletividade. Tal resultado mostra uma inconsistência física relativamente grande. Isso se dá devido ao rápido desenvolvimento de raios e sistemas tempestuosos. Onde os processos lagrangianos não permitem mais que existe uma boa relação entre as variáveis da árvore. Contudo, ainda existem alguns aspectos físicos presentes, como a necessidade altos valores de refletividade como possível sinal de raios em tempos mais distantes. Isto pode estar relacionado a eventos lentos com grande desenvolvimento vertical e severidade, provavelmente ligados a sistemas convectivos de mesoescala (HOUZE, 2014). Como exemplo podemos citar a segunda sub-árvore mais a esquerda e aquela mais a direita.

De modo geral as DTs conseguiram definir alguns critérios físicos importantes que podem ser usados tanto na classificação como na previsão. Principalmente no que diz nas questões de seleção de atributos. Tais informações serão levadas em consideração na seção posterior. Contudo, observou-se ainda que a maioria das variáveis conhecidas pela sua capacidade de melhor diferenciar tempestades elétricas, não foram selecionadas pelas DTs aqui analisadas. Como por exemplo, a altura do *echo top* de 35 dBZ e a altura de Waldvogel. Tal fato deve estar associado a uma menor co-localização desses dados em relação a outros, como as refletividades em diferentes níveis e a máxima refletividade. Estudos mais profundos precisam ser realizados para melhor entender e associar os dados, definindo com mais propriedade as DT e seu uso na classificação e previsão de raios.

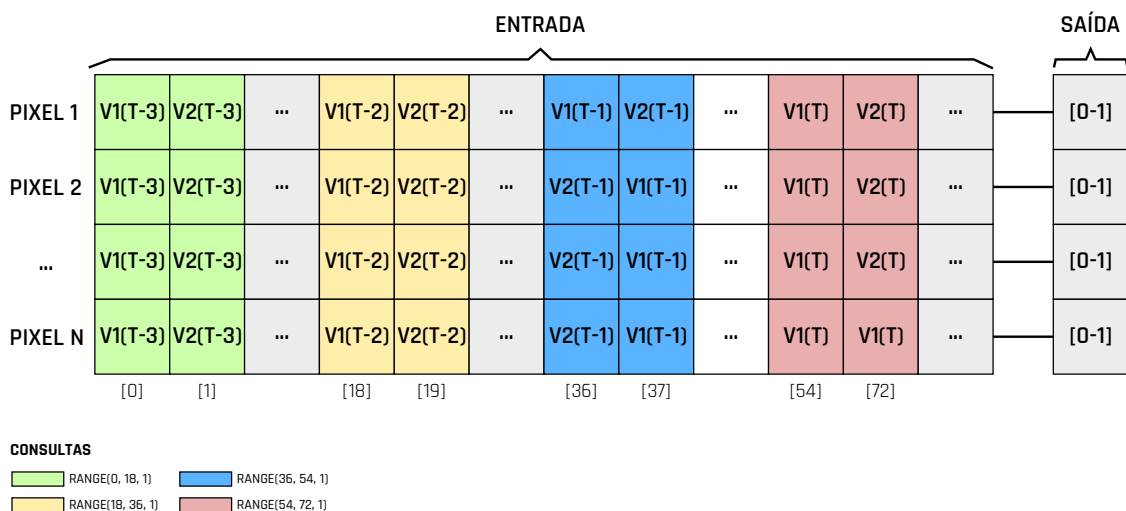
#### 4.5 Previsão de raios com *perceptron* multicamadas

Foram realizados diversos testes de sensibilidade, com base nas análises anteriores, em uma arquitetura de rede neural muito difundidas pela comunidade, a MLP (descrita na Seção 3.3). Sendo assim, para averiguar os impactos inerentes aos dados foram mantidas fixas as configurações das redes, conforme descritas na seção mencionada. Como citado na seção anterior, e reforçando novamente, ressalta-se que todas as análises foram baseadas nas medidas associadas ao IOP2 do projeto GoAmazon. Devido não só a maior quantidade de descargas observadas durante esse período, mas também por causa da necessidade de maior rapidez nos testes, usando assim um conjunto de dados menor para treinamento. Logo, é sabido que isso pode ter impacto direto na capacidade de aprendizado do modelo e as análises levam isto em consideração.

#### 4.5.0.1 Testes de sensibilidade para os dados de entrada

Como visto em seções anteriores, as variáveis do radar selecionadas para este trabalho apresentam certas características que podem ajudar a determinar a presença ou não de raios sobre a região de estudo. De modo a analisar os impactos destas variáveis em redes neurais simples, como a MLP, serão feitos testes de sensibilidade dos seus dados de entrada. As métricas aqui utilizadas são POD e FAR, descritos na Seção 3.3. A fim de garantir a homogeneidade dos dados de entrada nos diferentes testes, estes foram processados e agrupados previamente, conforme mostrado na Figura 4.26. Com isso, é mantido a referência dos pixels que serão utilizados no treinamento, validação e teste, mudando apenas a acervo de informações que serão acrescentadas na camada de entrada das diferentes RNAs, tais como: quantidade de variáveis meteorológicas e suas condições em tempos diferentes; dados de cobertura de solo; e ciclo diurno. Sendo assim, para cada teste de sensibilidade os dados de entrada são filtrados a partir das múltiplas seleções dos índices dessa estrutura de dados bi-dimensional. Os dados de entrada são definidos a partir da configuração da camada de entrada de cada uma das MLPs utilizadas. As descrições dos dados de entrada dessas primeiras configurações de cada rede são apresentadas a seguir.

Figura 4.26 - Esquema da preparação dos dados de entrada da RNA. Em que  $V_i$  são as variáveis utilizadas em cada instante de tempo  $T_j$ .



TREINAMENTO: 75%, VALIDAÇÃO: 15% E TESTE: 10%

Fonte: Produção Autor.

- MLP01 - Variáveis: Todos os valores de refletividade entre as alturas de 2km e 7km (11 níveis). Os tempos (instantes) utilizados: Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis em todos os tempos anteriores preenchidas.
- MLP02 - Variáveis: Todos os valores de refletividade entre as alturas de 2km e 7km (11 níveis). Os tempos (instantes) utilizados: O último tempo anterior ( $t-1$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis no tempo anterior preenchidas.
- MLP03 - Variáveis: Refletividade na altura de 2,5km. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis em todos os tempos anteriores preenchidas.
- MLP04 - Variáveis: Refletividade na altura de 2,5km. Os tempos (instantes) utilizados: O último tempo anterior ( $t-1$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis no tempo anterior preenchidas.
- MLP05 - Variáveis: Todas as variáveis do radar. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade em todos os tempos anteriores preenchidas.
- MLP06 - Variáveis: Todas as variáveis do radar. Os tempos (instantes) utilizados: O último tempo anterior ( $t-1$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade no tempo anterior preenchidas.
- MLP07 - Variáveis: Todas as variáveis do radar selecionadas pela DT com grau de importância maior que zero. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos duas entradas preenchidas.
- MLP08 - Variáveis: Todas as variáveis do radar selecionadas pela DT com grau de importância maior que zero. Os tempos (instantes) utilizados: O último tempo anterior ( $t-1$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos duas entradas preenchidas.

- MLP09 - Variáveis: Melhor conjunto de variáveis a partir do teste com todas as combinações possíveis (Altura de Waldvogel). Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.

Diante dessas configurações definidas a partir da análise exploratória de dados, seleção de atributos via DT e dos testes de sensibilidade, os melhores modelos foram selecionados para fazer outras parametrizações das variáveis de entrada, adicionando também informações de cobertura do solo e horário. Conforme será apresentado nas subseções seguintes, as configurações com melhor desempenho nos testes foram as configurações MLP05 e MLP09. Diante disso, novas configurações foram definidas, conforme descritas a seguir.

- MLP10 - Variáveis: Todas as variáveis do radar, dados de cobertura do solo e hora. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.
- MLP11 - Variáveis: Todas as variáveis do radar e dados de cobertura do solo. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.
- MLP12 - Variáveis: Todas as variáveis do radar e hora. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.
- MLP13 - Variáveis: Melhor conjunto de variáveis a partir do teste com todas as combinações possíveis (Altura de Waldvogel), dados de cobertura do solo e hora. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.
- MLP14 - Variáveis: Melhor conjunto de variáveis a partir do teste com todas as combinações possíveis (Altura de Waldvogel) e dados de cobertura do solo. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.

- MLP15 - Variáveis: Melhor conjunto de variáveis a partir do teste com todas as combinações possíveis (Altura de Waldvogel) e hora. Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Pelo menos a entrada do tempo anterior preenchida.

Por fim, outro conjunto de configurações foram definidas com o intuito de avaliar o desempenho das redes neurais na previsão de raios a partir de dados degradados. Uma vez que, esse teste visa avaliar o desempenho das previsões das redes neurais em dados de diferentes resoluções espaciais, as novas configurações foram baseadas na MLP05, que possui como entrada todas as variáveis nos três instantes de tempos anteriores à previsão. Além disso, a MLP05 foi uma das configurações que apresentaram melhor desempenho, conforme será apresentado em seguida. A lista a seguir descreve as configurações criadas para a realização desses testes.

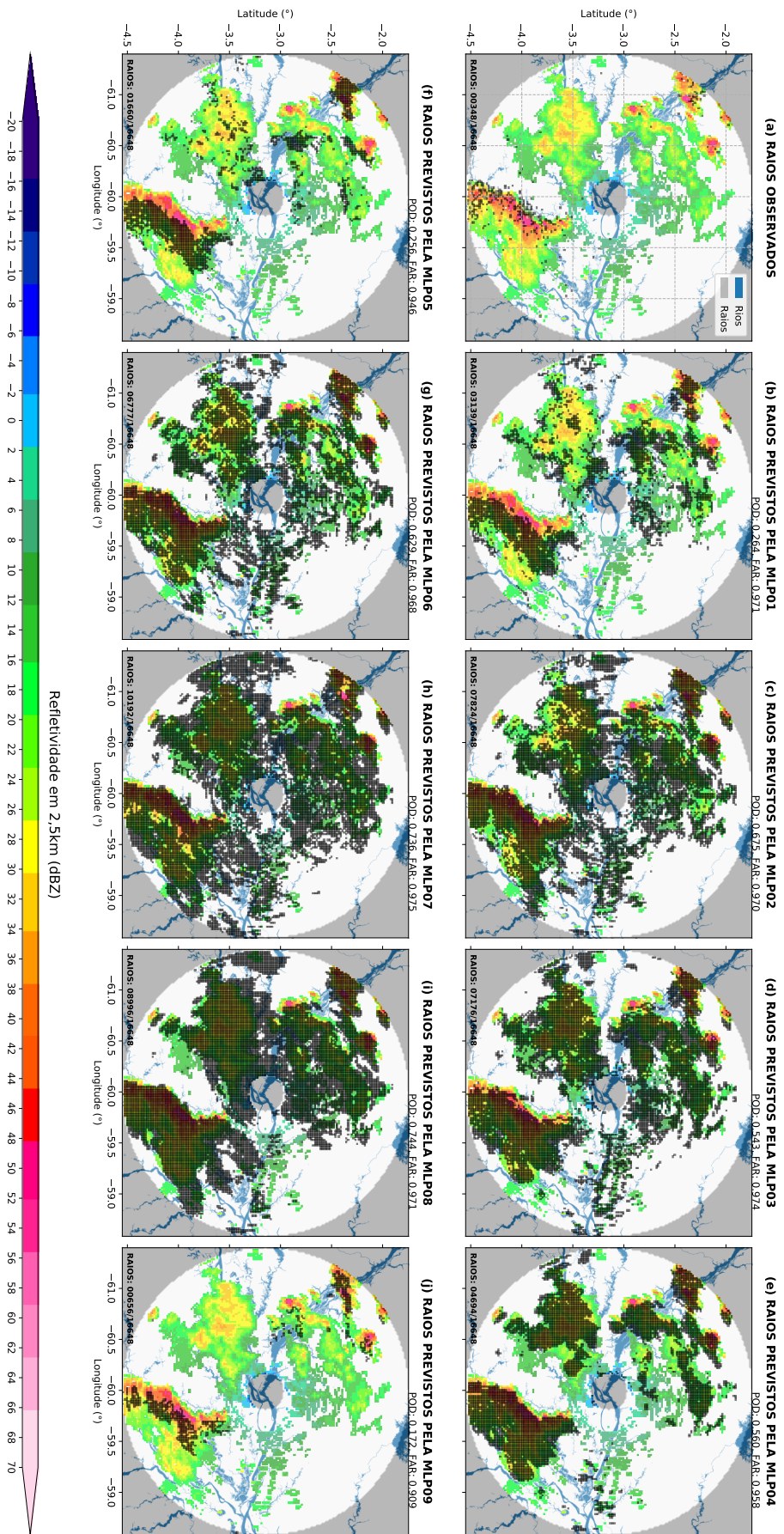
- MLP16 - Variáveis: Todas as variáveis do radar degradadas com *kernel*  $2 \times 2$ . Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade em todos os tempos anteriores preenchidas.
- MLP17 - Variáveis: Todas as variáveis do radar degradadas com *kernel*  $3 \times 3$ . Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade em todos os tempos anteriores preenchidas.
- MLP18 - Variáveis: Todas as variáveis do radar degradadas com *kernel*  $4 \times 4$ . Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade em todos os tempos anteriores preenchidas.
- MLP19 - Variáveis: Todas as variáveis do radar degradadas com *kernel*  $5 \times 5$ . Os três últimos tempos anteriores ( $t-1$ ,  $t-2$  e  $t-3$ ) à previsão ( $t$ ). Condição de seleção do pixel válido para o treinamento e validação: Todas as variáveis de refletividade em todos os tempos anteriores preenchidas.

As configurações das MLP01 a MLP08 foram definidas partir de informações obtidas na análise exploratória dos dados, relações síncronas/assíncronas e das interpretações feitas pelos resultados da DT. A MLP09 foi definida com base em um teste de sensibilidade mais simples e com menos rigor, em que foram testadas RNAs com todas as combinações possíveis de entrada em três tempos anteriores à previsão, totalizando 262143 RNAs. A fim aproveitar o máximo possível esse teste, foi aplicado nele um controle de qualidade nas amostras selecionadas e feito o seu balanceamento, de tal forma que os dados selecionados fossem o mais representativo possível para cada classe, já que os tamanhos das amostras eram pequenos em relação às populações (aproximadamente 7%). Com o intuito de agilizar o tempo de processamento, as amostras desses testes foram retiradas de 5000 sequências de imagens da estação seca dos anos de 2014 e 2015, em que as sequências pertenciam a eventos contínuos com duração de pelo menos duas horas. O seu balanceamento foi feito igualando a quantidade de dados da classe majoritária com os da minoritária. Nesses testes, a que teve melhor *skill* avaliada pelas métricas POD e FAR foi a combinação de uma única variável, a altura de Waldvogel.

Todas as análises a seguir se baseiam principalmente nos resultados mostrados na Tabela 4.3 e na Figura 4.27. A Tabela 4.3 mostra as *skills* gerais obtidas a partir dos testes com as configurações das RNAs descritas. Esses dados de testes representam os pixels extraídos dos 10% da população de imagens de radar observadas durante o IOP2, totalizando 1608593 pixels. A fim de analisar o erro de co-localização decorrente da propagação dos sistemas, foram aplicadas as previsões de todos os modelos para um evento acontecido em 6 de outubro de 2014, em que a presença de uma tempestade eletricamente ativa foi identificada. Este sistema apresentou características de deslocamento e desenvolvimento bem conhecidas na região, uma vez que o mesmo se deslocou de leste para oeste e começou no final da tarde e se estendeu pela noite. A Figura 4.27 mostra a observação e as previsões de raios por algumas das configurações de MLP apresentadas acima. A observação e as previsões são para o horário das 20:00h UTC e estão sobrepostas aos valores de refletividade em 2,5km nesse mesmo instante de tempo.

Figura 4.27 - Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade > 50%.

MANAUS-AM 2014-10-06 20:00 UTC



Fonte: Produção Autor.



Nesse mosaico (Figura 4.27) é possível observar que todos os modelos acabam superestimando, em menor ou maior proporção, o total de raios observados. Todas as metodologias aplicadas não foram capazes de acompanhar o deslocamento do sistema. Seus resultados apresentam um claro problema associado a propagação das tempestades. Todavia, as formas de algumas previsões apresentaram-se semelhantes à observação, o que mostra um certo *skill* em determinar a ocorrência total de descargas. Possíveis melhorias serão acrescentadas na seção de conclusões. Uma vez que o pre-processamento para uma melhor co-localização seja não trivial, preferiu-se apenas analisar os dados dispostos nas informações originais. Note que na figura é possível verificar os valores de FAR e POD para a imagem apresentada, assim como o total de pixels classificados com raios pelos modelos. Ressalta-se ainda que, as descargas que não estavam associadas aos valores de refletividade em qualquer nível foram eliminadas destas imagens, ou seja, os raios que não estão sobrepostos à qualquer das variáveis do perfil vertical de refletividade não são contabilizadas. Uma métrica que avalie os erros de co-localização entre os dados é necessária, entretanto, isso não foi aplicado nas avaliações gerais da Tabela 4.3.

Uma das previsões que se destaca é a realizada com a configuração da MLP09, que tem como entrada a altura de Waldvogel em três tempos anteriores à previsão, configuração que foi obtida a partir do teste de sensibilidade com todas as combinações e número de variáveis possíveis, conforme mencionado anteriormente. Essa RNA é a que proporcionou a menor superestimação nas previsões, apontando de forma bastante coerente as regiões mais ativas e em que foram observados raios, conforme pode ser visto no mosaico da Figura 4.27.

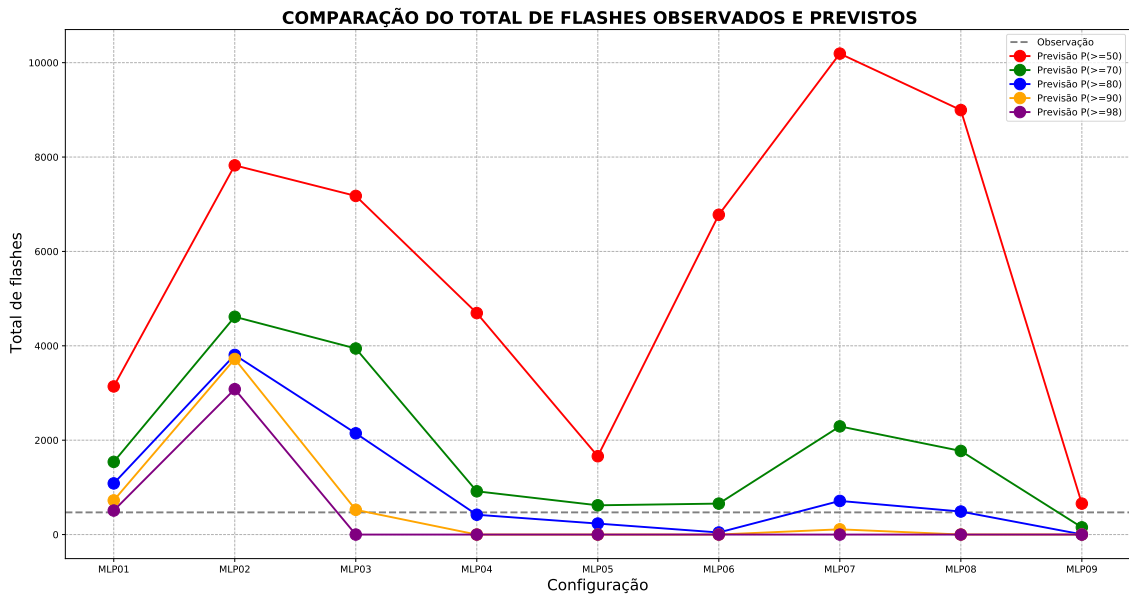
As análises das variações dos limiares de probabilidade nas previsões das diferentes configurações de redes partiram da Figura 4.28, que mostra o total de raios observados e previstos por cada MLP com os limiares de probabilidade para o caso de estudo exposto anteriormente. O total de raios observados é representado pela linha tracejada horizontal cinza, e as previsões com diferentes limiares de probabilidade na camada de saída são representadas pelas linhas contínuas e marcadores circulares, e são distinguidos pelas cores. Como previsível, as configurações com os limiares mais baixos tem uma superestimação na previsão dos raios. Em contrapartida, os limiares mais altos tendem a suavizar mais a sensibilidade na detecção de raios, em que para alguns casos não é capaz de prever a detecção, como é feito na maioria das configurações com os limiares  $>90\%$  e  $>98\%$ .

Tabela 4.3 - *Skills* dos testes realizados com as diferentes configurações da RNA baseadas na EDA.

CONF.	LIMIAR>50%		LIMIAR>70%		LIMIAR>80%		LIMIAR>90%		LIMIAR>98%	
	POD	FAR	POD	FAR	POD	FAR	POD	FAR	POD	FAR
MLP01-A	0.511	0.941	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MLP01-B	0.784	0.970	0.097	0.904	0.001	0.900	0.000	1.000	0.000	0.000
MLP01-C	0.853	0.976	0.227	0.930	0.009	0.930	0.000	1.000	0.000	1.000
MLP02-A	0.328	0.913	0.000	1.000	0.000	1.000	0.000	1.000	0.000	0.000
MLP02-B	0.647	0.961	0.011	0.897	0.000	0.976	0.000	1.000	0.000	0.000
MLP02-C	0.890	0.977	0.237	0.920	0.011	0.970	0.001	0.991	0.000	0.991
MLP03-A	0.758	0.975	0.186	0.914	0.000	1.000	0.000	0.000	0.000	0.000
MLP03-B	0.898	0.986	0.436	0.947	0.037	0.886	0.000	1.000	0.000	0.000
MLP03-C	0.950	0.991	0.746	0.977	0.463	0.953	0.114	0.912	0.002	0.973
MLP04-A	0.478	0.963	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MLP04-B	0.842	0.985	0.331	0.937	0.000	0.000	0.000	0.000	0.000	0.000
MLP04-C	0.914	0.989	0.473	0.962	0.246	0.924	0.000	0.000	0.000	0.000
MLP05-A	0.481	0.939	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MLP05-B	0.831	0.973	0.131	0.897	0.000	0.000	0.000	0.000	0.000	0.000
MLP05-C	0.879	0.980	0.551	0.950	0.131	0.913	0.000	0.000	0.000	0.000
MLP06-A	0.424	0.931	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MLP06-B	0.684	0.956	0.021	0.000	1.000	0.000	0.000	1.000	0.000	0.000
MLP06-C	0.924	0.979	0.369	0.921	0.001	0.929	0.000	0.000	0.000	0.000
MLP07-A	0.694	0.971	0.093	0.929	0.000	0.000	0.000	0.000	0.000	0.000
MLP07-B	0.819	0.982	0.357	0.946	0.004	0.958	0.000	0.000	0.000	0.000
MLP07-C	0.924	0.991	0.684	0.972	0.276	0.950	0.010	0.951	0.000	0.000
MLP08-A	0.633	0.969	0.114	0.876	0.000	0.900	0.000	0.000	0.000	0.000
MLP08-B	0.857	0.984	0.387	0.944	0.001	0.976	0.001	0.957	0.000	0.959
MLP08-C	0.934	0.989	0.585	0.966	0.233	0.920	0.001	0.864	0.000	0.000
MLP09-A	0.001	0.917	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MLP09-B	0.595	0.891	0.003	0.867	0.000	1.000	0.000	0.000	0.000	0.000
MLP09-C	1.000	0.915	0.025	0.812	0.007	0.868	0.003	0.882	0.000	1.000

As previsões das ocorrências de raios mostradas na Figura 4.27 foram feitas a partir do limiar de probabilidade padrão, ou seja,  $>50\%$ , os mosaicos com as previsões utilizando os limiares de probabilidade  $>70\%$ ,  $>80\%$ ,  $>90\%$  e  $>98\%$  não serão analisadas, mas podem ser consultados no Apêndice C.

Figura 4.28 - *Flashes* observados e previstos com as diferentes configurações de RNA.



Fonte: Produção Autor.

A fim de avaliar os impactos das variáveis de entrada na previsão de raios a partir da MLP, alguns testes foram discutidos com maior profundidade, e serão apresentados a seguir.

### Teste 1: balanceamento dos dados e probabilidade

O primeiro teste refere-se a um problema inerente às RNAs, o balanceamento de dados (HAIXIANG et al., 2017). Por ser um evento de alta variabilidade espacial e temporal, são raros os eventos com a detecção de raios. O impacto disso nos modelos de aprendizado de máquina está no desbalanceamento dos dados, onde os eventos não-raios são muito mais frequentes que aqueles associados à ocorrência de raios. Se não houver o balanceamento dos dados que favoreça a previsão do alvo, é muito provável que o modelo fique enviesado a generalizar somente os eventos da classe majoritária (TAHIR et al., 2009; HAIXIANG et al., 2017). Algumas abordagens foram

utilizadas para conter essa problemática, uma delas foi a customização da função utilizada para a atualização dos pesos das RNAs durante o *backpropagation*, sendo utilizada a *Focal Loss*. Devido ao desbalanceamento demasiado dos dados, somente a implementação da *Focal Loss* nos modelos não foi suficiente para que houvesse ganho na capacidade de generalização. Sendo assim, foi feito o equilíbrio das proporções dos dados de cada classe.

A fim de melhorar a representatividade das classes usadas no processo de treinamento/validação/teste das RNAs durante os testes de sensibilidade, foram aplicados para cada configuração os balanceamentos com as seguintes proporções de dados: A(Atividade: 25%, Inatividade: 75%); B(Atividade: 50%, Inatividade: 50%); e C(Atividade: 75%, Inatividade: 25%). A geração de dados para complementar a classe minoritária foi desconsiderada, uma vez que este trabalho propõe a parametrização de variáveis através de testes de sensibilidade sem intervenção direta. Sendo assim, a quantidade de dados da classe majoritária foi reduzida aleatoriamente para igualar às proporções citadas, abordagem conhecida como subamostragem aleatória (TAHIR et al., 2009).

Como dito anteriormente, a Tabela 4.3 mostra as *skills* (POD e FAR) para as configurações de todas RNAs analisadas no teste de sensibilidade com as variáveis de entrada dos modelos. Um dos principais ganhos que pode ser observado nessa tabela é o aumento de performance do modelo em relação ao balanceamento das classes. Em que para todas as configurações o POD aumenta significativamente ao incrementar a proporção de dados referente à classe de atividade, em contrapartida o aumento do FAR é mais modesto. Como já mencionado, essas métricas sofreram bastante com a falta da capacidade dos modelos em acompanhar o deslocamento dos sistemas de tempestades. Mesmo assim foi possível verificar um ganho no POD de um modo geral.

Todas as configurações utilizam na camada de saída uma função de ativação com probabilidades, a *softmax*. Sendo assim, alguns limiares de probabilidade de ocorrência de raios foram aplicados, a fim de avaliar a subestimação/superestimação em configurações que tem variáveis fortemente associadas aos raios. Os limiares da probabilidade aplicados foram >50%, >70%, >80%, >90% e >98%.

Com relação a estes limiares de probabilidade, notou-se que as *skills* através de sua variação sugere que o seu aumento influencia em uma maior subestimação na detecção dos raios. Onde o POD é penalizados em função de uma pequena melhoria do FAR em alguns casos. Uma vez que se limita ainda mais a detecção dos mesmos. Isso se dá pela alta variabilidade espacial dos raios e, principalmente, pelos problemas de co-localização decorrente do não acompanhamento da previsão em relação à propagação dos sistemas observados. Por outro lado estes resultados mostram que a função utilizada pela MLP para definir as probabilidades se mostra efetiva, mesmo com valores  $>50\%$ . Ou seja, essa mostra um certo grau de confiabilidade em seu uso.

Diante disso, verificou-se que o balanceamento das classes com a proporção C(Atividade: 75%, Inatividade: 25%) foi a que teve melhor aproveitamento por parte dos modelos de previsão. Em relação ao limiar de probabilidade, notou-se que o seu aumento prejudica a capacidade de detecção dos raios, aumentando a subestimação do evento, sendo assim, o limiar de probabilidade que melhor se adequou aos modelos foi  $>50\%$ . Sendo assim, para os demais testes, são essas parametrizações que serão levadas em consideração.

## **Teste 2: informações de instantes anteriores à previsão (t-1 vs t-1,t-2,t-3)**

Este teste foca no impacto dos dados provenientes de tempos anteriores. Como visto anteriormente, existem relações que podem ser assíncronas. Neste sentido, foram realizadas duas execuções cujas diferenças estavam apenas na inserção das informações das mesmas variáveis em tempos anteriores. Note que as variáveis escolhidas remetem ao perfil de refletividade radar para cada pixel. Na MLP01 foram selecionadas todas os valores de refletividade para os CAPPIs de 2 a 7km de altura em três intervalos de tempos que antecedem a previsão ( $t-1$ ,  $t-2$  e  $t-3$ ). Enquanto que, a MLP02 utiliza apenas os dados do instante t-1, com as mesmas variáveis. Como notado nas análises do EDA e da DT, os dados em diferentes perfis são importantes para definir alguns aspectos físicos e apresentam certa importância da definição das descargas em diferentes momentos. Nota-se na Figura 4.27 que a distribuição apresentada na MLP01 parece está mais bem definida. Uma possível resposta dos dados em três tempos anteriores à previsão, fazendo com que capture as tendências temporais do perfil durante esses intervalos. Contudo, a MLP01 apresenta um erro de co-localização evidente. Que como dito anteriormente pode estar associado ao rápido deslocamento do sistema observado. Em contrapartida, a MLP02 apresenta-se mais ruidosa, provavelmente um efeito da falta de informação de relações assíncronas. No

entanto, as bordas são melhor co-localizadas ao sistema em 2,5km. Isto pode ser um indicativo que a inclusão de informações assíncronas, como na MLP01, pode retardar os resultados das previsões de raios, no que tange ao deslocamento do sistema. Isso pode indicar que, as informações mais próximas ao evento pode inferir melhor deslocamento, ou seja, dando mais velocidade de propagação à previsão. Como pode ser observado, as previsões foram mais adiante em direção a observação a oeste.

Essa característica também pode ser evidenciada nas previsões da MLP05 e MLP06. A MLP05 é a configuração que tem como entrada todas as variáveis durante os três tempos anteriores à previsão e a MLP06 todas as variáveis, porém com um tempo anterior à previsão. Embora pareça ter uma co-localização melhor com as regiões mais ativas, a MLP06 apresenta uma distribuição mais ruidosa, assim como a MLP02, no entanto, com uma superestimativa menor, influenciada pela inclusão de mais variáveis. Já a MLP05 possui uma distribuição mais parecida com as regiões mais ativas, porém com erro de co-localização maior. Possivelmente associado pela rápida movimentação do sistema para oeste, assim como foi para a MLP01. Contudo, em termos de qualidade e valor da previsão, as MLP01 e 05 foram superiores. Mesmo que os *skills* apresentados fossem menores, como pode ser visto na Figura 4.27. Fica evidente pela Figura 4.28 que as previsões realizadas por estas duas últimas em termos de quantidade de raios foi similar às observações. Ou seja, o modelo consegue dizer que ocorrerá as descargas, contudo, erra em sua localização. Um problema que talvez esteja associado ao tipo de RNA escolhida. Talvez, uma MLP não seja a melhor escolha para este tipo de dado e comportamento dinâmico associado aos eventos a serem previstos.

Em termos de *skill* geral, como mostrado na Tabela 4.3 a MLP02 (MLP06) mostra uma leve superioridade em POD e FAR com relação a MLP01 (MLP05) que não chega a ser significante. Lembrando que nenhum teste estatístico de significância foi aplicado nessas médias, mas mesmo assim, é possível fazer tal afirmação, devido à proximidade dos valores.

Os resultados destes testes mostram um certo dilema, onde a escolha de informações anteriores pode dar qualidade à previsão, em termos de total de descargas previstas, mas em detrimento do deslocamento do sistema, que é melhor definido pela última informação. O melhor seria usar as duas abordagem em conjunto em redes integradas, dando peso as características mais relevantes de cada uma. Ou seja, usar o deslocamento de uma e a qualidade da outra.

### Teste 3: densidade de informações

Nestes testes são comparadas as RNAs com todos os valores de refletividade em relação aquelas com apenas um nível de informação. O intuito é verificar a capacidade de abstração de uma rede com menos informação com relação a uma outra com mais dados. Além disso, devido ao número de *inputs* há um impacto direto no tempo de processamento da RNA. Cabe ressaltar que, a utilização dos 40 níveis de refletividade disponíveis nos dados de radar foi dispensada, uma vez que desta forma uma grande quantidade de ruídos estavam sendo inclusos prejudicando análise e também o desempenho dos modelos. Diante disso, foi feita a utilização dos dados entre 2 km e 7 km, conforme mencionado no capítulo anterior. Já as redes mas simples utilizaram apenas a altura de 2,5 km, frequentemente aplicada na estimativa da chuva em superfície.

As configurações com o perfil vertical de refletividade (MLP01 e MLP02) na entrada foram capazes de se ajustar melhor aos núcleos mais ativos e capturar mais informações assíncronas, devido à variação existente nos valores da coluna vertical. Enquanto que, as configurações com os valores da refletividade em um nível na entrada (MLP03 e MLP04), embora com o processamento mais rápido (treinamento/validação/teste) acabaram superestimado mais a incidência de raios, marcando regiões ativas com limiares de refletividade mais modestos. Além disso, as configurações com apenas a refletividade em 2,5km em um instante de tempo anterior à previsão ( $t-1$ ) parece ser menos ruidosa que a configuração com os instantes de tempos  $t-1$ ,  $t-2$  e  $t-3$ . Isso pode está associado com o desenvolvimento vertical das nuvens, influenciado pelas correntes ascendentes.

Diante disso, é possível observar que as configurações com o perfil vertical de refletividade tem um melhor desempenho na previsão dos raios. Além disso, com a inclusão de dados de múltiplos instantes anteriores, também foram capazes de capturar informações assíncronas melhorando a sua performance. Já desempenho do modelo com apenas a refletividade em 2,5km em um instante de tempo (MLP04), apresentou resultados mais satisfatórios no que diz respeito ao caso da Figura 4.27 que a com todos os tempos anteriores (MLP03). Tal resultado pode estar associado a propagação de informações com baixa relevância pela rede em diferentes tempos para apenas uma variável. Isso mostra que mais variáveis podem inserir informações importantes as redes, principalmente se existir algum aspecto físico relacionado.

#### **Teste 4: as melhores variáveis selecionadas via DT**

A configuração MLP07, que usa as variáveis selecionadas pela DT, foi o modelo com a maior superestimação. Isto pode estar associado com critério de escolha das variáveis para *input* da RNA, que baseou-se na seleção daquelas com grau de importância maior que zero, e/ou a má definição de algumas variáveis pela DT. Isso indica que a escolha das variáveis de entrada baseada na seleção de atributos da DT precisa ser complementada pela definição de um limiar no grau de importância. Uma vez que, como apontado na Figura 4.21, algumas variáveis tiveram um grau de importância muito baixo na classificação, mas que foram incluídas na entrada da RNA, provavelmente aumentando as incertezas da rede à ocorrência de raios. Além disso, como foi observado nas DTs analisadas na Seção 4.4, alguns nós que foram utilizados não traziam informações relevantes, podendo a partir de uma análise subjetiva serem eliminados. Contudo, faz parte de escopo deste trabalho não interferir nas escolhas automáticas da DT.

Em estudos futuros, onde serão executadas análises mais aprofundadas a respeito do uso das DTs, escolhas mais pertinentes poderão ser aplicadas. Assim, espera-se que os resultados sejam mais satisfatórios. Cabe ressaltar que a MLP08, que também usa as variáveis selecionadas pela DT, porém, com o tempo mais próximo ( $t-1$ ), mostrou-se levemente melhor que a MLP07 em relação à superestimação. Tal resultado pode estar associado a melhor representatividade física mostrada pela DT para este tempo, como discutido em sua seção. Uma possível relação com o rápido deslocamento das tempestades, que conseqüentemente não deixaram muitos sinais de sua severidade em tempos anteriores nos dados de radar.

#### **Teste 5: os impactos de variáveis físicas conhecidas**

Outra análise baseada nas relações entre as MLP01/2 e MLP05/6 esta associada a inserção de dados conhecidos por sua relação com tempestades, como as variáveis calculadas em função do perfil de refletividade (e.g. VIL, Alturas de Walgvoegel e *echo top* de 35dBZ, entre outras). Neste teste foram inseridos todos os dados do radar utilizados neste estudo nas redes MLP05/6 e comparados com os resultados das MLP01/2. De um modo geral, nota-se que há um pequeno ganho com relação ao POD, entretanto, o FAR mostrou-se bem similar, com uma leve perda de *skill*. Contudo, em relação ao caso apresentado na Figura 4.27 e ao total de raios gerados pelas previsões observados na Figura 4.28, observa-se uma leve melhoria associada à MLP05 com relação as outras aqui analisadas.



## Teste 6: a melhor na matriz de combinações

Como mencionado anteriormente, foi realizado um teste prévio de sensibilidade em relação à quantidade de variáveis utilizadas como entrada, em todas as combinações possíveis. Esse foi um teste mais simplificado, que tinha como um dos propósitos avaliar a capacidade de generalização das redes neurais artificiais diante da inclusão de dados selecionados a partir de uma abordagem distinta, diminuindo a quantidade de informações utilizadas, porém com mais critérios de seleção. A diminuição drástica na amostragem se deu por questões de custo computacional, pois executar mais de 250 mil RNAs levaria bastante tempo. Além disso, esperava a partir desse teste encontrar uma combinação de variáveis que fossem melhor aproveitada pelo modelo. O critério utilizado para a escolha da melhor combinação foram os valores de POD e FAR mas para dados balanceados, como será descrito a seguir.

A Tabela 4.4 mostra as 3 melhores combinações de acordo com esse teste de sensibilidade. Onde é possível observar que a altura de Waldvogel foi superior a altura de 35 dBZ, e a combinação das duas. As variáveis pertencentes à essas combinações são associadas com o desenvolvimento vertical das nuvens, indicando assim, uma forte relação dessa característica de ocorrência de raios, podendo ser veemente consideradas em um sistema de previsão. Note que, para esses testes as *skills* tiveram valores muito melhores que os testes anteriores. Isso se deu devido a uma seleção mais rigorosa e controlada das amostras que foram utilizadas para o treinamento, validação e teste, em que foram selecionadas somente pixels em que haviam abundância de informações. Devido a isso, a propagação dos sistemas não provocaram grandes impactos nas *skills*, uma vez que foi levado em consideração a densidade de informações nas amostras selecionadas. Em suma, essas estatística não serviriam em um sistema real de validação de modelos de previsão, pois não consideram a totalidade da imagem prevista. Contudo, esses testes não tiveram o propósito de serem comparados aos testes apresentados até então, mas sim servir como alternativa na seleção de atributos a serem utilizadas nas configurações dos modelos que estão sendo avaliados. Uma espécie de mini-treinamento e validação em ambientes mais controlados.

Tendo em vista o exposto acima, a configuração do modelo MLP09 utiliza como variável de entrada a melhor combinação, neste caso a altura de Waldvogel, segundo esse teste de sensibilidade. Esta MLP foi treinada nas mesmas condições das redes dos testes anteriores, mas apenas usando esta variável. A Figura 4.27(j) mostra a representação visual da previsão utilizando esse modelo. De todos os outros modelos, este foi o que teve a melhor aproximação da quantidade de raios observados para o limiar de probabilidade que está sendo utilizado no caso de estudo, como também pode ser evidenciado na Figura 4.28. Assim como em outras configurações, o resultado dessa previsão também é impactado pelo erro de co-localização, porém reflete os aspectos presentes na observação. De todo modo fica claro que variáveis com melhor significado físico tendem a melhorar a RNA.

Tabela 4.4 - As 3 combinações com as melhores *skills* segundo teste de sensibilidade das variáveis em três instantes de tempos.

COMBINAÇÃO DE VARIÁVEIS	POD	FAR
Altura de Waldvogel	0.994	0.000
Altura máxima de 35dBZ	0.803	0.134
Altura máxima de 35dBZ; Altura de Waldvogel	0.799	0.046

### Teste 7: os efeitos do tipo de superfície e ciclo diurno

Como discutido em seções anteriores, as tempestades podem sofrer influência do tipo da superfície. Na análise exploratória dos dados não foram encontradas relações explícitas entre a ocorrência de descargas elétricas em diferentes topografias. Portanto, a inclusão dessa informação nos testes não foram consideradas. Como apontado por Machado et al. (2018), a orografia possui relações mais fortes com os regimes precipitantes, que são melhores definidos. Já com os raios, possivelmente influenciado pela sua alta variabilidade, não foi possível encontrar associações que motivassem alguma conclusão.

Através da análise espacial das descargas elétricas e da sua relação com os dados de cobertura de solo, foram encontrados alguns indícios que o tipo de solo pode exercer alguma influência sobre a ocorrência de raios. Sendo assim, foram realizados alguns testes com a inclusão das classes de cobertura de solo associadas aos dados de entrada das RNAs. A fim de avaliar os impactos dessas informações de forma mais objetiva, esses testes foram realizados apenas com as configurações dos modelos que

aproximaram melhor o total de raios da observação, ou seja, a MLP05 e MLP09.

Como visto em seções anteriores, a incidência de raios sofre forte influência do ciclo diurno, sendo mais ativa durante o final do dia. Além disso, como apontado por Morales (2018), a eficiência de detecção da STARNET sofre grande oscilação do dia para a noite. Sendo assim, a inclusão das informações de hora também foram incluídas nas configurações das MLPs mencionadas.

A Tabela 4.5 mostra a comparação entre as *skills* dos modelos com as primeiras configurações dos dados de entrada e com a inserção dos dados de cobertura do solo e hora. As variações da configuração MLP05, que tem como entrada todas as variáveis nos três instantes de tempos, tem uma mudança mais significativa em MLP10. Essa configuração tem os dois tipos de dados em questão como complemento, tendo um aumento de quase 4% no POD e um aumento mais modesto no FAR, de aproximadamente 0,2%. As configurações MLP11 e MLP12, que contém a inserção dos dados de cobertura do solo e hora, de forma separada, não tiveram mudanças significativas em relação à configuração base. A segunda parte da tabela diz respeito às *skills* das variações da MLP09, referente ao modelo definido com base nos testes de sensibilidade das combinações de variáveis, que tem como entrada a altura de Waldvogel em três instantes de tempo. Não houveram variações significativas com a inserção dos dados complementares para esse modelo.

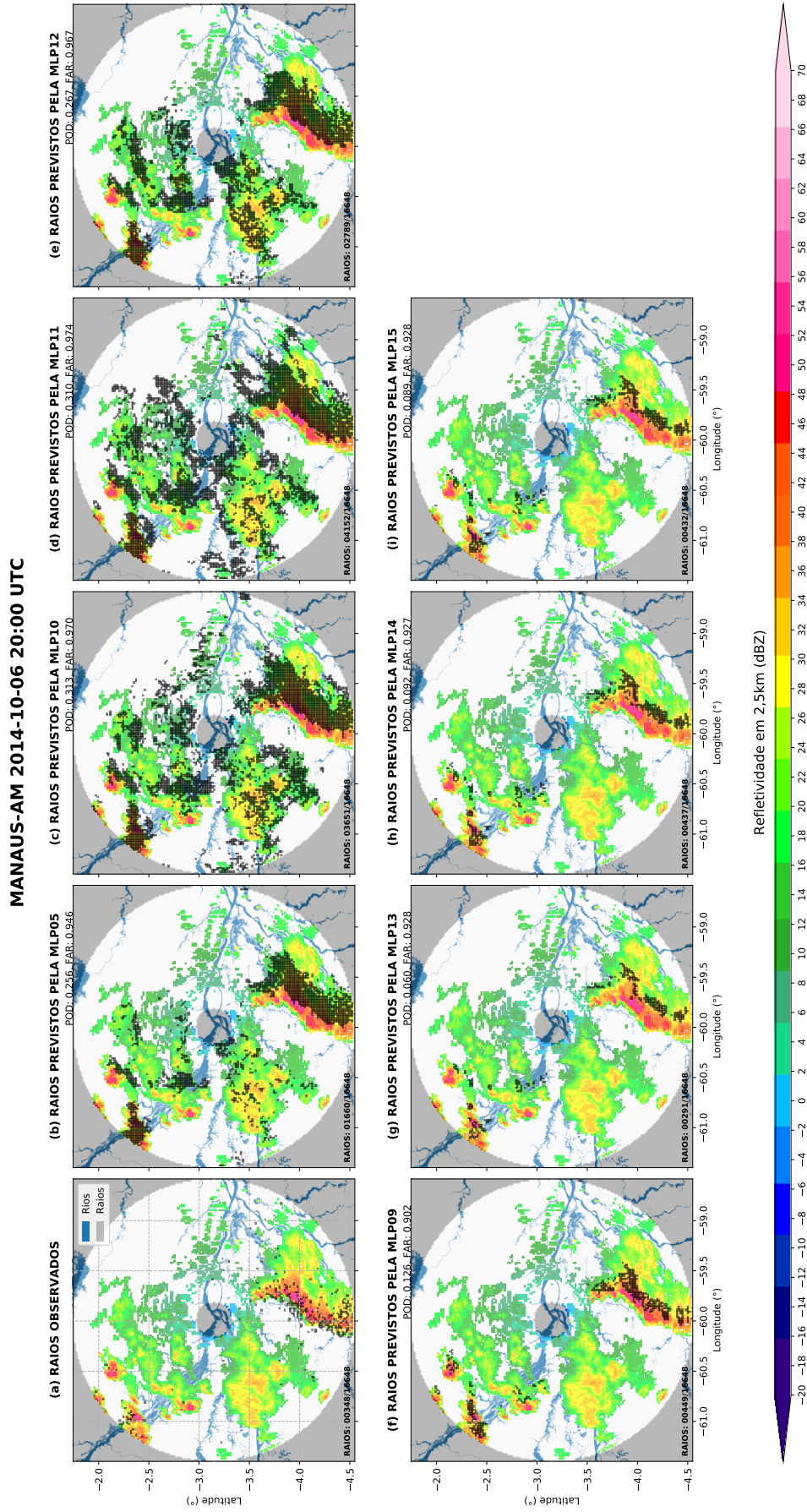
Tabela 4.5 - *Skills* dos testes com a inserção de dados de cobertura do solo e hora.

	MLP05		MLP10		MLP11		MLP12	
VARIAÇÕES DA MLP05	POD	FAR	POD	FAR	POD	FAR	POD	FAR
	0,879	0,980	0,917	0,982	0,905	0,984	0,897	0,980
	MLP09		MLP13		MLP14		MLP15	
VARIAÇÕES DA MLP09	POD	FAR	POD	FAR	POD	FAR	POD	FAR
	1,000	0,915	0,982	0,914	0,999	0,915	0,999	0,915

A fim de complementar a análise do impacto dos dados de cobertura do solo e hora, foi criado o mosaico com as previsões de raios para o evento utilizado como exemplo nas subseções anteriores. A Figura 4.29 mostra o mosaico com os dados de raios observados (a) e as previsões com as configurações MLP05, MLP10, MLP11, MLP12, MLP09, MLP13, MLP14 e MLP15, (b)-(i) respectivamente. A inclusão desses dados complementares sobre os modelos da configuração MLP05 influenciou na superestimação da ocorrência de raios, em que na inclusão somente dos dados cobertura do solo ficou ainda mais acentuada (Figura 4.29 (d)). Nota-se de um modo geral que existe uma tendência a gerar raios sobre os rios naquelas RNAs que levam em consideração esse tipo de informação. Isso se dá devido a alta relação entre esses eventos e os processos físicos associados a brisas fluviais, como comentado anteriormente na Seção 4.2. Contudo, este aspecto físico não foi incorporado pela rede, pelo contrário, causou ruídos nas saídas. Quanto as configurações baseadas na entrada da MLP09, observou-se que não houve tanto variação quanto nos modelos anteriores, ficando muito próximo da configuração somente com as entradas padrões. A configuração com o resultado mais discrepante é a que tem as duas variáveis complementares, que acaba limitando mais os pontos com previsão de raios, subestimando um pouco a quantidade observada, como pode ser visto na Figura 4.30.

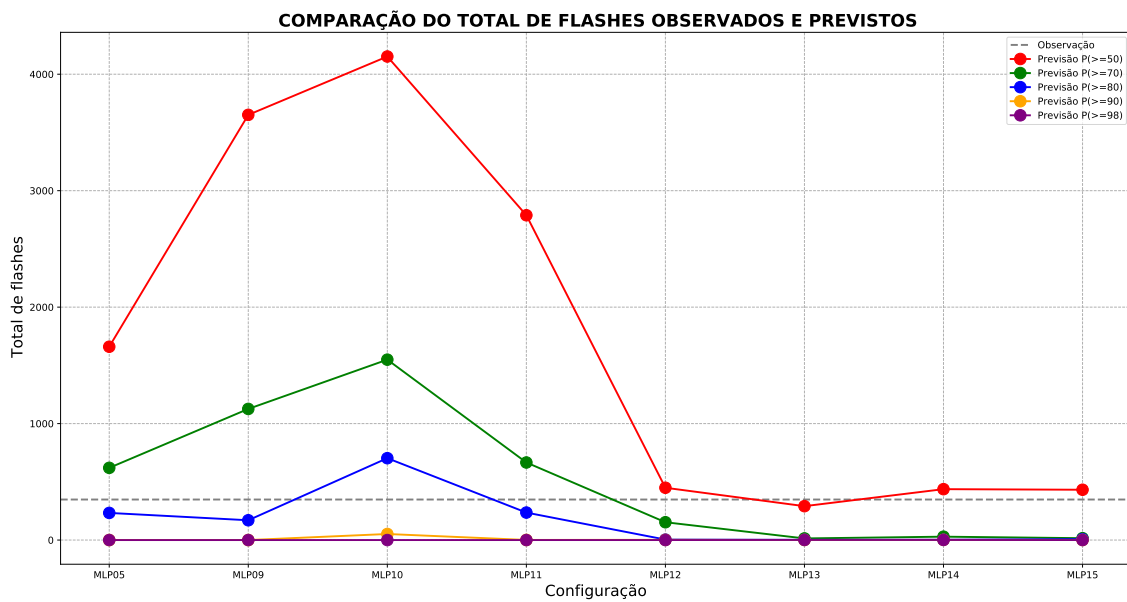
Com relação à análise pixel a pixel, não foi identificadas mudanças significativas com relação à inclusão dos dados de cobertura do solo e hora na previsão das descargas elétricas, o que, em partes, pode estar associado com a alta variabilidade do evento. Sendo assim, essas relações são melhores definidas a partir de análises mais globais, como as que foram feitas na análise exploratória dos dados. Diante disso, essas variáveis podem ter impactos mais expressivos perante um sistema previsão de maior escala e não pixel a pixel. Como sugestão para trabalhos futuros, o acompanhamento desses eventos usando um sistema de rastreamento que possa quantificar estes parâmetros em macroescala deve ser aplicado, a fim de definir um melhor sinal dos efeitos de superfície.

Figura 4.29 - Mosaico de previsão com as configurações criadas a partir da inclusão dos dados de cobertura do solo e hora.



Fonte: Produção Autor.

Figura 4.30 - *Flashes* observados e previstos com as configurações MLP05 e MLP09 com dados de cobertura do solo e hora.



Fonte: Produção Autor.

## Teste 8: os impactos da previsão com dados degradados

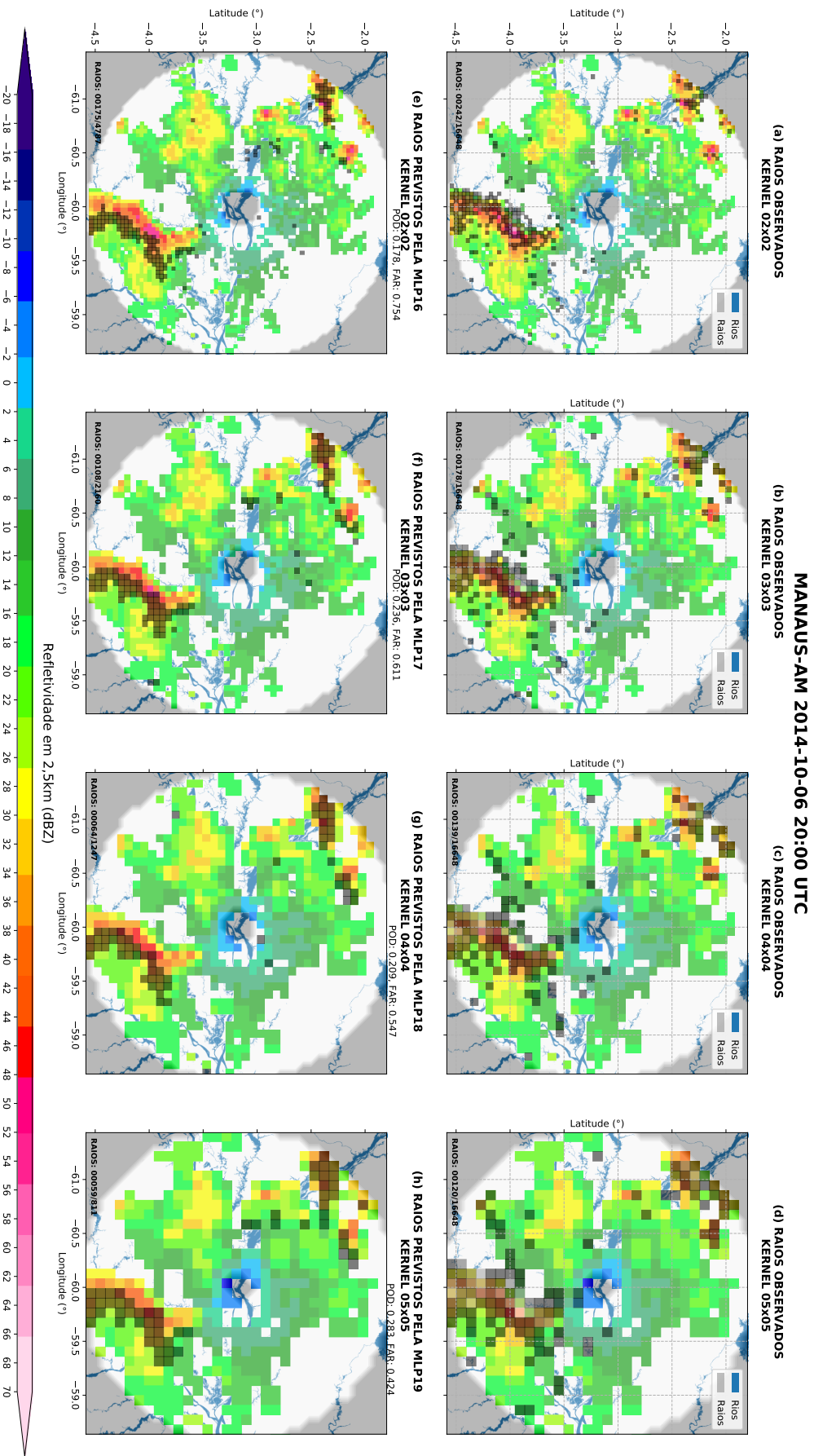
Um dos principais problemas enfrentados nos testes apresentados até aqui, estão associados com o erro de co-localização. Devido à rápida propagação de alguns sistemas, e o fato de que a previsão está sendo feita pixel a pixel, as diversas configurações testadas não foram capazes de acompanhar a referência espacial da observação, impactando diretamente na *skills* dos modelos. Diante disso, foram realizados os testes com os dados degradados, aumentando as áreas dos pixels. Esses testes foram realizados conforme as degradações dos dados descritas na Seção 3.3. Cabe ressaltar que apenas uma das funções de agregação sobre os dados de raios foi utilizada, enquanto a função de média foi aplicada as variáveis de entrada.

Como exposto anteriormente, o erro de co-localização é a principal problemática que está sendo encontrada na realização desses testes, e as taxas de falso alarme (FAR) indica isso. Quanto maior for esse valor, maior é erro de co-localização da previsão em relação à previsão, como foi apontado nas análises anteriores. A tentativa de diminuição desse erro sem a inserção de mais informações, foi feita com a degradação dos dados de entrada. Com o aumento dos *kernels* é possível notar que há uma diminuição na taxa de falso alarme, indicando que erro de co-localização também é atenuado. A Tabela 4.6 mostra as *skills* dos testes realizados com as as variações das configurações do modelo que usa como entrada todas as variáveis (MLP05) e as diferenças entre elas e os dados sem degradação. Nessa tabela é possível observar que há ganhos em ambas as *skills*, aproximadamente 9,6% no POD e 24% no FAR.

Tabela 4.6 - Comparação entre as *Skills* dos testes realizados com os modelos utilizando todas as variáveis sem e com as degradações nos diferentes *kernels* do caso de estudo.

CONFIGURAÇÃO	POD	FAR	DIFERENÇA POD	DIFERENÇA FAR
MLP05 (sem degradação)	0,879	0,980		
MLP16 (kernel 2×2)	0,964	0,865	0,085	0,115
MLP17 (kernel 3×3)	0,975	0,784	0,096	0,196
MLP18 (kernel 4×4)	0,951	0,747	0,072	0,233
MLP19 (kernel 5×5)	0,939	0,740	0,060	0,240

Figura 4.31 - Mosaico de previsão com as configurações criadas a partir da entrada de MLP05 com os dados degradados.





A Figura 4.31 mostra as observações e previsões com dados degradados para o evento que foi analisado anteriormente. Através dessa figura é possível observar que há uma melhora significativa nas *skills* de acordo com o aumento das áreas dos pixels, diminuindo assim o erro de co-localização. O FAR é a *skill* que tem a melhora mais significativa com o aumento do *kernel*. Diante disso, a falta de informações à respeito da propagação dos sistemas pode ser suprida pela degradação dos dados. Com isso, pode-se notar a significativa melhora na previsão, uma vez que a variabilidade espacial e falta de co-localização dos dados é atenuada com essa abordagem. A incerteza em relação à área aumenta, em contrapartida, a previsão se torna mais fidedigna.



## 5 CONCLUSÃO

Esta dissertação apresentou um estudo a respeito das informações meteorológicas inferidas a partir de radares que podem ser utilizadas para a implementação de sistemas de previsão de descargas elétricas baseados em técnicas de aprendizado de máquina. Além disso, uma investigação sobre as características espaciais, sazonais e temporais das descargas elétricas foi realizada complementando a análise exploratória dos dados sobre a região Central da Bacia Amazônica. Essas análises subsidiaram os testes de sensibilidades que foram feitos com a árvore de decisão e redes neurais artificiais durante o período de maior atividade das descargas elétricas, a estação seca. A partir desses testes foram discutidos os impactos das diferentes configurações dos dados de entrada sobre a previsão de raios na região.

A distribuição espacial das descargas elétricas sobre a área de estudo sofre variações significativas entre cada estação. O raio está diretamente associado com a presença de gelo no interior das nuvens de tempestades e isso pode ser intensificado a partir de influências termodinâmicas, dinâmicas e efeitos locais. Como já discutido, isso está associado com características inerentes da região, como as mudanças dos regimes precipitantes e as influências termodinâmicas provenientes da umidade da floresta e da brisa dos rios. Isso leva a considerar que os sistemas de previsão de raios desenvolvidos para a essa região podem ter seu desempenho afetado se aplicado em outras regiões do Brasil, uma vez que por ter uma grande extensão territorial e influência de sistemas de macroescala, possui condições locais muito distintas. Ainda com relação à distribuição dos raios, foi observado que as localidades muito próximas aos rios sofrem uma incidência maior ao longo de todas as estações, conforme apontado nos mapas temáticos agrupados por área e por municípios, que podem servir como subsídios na tomada decisões de autoridades competentes locais.

Sobre a influência da superfície, as relações encontradas mais evidentes foram com relação à cobertura do solo. Embora haja uma expressiva predominância da floresta como tipo de cobertura, as fortes relações das tempestades eletricamente ativas, já conhecidas, com os rios foram identificadas. Em que as regiões perto, ou sobre essas áreas tiveram alta incidência de raios. Já com relação à orografia, não foram identificadas relações significativas que pudessem motivar análises mais aprofundadas. Devido à alta dependência espacial desses dados e a não adição das informações de propagação dos sistemas, esses dados não proporcionaram ganhos significativos com as arquiteturas de redes neurais artificiais testadas. Entretanto, com a aplicação de diferentes abordagens que levam isso em consideração, esses dados podem

proporcionar algum ganho.

As relações entre a ocorrência de descargas elétricas e o ciclo diurno também foram identificadas nesse estudo. Como já sabido, esta é uma região que apresenta um ciclo de diurno da convecção bem definido. Diante disso, a incidência de raios durante todas as estações tendem a ter uma frequência maior começando pela manhã, tendo a sua máxima atividade aproximadamente entre as 12 h e 14 h local (UTC-4). Além disso, o sensor que está sendo utilizado para computar os raios observados (STARNET) possui uma melhor eficiência de detecção durante a noite.

O pré-processamento dos dados e sua análise são fundamentais no processo de construção de modelos baseados em aprendizado de máquina. Além de fornecer informações importantes para a escolha da melhor configuração do modelo a ser aplicado, através da EDA também é possível identificar limitações que os sistemas de previsão podem vir a ter. Através da análise exploratória dos dados foi possível compreender melhor as relações entre as variáveis meteorológicas usadas e os raios. Através das análises dos histogramas de frequências relativas, foi possível observar em quais variáveis e estações a separação das classes de atividade e inatividade de raios eram melhor definidas. Nessa tarefa, algumas variáveis com o perfil vertical de refletividade foram analisadas (i.e. 2km, 5km e 7km). A estação seca foi a que teve melhor tendência de separação, enquanto que a estação chuvosa foi a que teve a pior separação. Isso pode estar associado com as distintas características dos sistemas precipitantes, que predominantemente são estratiformes durante a chuvosa e convectivos durante a estação seca.

Além disso, o aumento dos limiares de atividade realçam a separação das classes, uma vez que as regiões mais ativas normalmente estão associadas à maior intensidade nos valores das variáveis. Embora haja uma sub-notificação associada com a co-localização dos dados, também é possível identificar tendências nos limiares de refletividade. Em suma, essas análises foram importantes para identificar o grau de separação entre as classes de inatividade e atividade e as variáveis com relações mais fortes com as descargas elétricas. Em que se destacaram principalmente aquelas associadas aos perfis verticais e suas derivações, como as alturas de Wadvogel e *echo top* de 35dBZ.

Além das relações síncronas, também foram realizadas análises das relações assíncronas entre as variáveis de radar e da STARNET. Essas análises foram realizadas com o intuito de identificar as variáveis que estão associadas com os raios em intervalos de tempos diferentes. Isso é importante para os sistemas de previsão, uma vez que os dados de anteriores podem trazer informações sobre tendências associadas a ocorrência ou não de raios. Para a inatividade não foram encontradas mudanças significativas. No entanto, os valores associados à classe de atividade tendem a se aproximarem dos valores da classe de inatividade a medida que se distanciam do instante de tempo atual. Sendo assim, a inclusão de informações muito passadas impactar negativamente na performance dos modelos de previsão.

Outra análise realizada foi em relação ao aumento das áreas dos pixels por meio das degradações com diferentes *kernels*. Observou-se que quanto maior o tamanho do pixel, maior é o grau de co-localização dos dados. Contudo, pior é a correlação entre eles para a maioria das variáveis. Este tipo de informação, como será abordado mais afrente, teve grande impacto nas redes neurais testadas.

Uma vez que existe uma relação de dependência entre as variáveis de radar e a ocorrência de raios, o uso de uma técnica de classificação para determinar possíveis relações e seus limiares se fez necessário. Neste sentido, uma técnica de aprendizado de máquina foi aplicada, a árvore de decisão. Através dela foi possível analisar as relações entre as variáveis e os raios em diferentes instantes de tempos. Embora houvesse uma dessincronização temporal e espacial das variáveis em relação aos raios, alguns limiares e associações dos processos físicos dos raios foram identificados através das representações das árvores de decisão. De modo geral, ficou evidentes que este tipo de técnica se bem aprofundada pode trazer soluções importantes para identificação de tempestades e raios. Além disso, a árvore de decisão foi utilizada com o intuito de fazer uma seleção de atributos, filtrando aqueles com maior relevância para a classificação dos dados com inatividade e atividade de raios. As variáveis selecionadas com a árvore de decisão foram utilizadas como entrada em algumas configurações de RNAs, conforme será apresentado em seguida.

Um dos objetivos principais desta pesquisa dar subsídios a criação de um sistemas de previsão de raios nuvem-solo. Sendo assim, o uso de RNAs foram aplicados. Algumas arquiteturas de RNAs possuem um certo grau de autonomia no seu processo interno de generalização entre as associações dos dados de entrada e saída, pois uma vez configurada a arquitetura e seus hiper-parâmetros, o treinamento do modelo acontece de modo independente. Dependendo da quantidade de parâmetros (i.e. nú-

mero de neurônios e camadas ocultas) definidos nas RNAs, a interpretação de como os processos de associações entre os dados de entrada e saída estão sendo realizados internamente vai sendo perdida durante o treinamento do modelo. Além disso, há um certo grau de subjetividade na escolha desses parâmetros de acordo com o tipo de problema e dados que estão sendo utilizados. Isso pode provocar uma certa resistência de alguns especialistas em utilizar as RNAs, apesar de seus ótimos resultados. Sendo assim, todas as análises e pré-processamento dos dados feitos antes da aplicação das redes neurais, tiveram como intuito avaliar a qualidade das informações que foram adicionados ao modelo, que é muitas vezes considerado uma "caixa-preta". Além disso, o treinamento das redes neurais artificiais sem um pré-processamento prévio dos dados pode afetar na capacidade de generalização dos modelos. Por isso, essa etapa de análise e processamento são de suma importância. Uma vez que elucidam processos físicos que podem impactar nos resultados, deixando a RNA mais transparente.

Com essas etapas anteriores ao treinamento das redes neurais cumpridas, foram testadas os modelos com variações nos dados de entrada para intervalos de tempos distintos anteriores à previsão. Dentre essas configurações de entrada, foram avaliados modelos com somente uma variável, com perfis verticais de refletividade, variáveis selecionadas pela árvore de decisão e com a inclusão de informações da cobertura do solo e hora. Também foram testados configurações em que os dados estavam com degradações. Além disso, testes complementares à essas configurações também foram feitos, variando a proporção do balanceamento dos dados e aplicando diferentes limiares na probabilidade de detecção dos raios.

Alguns resultados foram identificados nos testes realizados com diferentes configurações de entradas das RNAs. Aquelas que tiveram como entrada somente um instante de tempo, apresentaram comportamento muito dependente, superestimando a ocorrência de raios, porém tendo uma melhor co-localização com os dados observados. Já os modelos com as entradas dos três instantes anteriores à previsão, abstraíram um pouco dessa dependência e generalizaram melhor as tendências ao longo do tempo, aproximando mais à quantidade de raios observados em detrimento do aumento do erro de co-localização.

As configurações que tiveram melhor resultado foram aquelas com todas as variáveis do radar nos três instantes de tempo anterior à previsão e somente com a variável que apresentou o melhor desempenho em um teste massivo de combinações, a Altura de Waldvogel. Cabe ressaltar que com relação a este último, o *echo top* de 35 dBZ

também mostro resultados satisfatórios. Ou seja, variáveis que são conhecidas pela sua relação com a física das tempestades. Logo, os resultados dessas RNAs mostram que as características verticais das nuvens são de suma importância para definir uma melhor previsão de raios CG.

O balanceamento das entradas que proporcionou melhor capacidade de generalização dos modelos foram as proporções de 25% e 75% dos dados para as classes de inatividade e atividade, respectivamente. O aumento do limiar de probabilidade para considerar a ocorrência de raios aumentou o erro de co-localização e a subestimação das previsões, portanto a que melhor se ajustou foi o limiar  $>50\%$ . O que mostra que a metodologia utilizada para definir as probabilidades pela rede foram suficientes para definir o melhor critério.

Além disso, os desempenhos dessas configurações tiveram melhoras significativas ao aplicar a degradação dos dados. Isso aumenta um pouco da incerteza da localização do raios devido ao aumento das áreas dos pixels. Contudo, o ganho de informação, avaliado pelas *skills* POD e FAR, foi significativo, principalmente com as taxas de falso alarme. Diante disso, essa é uma abordagem que pode ser utilizada para suprir a falta de informações de deslocamento em sistemas de previsão de raios.

A inclusão dos dados de horários foram considerados em algumas configurações das redes neurais testadas. Assim como o uso de informações sobre o tipo de solo. Embora tenha havido mudanças com a inclusão dessas informações, não foram identificados ganhos em relação às *skills* dos modelos de previsão. Contudo, análises mais profundas precisam ser realizadas para averiguar os impactos em maior escala, uma vez que pixel-a-pixel não foram observadas mudança benéficas a RNA.

Apesar das limitações dos modelos simples aqui testados em acompanhar o deslocamento dos sistemas, as configurações de entradas testadas mostraram resultados promissores, e que podem ser aproveitados nos próximos sistemas de previsões de raios para essa região.

## 5.1 Trabalhos futuros

As análises das relações entre as descargas elétricas e as variáveis obtidas a partir de radar na região central da Bacia Amazônica apresentadas neste trabalho, podem servir com subsídios para a construção de sistemas de previsão de descargas elétricas nessa região. Além disso, as metodologias aplicadas podem ser extrapoladas para análises de mesmo propósito feitas em outras localidades. Sendo assim, a inclusão

de outras variáveis provenientes de outros sensores podem complementar essas análises, principalmente aquelas que estão associadas a perfis termodinâmicos ou de movimentos atmosférico.

Como apontado em diversas ocasiões ao longo deste trabalho, a abordagem utilizada em tratar as informações pixel a pixel, fez com que houvesse algumas limitações, a principal delas associadas à co-localização dos dados. O aumento das áreas que estão sendo analisadas favorece o ganho de informações, como visto ao aplicar as degradações. Partindo disso, abordagens diferentes para o aumento das áreas de análises podem ser aplicadas em estudos futuros, como por exemplo, análises baseadas em *clusters*. Além disso, a inclusão de informações complementares de deslocamento e propagação das nuvens podem ser incluídas, o que certamente aumentará o ganho de informações sobre a propagação dos sistemas.

Este trabalho fez uso de uma das redes neurais mais difundidas, a *perceptron* de múltiplas camadas. Durante a etapa de ajuste dos hiper-parâmetros, algumas problemáticas foram de encontro com as limitações da arquitetura da rede. A conversão dos dados de entrada de multi-dimensional para unidimensional sem a extração automática de características podem afetar a capacidade desse tipo de arquitetura. O aumento do número de neurônios, camadas ocultas e outros hiper-parâmetros podem até melhorar o desempenho, porém o custo computacional também aumenta de forma proporcional. Diante disso, trabalhos futuros deverão considerar testes com outras arquiteturas, como as redes neurais convolucionais, as quais possuem artefatos para extrair características de dados multi-dimensionais.

Como observado, a inclusão dos dados em diferentes tempos podem melhorar o desempenho dos modelos, removendo parte da dependência temporal das variáveis com a incidência de raios. A MLP pode tratar em até certo ponto essa relação temporal das variáveis com os raios, entretanto, com a inclusão de mais tempos anteriores à previsão a sua capacidade de associação pode diminuir, uma vez que não possui memórias de curto/longo prazo. Diante disso, as redes neurais recorrentes podem ser considerada em trabalhos futuros. Atualmente já existem arquiteturas que combinam as redes neurais convolucionais e recorrentes, essa pode ser um ponto de partida para a continuação deste trabalho.

Além disso, para diminuir a sub-notificação de raios sobre regiões convectivas mais intensas, é necessário aplicar algoritmos de busca e associação, como a máxima correlação cruzada. Tais técnicas podem associar de modo mais preciso os raios, que devido a dinâmica de medidas e física dos sistemas, podem não estar co-localizados.



Isto traria melhor relação entre as variáveis, o que ajudaria as RNAs a abstruírem as informações de maneiras mais correta.



## REFERÊNCIAS BIBLIOGRÁFICAS

- ABDULLAH, N. H.; ADNAN, R.; SAMAD, A. M.; RUSLAN, F. A. Lightning forecasting modelling using artificial neural network (ann): case study sultan abdul aziz shah airport or skypark subang. In: CONFERENCE ON SYSTEMS, PROCESS AND CONTROL, 7., 2018, Malaysia Melaka. **Proceedings...** [S.l.]: IEEE, 2018. p. 1–4. 3
- ALBRECHT, R. I.; GOODMAN, S. J.; BUECHLER, D. E.; BLAKESLEE, R. J.; CHRISTIAN, H. J. Where are the lightning hotspots on earth? **Bulletin of the American Meteorological Society**, v. 97, n. 11, p. 2051–2068, 2016. 1
- ALBRECHT, R. I.; MORALES, C. A.; DIAS, M. A. S. Electrification of precipitating systems over the amazon: physical processes of thunderstorm development. **Journal of Geophysical Research: Atmospheres**, v. 116, n. D8, 2011. 22
- ALLEN, G.; MARSHALL, J. L. An evaluation of neural networks and discriminant analysis methods for application in operational rain forecasting. **Australian Meteorological Magazine**, v. 43, n. 1, p. 17–28, 1994. 3
- ALMEIDA, A.; CALHEIROS, A.; ALBRECHT, R.; MACHADO, L.; SANTOS, R.; LEAL N. H.; SOUZA, F. 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. In: WORKSHOP DE COMPUTAÇÃO APLICADA À GESTÃO DO MEIO AMBIENTE E RECURSOS NATURAIS. **Anais...** Porto Alegre, 2020. p. 91–100. 58
- ALMEIDA, C. A. d.; COUTINHO, A. C.; ESQUERDO, J. C. D. M.; ADAMI, M.; VENTURIERI, A.; DINIZ, C. G.; DESSAY, N.; DURIEUX, L.; GOMES, A. R. High spatial resolution land use and land cover mapping of the brazilian legal amazon in 2008 using landsat-5/tm and modis data. **Acta Amazonica**, v. 46, n. 3, p. 291–302, 2016. 35
- AMBURN, S. A.; WOLF, P. L. Vil density as a hail indicator. **Weather and Forecasting**, v. 12, n. 3, p. 473–478, 1997. 20, 66
- AMERICAN METEOROLOGICAL SOCIETY (AMS). **Meteorology glossary**. 2012. Disponível em: <<https://glossary.ametsoc.org/wiki/Thunderstorm>>. 5

ANGELIS, C. F.; MCGREGOR, G. R.; KIDD, C. Diurnal cycle of rainfall over the brazilian Amazon. **Climate Research**, v. 26, n. 2, p. 139–149, 2004. 57

ASHFAQUE, J. M. **Introduction to support vector machines and kernel methods**. 2019. Disponível em: [https://www.researchgate.net/publication/332370436\\_Introduction\\_to\\_Support\\_Vector\\_Machines\\_and\\_Kernel\\_Methods](https://www.researchgate.net/publication/332370436_Introduction_to_Support_Vector_Machines_and_Kernel_Methods). Acesso em: 29 abr. 2020. 47, 48

ATKINSON, B. W. Precipitation. In: GREGORY, K. J.; WALLING, D. E. (Ed.). **Man and environmental processes**. [S.l.]: Routledge, 2019. 19

ATLAS, D. **Radar in meteorology: battan memorial and 40th anniversary Radar Meteorology Conference**. [S.l.]: American Meteorological Society, 2015. ISBN 9781935704157. 18

ATMOSPHERIC RADIATION MEASUREMENT. **SIPAM Manaus s-band radar CAPPI volumes v2.0a**: Campanha goamazon. 2018. Disponível em: <<https://www.arm.gov/research/campaigns/amf2014goamazon>>. 25, 26, 28

BARTHEM, R.; MARQUES, M.; CHARVET-ALMEIDA, P.; MONTAG, L. Amazon river basin: I—characterization and environmental impacts due to deforestation. **WIT Transactions on Ecology and the Environment**, v. 81, 2005. 23

BATES, B. C.; DOWDY, A. J.; CHANDLER, R. E. Lightning prediction for australia using multivariate analyses of large-scale atmospheric variables. **Journal of Applied Meteorology and Climatology**, v. 57, n. 3, p. 525–534, 2018. 15

BENGIO, Y. Gradient-Based Optimization of Hyperparameters. **Neural Computation**, v. 12, n. 8, p. 1889–1900, 2000. ISSN 0899-7667. 45

BENNETT, A.; ODAMS, P.; EDWARDS, D.; ARASON, P. Monitoring of lightning from the april–may 2010 eyjafjallajökull volcanic eruption using a very low frequency lightning location network. **Environmental Research Letters**, v. 5, n. 4, p. 044013, 2010. 9

BETZ, H. D.; SCHMIDT, K.; LAROCHE, P.; BLANCHET, P.; OETTINGER, W. P.; DEFER, E.; DZIEWIT, Z.; KONARSKI, J. Linet—an international lightning detection network in europe. **Atmospheric Research**, v. 91, n. 2-4, p. 564–573, 2009. 17

BISCARO, T. S. **Processos físicos que controlam a formação de nuvens rasas e sua evolução à convecção profunda na região central amazônica**. 163 p. Tese (Doutorado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2019. 27

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, Oct 2001. 14

BROOKS, C. E. P. The distribution of thunderstorms over the globe. **Geophysical Memoirs London**, v. 3, n. 24, p. 147–164, 1925. 9

BROWNING, K. A. Review lecture: local weather forecasting. **Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences**, v. 371, n. 1745, p. 179–211, 1980. 2, 10

BYERS, H. R.; BRAHAM, R. R. Thunderstorm structure and circulation. **Journal of Meteorology**, v. 5, n. 3, p. 71–86, 1948. 5, 6

CALHEIROS, A. J. P. **Sistema de previsão imediata da precipitação: o HydroTrack**. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos-SP, 2008. 11

CALHEIROS, A. J. P.; ENORÉ, D. P.; MATTOS, E. V.; COSTA, I. C. da; MACHADO, L. A. T. **Sistema de previsão imediata: descrição dos produtos**. São José dos Campos - SP, 2016. 2, 19

CAMPONOGARA, G.; DIAS, M. A. F. S.; CARRIÓ, G. G. Relationship between amazon biomass burning aerosols and rainfall over the la plata basin. **Atmospheric Chemistry and Physics**, v. 14, n. 9, p. 4397–4407, 2014. Disponível em: <<https://acp.copernicus.org/articles/14/4397/2014/>>. 54

CARANTI, J. Surface potential of ice in thunderstorm electrification. **Nature**, v. 284, p. 44, 1980. 8

CARDOSO, I.; PINTO JUNIOR, O.; PINTO, I.; HOLLE, R. Lightning casualty demographics in Brazil and their implications for safety rules. **Atmospheric Research**, v. 135, p. 374–379, 2014. 1

CARDOSO, M. C. **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, Rio de Janeiro-RJ, 2017. 1

CECCHINI, M. A.; MACHADO, L. A. T.; ANDREAE, M. O.; MARTIN, S. T.; ALBRECHT, R. I.; ARTAXO, P.; BARBOSA, H. M. J.; BORRMANN, S.; FÜTTERER, D.; JURKAT, T.; MAHNKE, C.; MINIKIN, A.; MOLLEKER, S.; PÖHLKER, M. L.; PÖSCHL, U.; ROSENFELD, D.; VOIGT, C.; WEINZIERL, B.; WENDISCH, M. Sensitivities of amazonian clouds to aerosols and updraft speed. **Atmospheric Chemistry and Physics**, v. 17, n. 16, p. 10037–10050, 2017. 56

CHAKRABARTY, H. et al. Application of machine learning technique to nowcast severe atmospheric event. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, COMPUTER VISION, AND PATTERN RECOGNITION. **Proceedings...** [S.l.], 2018. p. 137–141. 3

CHAUDHURI, S. Convective energies in forecasting severe thunderstorms with one hidden layer neural net and variable learning rate back propagation algorithm. **Asia-Pacific Journal of Atmospheric Sciences**, v. 46, n. 2, p. 173–183, 2010. 3

CHUNG, J.; GÜLÇEHRE, Ç.; CHO, K.; BENGIO, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. **Computing Research Repository – CoRR**, abs/1412.3555, 2014. 13

COLLINS, W. G.; TISSOT, P. Thunderstorm predictions using artificial neural networks. **Artificial Neural Networks - Models and Applications**, p. 251, 2016. 3, 5

COLQUHOUN, J. A decision tree method of forecasting thunderstorms, severe thunderstorms and tornadoes. **Weather and Forecasting**, v. 2, n. 4, p. 337–345, 1987. 14

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1995. 15

CUMMINS, K. L.; MURPHY, M. J.; BARDO, E. A.; HISCOX, W. L.; PYLE, R. B.; PIFER, A. E. A combined toa/mdf technology upgrade of the us national lightning detection network. **Journal of Geophysical Research: Atmospheres**, v. 103, n. D8, p. 9035–9044, 1998. 30, 31

DEIERLING, W.; PETERSEN, W. A. Total lightning activity as an indicator of updraft characteristics. **Journal of Geophysical Research: Atmospheres**, v. 113, n. D16, 2008. 9

- DEMETRIADES, N. W.; MURPHY, M. J.; CRAMER, J. A. Validation of vaisala's global lightning dataset (gld360) over the continental united states. In: CONFERENCE ON HURRICANES AND TROPICAL METEOROLOGY. **Proceedings**. Tucson, 2010. 17
- DESHPANDE, R. R. On the rainfall time series prediction using multilayer perceptron artificial neural network. **International Journal of Emerging Technology and Advanced Engineering**, v. 2, n. 1, p. 2250–2459, 2012. 3
- DIAS, M. S.; DIAS, P. S.; LONGO, M.; FITZJARRALD, D. R.; DENNING, A. S. River breeze circulation in eastern amazonia: observations and modelling results. **Theoretical and Applied Climatology**, v. 78, n. 1-3, p. 111–121, 2004. 21
- DIXON, M.; WIENER, G. Titan: thunderstorm identification, tracking, analysis, and nowcasting—a radar-based methodology. **Journal of Atmospheric and Oceanic Technology**, v. 10, n. 6, p. 785–797, 1993. 11
- DONALDSON, R. J. Radar reflectivity profiles in thunderstorms. **Journal of Atmospheric Sciences**, v. 18, n. 3, p. 292 – 305, 1961. Disponível em: <[https://journals.ametsoc.org/view/journals/atsc/18/3/1520-0469\\_1961\\_018\\_0292\\_rrpit\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/atsc/18/3/1520-0469_1961_018_0292_rrpit_2_0_co_2.xml)>. 61
- DUBROVSKÝ, M. The binary decision tree: the growing algorithm and application to thunderstorm forecasting. **Studia Geophysica et Geodaetica**, v. 39, n. 1, p. 84–100, 1995. 14
- DWYER, J. R.; UMAN, M. A. The physics of lightning. **Physics Reports**, v. 534, n. 4, p. 147–241, 2014. ISSN 0370-1573. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S037015731300375X>>. 85
- EMAMGHOREISHI, S.; MOINI, R.; SADEGHI, S. A modular neural network approach for locating cloud-to-ground lightning strokes. In: INTERNATIONAL SYMPOSIUM ON ELECTROMAGNETIC COMPATIBILITY. **Proceedings...** IEEE, 2001. p. 1042–1045. 3
- ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING. **Proceedings...** [S.l.], 1996. p. 226–231. 30
- ESTEVEVES, J. T.; ROLIM, G. de S.; FERRAUDO, A. S. Rainfall prediction methodology with binary multilayer perceptron neural networks. **Climate Dynamics**, v. 52, n. 3-4, p. 2319–2331, 2019. 3

EUROPEAN SPACE AGENCY - ESA. **ESA Climate change initiative – land cover led by UCLouvain (2017)**. 2017. Disponível em:

<<https://maps.elie.ucl.ac.be/CCI/viewer/download.php>>. 25

FAN, W.; MA, J.; ZHU, H. A thunderstorm forecast model based on weighted svm and data field. In: INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY. **Proceedings...** IEEE, 2009. p. 160–164. 15

FATHI, M.; KASHANI, M. H.; JAMEII, S. M.; MAHDIPOUR, E. Big data analytics in weather forecasting: a systematic review. **Archives of Computational Methods in Engineering**, 2021. ISSN 1886-1784. 10

FAZOLO, L. F. **Estudo de novas técnicas para previsão a curto prazo do deslocamento do campo horizontal da atividade elétrica em tempestades**. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos-SP, 2019. 7, 8, 17

FISCH, G.; MARENGO, J. A. A.; NOBRE, C. A. Uma revisão geral sobre o clima da amazônia. **Acta Amazonica**, v. 28, p. 101 – 101, 06 1998. ISSN 0044-5967. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0044-59671998000200101&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0044-59671998000200101&nrm=iso)>. 53

FRANKEL, D. S.; DRAPER, J. S. **Neural network for predicting lightning**. [S.l.]: Google Patents, ago. 18 1992. US Patent 5,140,523. 3

FREITAS, S. R.; PANETTA, J.; LONGO, K. M.; RODRIGUES, L. F.; MOREIRA, D. S.; ROSARIO, N. E.; DIAS, P. L. S.; DIAS, M. A. S.; SOUZA, E. P.; FREITAS, E. D. et al. The brazilian developments on the regional atmospheric modeling system (BRAMS): an integrated environmental model tuned for tropical areas. **Geoscientific Model Development**, v. 10, n. 1, p. 189–222, 2017. 10

GARDNER, M.; DORLING, S. Artificial neural networks (the multilayer perceptron): a review of applications in the atmospheric sciences. **Atmospheric Environment**, v. 32, n. 14, p. 2627–2636, 1998. ISSN 1352-2310. 41

GATLIN, P.; GOODMAN, S. Severe weather precursors in the lightning activity of tennessee valley thunderstorms. In: CONFERENCE ON METEOROLOGICAL APPLICATIONS OF LIGHTNING DATA, 3., 2008. **Proceedings...** New Orleans, 2008. v. 8. 9



- GENUER, R.; POGGI, J.-M.; TULEAU-MALOT, C. Variable selection using random forests. **Pattern Recognition Letters**, v. 31, n. 14, p. 2225 – 2236, 2010. 14
- GOODMAN, S. J.; BLAKESLEE, R. J.; KOSHAK, W. J.; MACH, D.; BAILEY, J.; BUECHLER, D.; CAREY, L.; SCHULTZ, C.; BATEMAN, M.; MCCAUL, E.; STANO, G. The GOES-R geostationary lightning mapper (GLM). **Atmospheric Research**, v. 125-126, p. 34–49, 2013. 2, 17
- GOODMAN, S. J.; BUECHLER, D. E.; WRIGHT, P. D.; RUST, W. D. Lightning and precipitation history of a microburst-producing storm. **Geophysical Research Letters**, v. 15, n. 11, p. 1185–1188, 1988. 19
- GOODMAN, S. J.; GURKA, J.; DEMARIA, M.; SCHMIT, T. J.; MOSTEK, A.; JEDLOVEC, G.; SIEWERT, C.; FELTZ, W.; GERTH, J.; BRUMMER, R. et al. The goes-r proving ground: accelerating user readiness for the next-generation geostationary environmental satellite system. **Bulletin of the American Meteorological Society**, v. 93, n. 7, p. 1029–1040, 2012. 2, 17
- GRECO, S.; ULANSKI, S.; GARSTANG, M.; HOUSTON, S. Low-level nocturnal wind maximum over the central amazon basin. **Boundary-Layer Meteorology**, v. 58, n. 1, p. 91–115, 1992. 53
- GREENE, D. R.; CLARK, R. A. Vertically integrated liquid water—a new analysis tool. **Monthly Weather Review**, v. 100, n. 7, p. 548–552, 1972. 20, 66
- GRENET, G. Essai d'explication de la charge electrique des nuages d'orages. **Annales Geophysicae**, v. 3, p. 306–307, 1947. 7
- GUILLOT, E. M.; SMITH, T. M.; LAKSHMANAN, V.; ELMORE, K. L.; BURGESS, D. W.; STUMPF, G. J. Tornado and severe thunderstorm warning forecast skill and its relationship to storm type. In: CONFERENCE ON INTERACTIVE INFORMATION PROCESSING SYSTEMS FOR METEOROLOGY, OCEANOGRAPHY, AND HYDROLOGY. **Proceedings...** American Meteorological Society, 2008. 14
- GUTTMAN, A. R-trees: a dynamic index structure for spatial searching. In: SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA. **Proceedings...** ACM, 1984. p. 47–57. 32
- HAIXIANG, G.; YIJING, L.; SHANG, J.; MINGYUN, G.; YUANYUE, H.; BING, G. Learning from class-imbalanced data: review of methods and applications.

**Expert Systems with Applications**, v. 73, p. 220–239, 2017. ISSN 0957-4174.

Disponível em: <[https:](https://www.sciencedirect.com/science/article/pii/S0957417416307175)

[//www.sciencedirect.com/science/article/pii/S0957417416307175](https://www.sciencedirect.com/science/article/pii/S0957417416307175)>. 97

HAN, L.; SUN, J.; ZHANG, W.; XIU, Y.; FENG, H.; LIN, Y. A machine learning nowcasting method based on real-time reanalysis data. **Journal of Geophysical Research: Atmospheres**, v. 122, n. 7, p. 4038–4051, 2017. 15

HASSOUN, M. H. **Fundamentals of artificial neural networks**. Cambridge, MA, USA: MIT Press, 1995. ISBN 026208239X. 12

HAYKIN, S. **Neural networks: a comprehensive foundation**. USA: Prentice Hall, 1994. ISBN 0023527617. 12

HILL, A. J.; HERMAN, G. R.; SCHUMACHER, R. S. Forecasting severe weather with random forests. **Monthly Weather Review**, v. 148, n. 5, p. 2135–2161, 2020. 15

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997. 13

HOUZE, R. A. **Cloud dynamics**. [s.n.], 2014. OCLC: 910878801. ISBN 9780080959955 9780080921464. Disponível em: <<https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1001667>>. 85, 88

HUAMÁN, J. H. **Uso de multi-canais do GOES-16 para previsão imediata de densidade de descargas elétricas**. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos-SP, 2018. 1, 17

HUNTER, S. M. Wsr-88d radar rainfall estimation: capabilities, limitations and potential improvements. **National Weather Digest**, v. 20, n. 4, p. 26–38, 1996. 19

HUSSAIN, J. **Deep learning black box problem**. 59 p. Dissertação (Mestrado) — Uppsala University, Uppsala, 2019. 14

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Missão, visão e valores**. 2021. Disponível em:

<[http://www.inpe.br/institucional/sobre\\_inpe/missao.php](http://www.inpe.br/institucional/sobre_inpe/missao.php)>. Acesso em: 20 maio 2021. 53

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial neural networks: a tutorial. **Computer**, v. 29, n. 3, p. 31–44, 1996. 12, 43, 45

JARVIS, A. et al. **Hole-filled srtm for the globe Version 4**. 2008. Disponível em: <<http://srtm.csi.cgiar.org>>. 25

KATONA, B.; MARKOWSKI, P.; ALEXANDER, C.; BENJAMIN, S. The influence of topography on convective storm environments in the eastern united states as deduced from the hrrr. **Weather and Forecasting**, Boston MA, USA, v. 31, n. 5, p. 1481 – 1490, 2016. Disponível em: <[https://journals.ametsoc.org/view/journals/wefo/31/5/waf-d-16-0038\\_1.xml](https://journals.ametsoc.org/view/journals/wefo/31/5/waf-d-16-0038_1.xml)>. 52

KLEINA, M.; MATIOLI, L. C.; LEITE, E. A. Identificação, monitoramento e previsão de tempestades elétricas utilizando métodos numéricos. **Boletim de Ciências Geodésicas**, v. 22, p. 589 – 612, 2016. 18

KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE. **Proceedings...** Montreal, Canada, 1995. p. 1137–1145. 47

KRIDER, E. P. Benjamin franklin and lightning rods. **Physics Today**, v. 59, n. 1, p. 42, 2006. 8

LAHIFF, C. T. **Vertically integrated liquid density and its associated hail size range across the Burlington, Vermont county warning area**. 2005. Disponível em: <<https://www.weather.gov/media/erh/ta2005-01.pdf>>. 20

LAMBERT, W.; WHEELER, M.; ROEDER, W. Objective lightning forecasting at kennedy space center and cape canaveral air force station using cloud-to-ground lightning surveillance system data. In: CONFERENCE ON METEOROLOGICAL APPLICATIONS OF LIGHTNING DATA. **Proceedings...** American Meteorological Society, 2005. 16

LATHAM, J. The electrification of thunderstorms. **Quarterly Journal of the Royal Meteorological Society**, v. 107, n. 452, p. 277–298, 1981. 8

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, 1998. 34

LEE, A. An experimental study of the remote location of lightning flashes using a vlf arrival time difference technique. **Quarterly Journal of the Royal Meteorological Society**, v. 112, n. 471, p. 203–229, 1986. 29

LEWIS, E.; HARVEY, R.; RASMUSSEN, J. Hyperbolic direction finding with spherics of transatlantic origin. **Journal of Geophysical Research**, v. 65, n. 7, p. 1879–1905, 1960. 29, 30

LHERMITTE, R.; WILLIAMS, E. Cloud electrification. **Reviews of Geophysics and Space Physics**, v. 21, n. 5, p. 984–992, 1983. 7, 19, 86

LIN, T.-Y.; GOYAL, P.; GIRSHICK, R.; HE, K.; DOLLÁR, P. **Focal loss for dense object detection**. 2018. Disponível em: <[arXiv:1708.02002](https://arxiv.org/abs/1708.02002)>. 46

LIU, C.; CECIL, D. J.; ZIPSER, E. J.; KRONFELD, K.; ROBERTSON, R. Relationships between lightning flash rates and radar reflectivity vertical structures in thunderstorms over the tropics and subtropics. **Journal of Geophysical Research: Atmospheres**, v. 117, n. D6, 2012. 19, 61

ŁOBODA, M. Lightning protection of structures. In: \_\_\_\_\_. **Lightning: principles, instruments and applications: review of modern lightning research**. Dordrecht: Springer Netherlands, 2009. p. 573–592. ISBN 978-1-4020-9079-0. Disponível em: <[https://doi.org/10.1007/978-1-4020-9079-0\\_27](https://doi.org/10.1007/978-1-4020-9079-0_27)>. 2

LOPES, C. d. C. **Microfísica, cinemática e eletrificação em tempestades tropicais que geram granizo durante o projeto SOS-CHUVA**. Dissertação (Mestrado em Meteorologia) — Universidade de São Paulo (USP), São Paulo, 2019. 30

LUMB, F. E. Topographic influences on thunderstorm activity near lake victoria. **Weather**, v. 25, n. 9, p. 404–410, 1970. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1477-8696.1970.tb04129.x>>. 52

LYNN, B. H.; YAIR, Y.; PRICE, C.; KELMAN, G.; CLARK, A. J. Predicting cloud-to-ground and intracloud lightning in weather forecast models. **Weather and Forecasting**, v. 27, n. 6, p. 1470–1488, 2012. 9

MACGORMAN, D. R.; NIELSEN, K. E. Cloud-to-ground lightning in a tornadic storm on 8 may 1986. **Monthly Weather Review**, v. 119, n. 7, p. 1557–1574, 1991. 9

MACHADO, L. A. T.; CALHEIROS, A. J. P.; BISCARO, T.; GIANGRANDE, S.; DIAS, M. A. F. S.; 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.; WENDISCH, M. Overview: precipitation characteristics and sensitivities to environmental conditions during goamazon2014/5 and acridicon-chuva. **Atmospheric Chemistry and Physics**, v. 18, p. 6461–6482, 2018. 2, 21, 24, 26, 36, 52, 54, 57, 104

MACHADO, L. A. T.; LAURENT, H. The convective system area expansion over amazonia and its relationships with convective system life duration and high-level wind divergence. **Monthly Weather Review**, v. 132, n. 3, p. 714–725, 2004. 19

MAI, X.; ZHONG, H.; LI, L. Using svm to provide precipitation nowcasting based on radar data. In: INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION, FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY. **Proceedings...** Springer, 2019. p. 349–356. 15

MANOLOPOULOS, Y.; THEODORIDIS, Y.; TSOTRAS, V. J. Spatial indexing techniques. In: \_\_\_\_\_. **Encyclopedia of database systems**. Boston, MA: Springer US, 2009. p. 2702–2707. ISBN 978-0-387-39940-9. Disponível em: <[https://doi.org/10.1007/978-0-387-39940-9\\_355](https://doi.org/10.1007/978-0-387-39940-9_355)>. 32

MANZATO, A. Sounding-derived indices for neural network based short-term thunderstorm and rainfall forecasts. **Atmospheric Research**, v. 83, n. 2-4, p. 349–365, 2007. 13

MAQSOOD, I.; KHAN, M. R.; ABRAHAM, A. An ensemble of neural networks for weather forecasting. **Neural Computing & Applications**, v. 13, n. 2, p. 112–122, 2004. 16

MARKET, P. S.; BECKER, A. E. A study of lightning flashes attending periods of banded snowfall. **Geophysical Research Letters**, v. 36, n. 1, 2009. 9

MARSHALL, J. S.; PALMER, W. M. K. The distribution of raindrops with size. **Journal of Meteorology**, v. 5, n. 4, p. 165–166, 1948. 19, 20

MARTIN, S.; ARTAXO, P.; MACHADO, L.; MANZI, A.; SOUZA, R.; SCHUMACHER, C.; WANG, J.; ANDREAE, M.; BARBOSA, H.; FAN, J.;

FISCH, G.; GOLDSTEIN, A.; GUENTHER, A.; JIMENEZ, J.; PÖSCHL, U.; DIAS, M.; SMITH, J.; WENDISCH, M. Introduction: observations and modeling of the green ocean amazon (goamazon2014/5). **Atmospheric Chemistry and Physics Discussions**, v. 15, p. 30175–30210, 11 2015. [2](#), [25](#), [56](#)

MARTIN, S. T. et al. The green ocean amazon experiment (goamazon2014/5) observes pollution affecting gases, aerosols, clouds, and rainfall over the rain forest. **Bulletin of the American Meteorological Society**, v. 98, n. 5, p. 981–997, 2017. [22](#)

MARZBAN, C.; WITT, A. A bayesian neural network for severe-hail size prediction. **Weather and Forecasting**, v. 16, n. 5, p. 600–610, 2001. [3](#)

MCCANN, D. W. A neural network short-term forecast of significant thunderstorms. **Weather and Forecasting**, v. 7, n. 3, p. 525–534, 1992. [11](#), [13](#)

MCGOVERN, A.; SUPINIE, T.; GAGNE, I.; DJ, T. N.; COLLIER, M.; BROWN, R. A.; BASARA, J.; WILLIAMS, J. Understanding severe weather processes through spatiotemporal relational random forests. In: NASA CONFERENCE ON INTELLIGENT DATA UNDERSTANDING. **Proceedings...** [S.l.], 2010. [15](#)

MENG, Q.; YAO, W.; XU, L. Development of lightning nowcasting and warning technique and its application. **Advances in Meteorology**, v. 2019, p. 2405936, Jan 2019. [14](#)

MINIERE, X.; PINÇON, J.-L.; LEFEUVRE, F. A neural network approach to the classification of electron and proton whistlers. **Journal of Atmospheric and Terrestrial Physics**, v. 58, n. 7, p. 911–924, 1996. [3](#)

MOLLER, A. R.; DOSWELL, C. A.; FOSTER, M. P.; WOODALL, G. R. The operational recognition of supercell thunderstorm environments and storm structures. **Weather and Forecasting**, v. 9, n. 3, p. 327–347, 1994. [5](#)

MONINGER, W. The artificial intelligence shootout: a comparison of severe-storm-forecasting systems. In: CONFERENCE ON SEVERE LOCAL STORMS. **Proceedings.** [S.l.], 1990. p. 1–6. [11](#)

MORALES, C. A. **Continuous thunderstorm monitoring: Retrieval of precipitation parameters from lightning observations.** Tese (Doutorado em Civil and Environmental Engineering) — The University of Connecticut, 2001. [29](#)

- MORALES, C. A. **Tempestades elétricas na América do Sul: monitoramento, climatologia e propriedades físicas**. Tese (Livre docência) – Universidade de São Paulo (USP), São Paulo-SP, 2018. 29, 105
- MORALES, C. A.; NEVES, J.; MOIMAZ, E.; CAMARA, K. Sferics timing and ranging network-starnet: 8 years of measurements in south of america. In: INTERNATIONAL CONFERENCE ON ATMOSPHERIC ELECTRICITY. **Proceedings...** [S.l.], 2014. p. 1–8. 17, 25, 29
- MORELL, D. A. **Estimativa da taxa de precipitação com variáveis polarimétricas do radar meteorológico MXPOL**. Dissertação (Mestrado em Meteorologia) — Universidade de São Paulo (USP), São Paulo, 2018. 18
- MOSMANN, R.; FALKENHEIN, F. U. H.; GONCALVES, A.; NEPOMUCENO FILHO, F. Oil and gas potential of the amazon paleozoic basins. In: FUTURE PETROLEUM PROVINCES OF THE WORLD. **Proceedings...** [S.l.]: American Association of Petroleum Geologists, 1986. 23
- MOURA, A. D. Von Neumann e a previsão numérica de tempo e clima. **Estudos Avançados**, v. 10, p. 227 – 236, 04 1996. 10
- MURPHY, M. J.; NAG, A. Cloud lightning performance and climatology of the us based on the upgraded us national lightning detection network. In: CONFERENCE ON THE METEOROLOGICAL APPLICATIONS OF LIGHTNING DATA. **Proceedings...** [S.l.], 2015. 30
- NACCARATO, K. P.; PINTO, O. Lightning detection in southeastern Brazil from the new brazilian total lightning network (brasildat). In: INTERNATIONAL CONFERENCE ON LIGHTNING PROTECTION. **Proceedings...** IEEE, 2012. p. 1–9. 17
- NATIONAL CENTER FOR ATMOSPHERIC RESEARCH. **A description of the advanced research WRF model version 4**: Ncar technical notes. Boulder, Colorado, USA, jul 2019. 165 p. 10
- NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA). **GOES-16 mission overview**. 2016. Disponível em: <<https://www.goes-r.gov/mission/mission.html>>. 2
- NAVON, I. M. Data assimilation for numerical weather prediction: a review. In: \_\_\_\_\_. **Data assimilation for atmospheric, oceanic and hydrologic applications**. Berlin, Heidelberg: Springer, 2009. p. 21–65. 10

- NUSSBAUM, M. A.; CHAFFIN, D. B. Evaluation of artificial neural network modelling to predict torso muscle activity. **Ergonomics**, v. 39, n. 12, p. 1430–1444, 1996. 12
- OBERNDORFER, C. **Research on new artificial intelligence based path planning algorithms with focus on autonomous driving**. Tese (Doutorado) — University of Applied Science Munich, 2017. 44
- OLIVEIRA, A. P.; FITZJARRALD, D. R. The amazon river breeze and the local boundary layer: I. observations. **Boundary-Layer Meteorology**, v. 63, n. 1, p. 141–162, 1993. 53
- ONODA, M.; EBECKEN, N. F. Implementação em java de um algoritmo de árvore de decisão acoplado a um sgbd relacional. In: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS, 16. **Anais...** [S.l.], 2001. p. 55–64. 39, 40
- PEREIRA, R. F. d. O. **Propriedades físicas da precipitação e eletrificação dos sistemas precipitantes observados durante o projeto CHUVA-Manaus e GoAmazon**. Dissertação (Mestrado em Meteorologia) — Universidade de São Paulo (USP), São Paulo, 2019. 2, 7, 8, 21, 22, 25, 54
- PEREIRA, R. F. de O.; ALBRECHT, R. Evolução temporal da precipitação e atividade elétrica de uma tempestade com ocorrência de tempo severo. **Ciência e Natura**, v. 38, p. 532–542, 2016. 8
- PETRE, E. G. A decision tree for weather prediction. **PP**, v. 77, p. 82, 2009. 14
- PIELKE, R. A. Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. **Reviews of Geophysics**, v. 39, n. 2, p. 151–177, 2001. 21
- PRUPPACHER, H. R.; KLETT, J. D. Microstructure of atmospheric clouds and precipitation. In: ROSEN, R. D. (Ed.). **Microphysics of clouds and precipitation**. [S.l.]: Springer, 2010. p. 10–73. 86
- QING, X.; NIU, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by lstm. **Energy**, v. 148, p. 461–468, 2018. 13
- QIU, G.; WU, Z.; LI, Z.; DU, Q. Application of least square support vector machine for thunderstorm prediction. In: IEEE. **2010 8th World Congress on Intelligent Control and Automation**. [S.l.], 2010. p. 345–349. 15



QUEIROZ, A. P. **Monitoramento e previsão imediata de tempestades severas usando dados de radar**. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos-SP, 2009. 19

QUINLAN, J. R. Induction of decision trees. **Machine Learning**, v. 1, n. 1, p. 81–106, 1986. 14, 38, 39

RADHIKA, Y.; SHASHI, M. Atmospheric temperature prediction using support vector machines. **International Journal of Computer Theory and Engineering**, v. 1, n. 1, p. 55, 2009. 15

RAKOV, V. A. Triggered lightning. In: \_\_\_\_\_. **Lightning: principles, instruments and applications: review of modern lightning research**. Dordrecht: Springer, 2009. p. 23–56. ISBN 978-1-4020-9079-0. Disponível em: <[https://doi.org/10.1007/978-1-4020-9079-0\\_2](https://doi.org/10.1007/978-1-4020-9079-0_2)>. 2

\_\_\_\_\_. Types of lightning discharges and lightning terminology. In: \_\_\_\_\_. **Fundamentals of Lightning**. [S.l.]: Cambridge University Press, 2016. p. 1–16. 9, 29

RAMZI, M. M.; ADNAN, R.; SAMAD, A. M.; RUSLAN, F. A. Lightning prediction modelling using MLPNN structure. case study: Kuala Lumpur International Airport (KLIA). In: INTERNATIONAL CONFERENCE ON AUTOMATIC CONTROL AND INTELLIGENT SYSTEMS, 2018. **Proceedings... IEEE**, 2018. p. 63–66. 3

RAVI, N.; MOHANTY, U.; MADAN, O.; PALIWAL, R. Forecasting of thunderstorms in the pre-monsoon season at Delhi. **Meteorological Applications**, v. 6, n. 1, p. 29–38, 1999. 5, 16

RENDA, A.; FRANKLE, J.; CARBIN, M. **Comparing rewinding and fine-tuning in neural network pruning**. 2020. Disponível em: <<https://arxiv.org/abs/2003.02389>>. 45

REYNOLDS, S.; BROOK, M.; GOURLEY, M. F. Thunderstorm charge separation. **Journal of Meteorology**, v. 14, n. 5, p. 426–436, 1957. 8

RICHARDSON, L. F. **Weather prediction by numerical process**. Cambridge: [s.n.], 1922. 10

RINEHART, R. E. **Radar for meteorologists**. [S.l.]: University of North Dakota, 1991. 27

ROGERS, R. R.; YAU, M. K. **A short course in cloud physics**. 3. ed. [S.l.]: Elsevier, 1989. 18

RUIZ, A.; VILLA, N. **Storms prediction : logistic regression vs random forest for unbalanced data**. 2008. Disponível em: <<https://arxiv.org/abs/0804.0650>>. 15

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986. 45

RYZHKOV, A. V.; ZRNIC, D. S. **Radar polarimetry for weather observations**. Cham: Springer, 2019. (Springer Atmospheric Sciences). ISBN 9783030050931. 61

SARAIVA, I.; DIAS, M. S.; MORALES, C.; SARAIVA, J. Regional variability of rain clouds in the amazon basin as seen by a network of weather radars. **Journal of Applied Meteorology and Climatology**, v. 55, n. 12, p. 2657–2675, 2016. 25, 26

SARTOR, J. D. Radio emission from clouds. **Journal of Geophysical Research**, v. 68, n. 18, p. 5169–5172, 1963. 8

SCHMIT, T.; LINDSTROM, S.; GERTH, J.; GUNSHOR, M. Applications of the 16 spectral bands on the advanced baseline imager (ABI). **Journal of Operational Meteorology**, v. 06, p. 33–46, 06 2018. 17

SCHMIT, T. J.; GUNSHOR, M. M.; MENZEL, W. P.; GURKA, J. J.; LI, J.; BACHMEIER, A. S. Introducing the next-generation advanced baseline imager on goes-r. **Bulletin of the American Meteorological Society**, v. 86, n. 8, p. 1079–1096, 2005. 2

SCHONLAND, B. F. J. The lightning discharge. **Handbuch der Physik**, v. 22, p. 576–628, 1956. 9

SEITY, Y.; SOULA, S.; SAUVAGEOT, H. Lightning and precipitation relationship in coastal thunderstorms. **Journal of Geophysical Research: Atmospheres**, v. 106, n. D19, p. 22801–22816, 2001. 19

SHAFER, M. A.; MACGORMAN, D. R.; CARR, F. H. Cloud-to-ground lightning throughout the lifetime of a severe storm system in oklahoma. **Monthly Weather Review**, v. 128, n. 6, p. 1798–1816, 2000. 9

SILVA, L.; ABAIDE, A.; NEGRI, V.; CAPELETTI, M.; LOPES, L.; CARDOSO, G. Diagnostic and input selection tool applied on weather variables for studies of short-term load forecasting. In: INTERNATIONAL CONFERENCE ON MODERN POWER SYSTEMS, 8., 2019. **Proceedings... IEEE**, 2019. p. 1–6. 2

SILVA, L. M. O. da. **Uma aplicação de árvores de decisão, redes neurais e KNN para a identificação de modelos arma não-sazonais e sazonais**. Tese (Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO, Rio de Janeiro – RJ, 2005. 39, 40

SILVA, T. G.; SIQUEIRA, P. H.; BENETI, C.; BUZZI, M.; CALVETTI, L. Identification of severe weather event with meteorological radar and multilayer perceptron. In: IBERIAN LATIN-AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING, 38., 2017. **Proceedings... Florianópolis-SC**, 2017. 3

SISSAKIAN, V.; AL-ANSARI, N.; KNUTSSON, S. Sand and dust storm events in iraq. **Journal of Natural Science**, v. 5, n. 10, p. 1084–1094, 2013. 9

SMITH, L. N. **A disciplined approach to neural network hyper-parameters: part 1 – learning rate, batch size, momentum, and weight decay**. 2018. Disponível em: <<https://arxiv.org/abs/1803.09820>>. 45

STANO, G. T. Using total lightning observations to enhance lightning safety. In: SYMPOSIUM ON POLICY AND SOCIO-ECONOMIC RESEARCH, 7., 2012. **Proceedings... [S.l.]**, 2012. 9

STEINER, M.; HOUZE, R. A.; YUTER, S. E. Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. **Journal of Applied Meteorology**, v. 34, n. 9, p. 1978–2007, 1995. 18

STENSÖ, I. N. **Predicting tropical thunderstorm trajectories using LSTM**. Dissertação (Mestrado) — School of Electrical Engineering and Computer Science, 2018. 13

STOLZENBURG, M. Observations of high ground flash densities of positive lightning in summertime thunderstorms. **Monthly Weather Review**, Boston MA, USA, v. 122, n. 8, p. 1740 – 1750, 01 Aug. 1994. Disponível em: <[https://journals.ametsoc.org/view/journals/mwre/122/8/1520-0493\\_1994\\_122\\_1740\\_oohgfd\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/122/8/1520-0493_1994_122_1740_oohgfd_2_0_co_2.xml)>. 64

- SUI, C.-H.; LI, X.; YANG, M.-J. On the definition of precipitation efficiency. **Journal of the Atmospheric Sciences**, v. 64, n. 12, p. 4506–4513, 12 2007. 19
- SUPINIE, T. A.; MCGOVERN, A.; WILLIAMS, J.; ABERNATHY, J. Spatiotemporal relational random forests. In: INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS, 2009. **Proceedings...** IEEE, 2009. p. 630–635. 15
- TAHIR, M. A.; KITTLER, J.; MIKOLAJCZYK, K.; YAN, F. A multiple expert approach to the class imbalance problem using inverse random under sampling. In: INTERNATIONAL WORKSHOP ON MULTIPLE CLASSIFIER SYSTEMS, 2009. **Proceedings...** Springer, 2009. p. 82–91. 97, 98
- TAYLOR, J. W.; BUIZZA, R. Neural network load forecasting with weather ensemble predictions. **IEEE Transactions on Power Systems**, v. 17, n. 3, p. 626–632, 2002. 16
- TORACINTA, E. R.; MOHR, K. I.; ZIPSER, E. J.; ORVILLE, R. E. A comparison of WSR-88D reflectivities, SSM/I brightness temperatures, and lightning for mesoscale convective systems in Texas. part i: radar reflectivity and lightning. **Journal of Applied Meteorology**, v. 35, n. 6, p. 902–918, 1996. 19
- UMAN, M. A.; MCLAIN, D. K. Magnetic field of lightning return stroke. **Journal of Geophysical Research**, v. 74, n. 28, p. 6899–6910, 1969. 9
- VELASCO, L. C. P.; SERQUIÑA, R. P.; ZAMAD, M. S. A. A.; JUANICO, B. F.; LOMOCOSO, J. C. Week-ahead rainfall forecasting using multilayer perceptron neural network. **Procedia Computer Science**, v. 161, p. 386–397, 2019. 3
- VICENTE, G. A.; SCOFIELD, R. A.; MENZEL, W. P. The operational goes infrared rainfall estimation technique. **Bulletin of the American Meteorological Society**, v. 79, n. 9, p. 1883–1898, 1998. 11
- VILA, D. A.; MACHADO, L. A. T.; LAURENT, H.; VELASCO, I. Forecast and tracking the evolution of cloud clusters (ForTraCC) using satellite infrared imagery: methodology and validation. **Weather and Forecasting**, v. 23, n. 2, p. 233–245, 2008. 11, 80
- VINCENT, B. R.; CAREY, L. D.; SCHNEIDER, D.; KEETER, K.; GONSKI, R. Using WSR-88D reflectivity data for the prediction of cloud-to-ground lightning: a central North Carolina study. **National Weather Digest**, v. 27, p. 35–44, 2003. 67

- VISACRO, S. **Descargas atmosféricas: uma abordagem de engenharia**. [S.l.]: Artliber, 2005. 9
- VONNEGUT, B. Possible mechanism for the formation of thunderstorm electricity. In: CONFERENCE ATMOSPHERIC ELECTRICITY PORTSMOUTH, 1995. **Proceedings...** [S.l.], 1955. p. 169–181. 7
- VOORMANSIK, T.; ROSSI, P. J.; MOISSEEV, D.; TANILSOO, T.; POST, P. Thunderstorm hail and lightning detection parameters based on dual-polarization doppler weather radar data. **Meteorological Applications**, v. 24, n. 3, p. 521–530, 2017. 18, 64
- WALDVOGEL, A.; FEDERER, B.; GRIMM, P. Criteria for the detection of hail cells. **Journal of Applied Meteorology**, v. 18, n. 12, p. 1521–1525, 1979. 19, 63
- WALLACE, J. M.; HOBBS, P. V. **Atmospheric science: an introductory survey**. [S.l.]: Elsevier, 2006. 5, 6
- WANG, D.; MIAO, J.; TAN, Z. Impacts of topography and land cover change on thunderstorm over the Huangshan (yellow mountain) area of China. **Natural Hazards**, v. 67, n. 2, p. 675–699, 2013. 21
- WANG, Y. Prediction of weather impacted airport capacity using ensemble learning. In: DIGITAL AVIONICS SYSTEMS CONFERENCE, 30., 2011. **Proceedings...** IEEE, 2011. p. 2D6–1. 16
- WANG, Z.; LIAO, R.; LI, J.; HE, J.; WANG, G.; XU, Z.; WANG, H. Thunderstorm identification algorithm research based on simulated airborne weather radar reflectivity data. **EURASIP Journal on Wireless Communications and Networking**, 2020. 5, 6
- WANG, Z.-h.; ZHANG, Y.; ZHU, J. A preliminary study on thunderstorm forecast with ls-svm method. **Journal of Tropical Meteorology**, v. 19, n. 1, p. 104, 2013. 15
- WATSON, A. I.; HOLLE, R. L.; LOPEZ, R. E. Lightning from two national detection networks related to vertically integrated liquid and echo-top information from wsr-88d radar. **Weather and Forecasting**, v. 10, n. 3, p. 592–605, 1995. 20
- WERBOS, P. **Beyond regression: new tools for prediction and analysis in the behavioral sciences**. Tese (Doutorado) — Harvard University, Cambridge, Massachusetts, USA, 1974. 45

- WILKS, D. **Statistical methods in the atmospheric sciences**. 3. ed. Cambridge - Massachusetts: Academic Press, 2011. 704 p. ISBN 9780123850232. 47, 48
- WILLIAMS, E. Lightning and climate: a review. **Atmospheric Research**, v. 76, n. 1-4, p. 272–287, 2005. 21
- WILLIAMS, E.; BOLDI, B.; MATLIN, A.; WEBER, M.; HODANISH, S.; SHARP, D.; GOODMAN, S.; RAGHAVAN, R.; BUECHLER, D. The behavior of total lightning activity in severe florida thunderstorms. **Atmospheric Research**, v. 51, n. 3, p. 245–265, 1999. ISSN 0169-8095. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169809599000113>>. 9, 67
- WILLIAMS, E. R. The electrification of severe storms. In: DOSWELL, C. A. (Ed.). **Severe convective storms**. [S.l.]: Springer, 2001. p. 527–561. 9
- \_\_\_\_\_. Predictable lightning paths? **Nature Geoscience**, v. 1, n. 4, p. 216–217, 2008. 16
- WU, F.; CUI, X.; ZHANG, D.-L. A lightning-based nowcast-warning approach for short-duration rainfall events: development and testing over Beijing during the warm seasons of 2006–2007. **Atmospheric Research**, v. 205, p. 2 – 17, 2018. ISSN 0169-8095. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169809517311936>>. 10
- XU, M.; WATANACHATURAPORN, P.; VARSHNEY, P. K.; ARORA, M. K. Decision tree regression for soft classification of remote sensing data. **Remote Sensing of Environment**, v. 97, n. 3, p. 322–336, 2005. 14
- YU, X.; YAO, X.; XIONG, T. **Principle and operational application of doppler weather radar [M]**. [S.l.]: China Meteorological Press, 2006. 5
- YUTER, S. E.; HOUZE, R. A. Three-dimensional kinematic and microphysical evolution of Florida Cumulonimbus. part ii: frequency distributions of vertical velocity, reflectivity, and differential reflectivity. **Monthly Weather Review**, Boston MA, USA, v. 123, n. 7, p. 1941 – 1963, 1995. Disponível em: <[https://journals.ametsoc.org/view/journals/mwre/123/7/1520-0493\\_1995\\_123\\_1941\\_tdkame\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/123/7/1520-0493_1995_123_1941_tdkame_2_0_co_2.xml)>. 62, 86
- ZIPSER, E. J.; LUTZ, K. R. The vertical profile of radar reflectivity of convective cells: A strong indicator of storm intensity and lightning probability? **Monthly Weather Review**, v. 122, n. 8, p. 1751–1759, 1994. 18, 19, 61

## APÊNDICE A - ESTATÍSTICAS SÍNCRONAS DAS VARIÁVEIS

A Tabela A.1 apresenta algumas informações estatísticas que complementam as análises assíncronas discutidas na Subseção 4.2.2 do Capítulo 4. Essas informações estão agrupadas por variável, atividade/inatividade e estação. São elas: assimetria, curtose, média, mediana, desvio-padrão, variância e se a hipótese nula foi aceita no teste de hipótese. Foi aplicado o teste t de *student* para verificar se as médias dos valores das variáveis na atividade são diferentes dos valores das variáveis na inatividade com o intervalo de confiança de 95%. Em que, quando a hipótese nula ( $H_0$ ) é aceita (Sim), as médias não apresentam diferenças estatisticamente significativas. Já quando for rejeitada (Não), essas diferenças são significantes.

Tabela A.1 - Estatísticas síncronas das variáveis e teste de hipótese.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t	Refletividade em 2.0 km	inatividade	chuvosa	0.175	-0.478	15.163	14.885	11.305	127.804	Não
t	Refletividade em 2.0 km	atividade	chuvosa	-0.401	-0.419	28.256	29.401	12.574	158.112	Não
t	Refletividade em 2.0 km	inatividade	seca	1.239	3.430	11.012	9.966	8.247	68.010	Não
t	Refletividade em 2.0 km	atividade	seca	-0.297	-0.658	30.637	31.507	13.321	177.453	Não
t	Refletividade em 2.0 km	inatividade	chuvosa-seca	0.741	0.378	12.241	10.070	10.765	115.882	Não
t	Refletividade em 2.0 km	atividade	chuvosa-seca	-0.378	-0.544	29.207	30.259	13.049	170.271	Não
t	Refletividade em 2.0 km	inatividade	seca-chuvosa	1.042	1.530	9.814	8.091	10.118	102.379	Não
t	Refletividade em 2.0 km	atividade	seca-chuvosa	-0.374	-0.652	28.312	29.808	14.084	198.350	Não
t	Refletividade em 2.5 km	inatividade	chuvosa	0.018	-0.336	16.201	16.303	11.149	124.300	Não
t	Refletividade em 2.5 km	atividade	chuvosa	-0.373	-0.371	28.578	29.531	12.187	148.520	Não
t	Refletividade em 2.5 km	inatividade	seca	1.071	2.774	11.347	10.467	8.693	75.564	Não
t	Refletividade em 2.5 km	atividade	seca	-0.278	-0.591	31.118	31.754	12.958	167.904	Não
t	Refletividade em 2.5 km	inatividade	chuvosa-seca	0.486	0.075	13.405	11.901	11.205	125.542	Não
t	Refletividade em 2.5 km	atividade	chuvosa-seca	-0.369	-0.446	29.661	30.483	12.623	159.341	Não
t	Refletividade em 2.5 km	inatividade	seca-chuvosa	0.832	0.981	10.657	9.142	10.655	113.520	Não
t	Refletividade em 2.5 km	atividade	seca-chuvosa	-0.379	-0.536	28.933	30.173	13.643	186.128	Não
t	Refletividade em 3.0 km	inatividade	chuvosa	-0.097	-0.136	17.073	17.220	10.853	117.786	Não
t	Refletividade em 3.0 km	atividade	chuvosa	-0.341	-0.338	28.634	29.386	11.908	141.798	Não
t	Refletividade em 3.0 km	inatividade	seca	0.877	2.166	12.035	11.383	9.137	83.479	Não
t	Refletividade em 3.0 km	atividade	seca	-0.231	-0.580	31.397	31.790	12.652	160.061	Não
t	Refletividade em 3.0 km	inatividade	chuvosa-seca	0.238	0.008	14.833	13.969	11.374	129.373	Não
t	Refletividade em 3.0 km	atividade	chuvosa-seca	-0.333	-0.416	29.859	30.501	12.267	150.487	Não
t	Refletividade em 3.0 km	inatividade	seca-chuvosa	0.630	0.602	11.740	10.486	11.036	121.797	Não
t	Refletividade em 3.0 km	atividade	seca-chuvosa	-0.363	-0.479	29.229	30.296	13.315	177.277	Não
t	Refletividade em 3.5 km	inatividade	chuvosa	-0.176	0.081	17.747	17.827	10.510	110.451	Não
t	Refletividade em 3.5 km	atividade	chuvosa	-0.297	-0.341	28.452	29.010	11.662	135.995	Não
t	Refletividade em 3.5 km	inatividade	seca	0.642	1.644	12.914	12.525	9.592	92.006	Não
t	Refletividade em 3.5 km	atividade	seca	-0.171	-0.599	31.396	31.504	12.423	154.327	Não
t	Refletividade em 3.5 km	inatividade	chuvosa-seca	0.041	0.149	16.110	15.550	11.255	126.676	Não
t	Refletividade em 3.5 km	atividade	chuvosa-seca	-0.287	-0.425	29.784	30.241	11.993	143.827	Não

Continua na próxima página



Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t	Refletividade em 3.5 km	inatividade	seca-chuvosa	0.453	0.421	12.662	11.616	11.229	126.097	Não
t	Refletividade em 3.5 km	atividade	seca-chuvosa	-0.325	-0.480	29.260	30.127	13.101	171.645	Não
t	Refletividade em 4.0 km	inatividade	chuvosa	-0.223	0.306	18.423	18.353	10.082	101.653	Não
t	Refletividade em 4.0 km	atividade	chuvosa	-0.241	-0.370	28.159	28.517	11.348	128.785	Não
t	Refletividade em 4.0 km	inatividade	seca	0.459	1.386	14.037	13.651	9.877	97.557	Não
t	Refletividade em 4.0 km	atividade	seca	-0.108	-0.635	31.260	31.146	12.207	149.015	Não
t	Refletividade em 4.0 km	inatividade	chuvosa-seca	-0.075	0.332	17.187	16.635	10.937	119.613	Não
t	Refletividade em 4.0 km	atividade	chuvosa-seca	-0.239	-0.458	29.550	29.903	11.720	137.358	Não
t	Refletividade em 4.0 km	inatividade	seca-chuvosa	0.300	0.369	13.782	12.696	11.242	126.384	Não
t	Refletividade em 4.0 km	atividade	seca-chuvosa	-0.281	-0.514	29.193	29.878	12.861	165.402	Não
t	Refletividade em 4.5 km	inatividade	chuvosa	-0.192	0.437	18.881	18.623	9.479	89.846	Não
t	Refletividade em 4.5 km	atividade	chuvosa	-0.165	-0.443	27.623	27.807	11.033	121.730	Não
t	Refletividade em 4.5 km	inatividade	seca	0.424	1.327	15.082	14.461	9.775	95.546	Não
t	Refletividade em 4.5 km	atividade	seca	-0.031	-0.688	30.901	30.552	11.996	143.898	Não
t	Refletividade em 4.5 km	inatividade	chuvosa-seca	-0.079	0.420	17.945	17.302	10.405	108.261	Não
t	Refletividade em 4.5 km	atividade	chuvosa-seca	-0.165	-0.540	29.089	29.265	11.428	130.598	Não
t	Refletividade em 4.5 km	inatividade	seca-chuvosa	0.250	0.426	14.883	13.597	10.875	118.274	Não
t	Refletividade em 4.5 km	atividade	seca-chuvosa	-0.209	-0.600	28.871	29.377	12.610	159.014	Não
t	Refletividade em 5.0 km	inatividade	chuvosa	-0.124	0.493	18.863	18.520	8.802	77.473	Não
t	Refletividade em 5.0 km	atividade	chuvosa	-0.081	-0.469	26.815	26.662	10.705	114.603	Não
t	Refletividade em 5.0 km	inatividade	seca	0.425	1.323	15.813	14.992	9.521	90.655	Não
t	Refletividade em 5.0 km	atividade	seca	0.060	-0.719	30.316	29.690	11.759	138.274	Não
t	Refletividade em 5.0 km	inatividade	chuvosa-seca	-0.049	0.502	18.311	17.635	9.775	95.554	Não
t	Refletividade em 5.0 km	atividade	chuvosa-seca	-0.083	-0.580	28.370	28.257	11.114	123.529	Não
t	Refletividade em 5.0 km	inatividade	seca-chuvosa	0.261	0.526	15.563	14.205	10.326	106.634	Não
t	Refletividade em 5.0 km	atividade	seca-chuvosa	-0.120	-0.688	28.301	28.502	12.337	152.213	Não
t	Refletividade em 5.5 km	inatividade	chuvosa	-0.052	0.492	18.439	18.138	8.190	67.079	Não
t	Refletividade em 5.5 km	atividade	chuvosa	0.009	-0.474	25.865	25.455	10.325	106.605	Não
t	Refletividade em 5.5 km	inatividade	seca	0.437	1.361	16.322	15.462	9.142	83.571	Não
t	Refletividade em 5.5 km	atividade	seca	0.137	-0.694	29.588	28.681	11.511	132.504	Não
t	Refletividade em 5.5 km	inatividade	chuvosa-seca	0.009	0.554	18.320	17.705	9.113	83.044	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t	Refletividade em 5.5 km	atividade	chuvosa-seca	0.001	-0.579	27.495	27.090	10.764	115.869	Não
t	Refletividade em 5.5 km	inatividade	seca-chuvosa	0.311	0.625	15.853	14.563	9.707	94.219	Não
t	Refletividade em 5.5 km	atividade	seca-chuvosa	-0.036	-0.728	27.539	27.387	12.049	145.177	Não
t	Refletividade em 6.0 km	inatividade	chuvosa	0.006	0.501	17.871	17.653	7.642	58.394	Não
t	Refletividade em 6.0 km	atividade	chuvosa	0.100	-0.480	24.846	24.197	9.924	98.485	Não
t	Refletividade em 6.0 km	inatividade	seca	0.457	1.432	16.707	15.884	8.711	75.885	Não
t	Refletividade em 6.0 km	atividade	seca	0.215	-0.653	28.788	27.648	11.222	125.926	Não
t	Refletividade em 6.0 km	inatividade	chuvosa-seca	0.072	0.602	18.110	17.588	8.484	71.971	Não
t	Refletividade em 6.0 km	atividade	chuvosa-seca	0.085	-0.560	26.533	25.858	10.377	107.691	Não
t	Refletividade em 6.0 km	inatividade	seca-chuvosa	0.362	0.715	15.907	14.754	9.123	83.238	Não
t	Refletividade em 6.0 km	atividade	seca-chuvosa	0.044	-0.726	26.730	26.259	11.696	136.797	Não
t	Refletividade em 6.5 km	inatividade	chuvosa	0.009	0.564	17.234	17.126	7.227	52.234	Não
t	Refletividade em 6.5 km	atividade	chuvosa	0.159	-0.425	23.818	22.991	9.547	91.144	Não
t	Refletividade em 6.5 km	inatividade	seca	0.417	1.437	16.852	16.176	8.419	70.876	Não
t	Refletividade em 6.5 km	atividade	seca	0.279	-0.585	27.982	26.666	10.906	118.936	Não
t	Refletividade em 6.5 km	inatividade	chuvosa-seca	0.079	0.684	17.701	17.338	8.028	64.443	Não
t	Refletividade em 6.5 km	atividade	chuvosa-seca	0.151	-0.500	25.562	24.718	9.986	99.727	Não
t	Refletividade em 6.5 km	inatividade	seca-chuvosa	0.359	0.785	15.726	14.787	8.691	75.542	Não
t	Refletividade em 6.5 km	atividade	seca-chuvosa	0.113	-0.696	25.865	25.098	11.322	128.181	Não
t	Refletividade em 7.0 km	inatividade	chuvosa	0.008	0.607	16.640	16.637	6.796	46.186	Não
t	Refletividade em 7.0 km	atividade	chuvosa	0.214	-0.352	22.822	21.873	9.113	83.046	Não
t	Refletividade em 7.0 km	inatividade	seca	0.415	1.553	16.976	16.474	7.943	63.094	Não
t	Refletividade em 7.0 km	atividade	seca	0.343	-0.502	27.169	25.716	10.537	111.018	Não
t	Refletividade em 7.0 km	inatividade	chuvosa-seca	0.108	0.750	17.241	17.031	7.508	56.367	Não
t	Refletividade em 7.0 km	atividade	chuvosa-seca	0.209	-0.429	24.593	23.596	9.561	91.421	Não
t	Refletividade em 7.0 km	inatividade	seca-chuvosa	0.377	0.872	15.500	14.753	8.181	66.931	Não
t	Refletividade em 7.0 km	atividade	seca-chuvosa	0.177	-0.642	25.003	24.003	10.901	118.835	Não
t	Máxima refletividade	inatividade	chuvosa	0.161	-0.338	16.216	15.985	10.848	117.688	Não
t	Máxima refletividade	atividade	chuvosa	-0.301	-0.451	28.733	29.345	12.220	149.320	Não
t	Máxima refletividade	inatividade	seca	1.165	3.194	12.044	10.989	8.238	67.861	Não
t	Máxima refletividade	atividade	seca	-0.138	-0.681	31.117	30.996	13.054	170.396	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t	Máxima refletividade	inatividade	chuvosa-seca	0.652	0.316	13.378	11.596	10.569	111.708	Não
t	Máxima refletividade	atividade	chuvosa-seca	-0.268	-0.556	29.755	30.172	12.604	158.857	Não
t	Máxima refletividade	inatividade	seca-chuvosa	0.956	1.393	11.184	9.614	9.894	97.896	Não
t	Máxima refletividade	atividade	seca-chuvosa	-0.273	-0.620	29.330	30.032	13.513	182.612	Não
t	Altura máxima de 35 dbz	inatividade	chuvosa	0.128	0.371	4.341	4.500	1.608	2.584	Não
t	Altura máxima de 35 dbz	atividade	chuvosa	0.579	0.746	5.517	5.500	2.432	5.916	Não
t	Altura máxima de 35 dbz	inatividade	seca	0.786	1.361	5.149	5.000	2.309	5.331	Não
t	Altura máxima de 35 dbz	atividade	seca	0.468	-0.210	7.254	7.000	3.416	11.666	Não
t	Altura máxima de 35 dbz	inatividade	chuvosa-seca	0.241	0.309	4.591	4.500	1.790	3.204	Não
t	Altura máxima de 35 dbz	atividade	chuvosa-seca	0.394	0.199	5.885	5.500	2.586	6.688	Não
t	Altura máxima de 35 dbz	inatividade	seca-chuvosa	0.521	1.007	5.087	5.000	2.066	4.267	Não
t	Altura máxima de 35 dbz	atividade	seca-chuvosa	0.474	0.067	6.934	6.500	3.128	9.782	Não
t	Taxa de precipitação	inatividade	chuvosa	91.124	102079.629	1.434	0.405	3.603	12.982	Não
t	Taxa de precipitação	atividade	chuvosa	2.765	10.747	8.251	2.854	12.569	157.989	Não
t	Taxa de precipitação	inatividade	seca	166.779	240936.828	0.835	0.171	4.526	20.481	Não
t	Taxa de precipitação	atividade	seca	2.838	11.426	13.994	3.962	22.435	503.344	Não
t	Taxa de precipitação	inatividade	chuvosa-seca	11.901	1187.559	1.281	0.212	4.002	16.018	Não
t	Taxa de precipitação	atividade	chuvosa-seca	2.537	8.690	10.287	3.285	15.522	240.945	Não
t	Taxa de precipitação	inatividade	seca-chuvosa	12.369	335.638	1.016	0.141	4.249	18.052	Não
t	Taxa de precipitação	atividade	seca-chuvosa	3.168	20.212	10.788	3.138	17.354	301.164	Não
t	Vil	inatividade	chuvosa	11.020	2811.173	0.261	0.089	0.519	0.269	Não
t	Vil	atividade	chuvosa	2.593	8.655	1.292	0.503	1.879	3.529	Não
t	Vil	inatividade	seca	16.139	862.634	0.140	0.022	0.653	0.427	Não
t	Vil	atividade	seca	2.921	11.525	2.248	0.651	3.651	13.333	Não
t	Vil	inatividade	chuvosa-seca	7.380	108.377	0.216	0.035	0.585	0.342	Não
t	Vil	atividade	chuvosa-seca	2.578	9.088	1.567	0.561	2.314	5.355	Não
t	Vil	inatividade	seca-chuvosa	10.848	220.094	0.173	0.021	0.639	0.409	Não
t	Vil	atividade	seca-chuvosa	2.977	16.273	1.797	0.572	2.805	7.871	Não
t	Densidade do vil	inatividade	chuvosa	42.943	23969.500	0.026	0.011	0.068	0.005	Não
t	Densidade do vil	atividade	chuvosa	30.376	1625.888	0.057	0.025	0.172	0.030	Não
t	Densidade do vil	inatividade	seca	170.069	81059.575	0.014	0.008	0.063	0.004	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t	Densidade do vil	atividade	seca	79.392	9972.605	0.084	0.024	0.573	0.328	Não
t	Densidade do vil	inatividade	chuvosa-seca	48.021	10030.021	0.018	0.008	0.058	0.003	Não
t	Densidade do vil	atividade	chuvosa-seca	37.194	2270.584	0.064	0.023	0.283	0.080	Não
t	Densidade do vil	inatividade	seca-chuvosa	350.633	436542.185	0.015	0.006	0.070	0.005	Não
t	Densidade do vil	atividade	seca-chuvosa	34.207	2170.110	0.066	0.021	0.292	0.085	Não
t	Altura de waldvogel	inatividade	chuvosa	0.216	-0.132	-1.160	-1.000	1.411	1.992	Não
t	Altura de waldvogel	atividade	chuvosa	0.152	-0.239	-0.479	-0.500	1.645	2.707	Não
t	Altura de waldvogel	inatividade	seca	0.469	0.548	-0.031	0.000	1.926	3.708	Não
t	Altura de waldvogel	atividade	seca	0.554	0.560	0.827	0.500	2.379	5.658	Não
t	Altura de waldvogel	inatividade	chuvosa-seca	0.254	-0.019	-0.835	-1.000	1.522	2.318	Não
t	Altura de waldvogel	atividade	chuvosa-seca	0.150	-0.187	-0.201	0.000	1.765	3.116	Não
t	Altura de waldvogel	inatividade	seca-chuvosa	0.369	0.373	-0.249	-0.500	1.771	3.138	Não
t	Altura de waldvogel	atividade	seca-chuvosa	0.387	0.499	0.504	0.500	2.092	4.377	Não
t	Varição do topo da nuvem	inatividade	chuvosa	-0.017	0.592	4.339	4.294	3.598	12.945	Não
t	Varição do topo da nuvem	atividade	chuvosa	0.138	-0.431	7.800	7.444	4.717	22.252	Não
t	Varição do topo da nuvem	inatividade	seca	0.359	1.627	4.214	3.985	4.132	17.075	Não
t	Varição do topo da nuvem	atividade	seca	0.218	-0.609	10.118	9.649	5.403	29.188	Não
t	Varição do topo da nuvem	inatividade	chuvosa-seca	0.010	0.680	4.819	4.666	4.065	16.528	Não
t	Varição do topo da nuvem	atividade	chuvosa-seca	0.098	-0.461	8.988	8.673	4.918	24.191	Não
t	Varição do topo da nuvem	inatividade	seca-chuvosa	0.344	0.886	3.528	3.123	4.301	18.495	Não
t	Varição do topo da nuvem	atividade	seca-chuvosa	0.079	-0.710	8.830	8.580	5.624	31.629	Não
t-1	Refletividade em 2.0 km	inatividade	chuvosa	0.176	-0.477	15.167	14.880	11.307	127.852	Não
t-1	Refletividade em 2.0 km	atividade	chuvosa	-0.331	-0.514	27.035	27.825	12.853	165.192	Não
t-1	Refletividade em 2.0 km	inatividade	seca	1.249	3.438	11.005	9.942	8.278	68.527	Não
t-1	Refletividade em 2.0 km	atividade	seca	-0.209	-0.746	29.165	29.558	13.681	187.169	Não
t-1	Refletividade em 2.0 km	inatividade	chuvosa-seca	0.744	0.385	12.211	10.034	10.768	115.947	Não
t-1	Refletividade em 2.0 km	atividade	chuvosa-seca	-0.280	-0.663	27.824	28.469	13.465	181.293	Não
t-1	Refletividade em 2.0 km	inatividade	seca-chuvosa	1.039	1.518	9.733	8.005	10.138	102.777	Não
t-1	Refletividade em 2.0 km	atividade	seca-chuvosa	-0.266	-0.789	27.006	27.956	14.661	214.945	Não
t-1	Refletividade em 2.5 km	inatividade	chuvosa	0.019	-0.337	16.204	16.302	11.152	124.360	Não
t-1	Refletividade em 2.5 km	atividade	chuvosa	-0.322	-0.400	27.446	28.030	12.441	154.769	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-1	Refletividade em 2.5 km	inatividade	seca	1.080	2.780	11.340	10.444	8.725	76.122	Não
t-1	Refletividade em 2.5 km	atividade	seca	-0.203	-0.647	29.814	29.964	13.249	175.541	Não
t-1	Refletividade em 2.5 km	inatividade	chuvosa-seca	0.489	0.079	13.374	11.860	11.210	125.670	Não
t-1	Refletividade em 2.5 km	atividade	chuvosa-seca	-0.283	-0.532	28.389	28.818	13.001	169.038	Não
t-1	Refletividade em 2.5 km	inatividade	seca-chuvosa	0.831	0.979	10.588	9.066	10.664	113.726	Não
t-1	Refletividade em 2.5 km	atividade	seca-chuvosa	-0.291	-0.643	27.831	28.632	14.152	200.292	Não
t-1	Refletividade em 3.0 km	inatividade	chuvosa	-0.095	-0.138	17.077	17.224	10.856	117.858	Não
t-1	Refletividade em 3.0 km	atividade	chuvosa	-0.283	-0.365	27.563	27.925	12.084	146.024	Não
t-1	Refletividade em 3.0 km	inatividade	seca	0.885	2.170	12.031	11.364	9.169	84.073	Não
t-1	Refletividade em 3.0 km	atividade	seca	-0.161	-0.607	30.223	30.112	12.870	165.635	Não
t-1	Refletividade em 3.0 km	inatividade	chuvosa-seca	0.241	0.009	14.802	13.932	11.383	129.578	Não
t-1	Refletividade em 3.0 km	atividade	chuvosa-seca	-0.251	-0.474	28.672	28.864	12.596	158.657	Não
t-1	Refletividade em 3.0 km	inatividade	seca-chuvosa	0.632	0.609	11.668	10.415	11.033	121.717	Não
t-1	Refletividade em 3.0 km	atividade	seca-chuvosa	-0.279	-0.576	28.233	28.796	13.791	190.178	Não
t-1	Refletividade em 3.5 km	inatividade	chuvosa	-0.175	0.078	17.753	17.835	10.513	110.529	Não
t-1	Refletividade em 3.5 km	atividade	chuvosa	-0.247	-0.339	27.385	27.535	11.835	140.075	Não
t-1	Refletividade em 3.5 km	inatividade	seca	0.649	1.645	12.914	12.511	9.624	92.613	Não
t-1	Refletividade em 3.5 km	atividade	seca	-0.107	-0.591	30.315	29.887	12.599	158.736	Não
t-1	Refletividade em 3.5 km	inatividade	chuvosa-seca	0.043	0.147	16.081	15.517	11.267	126.944	Não
t-1	Refletividade em 3.5 km	atividade	chuvosa-seca	-0.200	-0.471	28.617	28.556	12.288	150.987	Não
t-1	Refletividade em 3.5 km	inatividade	seca-chuvosa	0.458	0.432	12.580	11.530	11.217	125.814	Não
t-1	Refletividade em 3.5 km	atividade	seca-chuvosa	-0.247	-0.563	28.359	28.679	13.551	183.624	Não
t-1	Refletividade em 4.0 km	inatividade	chuvosa	-0.222	0.302	18.432	18.364	10.087	101.739	Não
t-1	Refletividade em 4.0 km	atividade	chuvosa	-0.196	-0.354	27.159	27.096	11.508	132.432	Não
t-1	Refletividade em 4.0 km	inatividade	seca	0.467	1.386	14.039	13.636	9.907	98.151	Não
t-1	Refletividade em 4.0 km	atividade	seca	-0.052	-0.604	30.283	29.679	12.369	152.994	Não
t-1	Refletividade em 4.0 km	inatividade	chuvosa-seca	-0.073	0.329	17.162	16.608	10.950	119.903	Não
t-1	Refletividade em 4.0 km	atividade	chuvosa-seca	-0.152	-0.486	28.407	28.213	12.006	144.132	Não
t-1	Refletividade em 4.0 km	inatividade	seca-chuvosa	0.304	0.382	13.698	12.605	11.226	126.013	Não
t-1	Refletividade em 4.0 km	atividade	seca-chuvosa	-0.217	-0.560	28.368	28.561	13.328	177.641	Não
t-1	Refletividade em 4.5 km	inatividade	chuvosa	-0.191	0.434	18.893	18.637	9.482	89.917	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-1	Refletividade em 4.5 km	atividade	chuvosa	-0.119	-0.416	26.714	26.484	11.146	124.237	Não
t-1	Refletividade em 4.5 km	inatividade	seca	0.431	1.326	15.089	14.448	9.803	96.101	Não
t-1	Refletividade em 4.5 km	atividade	seca	0.021	-0.654	30.052	29.290	12.136	147.272	Não
t-1	Refletividade em 4.5 km	inatividade	chuvosa-seca	-0.078	0.417	17.923	17.279	10.418	108.530	Não
t-1	Refletividade em 4.5 km	atividade	chuvosa-seca	-0.077	-0.555	28.010	27.626	11.686	136.565	Não
t-1	Refletividade em 4.5 km	inatividade	seca-chuvosa	0.254	0.443	14.798	13.509	10.854	117.813	Não
t-1	Refletividade em 4.5 km	atividade	seca-chuvosa	-0.150	-0.632	28.172	28.194	13.052	170.352	Não
t-1	Refletividade em 5.0 km	inatividade	chuvosa	-0.123	0.492	18.876	18.534	8.804	77.508	Não
t-1	Refletividade em 5.0 km	atividade	chuvosa	-0.031	-0.467	25.962	25.521	10.797	116.570	Não
t-1	Refletividade em 5.0 km	inatividade	seca	0.432	1.322	15.824	14.982	9.548	91.157	Não
t-1	Refletividade em 5.0 km	atividade	seca	0.111	-0.683	29.579	28.568	11.876	141.035	Não
t-1	Refletividade em 5.0 km	inatividade	chuvosa-seca	-0.048	0.500	18.293	17.615	9.787	95.793	Não
t-1	Refletividade em 5.0 km	atividade	chuvosa-seca	0.010	-0.591	27.375	26.731	11.332	128.410	Não
t-1	Refletividade em 5.0 km	inatividade	seca-chuvosa	0.267	0.547	15.475	14.114	10.301	106.102	Não
t-1	Refletividade em 5.0 km	atividade	seca-chuvosa	-0.057	-0.721	27.728	27.492	12.731	162.067	Não
t-1	Refletividade em 5.5 km	inatividade	chuvosa	-0.052	0.491	18.451	18.151	8.192	67.109	Não
t-1	Refletividade em 5.5 km	atividade	chuvosa	0.062	-0.478	25.082	24.398	10.382	107.795	Não
t-1	Refletividade em 5.5 km	inatividade	seca	0.445	1.362	16.335	15.454	9.166	84.020	Não
t-1	Refletividade em 5.5 km	atividade	seca	0.192	-0.661	28.934	27.654	11.602	134.608	Não
t-1	Refletividade em 5.5 km	inatividade	chuvosa-seca	0.010	0.552	18.305	17.686	9.125	83.258	Não
t-1	Refletividade em 5.5 km	atividade	chuvosa-seca	0.099	-0.584	26.562	25.630	10.947	119.838	Não
t-1	Refletividade em 5.5 km	inatividade	seca-chuvosa	0.319	0.646	15.759	14.468	9.683	93.751	Não
t-1	Refletividade em 5.5 km	atividade	seca-chuvosa	0.029	-0.761	27.061	26.429	12.407	153.936	Não
t-1	Refletividade em 6.0 km	inatividade	chuvosa	0.007	0.500	17.883	17.664	7.644	58.431	Não
t-1	Refletividade em 6.0 km	atividade	chuvosa	0.139	-0.443	24.118	23.237	9.965	99.310	Não
t-1	Refletividade em 6.0 km	inatividade	seca	0.465	1.434	16.723	15.879	8.734	76.291	Não
t-1	Refletividade em 6.0 km	atividade	seca	0.267	-0.608	28.203	26.722	11.306	127.831	Não
t-1	Refletividade em 6.0 km	inatividade	chuvosa-seca	0.075	0.601	18.097	17.571	8.495	72.171	Não
t-1	Refletividade em 6.0 km	atividade	chuvosa-seca	0.182	-0.543	25.651	24.515	10.546	111.218	Não
t-1	Refletividade em 6.0 km	inatividade	seca-chuvosa	0.372	0.736	15.809	14.648	9.103	82.869	Não
t-1	Refletividade em 6.0 km	atividade	seca-chuvosa	0.103	-0.753	26.261	25.311	12.063	145.507	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-1	Refletividade em 6.5 km	inatividade	chuvosa	0.010	0.564	17.244	17.135	7.230	52.268	Não
t-1	Refletividade em 6.5 km	atividade	chuvosa	0.191	-0.373	23.150	22.145	9.564	91.478	Não
t-1	Refletividade em 6.5 km	inatividade	seca	0.426	1.441	16.870	16.175	8.441	71.258	Não
t-1	Refletividade em 6.5 km	atividade	seca	0.327	-0.527	27.439	25.818	10.998	120.962	Não
t-1	Refletividade em 6.5 km	inatividade	chuvosa-seca	0.082	0.682	17.689	17.321	8.040	64.639	Não
t-1	Refletividade em 6.5 km	atividade	chuvosa-seca	0.245	-0.462	24.718	23.466	10.144	102.892	Não
t-1	Refletividade em 6.5 km	inatividade	seca-chuvosa	0.368	0.803	15.629	14.680	8.675	75.259	Não
t-1	Refletividade em 6.5 km	atividade	seca-chuvosa	0.168	-0.710	25.412	24.207	11.689	136.643	Não
t-1	Refletividade em 7.0 km	inatividade	chuvosa	0.009	0.608	16.648	16.646	6.799	46.221	Não
t-1	Refletividade em 7.0 km	atividade	chuvosa	0.242	-0.297	22.165	21.091	9.115	83.083	Não
t-1	Refletividade em 7.0 km	inatividade	seca	0.426	1.559	16.995	16.475	7.966	63.460	Não
t-1	Refletividade em 7.0 km	atividade	seca	0.396	-0.427	26.627	24.921	10.622	112.832	Não
t-1	Refletividade em 7.0 km	inatividade	chuvosa-seca	0.111	0.747	17.230	17.015	7.520	56.556	Não
t-1	Refletividade em 7.0 km	atividade	chuvosa-seca	0.305	-0.363	23.776	22.443	9.687	93.834	Não
t-1	Refletividade em 7.0 km	inatividade	seca-chuvosa	0.386	0.886	15.404	14.649	8.171	66.758	Não
t-1	Refletividade em 7.0 km	atividade	seca-chuvosa	0.236	-0.653	24.552	23.145	11.249	126.536	Não
t-1	Máxima refletividade	inatividade	chuvosa	0.161	-0.338	16.226	15.990	10.850	117.727	Não
t-1	Máxima refletividade	atividade	chuvosa	-0.234	-0.478	27.538	27.663	12.430	154.509	Não
t-1	Máxima refletividade	inatividade	seca	1.173	3.205	12.037	10.971	8.264	68.298	Não
t-1	Máxima refletividade	atividade	seca	-0.069	-0.678	29.821	29.108	13.299	176.871	Não
t-1	Máxima refletividade	inatividade	chuvosa-seca	0.654	0.320	13.353	11.564	10.573	111.798	Não
t-1	Máxima refletividade	atividade	chuvosa-seca	-0.159	-0.622	28.372	28.172	12.915	166.795	Não
t-1	Máxima refletividade	inatividade	seca-chuvosa	0.952	1.388	11.124	9.559	9.900	98.009	Não
t-1	Máxima refletividade	atividade	seca-chuvosa	-0.179	-0.702	28.200	28.212	14.019	196.535	Não
t-1	Altura máxima de 35 dbz	inatividade	chuvosa	0.138	0.417	4.345	4.500	1.610	2.593	Não
t-1	Altura máxima de 35 dbz	atividade	chuvosa	0.420	0.406	5.331	5.000	2.304	5.307	Não
t-1	Altura máxima de 35 dbz	inatividade	seca	0.827	1.452	5.182	5.000	2.347	5.507	Não
t-1	Altura máxima de 35 dbz	atividade	seca	0.530	-0.018	7.026	6.500	3.332	11.104	Não
t-1	Altura máxima de 35 dbz	inatividade	chuvosa-seca	0.256	0.363	4.601	4.500	1.796	3.225	Não
t-1	Altura máxima de 35 dbz	atividade	chuvosa-seca	0.359	0.171	5.766	5.500	2.529	6.397	Não
t-1	Altura máxima de 35 dbz	inatividade	seca-chuvosa	0.552	1.068	5.109	5.000	2.086	4.351	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-1	Altura máxima de 35 dbz	atividade	seca-chuvosa	0.497	0.162	6.903	6.500	3.149	9.917	Não
t-1	Taxa de precipitação	inatividade	chuvosa	92.220	102995.380	1.437	0.405	3.618	13.089	Não
t-1	Taxa de precipitação	atividade	chuvosa	2.880	11.215	7.483	2.287	11.962	143.079	Não
t-1	Taxa de precipitação	inatividade	seca	163.864	234527.495	0.845	0.171	4.574	20.923	Não
t-1	Taxa de precipitação	atividade	seca	2.987	12.597	12.635	3.043	21.322	454.642	Não
t-1	Taxa de precipitação	inatividade	chuvosa-seca	11.985	1201.767	1.280	0.210	4.013	16.101	Não
t-1	Taxa de precipitação	atividade	chuvosa-seca	2.639	8.914	9.461	2.569	15.123	228.717	Não
t-1	Taxa de precipitação	inatividade	seca-chuvosa	12.348	345.655	1.008	0.139	4.205	17.681	Não
t-1	Taxa de precipitação	atividade	seca-chuvosa	3.138	17.042	10.459	2.500	17.787	316.364	Não
t-1	Vil	inatividade	chuvosa	11.108	2837.789	0.261	0.089	0.521	0.271	Não
t-1	Vil	atividade	chuvosa	2.701	9.237	1.161	0.401	1.774	3.145	Não
t-1	Vil	inatividade	seca	15.861	828.316	0.141	0.021	0.661	0.437	Não
t-1	Vil	atividade	seca	3.142	13.806	2.048	0.508	3.520	12.393	Não
t-1	Vil	inatividade	chuvosa-seca	7.422	109.841	0.216	0.035	0.587	0.344	Não
t-1	Vil	atividade	chuvosa-seca	2.705	9.504	1.430	0.430	2.260	5.108	Não
t-1	Vil	inatividade	seca-chuvosa	10.730	217.460	0.172	0.021	0.633	0.400	Não
t-1	Vil	atividade	seca-chuvosa	3.016	14.809	1.749	0.451	2.897	8.390	Não
t-1	Densidade do vil	inatividade	chuvosa	43.232	24233.605	0.026	0.011	0.068	0.005	Não
t-1	Densidade do vil	atividade	chuvosa	35.993	2104.832	0.054	0.023	0.177	0.031	Não
t-1	Densidade do vil	inatividade	seca	189.137	111479.600	0.014	0.008	0.063	0.004	Não
t-1	Densidade do vil	atividade	seca	66.967	7696.966	0.083	0.023	0.552	0.305	Não
t-1	Densidade do vil	inatividade	chuvosa-seca	46.910	9171.483	0.018	0.008	0.058	0.003	Não
t-1	Densidade do vil	atividade	chuvosa-seca	39.637	2982.808	0.064	0.022	0.295	0.087	Não
t-1	Densidade do vil	inatividade	seca-chuvosa	367.690	459357.281	0.015	0.006	0.070	0.005	Não
t-1	Densidade do vil	atividade	seca-chuvosa	18.951	566.530	0.064	0.019	0.262	0.069	Não
t-1	Altura de waldvogel	inatividade	chuvosa	0.219	-0.124	-1.155	-1.000	1.414	2.001	Não
t-1	Altura de waldvogel	atividade	chuvosa	0.128	-0.267	-0.541	-0.500	1.623	2.635	Não
t-1	Altura de waldvogel	inatividade	seca	0.495	0.625	-0.014	0.000	1.942	3.772	Não
t-1	Altura de waldvogel	atividade	seca	0.534	0.573	0.838	0.500	2.354	5.539	Não
t-1	Altura de waldvogel	inatividade	chuvosa-seca	0.248	-0.044	-0.828	-1.000	1.524	2.322	Não
t-1	Altura de waldvogel	atividade	chuvosa-seca	0.196	-0.058	-0.169	0.000	1.815	3.293	Não

Continua na próxima página



Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-1	Altura de waldvogel	inatividade	seca-chuvosa	0.360	0.347	-0.251	-0.500	1.768	3.125	Não
t-1	Altura de waldvogel	atividade	seca-chuvosa	0.345	0.485	0.619	0.500	2.115	4.475	Não
t-1	Varição do topo da nuvem	inatividade	chuvosa	-0.016	0.592	4.344	4.300	3.599	12.951	Não
t-1	Varição do topo da nuvem	atividade	chuvosa	0.168	-0.387	7.551	7.111	4.747	22.535	Não
t-1	Varição do topo da nuvem	inatividade	seca	0.369	1.627	4.222	3.984	4.146	17.189	Não
t-1	Varição do topo da nuvem	atividade	seca	0.225	-0.605	10.029	9.488	5.514	30.406	Não
t-1	Varição do topo da nuvem	inatividade	chuvosa-seca	0.012	0.679	4.811	4.657	4.069	16.557	Não
t-1	Varição do topo da nuvem	atividade	chuvosa-seca	0.163	-0.460	8.739	8.279	5.012	25.119	Não
t-1	Varição do topo da nuvem	inatividade	seca-chuvosa	0.352	0.899	3.488	3.077	4.296	18.456	Não
t-1	Varição do topo da nuvem	atividade	seca-chuvosa	0.106	-0.742	8.738	8.353	5.846	34.180	Não
t-2	Refletividade em 2.0 km	inatividade	chuvosa	0.176	-0.476	15.190	14.903	11.309	127.897	Não
t-2	Refletividade em 2.0 km	atividade	chuvosa	-0.230	-0.573	25.232	25.650	13.005	169.130	Não
t-2	Refletividade em 2.0 km	inatividade	seca	1.260	3.454	10.995	9.915	8.312	69.087	Não
t-2	Refletividade em 2.0 km	atividade	seca	-0.095	-0.771	27.029	26.800	13.797	190.345	Não
t-2	Refletividade em 2.0 km	inatividade	chuvosa-seca	0.746	0.391	12.193	10.013	10.774	116.070	Não
t-2	Refletividade em 2.0 km	atividade	chuvosa-seca	-0.149	-0.719	25.462	25.493	13.565	184.001	Não
t-2	Refletividade em 2.0 km	inatividade	seca-chuvosa	1.041	1.515	9.676	7.941	10.175	103.522	Não
t-2	Refletividade em 2.0 km	atividade	seca-chuvosa	-0.126	-0.829	24.782	24.814	14.801	219.062	Não
t-2	Refletividade em 2.5 km	inatividade	chuvosa	0.020	-0.337	16.224	16.322	11.153	124.386	Não
t-2	Refletividade em 2.5 km	atividade	chuvosa	-0.237	-0.429	25.771	26.022	12.560	157.744	Não
t-2	Refletividade em 2.5 km	inatividade	seca	1.090	2.793	11.331	10.416	8.759	76.728	Não
t-2	Refletividade em 2.5 km	atividade	seca	-0.111	-0.629	27.904	27.466	13.316	177.310	Não
t-2	Refletividade em 2.5 km	inatividade	chuvosa-seca	0.491	0.083	13.357	11.836	11.218	125.845	Não
t-2	Refletividade em 2.5 km	atividade	chuvosa-seca	-0.179	-0.532	26.295	26.095	13.035	169.899	Não
t-2	Refletividade em 2.5 km	inatividade	seca-chuvosa	0.834	0.984	10.543	9.010	10.691	114.289	Não
t-2	Refletividade em 2.5 km	atividade	seca-chuvosa	-0.174	-0.662	25.763	25.701	14.317	204.976	Não
t-2	Refletividade em 3.0 km	inatividade	chuvosa	-0.095	-0.139	17.097	17.244	10.857	117.884	Não
t-2	Refletividade em 3.0 km	atividade	chuvosa	-0.214	-0.336	25.968	25.931	12.149	147.586	Não
t-2	Refletividade em 3.0 km	inatividade	seca	0.895	2.179	12.026	11.340	9.205	84.735	Não
t-2	Refletividade em 3.0 km	atividade	seca	-0.081	-0.539	28.468	27.761	12.862	165.441	Não
t-2	Refletividade em 3.0 km	inatividade	chuvosa-seca	0.243	0.011	14.787	13.910	11.393	129.808	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-2	Refletividade em 3.0 km	atividade	chuvosa-seca	-0.160	-0.426	26.723	26.307	12.570	158.003	Não
t-2	Refletividade em 3.0 km	inatividade	seca-chuvosa	0.639	0.622	11.622	10.354	11.046	122.010	Não
t-2	Refletividade em 3.0 km	atividade	seca-chuvosa	-0.173	-0.564	26.351	26.096	13.889	192.916	Não
t-2	Refletividade em 3.5 km	inatividade	chuvosa	-0.174	0.077	17.773	17.856	10.515	110.565	Não
t-2	Refletividade em 3.5 km	atividade	chuvosa	-0.176	-0.301	25.842	25.571	11.824	139.800	Não
t-2	Refletividade em 3.5 km	inatividade	seca	0.659	1.651	12.912	12.499	9.660	93.324	Não
t-2	Refletividade em 3.5 km	atividade	seca	-0.030	-0.482	28.640	27.651	12.526	156.907	Não
t-2	Refletividade em 3.5 km	inatividade	chuvosa-seca	0.046	0.148	16.067	15.500	11.279	127.216	Não
t-2	Refletividade em 3.5 km	atividade	chuvosa-seca	-0.116	-0.368	26.753	26.068	12.190	148.592	Não
t-2	Refletividade em 3.5 km	inatividade	seca-chuvosa	0.466	0.450	12.525	11.471	11.222	125.930	Não
t-2	Refletividade em 3.5 km	atividade	seca-chuvosa	-0.146	-0.514	26.574	26.050	13.603	185.053	Não
t-2	Refletividade em 4.0 km	inatividade	chuvosa	-0.222	0.301	18.451	18.386	10.089	101.790	Não
t-2	Refletividade em 4.0 km	atividade	chuvosa	-0.134	-0.284	25.646	25.203	11.466	131.471	Não
t-2	Refletividade em 4.0 km	inatividade	seca	0.477	1.389	14.042	13.618	9.945	98.895	Não
t-2	Refletividade em 4.0 km	atividade	seca	0.027	-0.484	28.703	27.524	12.241	149.841	Não
t-2	Refletividade em 4.0 km	inatividade	chuvosa-seca	-0.071	0.329	17.150	16.592	10.963	120.184	Não
t-2	Refletividade em 4.0 km	atividade	chuvosa-seca	-0.065	-0.367	26.634	25.822	11.846	140.333	Não
t-2	Refletividade em 4.0 km	inatividade	seca-chuvosa	0.313	0.401	13.642	12.533	11.228	126.059	Não
t-2	Refletividade em 4.0 km	atividade	seca-chuvosa	-0.113	-0.512	26.710	26.088	13.327	177.597	Não
t-2	Refletividade em 4.5 km	inatividade	chuvosa	-0.191	0.433	18.912	18.660	9.485	89.961	Não
t-2	Refletividade em 4.5 km	atividade	chuvosa	-0.061	-0.323	25.225	24.641	11.075	122.664	Não
t-2	Refletividade em 4.5 km	inatividade	seca	0.441	1.328	15.096	14.434	9.841	96.838	Não
t-2	Refletividade em 4.5 km	atividade	seca	0.101	-0.522	28.558	27.241	11.961	143.077	Não
t-2	Refletividade em 4.5 km	inatividade	chuvosa-seca	-0.076	0.416	17.914	17.265	10.430	108.790	Não
t-2	Refletividade em 4.5 km	atividade	chuvosa-seca	0.015	-0.420	26.276	25.314	11.487	131.953	Não
t-2	Refletividade em 4.5 km	inatividade	seca-chuvosa	0.264	0.465	14.741	13.445	10.853	117.787	Não
t-2	Refletividade em 4.5 km	atividade	seca-chuvosa	-0.046	-0.567	26.593	25.811	13.031	169.817	Não
t-2	Refletividade em 5.0 km	inatividade	chuvosa	-0.123	0.492	18.895	18.555	8.805	77.533	Não
t-2	Refletividade em 5.0 km	atividade	chuvosa	0.029	-0.356	24.578	23.808	10.646	113.336	Não
t-2	Refletividade em 5.0 km	inatividade	seca	0.442	1.326	15.836	14.970	9.585	91.867	Não
t-2	Refletividade em 5.0 km	atividade	seca	0.188	-0.537	28.207	26.736	11.656	135.873	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-2	Refletividade em 5.0 km	inatividade	chuvosa-seca	-0.045	0.499	18.285	17.602	9.799	96.029	Não
t-2	Refletividade em 5.0 km	atividade	chuvosa-seca	0.096	-0.433	25.683	24.603	11.107	123.372	Não
t-2	Refletividade em 5.0 km	inatividade	seca-chuvosa	0.279	0.574	15.417	14.044	10.296	106.010	Não
t-2	Refletividade em 5.0 km	atividade	seca-chuvosa	0.036	-0.621	26.207	25.223	12.712	161.604	Não
t-2	Refletividade em 5.5 km	inatividade	chuvosa	-0.051	0.492	18.469	18.170	8.194	67.134	Não
t-2	Refletividade em 5.5 km	atividade	chuvosa	0.117	-0.348	23.778	22.844	10.175	103.521	Não
t-2	Refletividade em 5.5 km	inatividade	seca	0.456	1.370	16.351	15.446	9.202	84.684	Não
t-2	Refletividade em 5.5 km	atividade	seca	0.269	-0.512	27.661	26.014	11.348	128.767	Não
t-2	Refletividade em 5.5 km	inatividade	chuvosa-seca	0.014	0.553	18.299	17.674	9.137	83.479	Não
t-2	Refletividade em 5.5 km	atividade	chuvosa-seca	0.184	-0.417	24.925	23.655	10.684	114.153	Não
t-2	Refletividade em 5.5 km	inatividade	seca-chuvosa	0.333	0.676	15.696	14.394	9.680	93.695	Não
t-2	Refletividade em 5.5 km	atividade	seca-chuvosa	0.122	-0.635	25.605	24.378	12.358	152.721	Não
t-2	Refletividade em 6.0 km	inatividade	chuvosa	0.007	0.502	17.899	17.681	7.646	58.466	Não
t-2	Refletividade em 6.0 km	atividade	chuvosa	0.211	-0.334	22.902	21.838	9.673	93.569	Não
t-2	Refletividade em 6.0 km	inatividade	seca	0.478	1.446	16.742	15.874	8.770	76.920	Não
t-2	Refletividade em 6.0 km	atividade	seca	0.344	-0.452	27.017	25.224	11.021	121.455	Não
t-2	Refletividade em 6.0 km	inatividade	chuvosa-seca	0.079	0.602	18.092	17.560	8.508	72.383	Não
t-2	Refletividade em 6.0 km	atividade	chuvosa-seca	0.268	-0.365	24.073	22.707	10.239	104.843	Não
t-2	Refletividade em 6.0 km	inatividade	seca-chuvosa	0.387	0.768	15.744	14.569	9.103	82.865	Não
t-2	Refletividade em 6.0 km	atividade	seca-chuvosa	0.202	-0.619	24.877	23.392	11.978	143.462	Não
t-2	Refletividade em 6.5 km	inatividade	chuvosa	0.010	0.566	17.258	17.150	7.232	52.307	Não
t-2	Refletividade em 6.5 km	atividade	chuvosa	0.262	-0.238	21.982	20.878	9.228	85.156	Não
t-2	Refletividade em 6.5 km	inatividade	seca	0.441	1.457	16.892	16.176	8.477	71.862	Não
t-2	Refletividade em 6.5 km	atividade	seca	0.402	-0.361	26.306	24.504	10.707	114.646	Não
t-2	Refletividade em 6.5 km	inatividade	chuvosa-seca	0.086	0.685	17.684	17.309	8.053	64.846	Não
t-2	Refletividade em 6.5 km	atividade	chuvosa-seca	0.331	-0.271	23.191	21.790	9.808	96.204	Não
t-2	Refletividade em 6.5 km	inatividade	seca-chuvosa	0.385	0.836	15.565	14.601	8.678	75.314	Não
t-2	Refletividade em 6.5 km	atividade	seca-chuvosa	0.270	-0.572	24.076	22.404	11.583	134.158	Não
t-2	Refletividade em 7.0 km	inatividade	chuvosa	0.010	0.611	16.660	16.657	6.802	46.266	Não
t-2	Refletividade em 7.0 km	atividade	chuvosa	0.308	-0.141	21.051	19.991	8.744	76.451	Não
t-2	Refletividade em 7.0 km	inatividade	seca	0.442	1.580	17.019	16.477	8.002	64.025	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-2	Refletividade em 7.0 km	atividade	seca	0.466	-0.240	25.523	23.714	10.317	106.435	Não
t-2	Refletividade em 7.0 km	inatividade	chuvosa-seca	0.116	0.752	17.225	17.004	7.533	56.752	Não
t-2	Refletividade em 7.0 km	atividade	chuvosa-seca	0.379	-0.147	22.317	20.965	9.336	87.154	Não
t-2	Refletividade em 7.0 km	inatividade	seca-chuvosa	0.402	0.920	15.341	14.572	8.178	66.885	Não
t-2	Refletividade em 7.0 km	atividade	seca-chuvosa	0.341	-0.503	23.244	21.513	11.122	123.698	Não
t-2	Máxima refletividade	inatividade	chuvosa	0.161	-0.337	16.250	16.008	10.852	117.766	Não
t-2	Máxima refletividade	atividade	chuvosa	-0.135	-0.478	25.691	25.293	12.493	156.074	Não
t-2	Máxima refletividade	inatividade	seca	1.184	3.223	12.029	10.949	8.295	68.800	Não
t-2	Máxima refletividade	atividade	seca	0.016	-0.613	27.816	26.650	13.309	177.137	Não
t-2	Máxima refletividade	inatividade	chuvosa-seca	0.656	0.325	13.340	11.548	10.580	111.932	Não
t-2	Máxima refletividade	atividade	chuvosa-seca	-0.049	-0.579	26.119	25.263	12.903	166.483	Não
t-2	Máxima refletividade	inatividade	seca-chuvosa	0.953	1.393	11.085	9.518	9.922	98.441	Não
t-2	Máxima refletividade	atividade	seca-chuvosa	-0.057	-0.677	26.093	25.221	14.100	198.803	Não
t-2	Altura máxima de 35 dbz	inatividade	chuvosa	0.143	0.433	4.347	4.500	1.613	2.602	Não
t-2	Altura máxima de 35 dbz	atividade	chuvosa	0.451	0.696	5.087	5.000	2.190	4.798	Não
t-2	Altura máxima de 35 dbz	inatividade	seca	0.861	1.524	5.219	5.000	2.385	5.690	Não
t-2	Altura máxima de 35 dbz	atividade	seca	0.603	0.271	6.598	6.500	3.160	9.986	Não
t-2	Altura máxima de 35 dbz	inatividade	chuvosa-seca	0.271	0.412	4.610	4.500	1.802	3.248	Não
t-2	Altura máxima de 35 dbz	atividade	chuvosa-seca	0.362	0.177	5.447	5.500	2.386	5.694	Não
t-2	Altura máxima de 35 dbz	inatividade	seca-chuvosa	0.606	1.203	5.139	5.000	2.123	4.506	Não
t-2	Altura máxima de 35 dbz	atividade	seca-chuvosa	0.601	0.429	6.600	6.500	3.076	9.459	Não
t-2	Taxa de precipitação	inatividade	chuvosa	92.907	103290.889	1.442	0.406	3.637	13.230	Não
t-2	Taxa de precipitação	atividade	chuvosa	3.351	16.388	6.284	1.700	10.886	118.513	Não
t-2	Taxa de precipitação	inatividade	seca	158.874	223394.447	0.858	0.170	4.648	21.601	Não
t-2	Taxa de precipitação	atividade	seca	3.367	16.836	10.337	2.104	18.923	358.088	Não
t-2	Taxa de precipitação	inatividade	chuvosa-seca	12.006	1198.703	1.283	0.210	4.035	16.284	Não
t-2	Taxa de precipitação	atividade	chuvosa-seca	3.100	12.864	7.521	1.719	13.235	175.161	Não
t-2	Taxa de precipitação	inatividade	seca-chuvosa	12.066	322.091	1.012	0.138	4.231	17.904	Não
t-2	Taxa de precipitação	atividade	seca-chuvosa	3.349	15.027	8.707	1.622	16.324	266.485	Não
t-2	Vil	inatividade	chuvosa	11.180	2858.427	0.262	0.089	0.523	0.274	Não
t-2	Vil	atividade	chuvosa	3.126	13.217	0.951	0.294	1.568	2.458	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-2	Vil	inatividade	seca	15.795	793.645	0.143	0.021	0.676	0.456	Não
t-2	Vil	atividade	seca	3.425	15.849	1.661	0.366	3.083	9.502	Não
t-2	Vil	inatividade	chuvosa-seca	7.382	101.575	0.216	0.035	0.590	0.349	Não
t-2	Vil	atividade	chuvosa-seca	3.116	12.849	1.123	0.297	1.943	3.773	Não
t-2	Vil	inatividade	seca-chuvosa	10.468	190.108	0.172	0.022	0.638	0.406	Não
t-2	Vil	atividade	seca-chuvosa	3.292	14.776	1.456	0.305	2.661	7.081	Não
t-2	Densidade do vil	inatividade	chuvosa	43.431	24487.014	0.026	0.011	0.068	0.005	Não
t-2	Densidade do vil	atividade	chuvosa	46.412	3397.104	0.051	0.022	0.188	0.036	Não
t-2	Densidade do vil	inatividade	seca	218.815	149398.945	0.014	0.008	0.065	0.004	Não
t-2	Densidade do vil	atividade	seca	31.460	1425.542	0.076	0.021	0.423	0.179	Não
t-2	Densidade do vil	inatividade	chuvosa-seca	43.825	7891.758	0.018	0.008	0.058	0.003	Não
t-2	Densidade do vil	atividade	chuvosa-seca	45.890	3476.655	0.058	0.020	0.298	0.089	Não
t-2	Densidade do vil	inatividade	seca-chuvosa	374.429	465185.172	0.015	0.006	0.071	0.005	Não
t-2	Densidade do vil	atividade	seca-chuvosa	21.791	868.830	0.059	0.018	0.237	0.056	Não
t-2	Altura de waldvogel	inatividade	chuvosa	0.223	-0.116	-1.150	-1.000	1.417	2.009	Não
t-2	Altura de waldvogel	atividade	chuvosa	0.168	-0.271	-0.666	-0.500	1.581	2.498	Não
t-2	Altura de waldvogel	inatividade	seca	0.522	0.698	0.021	0.000	1.969	3.875	Não
t-2	Altura de waldvogel	atividade	seca	0.551	0.645	0.695	0.500	2.286	5.227	Não
t-2	Altura de waldvogel	inatividade	chuvosa-seca	0.260	-0.016	-0.813	-1.000	1.535	2.355	Não
t-2	Altura de waldvogel	atividade	chuvosa-seca	0.192	-0.091	-0.303	-0.500	1.743	3.039	Não
t-2	Altura de waldvogel	inatividade	seca-chuvosa	0.355	0.307	-0.227	-0.500	1.777	3.157	Não
t-2	Altura de waldvogel	atividade	seca-chuvosa	0.359	0.504	0.612	0.500	2.131	4.540	Não
t-2	Variação do topo da nuvem	inatividade	chuvosa	-0.015	0.594	4.349	4.305	3.600	12.959	Não
t-2	Variação do topo da nuvem	atividade	chuvosa	0.223	-0.274	6.998	6.529	4.589	21.055	Não
t-2	Variação do topo da nuvem	inatividade	seca	0.386	1.643	4.236	3.987	4.165	17.348	Não
t-2	Variação do topo da nuvem	atividade	seca	0.282	-0.489	9.516	8.838	5.442	29.619	Não
t-2	Variação do topo da nuvem	inatividade	chuvosa-seca	0.016	0.681	4.808	4.651	4.075	16.603	Não
t-2	Variação do topo da nuvem	atividade	chuvosa-seca	0.229	-0.306	8.081	7.527	4.909	24.102	Não
t-2	Variação do topo da nuvem	inatividade	seca-chuvosa	0.369	0.927	3.468	3.047	4.303	18.513	Não
t-2	Variação do topo da nuvem	atividade	seca-chuvosa	0.200	-0.654	8.109	7.493	5.854	34.266	Não
t-3	Refletividade em 2.0 km	inatividade	chuvosa	0.176	-0.476	15.209	14.922	11.310	127.919	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-3	Refletividade em 2.0 km	atividade	chuvosa	-0.123	-0.596	23.308	23.175	13.076	170.992	Não
t-3	Refletividade em 2.0 km	inatividade	seca	1.267	3.464	10.993	9.901	8.336	69.490	Não
t-3	Refletividade em 2.0 km	atividade	seca	0.013	-0.743	24.754	24.045	13.755	189.188	Não
t-3	Refletividade em 2.0 km	inatividade	chuvosa-seca	0.748	0.396	12.176	10.000	10.778	116.156	Não
t-3	Refletividade em 2.0 km	atividade	chuvosa-seca	-0.013	-0.726	22.930	22.423	13.518	182.748	Não
t-3	Refletividade em 2.0 km	inatividade	seca-chuvosa	1.040	1.498	9.641	7.882	10.224	104.539	Não
t-3	Refletividade em 2.0 km	atividade	seca-chuvosa	-0.022	-0.787	22.292	21.759	14.628	213.977	Não
t-3	Refletividade em 2.5 km	inatividade	chuvosa	0.020	-0.336	16.241	16.338	11.153	124.393	Não
t-3	Refletividade em 2.5 km	atividade	chuvosa	-0.162	-0.399	24.028	23.707	12.614	159.102	Não
t-3	Refletividade em 2.5 km	inatividade	seca	1.097	2.800	11.331	10.402	8.785	77.171	Não
t-3	Refletividade em 2.5 km	atividade	seca	-0.042	-0.549	25.948	25.109	13.241	175.323	Não
t-3	Refletividade em 2.5 km	inatividade	chuvosa-seca	0.493	0.086	13.341	11.816	11.224	125.968	Não
t-3	Refletividade em 2.5 km	atividade	chuvosa-seca	-0.091	-0.492	24.017	23.430	13.014	169.370	Não
t-3	Refletividade em 2.5 km	inatividade	seca-chuvosa	0.834	0.977	10.523	8.982	10.730	115.123	Não
t-3	Refletividade em 2.5 km	atividade	seca-chuvosa	-0.103	-0.585	23.520	22.913	14.145	200.074	Não
t-3	Refletividade em 3.0 km	inatividade	chuvosa	-0.094	-0.139	17.112	17.260	10.858	117.892	Não
t-3	Refletividade em 3.0 km	atividade	chuvosa	-0.158	-0.251	24.392	23.885	12.155	147.734	Não
t-3	Refletividade em 3.0 km	inatividade	seca	0.902	2.184	12.029	11.330	9.231	85.216	Não
t-3	Refletividade em 3.0 km	atividade	seca	-0.032	-0.410	26.682	25.600	12.756	162.703	Não
t-3	Refletividade em 3.0 km	inatividade	chuvosa-seca	0.244	0.012	14.772	13.892	11.401	129.972	Não
t-3	Refletividade em 3.0 km	atividade	chuvosa-seca	-0.106	-0.307	24.656	23.898	12.521	156.774	Não
t-3	Refletividade em 3.0 km	inatividade	seca-chuvosa	0.642	0.625	11.603	10.317	11.070	122.552	Não
t-3	Refletividade em 3.0 km	atividade	seca-chuvosa	-0.110	-0.454	24.170	23.403	13.727	188.421	Não
t-3	Refletividade em 3.5 km	inatividade	chuvosa	-0.174	0.077	17.788	17.873	10.516	110.579	Não
t-3	Refletividade em 3.5 km	atividade	chuvosa	-0.115	-0.184	24.392	23.669	11.735	137.717	Não
t-3	Refletividade em 3.5 km	inatividade	seca	0.666	1.654	12.919	12.494	9.687	93.830	Não
t-3	Refletividade em 3.5 km	atividade	seca	0.009	-0.313	26.980	25.659	12.373	153.088	Não
t-3	Refletividade em 3.5 km	inatividade	chuvosa-seca	0.047	0.148	16.053	15.488	11.288	127.410	Não
t-3	Refletividade em 3.5 km	atividade	chuvosa-seca	-0.073	-0.193	24.873	23.901	12.063	145.507	Não
t-3	Refletividade em 3.5 km	inatividade	seca-chuvosa	0.472	0.459	12.500	11.430	11.236	126.240	Não
t-3	Refletividade em 3.5 km	atividade	seca-chuvosa	-0.085	-0.373	24.465	23.480	13.409	179.809	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-3	Refletividade em 4.0 km	inatividade	chuvosa	-0.221	0.300	18.466	18.403	10.091	101.819	Não
t-3	Refletividade em 4.0 km	atividade	chuvosa	-0.082	-0.125	24.271	23.426	11.343	128.657	Não
t-3	Refletividade em 4.0 km	inatividade	seca	0.483	1.390	14.052	13.613	9.970	99.404	Não
t-3	Refletividade em 4.0 km	atividade	seca	0.067	-0.287	27.166	25.711	12.025	144.594	Não
t-3	Refletividade em 4.0 km	inatividade	chuvosa-seca	-0.070	0.328	17.137	16.578	10.972	120.388	Não
t-3	Refletividade em 4.0 km	atividade	chuvosa-seca	-0.021	-0.161	24.874	23.793	11.646	135.639	Não
t-3	Refletividade em 4.0 km	inatividade	seca-chuvosa	0.319	0.414	13.614	12.493	11.236	126.240	Não
t-3	Refletividade em 4.0 km	atividade	seca-chuvosa	-0.045	-0.355	24.668	23.584	13.103	171.692	Não
t-3	Refletividade em 4.5 km	inatividade	chuvosa	-0.190	0.433	18.927	18.677	9.486	89.984	Não
t-3	Refletividade em 4.5 km	atividade	chuvosa	-0.006	-0.162	23.932	23.034	10.913	119.090	Não
t-3	Refletividade em 4.5 km	inatividade	seca	0.448	1.330	15.109	14.430	9.866	97.329	Não
t-3	Refletividade em 4.5 km	atividade	seca	0.150	-0.317	27.126	25.568	11.692	136.698	Não
t-3	Refletividade em 4.5 km	inatividade	chuvosa-seca	-0.074	0.416	17.902	17.252	10.439	108.973	Não
t-3	Refletividade em 4.5 km	atividade	chuvosa-seca	0.059	-0.201	24.645	23.471	11.238	126.284	Não
t-3	Refletividade em 4.5 km	inatividade	seca-chuvosa	0.271	0.483	14.712	13.402	10.856	117.859	Não
t-3	Refletividade em 4.5 km	atividade	seca-chuvosa	0.034	-0.399	24.682	23.453	12.745	162.446	Não
t-3	Refletividade em 5.0 km	inatividade	chuvosa	-0.122	0.492	18.910	18.572	8.806	77.540	Não
t-3	Refletividade em 5.0 km	atividade	chuvosa	0.087	-0.188	23.380	22.350	10.437	108.938	Não
t-3	Refletividade em 5.0 km	inatividade	seca	0.449	1.329	15.851	14.968	9.609	92.333	Não
t-3	Refletividade em 5.0 km	atividade	seca	0.245	-0.328	26.864	25.194	11.345	128.710	Não
t-3	Refletividade em 5.0 km	inatividade	chuvosa-seca	-0.043	0.500	18.274	17.590	9.808	96.201	Não
t-3	Refletividade em 5.0 km	atividade	chuvosa-seca	0.143	-0.209	24.130	22.828	10.814	116.934	Não
t-3	Refletividade em 5.0 km	inatividade	seca-chuvosa	0.289	0.596	15.385	14.002	10.296	106.006	Não
t-3	Refletividade em 5.0 km	atividade	seca-chuvosa	0.120	-0.430	24.407	22.982	12.386	153.413	Não
t-3	Refletividade em 5.5 km	inatividade	chuvosa	-0.051	0.492	18.482	18.184	8.194	67.144	Não
t-3	Refletividade em 5.5 km	atividade	chuvosa	0.182	-0.194	22.662	21.599	9.929	98.576	Não
t-3	Refletividade em 5.5 km	inatividade	seca	0.464	1.374	16.369	15.446	9.226	85.122	Não
t-3	Refletividade em 5.5 km	atividade	seca	0.331	-0.291	26.412	24.635	11.002	121.039	Não
t-3	Refletividade em 5.5 km	inatividade	chuvosa-seca	0.016	0.554	18.288	17.661	9.145	83.636	Não
t-3	Refletividade em 5.5 km	atividade	chuvosa-seca	0.230	-0.191	23.483	22.140	10.344	106.998	Não
t-3	Refletividade em 5.5 km	inatividade	seca-chuvosa	0.345	0.702	15.657	14.346	9.680	93.710	Não

Continua na próxima página

Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-3	Refletividade em 5.5 km	atividade	seca-chuvosa	0.217	-0.434	23.962	22.369	11.957	142.973	Não
t-3	Refletividade em 6.0 km	inatividade	chuvosa	0.008	0.502	17.910	17.693	7.647	58.484	Não
t-3	Refletividade em 6.0 km	atividade	chuvosa	0.266	-0.161	21.869	20.737	9.436	89.034	Não
t-3	Refletividade em 6.0 km	inatividade	seca	0.487	1.453	16.761	15.876	8.794	77.330	Não
t-3	Refletividade em 6.0 km	atividade	seca	0.410	-0.225	25.870	24.038	10.651	113.443	Não
t-3	Refletividade em 6.0 km	inatividade	chuvosa-seca	0.081	0.603	18.082	17.546	8.517	72.533	Não
t-3	Refletividade em 6.0 km	atividade	chuvosa-seca	0.310	-0.143	22.721	21.362	9.886	97.739	Não
t-3	Refletividade em 6.0 km	inatividade	seca-chuvosa	0.401	0.796	15.701	14.516	9.106	82.924	Não
t-3	Refletividade em 6.0 km	atividade	seca-chuvosa	0.302	-0.404	23.327	21.585	11.552	133.452	Não
t-3	Refletividade em 6.5 km	inatividade	chuvosa	0.011	0.568	17.267	17.160	7.234	52.331	Não
t-3	Refletividade em 6.5 km	atividade	chuvosa	0.305	-0.058	20.985	19.845	9.028	81.503	Não
t-3	Refletividade em 6.5 km	inatividade	seca	0.450	1.466	16.913	16.180	8.500	72.251	Não
t-3	Refletividade em 6.5 km	atividade	seca	0.463	-0.132	25.204	23.352	10.342	106.954	Não
t-3	Refletividade em 6.5 km	inatividade	chuvosa-seca	0.089	0.686	17.673	17.294	8.062	64.997	Não
t-3	Refletividade em 6.5 km	atividade	chuvosa-seca	0.356	-0.048	21.925	20.644	9.463	89.556	Não
t-3	Refletividade em 6.5 km	inatividade	seca-chuvosa	0.399	0.864	15.521	14.545	8.684	75.410	Não
t-3	Refletividade em 6.5 km	atividade	seca-chuvosa	0.369	-0.341	22.562	20.763	11.169	124.743	Não
t-3	Refletividade em 7.0 km	inatividade	chuvosa	0.011	0.612	16.668	16.666	6.804	46.292	Não
t-3	Refletividade em 7.0 km	atividade	chuvosa	0.347	0.032	20.110	19.066	8.569	73.431	Não
t-3	Refletividade em 7.0 km	inatividade	seca	0.453	1.592	17.039	16.482	8.024	64.388	Não
t-3	Refletividade em 7.0 km	atividade	seca	0.528	-0.008	24.488	22.695	9.947	98.933	Não
t-3	Refletividade em 7.0 km	inatividade	chuvosa-seca	0.119	0.753	17.215	16.992	7.543	56.898	Não
t-3	Refletividade em 7.0 km	atividade	chuvosa-seca	0.408	0.063	21.155	19.963	8.980	80.636	Não
t-3	Refletividade em 7.0 km	inatividade	seca-chuvosa	0.417	0.949	15.296	14.516	8.188	67.045	Não
t-3	Refletividade em 7.0 km	atividade	seca-chuvosa	0.436	-0.247	21.789	19.993	10.704	114.572	Não
t-3	Máxima refletividade	inatividade	chuvosa	0.161	-0.337	16.269	16.026	10.853	117.781	Não
t-3	Máxima refletividade	atividade	chuvosa	-0.055	-0.424	23.764	22.964	12.511	156.531	Não
t-3	Máxima refletividade	inatividade	seca	1.191	3.236	12.029	10.938	8.316	69.161	Não
t-3	Máxima refletividade	atividade	seca	0.081	-0.520	25.773	24.574	13.215	174.640	Não
t-3	Máxima refletividade	inatividade	chuvosa-seca	0.657	0.328	13.327	11.533	10.584	112.025	Não
t-3	Máxima refletividade	atividade	chuvosa-seca	0.042	-0.510	23.759	22.776	12.824	164.452	Não

Continua na próxima página



Tabela A.1: Continuação.

TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-3	Máxima refletividade	inatividade	seca-chuvosa	0.952	1.385	11.066	9.499	9.955	99.101	Não
t-3	Máxima refletividade	atividade	seca-chuvosa	0.031	-0.564	23.711	22.577	13.900	193.217	Não
t-3	Altura máxima de 35 dbz	inatividade	chuvosa	0.147	0.442	4.348	4.500	1.615	2.608	Não
t-3	Altura máxima de 35 dbz	atividade	chuvosa	0.504	0.831	4.967	5.000	2.171	4.713	Não
t-3	Altura máxima de 35 dbz	inatividade	seca	0.873	1.540	5.238	5.000	2.405	5.783	Não
t-3	Altura máxima de 35 dbz	atividade	seca	0.649	0.463	6.201	6.000	2.966	8.799	Não
t-3	Altura máxima de 35 dbz	inatividade	chuvosa-seca	0.277	0.426	4.615	4.500	1.806	3.261	Não
t-3	Altura máxima de 35 dbz	atividade	chuvosa-seca	0.404	0.271	5.189	5.000	2.285	5.222	Não
t-3	Altura máxima de 35 dbz	inatividade	seca-chuvosa	0.651	1.345	5.161	5.000	2.150	4.624	Não
t-3	Altura máxima de 35 dbz	atividade	seca-chuvosa	0.630	0.552	6.272	6.000	2.950	8.703	Não
t-3	Taxa de precipitação	inatividade	chuvosa	93.527	103788.820	1.446	0.407	3.653	13.342	Não
t-3	Taxa de precipitação	atividade	chuvosa	3.578	17.748	5.187	1.208	9.649	93.094	Não
t-3	Taxa de precipitação	inatividade	seca	155.363	215613.564	0.867	0.170	4.702	22.113	Não
t-3	Taxa de precipitação	atividade	seca	3.802	22.021	8.181	1.486	16.246	263.928	Não
t-3	Taxa de precipitação	inatividade	chuvosa-seca	12.064	1206.210	1.283	0.209	4.049	16.393	Não
t-3	Taxa de precipitação	atividade	chuvosa-seca	3.608	17.949	5.768	1.160	11.178	124.947	Não
t-3	Taxa de precipitação	inatividade	seca-chuvosa	11.714	278.876	1.021	0.137	4.269	18.228	Não
t-3	Taxa de precipitação	atividade	seca-chuvosa	3.951	22.272	6.568	1.075	13.488	181.923	Não
t-3	Vil	inatividade	chuvosa	11.243	2882.685	0.263	0.090	0.525	0.275	Não
t-3	Vil	atividade	chuvosa	3.399	15.471	0.778	0.218	1.381	1.907	Não
t-3	Vil	inatividade	seca	15.673	768.307	0.145	0.021	0.685	0.469	Não
t-3	Vil	atividade	seca	3.967	22.611	1.322	0.279	2.659	7.072	Não
t-3	Vil	inatividade	chuvosa-seca	7.418	102.529	0.216	0.035	0.593	0.351	Não
t-3	Vil	atividade	chuvosa-seca	3.569	16.813	0.861	0.212	1.625	2.642	Não
t-3	Vil	inatividade	seca-chuvosa	10.319	169.040	0.174	0.022	0.645	0.415	Não
t-3	Vil	atividade	seca-chuvosa	3.773	19.016	1.109	0.213	2.230	4.972	Não
t-3	Densidade do vil	inatividade	chuvosa	44.134	24760.070	0.026	0.011	0.068	0.005	Não
t-3	Densidade do vil	atividade	chuvosa	19.878	698.430	0.045	0.020	0.124	0.015	Não
t-3	Densidade do vil	inatividade	seca	194.728	114212.574	0.014	0.008	0.065	0.004	Não
t-3	Densidade do vil	atividade	seca	88.256	12887.988	0.069	0.020	0.485	0.235	Não

Continua na próxima página

Tabela A.1: Conclusão.

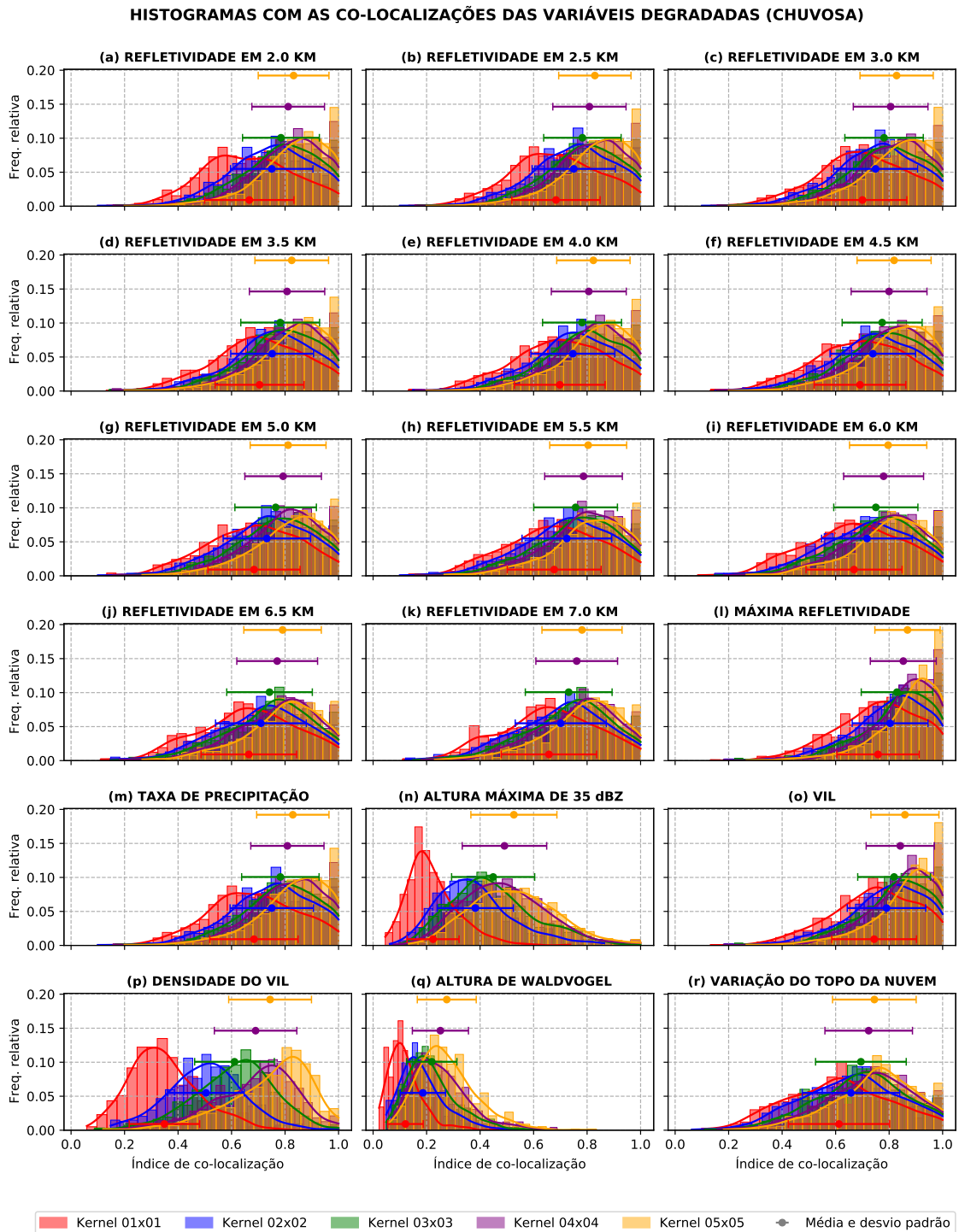
TEMPO	VARIÁVEL	CONDIÇÃO	ESTAÇÃO	ASSIMETRIA	CURTOSE	MÉDIA	MEDIANA	DES. PADRÃO	VARIÂNCIA	H <sub>0</sub>
t-3	Densidade do vil	inatividade	chuvosa-seca	44.379	8029.280	0.018	0.008	0.058	0.003	Não
t-3	Densidade do vil	atividade	chuvosa-seca	33.319	1796.306	0.051	0.018	0.234	0.055	Não
t-3	Densidade do vil	inatividade	seca-chuvosa	381.933	480579.617	0.015	0.006	0.071	0.005	Não
t-3	Densidade do vil	atividade	seca-chuvosa	20.499	681.101	0.052	0.017	0.206	0.042	Não
t-3	Altura de waldvogel	inatividade	chuvosa	0.224	-0.115	-1.148	-1.000	1.418	2.012	Não
t-3	Altura de waldvogel	atividade	chuvosa	0.125	-0.336	-0.691	-0.500	1.558	2.429	Não
t-3	Altura de waldvogel	inatividade	seca	0.533	0.717	0.040	0.000	1.980	3.921	Não
t-3	Altura de waldvogel	atividade	seca	0.529	0.569	0.503	0.500	2.260	5.108	Não
t-3	Altura de waldvogel	inatividade	chuvosa-seca	0.262	-0.009	-0.804	-1.000	1.539	2.369	Não
t-3	Altura de waldvogel	atividade	chuvosa-seca	0.215	-0.189	-0.407	-0.500	1.727	2.982	Não
t-3	Altura de waldvogel	inatividade	seca-chuvosa	0.371	0.359	-0.196	-0.500	1.797	3.229	Não
t-3	Altura de waldvogel	atividade	seca-chuvosa	0.411	0.534	0.506	0.500	2.091	4.374	Não
t-3	Variação do topo da nuvem	inatividade	chuvosa	-0.014	0.595	4.351	4.307	3.601	12.966	Não
t-3	Variação do topo da nuvem	atividade	chuvosa	0.264	-0.119	6.492	5.974	4.523	20.458	Não
t-3	Variação do topo da nuvem	inatividade	seca	0.396	1.648	4.243	3.988	4.176	17.436	Não
t-3	Variação do topo da nuvem	atividade	seca	0.352	-0.302	8.943	8.140	5.320	28.298	Não
t-3	Variação do topo da nuvem	inatividade	chuvosa-seca	0.019	0.682	4.798	4.639	4.080	16.643	Não
t-3	Variação do topo da nuvem	atividade	chuvosa-seca	0.261	-0.090	7.448	6.909	4.776	22.806	Não
t-3	Variação do topo da nuvem	inatividade	seca-chuvosa	0.380	0.945	3.467	3.039	4.310	18.574	Não
t-3	Variação do topo da nuvem	atividade	seca-chuvosa	0.305	-0.455	7.340	6.511	5.692	32.404	Não

## APÊNDICE B - ESTATÍSTICAS ÍNDICE DE CO-LOCALIZAÇÃO

As Figuras B.1, B.2, B.3 e B.4, mostram as frequências relativa das médias do índice de co-localização e seus respectivos desvios-padrão para todas as variáveis com diferentes tamanhos de *kernels* de degradação, respectivamente, na estação chuvosa, chuvosa-seca, seca e seca-chuvosa. Essas figuras complementam as análises discutidas na Seção 4.3 do Capítulo 4.

A Tabela B.1 apresenta algumas informações estatísticas dos valores do índice de co-localização, conforme discutido na Seção 4.3 do Capítulo 4. Essas informações estão agrupadas por estação, variável e tamanho do *kernel* usado na degradação. São elas: média, desvio padrão, mediana, percentis (de 33, 66 e 99), assimetria e curtose.

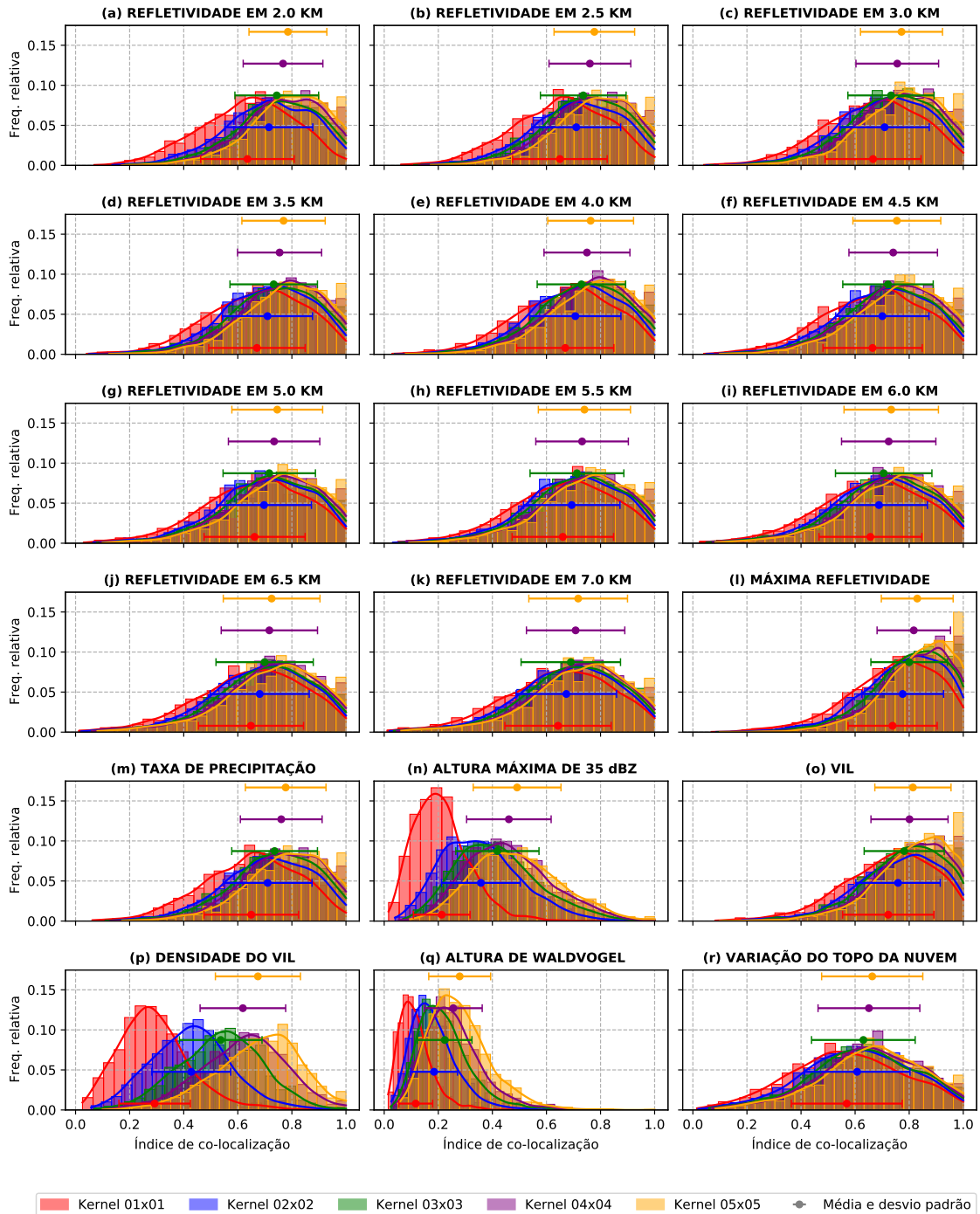
Figura B.1 - Histogramas da estação chuvosa.



Fonte: Produção Autor.

Figura B.2 - Histogramas da transição entre a estação chuvosa e seca.

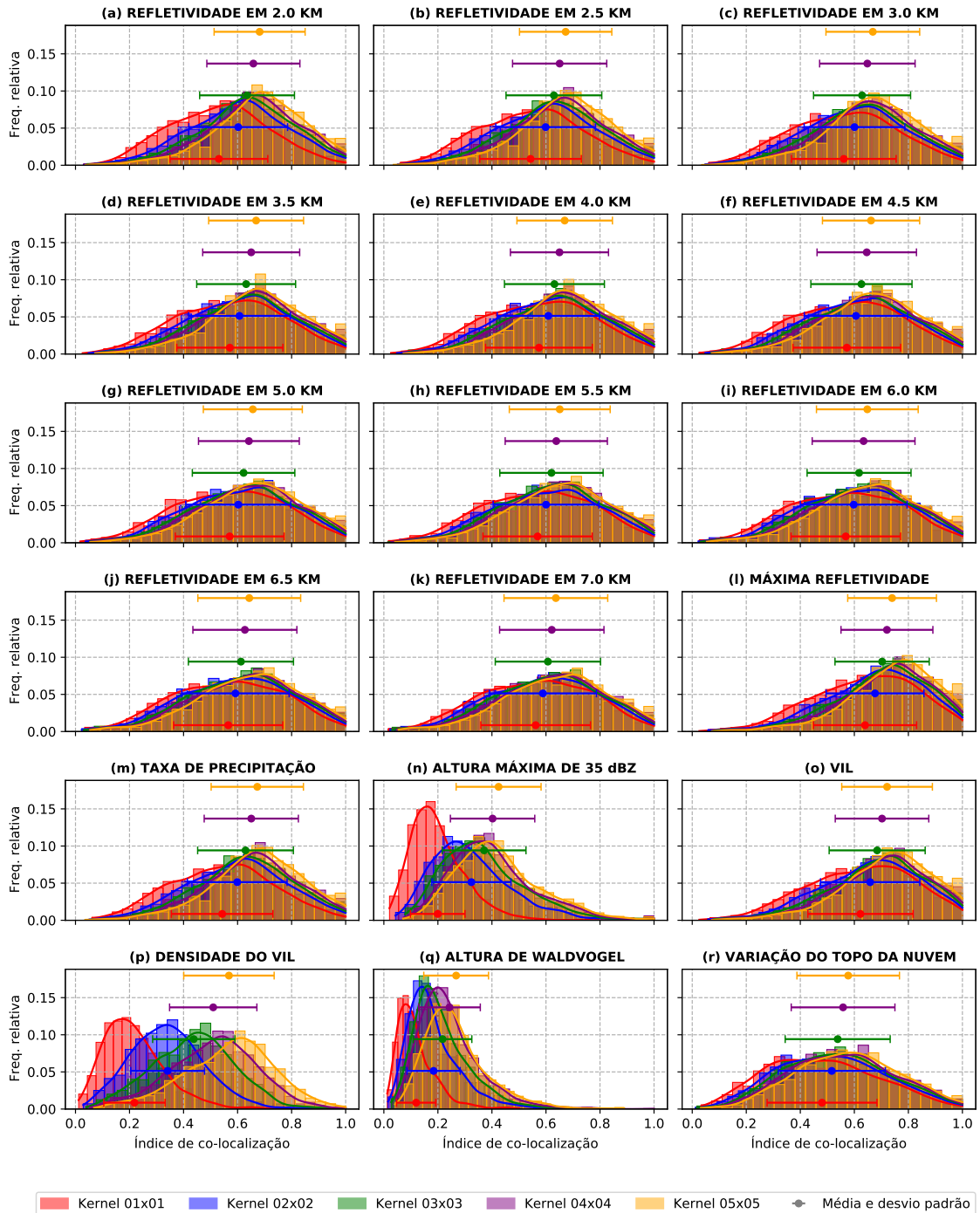
**HISTOGRAMAS COM AS CO-LOCALIZAÇÕES DAS VARIÁVEIS DEGRADADAS (CHUVOSA-SECA)**



Fonte: Produção Autor.

Figura B.3 - Histogramas da estação seca.

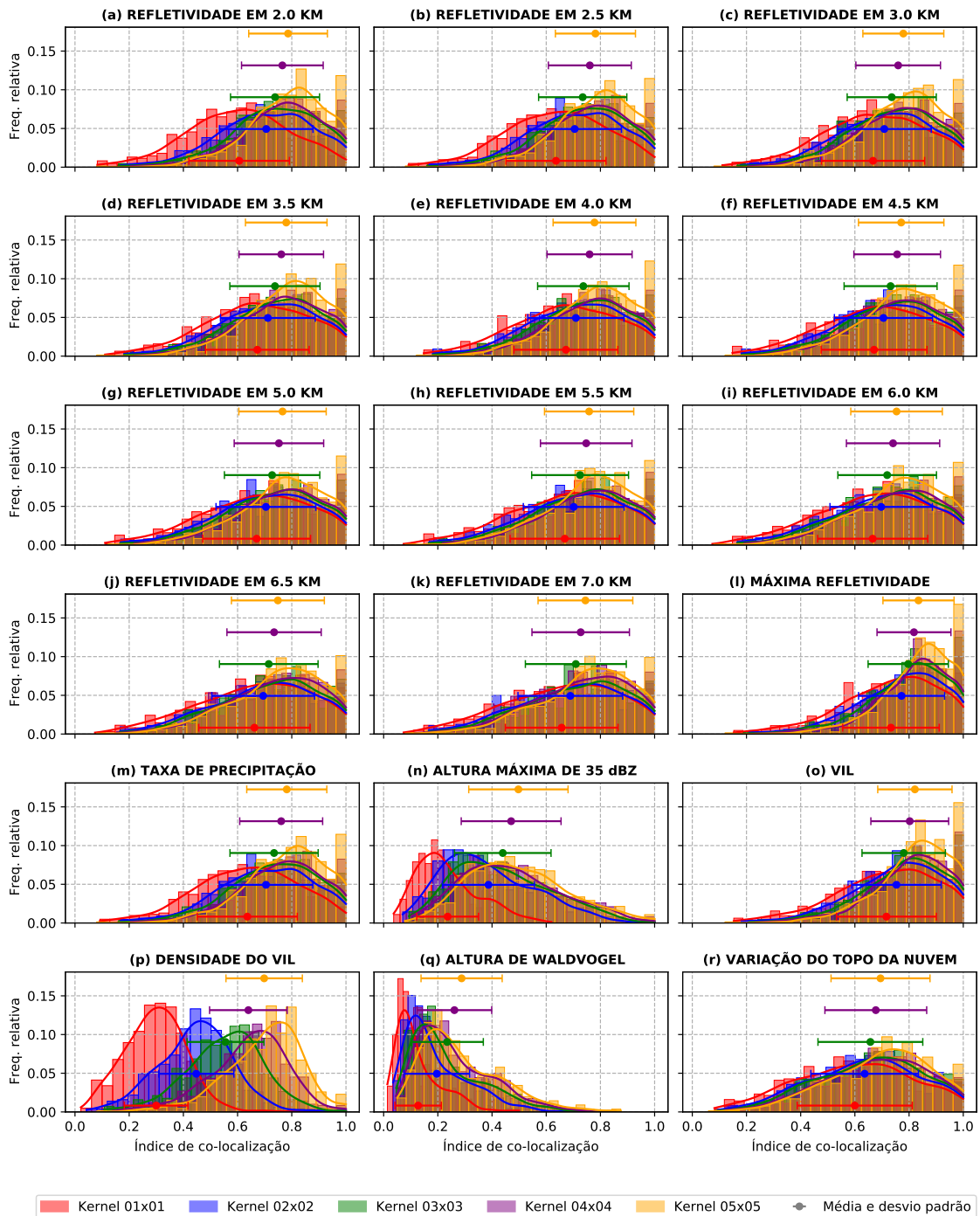
**HISTOGRAMAS COM AS CO-LOCALIZAÇÕES DAS VARIÁVEIS DEGRADADAS (SECA)**



Fonte: Produção Autor.

Figura B.4 - Histogramas da transição entre a estação seca e chuvosa.

**HISTOGRAMAS COM AS CO-LOCALIZAÇÕES DAS VARIÁVEIS DEGRADADAS (SECA-CHUVOSA)**



Fonte: Produção Autor.

Tabela B.1 - Estatísticas dos os índices de co-localização das variáveis em diferentes com degradações em diferentes *kernels*.

ESTACÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA		PERCENTIL		ASSIMETRIA	CURTOSE
					P33	P66	P99			
chuvosa	Densidade do vil	1×1	0.349	0.132	0.333	0.282	0.382	0.692	0.770	1.276
chuvosa	Densidade do vil	2×2	0.504	0.143	0.505	0.442	0.559	0.858	0.185	0.489
chuvosa	Densidade do vil	3×3	0.611	0.149	0.625	0.556	0.676	0.915	-0.343	0.279
chuvosa	Densidade do vil	4×4	0.690	0.154	0.714	0.643	0.769	1.000	-0.599	0.237
chuvosa	Densidade do vil	5×5	0.744	0.155	0.778	0.700	0.829	1.000	-0.903	0.758
chuvosa	Varição do topo da nuvem	1×1	0.612	0.190	0.612	0.526	0.684	1.000	0.014	-0.551
chuvosa	Varição do topo da nuvem	2×2	0.657	0.182	0.667	0.578	0.731	1.000	-0.191	-0.358
chuvosa	Varição do topo da nuvem	3×3	0.694	0.170	0.708	0.637	0.768	1.000	-0.343	-0.116
chuvosa	Varição do topo da nuvem	4×4	0.724	0.164	0.739	0.667	0.800	1.000	-0.489	0.033
chuvosa	Varição do topo da nuvem	5×5	0.745	0.156	0.759	0.700	0.816	1.000	-0.568	0.159
chuvosa	Altura máxima de 35 dbz	1×1	0.224	0.097	0.203	0.173	0.240	0.514	1.232	2.594
chuvosa	Altura máxima de 35 dbz	2×2	0.382	0.142	0.364	0.309	0.417	0.800	0.810	0.667
chuvosa	Altura máxima de 35 dbz	3×3	0.449	0.155	0.429	0.374	0.500	0.852	0.570	0.037
chuvosa	Altura máxima de 35 dbz	4×4	0.491	0.158	0.483	0.413	0.545	0.934	0.479	0.276
chuvosa	Altura máxima de 35 dbz	5×5	0.526	0.161	0.520	0.442	0.588	0.955	0.341	-0.047
chuvosa	Máxima reffetividade	1×1	0.758	0.154	0.771	0.702	0.833	1.000	-0.484	-0.122
chuvosa	Máxima reffetividade	2×2	0.804	0.142	0.825	0.757	0.875	1.000	-0.886	0.955
chuvosa	Máxima reffetividade	3×3	0.829	0.134	0.854	0.800	0.900	1.000	-1.076	1.464
chuvosa	Máxima reffetividade	4×4	0.853	0.123	0.875	0.831	0.918	1.000	-1.320	2.679
chuvosa	Máxima reffetividade	5×5	0.869	0.122	0.900	0.846	0.935	1.000	-1.472	2.670
chuvosa	Taxa de precipitação	1×1	0.684	0.165	0.681	0.605	0.750	1.000	-0.096	-0.438
chuvosa	Taxa de precipitação	2×2	0.750	0.156	0.765	0.686	0.821	1.000	-0.517	0.213
chuvosa	Taxa de precipitação	3×3	0.782	0.145	0.795	0.732	0.857	1.000	-0.695	0.667
chuvosa	Taxa de precipitação	4×4	0.809	0.137	0.827	0.764	0.879	1.000	-0.874	1.204
chuvosa	Taxa de precipitação	5×5	0.829	0.135	0.853	0.794	0.903	1.000	-1.062	1.338
chuvosa	Reffetividade em 2.0 km	1×1	0.666	0.167	0.662	0.581	0.739	1.000	0.028	-0.540
chuvosa	Reffetividade em 2.0 km	2×2	0.750	0.154	0.760	0.686	0.822	1.000	-0.483	0.196

Continua na próxima página



Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
chuvosa	Refletividade em 2.0 km	3×3	0.785	0.144	0.794	0.734	0.857	1.000	-0.670	0.510
chuvosa	Refletividade em 2.0 km	4×4	0.812	0.136	0.833	0.768	0.882	1.000	-0.900	1.321
chuvosa	Refletividade em 2.0 km	5×5	0.832	0.132	0.856	0.797	0.903	1.000	-1.054	1.383
chuvosa	Refletividade em 2.5 km	1×1	0.684	0.165	0.681	0.605	0.750	1.000	-0.096	-0.438
chuvosa	Refletividade em 2.5 km	2×2	0.750	0.156	0.765	0.686	0.821	1.000	-0.517	0.213
chuvosa	Refletividade em 2.5 km	3×3	0.782	0.145	0.795	0.732	0.857	1.000	-0.695	0.667
chuvosa	Refletividade em 2.5 km	4×4	0.809	0.137	0.827	0.764	0.879	1.000	-0.874	1.204
chuvosa	Refletividade em 2.5 km	5×5	0.829	0.135	0.853	0.794	0.903	1.000	-1.062	1.338
chuvosa	Refletividade em 3.0 km	1×1	0.700	0.166	0.704	0.627	0.775	1.000	-0.226	-0.309
chuvosa	Refletividade em 3.0 km	2×2	0.749	0.157	0.765	0.688	0.818	1.000	-0.559	0.288
chuvosa	Refletividade em 3.0 km	3×3	0.781	0.147	0.793	0.729	0.853	1.000	-0.685	0.579
chuvosa	Refletividade em 3.0 km	4×4	0.806	0.140	0.825	0.757	0.879	1.000	-0.874	1.093
chuvosa	Refletividade em 3.0 km	5×5	0.828	0.137	0.853	0.789	0.903	1.000	-1.024	1.187
chuvosa	Refletividade em 3.5 km	1×1	0.705	0.166	0.708	0.636	0.787	1.000	-0.270	-0.318
chuvosa	Refletividade em 3.5 km	2×2	0.752	0.155	0.763	0.698	0.824	1.000	-0.577	0.298
chuvosa	Refletividade em 3.5 km	3×3	0.782	0.147	0.794	0.730	0.857	1.000	-0.736	0.723
chuvosa	Refletividade em 3.5 km	4×4	0.807	0.141	0.828	0.763	0.881	1.000	-0.993	1.547
chuvosa	Refletividade em 3.5 km	5×5	0.825	0.138	0.853	0.788	0.900	1.000	-1.054	1.304
chuvosa	Refletividade em 4.0 km	1×1	0.698	0.170	0.704	0.625	0.782	1.000	-0.259	-0.403
chuvosa	Refletividade em 4.0 km	2×2	0.747	0.157	0.757	0.692	0.819	1.000	-0.536	0.154
chuvosa	Refletividade em 4.0 km	3×3	0.781	0.148	0.794	0.728	0.853	1.000	-0.628	0.255
chuvosa	Refletividade em 4.0 km	4×4	0.806	0.140	0.827	0.760	0.875	1.000	-0.913	1.203
chuvosa	Refletividade em 4.0 km	5×5	0.824	0.137	0.846	0.789	0.895	1.000	-0.995	1.051
chuvosa	Refletividade em 4.5 km	1×1	0.691	0.171	0.702	0.613	0.775	1.000	-0.246	-0.420
chuvosa	Refletividade em 4.5 km	2×2	0.739	0.160	0.750	0.681	0.817	1.000	-0.502	0.042
chuvosa	Refletividade em 4.5 km	3×3	0.774	0.150	0.789	0.720	0.849	1.000	-0.615	0.197
chuvosa	Refletividade em 4.5 km	4×4	0.800	0.142	0.818	0.750	0.874	1.000	-0.846	1.009
chuvosa	Refletividade em 4.5 km	5×5	0.819	0.139	0.840	0.778	0.893	1.000	-0.952	0.914
chuvosa	Refletividade em 5.0 km	1×1	0.684	0.172	0.693	0.605	0.765	1.000	-0.245	-0.403
chuvosa	Refletividade em 5.0 km	2×2	0.732	0.163	0.742	0.672	0.814	1.000	-0.492	0.041

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
chuvosa	Refletividade em 5.0 km	3×3	0.765	0.152	0.778	0.714	0.841	1.000	-0.587	0.162
chuvosa	Refletividade em 5.0 km	4×4	0.793	0.143	0.810	0.750	0.865	1.000	-0.796	0.828
chuvosa	Refletividade em 5.0 km	5×5	0.811	0.142	0.833	0.769	0.886	1.000	-0.907	0.731
chuvosa	Refletividade em 5.5 km	1×1	0.677	0.176	0.684	0.604	0.759	1.000	-0.218	-0.491
chuvosa	Refletividade em 5.5 km	2×2	0.724	0.168	0.737	0.667	0.805	1.000	-0.502	0.025
chuvosa	Refletividade em 5.5 km	3×3	0.757	0.157	0.775	0.709	0.833	1.000	-0.636	0.225
chuvosa	Refletividade em 5.5 km	4×4	0.787	0.145	0.800	0.741	0.863	1.000	-0.741	0.607
chuvosa	Refletividade em 5.5 km	5×5	0.804	0.144	0.824	0.765	0.882	1.000	-0.879	0.709
chuvosa	Refletividade em 6.0 km	1×1	0.669	0.180	0.673	0.596	0.750	1.000	-0.202	-0.508
chuvosa	Refletividade em 6.0 km	2×2	0.717	0.170	0.733	0.650	0.800	1.000	-0.474	-0.042
chuvosa	Refletividade em 6.0 km	3×3	0.750	0.157	0.770	0.698	0.833	1.000	-0.562	0.023
chuvosa	Refletividade em 6.0 km	4×4	0.779	0.150	0.795	0.727	0.857	1.000	-0.715	0.454
chuvosa	Refletividade em 6.0 km	5×5	0.796	0.145	0.818	0.757	0.871	1.000	-0.824	0.515
chuvosa	Refletividade em 6.5 km	1×1	0.665	0.179	0.667	0.592	0.749	1.000	-0.189	-0.458
chuvosa	Refletividade em 6.5 km	2×2	0.710	0.170	0.726	0.647	0.791	1.000	-0.469	-0.035
chuvosa	Refletividade em 6.5 km	3×3	0.742	0.160	0.762	0.691	0.818	1.000	-0.573	0.071
chuvosa	Refletividade em 6.5 km	4×4	0.771	0.151	0.792	0.720	0.848	1.000	-0.694	0.414
chuvosa	Refletividade em 6.5 km	5×5	0.791	0.145	0.812	0.750	0.864	1.000	-0.789	0.422
chuvosa	Refletividade em 7.0 km	1×1	0.657	0.178	0.663	0.588	0.733	1.000	-0.169	-0.481
chuvosa	Refletividade em 7.0 km	2×2	0.702	0.170	0.722	0.636	0.775	1.000	-0.451	-0.043
chuvosa	Refletividade em 7.0 km	3×3	0.731	0.162	0.750	0.681	0.808	1.000	-0.588	0.165
chuvosa	Refletividade em 7.0 km	4×4	0.761	0.153	0.786	0.714	0.835	1.000	-0.664	0.369
chuvosa	Refletividade em 7.0 km	5×5	0.781	0.150	0.800	0.738	0.857	1.000	-0.833	0.605
chuvosa	Vil	1×1	0.744	0.159	0.750	0.684	0.821	1.000	-0.447	-0.196
chuvosa	Vil	2×2	0.790	0.147	0.810	0.740	0.867	1.000	-0.810	0.654
chuvosa	Vil	3×3	0.819	0.136	0.839	0.784	0.889	1.000	-0.952	1.136
chuvosa	Vil	4×4	0.842	0.128	0.866	0.809	0.909	1.000	-1.140	1.856
chuvosa	Vil	5×5	0.859	0.127	0.889	0.834	0.928	1.000	-1.364	2.190
chuvosa	Altura de waldvogel	1×1	0.121	0.066	0.108	0.088	0.132	0.352	1.649	4.398
chuvosa	Altura de waldvogel	2×2	0.186	0.085	0.169	0.144	0.203	0.452	1.064	1.387

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
chuvosa	Altura de waldvogel	3×3	0.219	0.094	0.204	0.169	0.240	0.494	0.986	1.127
chuvosa	Altura de waldvogel	4×4	0.252	0.105	0.237	0.192	0.278	0.560	1.019	1.421
chuvosa	Altura de waldvogel	5×5	0.275	0.110	0.258	0.221	0.306	0.600	0.939	1.677
chuvosa-seca	Densidade do vil	1×1	0.292	0.132	0.278	0.228	0.331	0.695	0.864	1.549
chuvosa-seca	Densidade do vil	2×2	0.428	0.146	0.425	0.363	0.485	0.804	0.242	0.126
chuvosa-seca	Densidade do vil	3×3	0.537	0.153	0.543	0.476	0.600	0.875	-0.069	-0.049
chuvosa-seca	Densidade do vil	4×4	0.619	0.158	0.629	0.557	0.689	1.000	-0.247	-0.024
chuvosa-seca	Densidade do vil	5×5	0.674	0.157	0.689	0.615	0.750	1.000	-0.423	0.038
chuvosa-seca	Varição do topo da nuvem	1×1	0.569	0.205	0.566	0.478	0.660	1.000	-0.071	-0.512
chuvosa-seca	Varição do topo da nuvem	2×2	0.608	0.195	0.611	0.523	0.696	1.000	-0.220	-0.361
chuvosa-seca	Varição do topo da nuvem	3×3	0.631	0.192	0.636	0.550	0.714	1.000	-0.275	-0.228
chuvosa-seca	Varição do topo da nuvem	4×4	0.651	0.188	0.667	0.579	0.739	1.000	-0.382	-0.124
chuvosa-seca	Varição do topo da nuvem	5×5	0.663	0.187	0.672	0.590	0.750	1.000	-0.397	-0.102
chuvosa-seca	Altura máxima de 35 dbz	1×1	0.213	0.105	0.200	0.160	0.240	0.537	1.219	3.915
chuvosa-seca	Altura máxima de 35 dbz	2×2	0.357	0.146	0.345	0.277	0.407	0.750	0.553	0.183
chuvosa-seca	Altura máxima de 35 dbz	3×3	0.419	0.154	0.403	0.333	0.467	0.823	0.550	0.126
chuvosa-seca	Altura máxima de 35 dbz	4×4	0.461	0.156	0.444	0.381	0.510	0.850	0.457	-0.006
chuvosa-seca	Altura máxima de 35 dbz	5×5	0.491	0.162	0.480	0.407	0.550	0.900	0.376	-0.099
chuvosa-seca	Máxima reftetividade	1×1	0.739	0.165	0.762	0.681	0.824	1.000	-0.697	0.311
chuvosa-seca	Máxima reftetividade	2×2	0.776	0.151	0.797	0.722	0.857	1.000	-0.786	0.587
chuvosa-seca	Máxima reftetividade	3×3	0.800	0.141	0.818	0.750	0.875	1.000	-0.820	0.690
chuvosa-seca	Máxima reftetividade	4×4	0.817	0.136	0.833	0.774	0.895	1.000	-0.938	1.088
chuvosa-seca	Máxima reftetividade	5×5	0.830	0.133	0.850	0.788	0.905	1.000	-1.000	1.268
chuvosa-seca	Taxa de precipitação	1×1	0.650	0.175	0.660	0.579	0.729	1.000	-0.338	-0.241
chuvosa-seca	Taxa de precipitação	2×2	0.710	0.164	0.721	0.647	0.793	1.000	-0.453	-0.101
chuvosa-seca	Taxa de precipitação	3×3	0.736	0.158	0.750	0.672	0.815	1.000	-0.554	0.175
chuvosa-seca	Taxa de precipitação	4×4	0.761	0.151	0.772	0.700	0.840	1.000	-0.587	0.205
chuvosa-seca	Taxa de precipitação	5×5	0.777	0.149	0.791	0.727	0.857	1.000	-0.689	0.408
chuvosa-seca	Reftetividade em 2.0 km	1×1	0.636	0.173	0.646	0.569	0.717	1.000	-0.277	-0.342
chuvosa-seca	Reftetividade em 2.0 km	2×2	0.715	0.162	0.724	0.650	0.798	1.000	-0.434	-0.154

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
chuvosa-seca	Refletividade em 2.0 km	3×3	0.744	0.155	0.755	0.685	0.822	1.000	-0.544	0.081
chuvosa-seca	Refletividade em 2.0 km	4×4	0.767	0.147	0.778	0.712	0.846	1.000	-0.557	0.083
chuvosa-seca	Refletividade em 2.0 km	5×5	0.785	0.144	0.800	0.733	0.859	1.000	-0.692	0.418
chuvosa-seca	Refletividade em 2.5 km	1×1	0.650	0.175	0.660	0.579	0.729	1.000	-0.338	-0.241
chuvosa-seca	Refletividade em 2.5 km	2×2	0.710	0.164	0.721	0.647	0.793	1.000	-0.453	-0.101
chuvosa-seca	Refletividade em 2.5 km	3×3	0.736	0.158	0.750	0.672	0.815	1.000	-0.554	0.175
chuvosa-seca	Refletividade em 2.5 km	4×4	0.761	0.151	0.772	0.700	0.840	1.000	-0.587	0.205
chuvosa-seca	Refletividade em 2.5 km	5×5	0.777	0.149	0.791	0.727	0.857	1.000	-0.689	0.408
chuvosa-seca	Refletividade em 3.0 km	1×1	0.666	0.177	0.678	0.591	0.750	1.000	-0.367	-0.232
chuvosa-seca	Refletividade em 3.0 km	2×2	0.709	0.165	0.721	0.639	0.795	1.000	-0.448	-0.072
chuvosa-seca	Refletividade em 3.0 km	3×3	0.733	0.159	0.745	0.667	0.812	1.000	-0.513	0.056
chuvosa-seca	Refletividade em 3.0 km	4×4	0.756	0.153	0.769	0.695	0.833	1.000	-0.565	0.205
chuvosa-seca	Refletividade em 3.0 km	5×5	0.772	0.152	0.786	0.714	0.850	1.000	-0.659	0.335
chuvosa-seca	Refletividade em 3.5 km	1×1	0.671	0.178	0.684	0.597	0.758	1.000	-0.389	-0.210
chuvosa-seca	Refletividade em 3.5 km	2×2	0.710	0.167	0.724	0.643	0.794	1.000	-0.479	0.034
chuvosa-seca	Refletividade em 3.5 km	3×3	0.733	0.162	0.747	0.667	0.814	1.000	-0.548	0.158
chuvosa-seca	Refletividade em 3.5 km	4×4	0.755	0.156	0.773	0.694	0.833	1.000	-0.607	0.260
chuvosa-seca	Refletividade em 3.5 km	5×5	0.769	0.154	0.786	0.714	0.848	1.000	-0.677	0.375
chuvosa-seca	Refletividade em 4.0 km	1×1	0.670	0.180	0.686	0.593	0.755	1.000	-0.393	-0.207
chuvosa-seca	Refletividade em 4.0 km	2×2	0.707	0.169	0.722	0.636	0.792	1.000	-0.483	0.028
chuvosa-seca	Refletividade em 4.0 km	3×3	0.729	0.164	0.741	0.667	0.811	1.000	-0.536	0.128
chuvosa-seca	Refletividade em 4.0 km	4×4	0.750	0.159	0.769	0.690	0.831	1.000	-0.619	0.302
chuvosa-seca	Refletividade em 4.0 km	5×5	0.763	0.159	0.783	0.709	0.846	1.000	-0.728	0.536
chuvosa-seca	Refletividade em 4.5 km	1×1	0.665	0.184	0.679	0.588	0.750	1.000	-0.395	-0.240
chuvosa-seca	Refletividade em 4.5 km	2×2	0.701	0.172	0.714	0.628	0.789	1.000	-0.493	-0.003
chuvosa-seca	Refletividade em 4.5 km	3×3	0.722	0.167	0.734	0.660	0.803	1.000	-0.553	0.150
chuvosa-seca	Refletividade em 4.5 km	4×4	0.741	0.164	0.755	0.684	0.824	1.000	-0.605	0.195
chuvosa-seca	Refletividade em 4.5 km	5×5	0.754	0.162	0.773	0.700	0.836	1.000	-0.679	0.323
chuvosa-seca	Refletividade em 5.0 km	1×1	0.662	0.187	0.677	0.583	0.750	1.000	-0.426	-0.158
chuvosa-seca	Refletividade em 5.0 km	2×2	0.697	0.175	0.711	0.624	0.786	1.000	-0.532	0.089

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			
chuvosa-seca	Refletividade em 5.0 km	3×3	0.716	0.171	0.727	0.653	0.800	1.000	-0.562	0.181
chuvosa-seca	Refletividade em 5.0 km	4×4	0.734	0.169	0.750	0.674	0.820	1.000	-0.667	0.377
chuvosa-seca	Refletividade em 5.0 km	5×5	0.746	0.168	0.765	0.688	0.830	1.000	-0.703	0.386
chuvosa-seca	Refletividade em 5.5 km	1×1	0.661	0.188	0.675	0.585	0.750	1.000	-0.420	-0.180
chuvosa-seca	Refletividade em 5.5 km	2×2	0.693	0.179	0.710	0.620	0.782	1.000	-0.568	0.194
chuvosa-seca	Refletividade em 5.5 km	3×3	0.713	0.173	0.727	0.645	0.796	1.000	-0.563	0.166
chuvosa-seca	Refletividade em 5.5 km	4×4	0.732	0.171	0.750	0.667	0.818	1.000	-0.697	0.466
chuvosa-seca	Refletividade em 5.5 km	5×5	0.740	0.170	0.760	0.679	0.824	1.000	-0.692	0.361
chuvosa-seca	Refletividade em 6.0 km	1×1	0.657	0.190	0.670	0.582	0.750	1.000	-0.429	-0.147
chuvosa-seca	Refletividade em 6.0 km	2×2	0.688	0.179	0.704	0.616	0.775	1.000	-0.531	0.121
chuvosa-seca	Refletividade em 6.0 km	3×3	0.706	0.178	0.723	0.640	0.792	1.000	-0.620	0.289
chuvosa-seca	Refletividade em 6.0 km	4×4	0.724	0.174	0.743	0.667	0.811	1.000	-0.716	0.495
chuvosa-seca	Refletividade em 6.0 km	5×5	0.734	0.175	0.756	0.667	0.821	1.000	-0.731	0.445
chuvosa-seca	Refletividade em 6.5 km	1×1	0.649	0.195	0.667	0.575	0.747	1.000	-0.449	-0.148
chuvosa-seca	Refletividade em 6.5 km	2×2	0.681	0.183	0.699	0.611	0.770	1.000	-0.576	0.227
chuvosa-seca	Refletividade em 6.5 km	3×3	0.699	0.180	0.714	0.633	0.786	1.000	-0.622	0.311
chuvosa-seca	Refletividade em 6.5 km	4×4	0.717	0.178	0.735	0.667	0.805	1.000	-0.746	0.587
chuvosa-seca	Refletividade em 6.5 km	5×5	0.725	0.179	0.750	0.667	0.815	1.000	-0.740	0.457
chuvosa-seca	Refletividade em 7.0 km	1×1	0.643	0.197	0.658	0.571	0.739	1.000	-0.444	-0.118
chuvosa-seca	Refletividade em 7.0 km	2×2	0.674	0.186	0.692	0.600	0.764	1.000	-0.574	0.229
chuvosa-seca	Refletividade em 7.0 km	3×3	0.690	0.184	0.706	0.625	0.783	1.000	-0.641	0.384
chuvosa-seca	Refletividade em 7.0 km	4×4	0.708	0.182	0.724	0.652	0.800	1.000	-0.753	0.590
chuvosa-seca	Refletividade em 7.0 km	5×5	0.717	0.183	0.743	0.658	0.809	1.000	-0.761	0.593
chuvosa-seca	Vil	1×1	0.722	0.169	0.742	0.667	0.808	1.000	-0.618	0.170
chuvosa-seca	Vil	2×2	0.759	0.156	0.778	0.699	0.840	1.000	-0.703	0.375
chuvosa-seca	Vil	3×3	0.781	0.148	0.800	0.730	0.857	1.000	-0.742	0.511
chuvosa-seca	Vil	4×4	0.801	0.143	0.821	0.750	0.879	1.000	-0.866	0.866
chuvosa-seca	Vil	5×5	0.814	0.141	0.833	0.769	0.889	1.000	-0.955	1.169
chuvosa-seca	Altura de waldvogel	1×1	0.117	0.063	0.105	0.082	0.129	0.326	1.392	3.217
chuvosa-seca	Altura de waldvogel	2×2	0.184	0.087	0.170	0.139	0.206	0.466	1.230	2.735

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
chuvosa-seca	Altura de waldvogel	3×3	0.224	0.100	0.208	0.170	0.250	0.571	1.211	2.436
chuvosa-seca	Altura de waldvogel	4×4	0.255	0.107	0.240	0.200	0.282	0.600	0.986	1.764
chuvosa-seca	Altura de waldvogel	5×5	0.279	0.114	0.264	0.219	0.310	0.624	1.013	2.092
seca	Densidade do vil	1×1	0.218	0.113	0.201	0.157	0.249	0.560	1.029	1.690
seca	Densidade do vil	2×2	0.341	0.137	0.336	0.275	0.391	0.689	0.444	0.522
seca	Densidade do vil	3×3	0.438	0.152	0.441	0.375	0.500	0.815	0.072	-0.013
seca	Densidade do vil	4×4	0.510	0.162	0.521	0.444	0.578	0.879	-0.138	-0.070
seca	Densidade do vil	5×5	0.568	0.167	0.584	0.508	0.643	0.931	-0.286	-0.061
seca	Varição do topo da nuvem	1×1	0.480	0.203	0.472	0.370	0.561	0.944	0.236	-0.600
seca	Varição do topo da nuvem	2×2	0.516	0.198	0.509	0.417	0.600	0.975	0.146	-0.541
seca	Varição do topo da nuvem	3×3	0.538	0.194	0.535	0.443	0.618	0.978	0.106	-0.498
seca	Varição do topo da nuvem	4×4	0.558	0.192	0.556	0.461	0.638	1.000	0.065	-0.521
seca	Varição do topo da nuvem	5×5	0.577	0.190	0.576	0.486	0.657	1.000	-0.003	-0.457
seca	Altura máxima de 35 dbz	1×1	0.199	0.101	0.182	0.146	0.222	0.507	1.197	2.807
seca	Altura máxima de 35 dbz	2×2	0.324	0.147	0.300	0.246	0.361	0.775	0.895	0.917
seca	Altura máxima de 35 dbz	3×3	0.371	0.155	0.348	0.290	0.407	0.843	0.894	1.049
seca	Altura máxima de 35 dbz	4×4	0.402	0.156	0.383	0.321	0.438	0.857	0.849	0.933
seca	Altura máxima de 35 dbz	5×5	0.425	0.157	0.404	0.344	0.464	0.875	0.732	0.596
seca	Máxima refletividade	1×1	0.640	0.191	0.655	0.557	0.736	1.000	-0.316	-0.430
seca	Máxima refletividade	2×2	0.677	0.181	0.694	0.614	0.764	1.000	-0.506	-0.014
seca	Máxima refletividade	3×3	0.703	0.174	0.724	0.645	0.786	1.000	-0.636	0.277
seca	Máxima refletividade	4×4	0.721	0.170	0.742	0.667	0.800	1.000	-0.708	0.478
seca	Máxima refletividade	5×5	0.740	0.164	0.760	0.686	0.818	1.000	-0.776	0.733
seca	Taxa de precipitação	1×1	0.543	0.188	0.547	0.457	0.626	0.959	0.005	-0.496
seca	Taxa de precipitação	2×2	0.599	0.185	0.608	0.519	0.678	1.000	-0.129	-0.357
seca	Taxa de precipitação	3×3	0.629	0.177	0.634	0.560	0.704	1.000	-0.217	-0.191
seca	Taxa de precipitação	4×4	0.651	0.174	0.659	0.584	0.725	1.000	-0.314	-0.010
seca	Taxa de precipitação	5×5	0.673	0.171	0.681	0.610	0.746	1.000	-0.416	0.225
seca	Refletividade em 2.0 km	1×1	0.531	0.181	0.535	0.445	0.606	0.938	0.055	-0.443
seca	Refletividade em 2.0 km	2×2	0.602	0.183	0.610	0.529	0.681	1.000	-0.163	-0.283

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
seca	Refletividade em 2.0 km	3×3	0.635	0.176	0.640	0.571	0.710	1.000	-0.290	-0.038
seca	Refletividade em 2.0 km	4×4	0.658	0.172	0.667	0.597	0.733	1.000	-0.388	0.120
seca	Refletividade em 2.0 km	5×5	0.682	0.168	0.689	0.625	0.755	1.000	-0.490	0.354
seca	Refletividade em 2.5 km	1×1	0.543	0.188	0.547	0.457	0.626	0.959	0.005	-0.496
seca	Refletividade em 2.5 km	2×2	0.599	0.185	0.608	0.519	0.678	1.000	-0.129	-0.357
seca	Refletividade em 2.5 km	3×3	0.629	0.177	0.634	0.560	0.704	1.000	-0.217	-0.191
seca	Refletividade em 2.5 km	4×4	0.651	0.174	0.659	0.584	0.725	1.000	-0.314	-0.010
seca	Refletividade em 2.5 km	5×5	0.673	0.171	0.681	0.610	0.746	1.000	-0.416	0.225
seca	Refletividade em 3.0 km	1×1	0.561	0.194	0.569	0.468	0.651	0.981	-0.039	-0.564
seca	Refletividade em 3.0 km	2×2	0.601	0.187	0.609	0.517	0.684	1.000	-0.143	-0.377
seca	Refletividade em 3.0 km	3×3	0.629	0.180	0.634	0.554	0.705	1.000	-0.206	-0.253
seca	Refletividade em 3.0 km	4×4	0.648	0.177	0.654	0.579	0.725	1.000	-0.278	-0.098
seca	Refletividade em 3.0 km	5×5	0.668	0.173	0.671	0.600	0.745	1.000	-0.369	0.066
seca	Refletividade em 3.5 km	1×1	0.571	0.196	0.579	0.476	0.663	0.985	-0.084	-0.575
seca	Refletividade em 3.5 km	2×2	0.607	0.188	0.618	0.525	0.691	1.000	-0.177	-0.403
seca	Refletividade em 3.5 km	3×3	0.632	0.183	0.640	0.557	0.712	1.000	-0.231	-0.280
seca	Refletividade em 3.5 km	4×4	0.650	0.180	0.659	0.583	0.727	1.000	-0.287	-0.159
seca	Refletividade em 3.5 km	5×5	0.669	0.176	0.675	0.600	0.746	1.000	-0.349	-0.014
seca	Refletividade em 4.0 km	1×1	0.574	0.198	0.582	0.476	0.670	1.000	-0.078	-0.603
seca	Refletividade em 4.0 km	2×2	0.609	0.189	0.617	0.524	0.695	1.000	-0.176	-0.453
seca	Refletividade em 4.0 km	3×3	0.631	0.185	0.640	0.556	0.712	1.000	-0.235	-0.334
seca	Refletividade em 4.0 km	4×4	0.650	0.182	0.658	0.579	0.731	1.000	-0.288	-0.221
seca	Refletividade em 4.0 km	5×5	0.669	0.177	0.674	0.600	0.750	1.000	-0.329	-0.103
seca	Refletividade em 4.5 km	1×1	0.572	0.200	0.583	0.474	0.667	1.000	-0.089	-0.614
seca	Refletividade em 4.5 km	2×2	0.606	0.191	0.614	0.519	0.695	1.000	-0.175	-0.476
seca	Refletividade em 4.5 km	3×3	0.626	0.187	0.636	0.549	0.711	1.000	-0.237	-0.351
seca	Refletividade em 4.5 km	4×4	0.646	0.184	0.653	0.574	0.727	1.000	-0.289	-0.235
seca	Refletividade em 4.5 km	5×5	0.662	0.180	0.667	0.593	0.740	1.000	-0.325	-0.149
seca	Refletividade em 5.0 km	1×1	0.571	0.201	0.578	0.472	0.667	0.989	-0.107	-0.598
seca	Refletividade em 5.0 km	2×2	0.602	0.194	0.613	0.512	0.696	1.000	-0.191	-0.481

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
seca	Refletividade em 5.0 km	3×3	0.622	0.190	0.632	0.543	0.706	1.000	-0.247	-0.336
seca	Refletividade em 5.0 km	4×4	0.642	0.187	0.650	0.566	0.728	1.000	-0.296	-0.248
seca	Refletividade em 5.0 km	5×5	0.656	0.183	0.667	0.583	0.739	1.000	-0.313	-0.196
seca	Refletividade em 5.5 km	1×1	0.569	0.202	0.579	0.471	0.667	0.985	-0.105	-0.608
seca	Refletividade em 5.5 km	2×2	0.600	0.196	0.609	0.512	0.693	1.000	-0.180	-0.506
seca	Refletividade em 5.5 km	3×3	0.620	0.191	0.629	0.541	0.708	1.000	-0.239	-0.341
seca	Refletividade em 5.5 km	4×4	0.638	0.189	0.649	0.562	0.727	1.000	-0.316	-0.229
seca	Refletividade em 5.5 km	5×5	0.651	0.187	0.666	0.576	0.733	1.000	-0.307	-0.218
seca	Refletividade em 6.0 km	1×1	0.568	0.202	0.574	0.469	0.667	0.986	-0.090	-0.626
seca	Refletividade em 6.0 km	2×2	0.598	0.198	0.604	0.509	0.694	1.000	-0.185	-0.483
seca	Refletividade em 6.0 km	3×3	0.617	0.192	0.627	0.535	0.707	1.000	-0.248	-0.344
seca	Refletividade em 6.0 km	4×4	0.634	0.190	0.647	0.556	0.723	1.000	-0.300	-0.277
seca	Refletividade em 6.0 km	5×5	0.648	0.188	0.661	0.571	0.733	1.000	-0.328	-0.182
seca	Refletividade em 6.5 km	1×1	0.565	0.202	0.571	0.463	0.662	0.986	-0.087	-0.636
seca	Refletividade em 6.5 km	2×2	0.592	0.199	0.600	0.500	0.686	1.000	-0.196	-0.461
seca	Refletividade em 6.5 km	3×3	0.612	0.195	0.623	0.529	0.702	1.000	-0.250	-0.333
seca	Refletividade em 6.5 km	4×4	0.627	0.192	0.641	0.545	0.718	1.000	-0.307	-0.286
seca	Refletividade em 6.5 km	5×5	0.643	0.190	0.654	0.565	0.729	1.000	-0.347	-0.170
seca	Refletividade em 7.0 km	1×1	0.562	0.203	0.570	0.463	0.660	0.984	-0.114	-0.607
seca	Refletividade em 7.0 km	2×2	0.589	0.198	0.600	0.500	0.684	1.000	-0.193	-0.473
seca	Refletividade em 7.0 km	3×3	0.607	0.195	0.618	0.524	0.697	1.000	-0.247	-0.305
seca	Refletividade em 7.0 km	4×4	0.622	0.193	0.632	0.542	0.713	1.000	-0.294	-0.286
seca	Refletividade em 7.0 km	5×5	0.637	0.192	0.646	0.556	0.725	1.000	-0.335	-0.169
seca	Vil	1×1	0.623	0.195	0.639	0.535	0.720	1.000	-0.250	-0.535
seca	Vil	2×2	0.659	0.184	0.673	0.588	0.749	1.000	-0.397	-0.206
seca	Vil	3×3	0.684	0.178	0.704	0.621	0.765	1.000	-0.519	0.057
seca	Vil	4×4	0.702	0.173	0.722	0.645	0.783	1.000	-0.569	0.211
seca	Vil	5×5	0.721	0.168	0.737	0.667	0.800	1.000	-0.635	0.407
seca	Altura de waldvogel	1×1	0.120	0.073	0.103	0.080	0.128	0.369	1.658	3.958
seca	Altura de waldvogel	2×2	0.184	0.098	0.162	0.132	0.197	0.512	1.425	3.069

Continua na próxima página



Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			
seca	Altura de waldvogel	3×3	0.217	0.108	0.194	0.158	0.234	0.571	1.495	3.838
seca	Altura de waldvogel	4×4	0.243	0.114	0.222	0.184	0.263	0.614	1.362	3.226
seca	Altura de waldvogel	5×5	0.268	0.120	0.250	0.205	0.289	0.646	1.258	2.529
seca-chuvosa	Densidade do vil	1×1	0.300	0.118	0.300	0.253	0.345	0.582	0.769	3.304
seca-chuvosa	Densidade do vil	2×2	0.448	0.139	0.456	0.400	0.503	0.817	0.115	0.899
seca-chuvosa	Densidade do vil	3×3	0.555	0.142	0.566	0.500	0.621	0.855	-0.275	0.391
seca-chuvosa	Densidade do vil	4×4	0.640	0.143	0.656	0.595	0.709	0.945	-0.540	0.503
seca-chuvosa	Densidade do vil	5×5	0.698	0.141	0.719	0.656	0.772	1.000	-0.699	0.985
seca-chuvosa	Varição do topo da nuvem	1×1	0.600	0.212	0.610	0.509	0.700	1.000	-0.146	-0.675
seca-chuvosa	Varição do topo da nuvem	2×2	0.636	0.199	0.649	0.557	0.731	1.000	-0.240	-0.556
seca-chuvosa	Varição do topo da nuvem	3×3	0.657	0.194	0.671	0.580	0.757	1.000	-0.287	-0.543
seca-chuvosa	Varição do topo da nuvem	4×4	0.677	0.188	0.687	0.600	0.778	1.000	-0.402	-0.372
seca-chuvosa	Varição do topo da nuvem	5×5	0.695	0.183	0.714	0.627	0.786	1.000	-0.461	-0.220
seca-chuvosa	Altura máxima de 35 dbz	1×1	0.236	0.114	0.211	0.173	0.260	0.546	0.792	0.084
seca-chuvosa	Altura máxima de 35 dbz	2×2	0.387	0.170	0.357	0.287	0.433	0.789	0.597	-0.181
seca-chuvosa	Altura máxima de 35 dbz	3×3	0.440	0.178	0.418	0.337	0.500	0.858	0.490	-0.370
seca-chuvosa	Altura máxima de 35 dbz	4×4	0.471	0.184	0.455	0.369	0.533	0.900	0.442	-0.307
seca-chuvosa	Altura máxima de 35 dbz	5×5	0.497	0.183	0.476	0.400	0.561	0.979	0.464	-0.227
seca-chuvosa	Máxima refletividade	1×1	0.733	0.178	0.752	0.667	0.833	1.000	-0.644	0.035
seca-chuvosa	Máxima refletividade	2×2	0.772	0.159	0.789	0.718	0.857	1.000	-0.732	0.313
seca-chuvosa	Máxima refletividade	3×3	0.797	0.148	0.818	0.758	0.875	1.000	-0.890	0.799
seca-chuvosa	Máxima refletividade	4×4	0.818	0.136	0.838	0.781	0.889	1.000	-0.989	1.204
seca-chuvosa	Máxima refletividade	5×5	0.835	0.131	0.857	0.800	0.900	1.000	-1.105	1.847
seca-chuvosa	Taxa de precipitação	1×1	0.636	0.185	0.639	0.551	0.723	0.996	-0.266	-0.335
seca-chuvosa	Taxa de precipitação	2×2	0.705	0.173	0.713	0.636	0.797	1.000	-0.433	-0.149
seca-chuvosa	Taxa de precipitação	3×3	0.735	0.163	0.744	0.671	0.818	1.000	-0.528	0.052
seca-chuvosa	Taxa de precipitação	4×4	0.761	0.153	0.773	0.700	0.836	1.000	-0.598	0.262
seca-chuvosa	Taxa de precipitação	5×5	0.782	0.148	0.800	0.724	0.855	1.000	-0.684	0.542
seca-chuvosa	Refletividade em 2.0 km	1×1	0.606	0.186	0.611	0.519	0.688	0.996	-0.162	-0.304
seca-chuvosa	Refletividade em 2.0 km	2×2	0.706	0.173	0.713	0.634	0.794	1.000	-0.393	-0.184

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
seca-chuvosa	Refletividade em 2.0 km	3×3	0.738	0.164	0.750	0.673	0.821	1.000	-0.579	0.195
seca-chuvosa	Refletividade em 2.0 km	4×4	0.765	0.151	0.778	0.709	0.834	1.000	-0.618	0.355
seca-chuvosa	Refletividade em 2.0 km	5×5	0.786	0.145	0.804	0.732	0.857	1.000	-0.721	0.688
seca-chuvosa	Refletividade em 2.5 km	1×1	0.636	0.185	0.639	0.551	0.723	0.996	-0.266	-0.335
seca-chuvosa	Refletividade em 2.5 km	2×2	0.705	0.173	0.713	0.636	0.797	1.000	-0.433	-0.149
seca-chuvosa	Refletividade em 2.5 km	3×3	0.735	0.163	0.744	0.671	0.818	1.000	-0.528	0.052
seca-chuvosa	Refletividade em 2.5 km	4×4	0.761	0.153	0.773	0.700	0.836	1.000	-0.598	0.262
seca-chuvosa	Refletividade em 2.5 km	5×5	0.782	0.148	0.800	0.724	0.855	1.000	-0.684	0.542
seca-chuvosa	Refletividade em 3.0 km	1×1	0.667	0.190	0.670	0.589	0.762	1.000	-0.326	-0.368
seca-chuvosa	Refletividade em 3.0 km	2×2	0.709	0.173	0.723	0.636	0.798	1.000	-0.415	-0.241
seca-chuvosa	Refletividade em 3.0 km	3×3	0.736	0.165	0.750	0.667	0.822	1.000	-0.527	0.031
seca-chuvosa	Refletividade em 3.0 km	4×4	0.760	0.156	0.772	0.700	0.841	1.000	-0.577	0.107
seca-chuvosa	Refletividade em 3.0 km	5×5	0.779	0.149	0.800	0.724	0.857	1.000	-0.659	0.415
seca-chuvosa	Refletividade em 3.5 km	1×1	0.673	0.191	0.674	0.597	0.766	1.000	-0.324	-0.421
seca-chuvosa	Refletividade em 3.5 km	2×2	0.712	0.175	0.725	0.634	0.806	1.000	-0.425	-0.278
seca-chuvosa	Refletividade em 3.5 km	3×3	0.738	0.167	0.755	0.672	0.826	1.000	-0.530	-0.066
seca-chuvosa	Refletividade em 3.5 km	4×4	0.761	0.156	0.774	0.704	0.841	1.000	-0.528	-0.086
seca-chuvosa	Refletividade em 3.5 km	5×5	0.780	0.151	0.800	0.726	0.857	1.000	-0.657	0.387
seca-chuvosa	Refletividade em 4.0 km	1×1	0.673	0.192	0.679	0.596	0.767	1.000	-0.325	-0.452
seca-chuvosa	Refletividade em 4.0 km	2×2	0.709	0.178	0.725	0.636	0.806	1.000	-0.442	-0.265
seca-chuvosa	Refletividade em 4.0 km	3×3	0.737	0.169	0.757	0.667	0.825	1.000	-0.521	-0.101
seca-chuvosa	Refletividade em 4.0 km	4×4	0.760	0.157	0.778	0.700	0.839	1.000	-0.530	-0.118
seca-chuvosa	Refletividade em 4.0 km	5×5	0.778	0.153	0.797	0.726	0.856	1.000	-0.643	0.256
seca-chuvosa	Refletividade em 4.5 km	1×1	0.671	0.196	0.684	0.590	0.764	1.000	-0.338	-0.483
seca-chuvosa	Refletividade em 4.5 km	2×2	0.705	0.181	0.724	0.633	0.800	1.000	-0.442	-0.299
seca-chuvosa	Refletividade em 4.5 km	3×3	0.732	0.172	0.750	0.667	0.821	1.000	-0.505	-0.171
seca-chuvosa	Refletividade em 4.5 km	4×4	0.756	0.160	0.772	0.700	0.841	1.000	-0.542	-0.110
seca-chuvosa	Refletividade em 4.5 km	5×5	0.771	0.158	0.783	0.717	0.852	1.000	-0.640	0.181
seca-chuvosa	Refletividade em 5.0 km	1×1	0.670	0.199	0.684	0.590	0.773	1.000	-0.365	-0.470
seca-chuvosa	Refletividade em 5.0 km	2×2	0.705	0.184	0.721	0.633	0.800	1.000	-0.457	-0.288

Continua na próxima página

Tabela B.1: Continuação.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			
seca-chuvosa	Refletividade em 5.0 km	3×3	0.727	0.176	0.749	0.658	0.820	1.000	-0.519	-0.153
seca-chuvosa	Refletividade em 5.0 km	4×4	0.753	0.165	0.770	0.688	0.839	1.000	-0.583	-0.009
seca-chuvosa	Refletividade em 5.0 km	5×5	0.766	0.161	0.781	0.714	0.850	1.000	-0.649	0.147
seca-chuvosa	Refletividade em 5.5 km	1×1	0.669	0.202	0.684	0.587	0.771	1.000	-0.388	-0.471
seca-chuvosa	Refletividade em 5.5 km	2×2	0.701	0.186	0.721	0.630	0.798	1.000	-0.467	-0.307
seca-chuvosa	Refletividade em 5.5 km	3×3	0.725	0.179	0.748	0.656	0.815	1.000	-0.535	-0.157
seca-chuvosa	Refletividade em 5.5 km	4×4	0.748	0.169	0.768	0.684	0.838	1.000	-0.589	-0.049
seca-chuvosa	Refletividade em 5.5 km	5×5	0.758	0.165	0.778	0.711	0.840	1.000	-0.664	0.203
seca-chuvosa	Refletividade em 6.0 km	1×1	0.666	0.203	0.687	0.587	0.769	1.000	-0.379	-0.495
seca-chuvosa	Refletividade em 6.0 km	2×2	0.697	0.189	0.722	0.625	0.795	1.000	-0.466	-0.357
seca-chuvosa	Refletividade em 6.0 km	3×3	0.719	0.182	0.748	0.644	0.814	1.000	-0.520	-0.266
seca-chuvosa	Refletividade em 6.0 km	4×4	0.741	0.172	0.758	0.677	0.833	1.000	-0.572	-0.105
seca-chuvosa	Refletividade em 6.0 km	5×5	0.754	0.169	0.772	0.700	0.844	1.000	-0.656	0.101
seca-chuvosa	Refletividade em 6.5 km	1×1	0.662	0.206	0.684	0.581	0.773	1.000	-0.394	-0.537
seca-chuvosa	Refletividade em 6.5 km	2×2	0.695	0.190	0.714	0.625	0.795	1.000	-0.443	-0.418
seca-chuvosa	Refletividade em 6.5 km	3×3	0.715	0.182	0.735	0.644	0.811	1.000	-0.498	-0.316
seca-chuvosa	Refletividade em 6.5 km	4×4	0.735	0.174	0.755	0.667	0.831	1.000	-0.563	-0.142
seca-chuvosa	Refletividade em 6.5 km	5×5	0.749	0.172	0.771	0.697	0.840	1.000	-0.650	0.055
seca-chuvosa	Refletividade em 7.0 km	1×1	0.657	0.208	0.678	0.572	0.769	1.000	-0.373	-0.582
seca-chuvosa	Refletividade em 7.0 km	2×2	0.689	0.193	0.713	0.611	0.789	1.000	-0.434	-0.471
seca-chuvosa	Refletividade em 7.0 km	3×3	0.709	0.186	0.729	0.632	0.809	1.000	-0.493	-0.378
seca-chuvosa	Refletividade em 7.0 km	4×4	0.727	0.180	0.746	0.667	0.826	1.000	-0.582	-0.116
seca-chuvosa	Refletividade em 7.0 km	5×5	0.745	0.175	0.769	0.694	0.839	1.000	-0.695	0.155
seca-chuvosa	Vil	1×1	0.716	0.184	0.735	0.649	0.818	1.000	-0.558	-0.154
seca-chuvosa	Vil	2×2	0.753	0.165	0.771	0.697	0.838	1.000	-0.684	0.212
seca-chuvosa	Vil	3×3	0.780	0.154	0.800	0.733	0.860	1.000	-0.752	0.370
seca-chuvosa	Vil	4×4	0.803	0.144	0.822	0.759	0.875	1.000	-0.881	0.808
seca-chuvosa	Vil	5×5	0.821	0.137	0.843	0.782	0.893	1.000	-0.970	1.269
seca-chuvosa	Altura de waldvogel	1×1	0.127	0.086	0.100	0.076	0.130	0.396	1.418	2.020

Continua na próxima página

Tabela B.1: Conclusão.

ESTAÇÃO	VARIÁVEL	KERNEL	MÉDIA	DESVIO PADRÃO	MEDIANA	PERCENTIL		ASSIMETRIA	CURTOSE	
						P33	P66			P99
seca-chuvosa	Altura de waldvogel	2×2	0.196	0.121	0.157	0.120	0.215	0.543	1.262	1.503
seca-chuvosa	Altura de waldvogel	3×3	0.234	0.134	0.190	0.154	0.259	0.616	1.092	0.733
seca-chuvosa	Altura de waldvogel	4×4	0.261	0.138	0.227	0.178	0.295	0.649	0.989	0.704
seca-chuvosa	Altura de waldvogel	5×5	0.288	0.150	0.250	0.197	0.315	0.731	1.014	0.840

## **APÊNDICE C - MOSAICOS COM AS PREVISÕES DE RAIOS DIFERENTES LIMIARES DE PROBABILIDADE**

As figuras apresentadas neste apêndice complementam as análises discutidas na Seção 4.5 do Capítulo 4. As figuras são referentes aos mosaicos de previsão de diferentes configurações de RNAs usando no pós-processamento os limiares de probabilidade >70%, >80%, >90% e >98%.

Figura C.1 - Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade >70%.

**MANAUS-AM 2014-10-06 20:00 UTC**

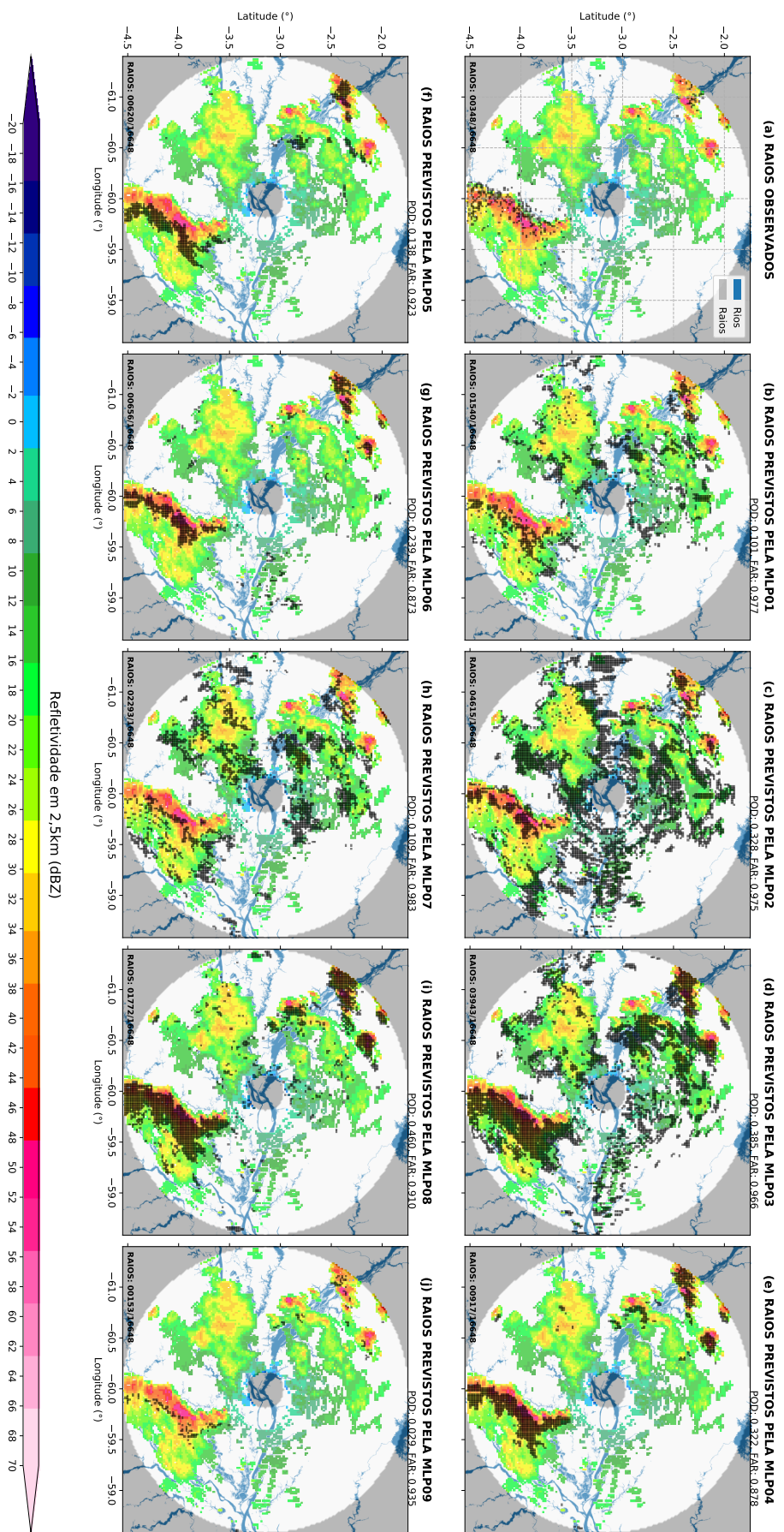
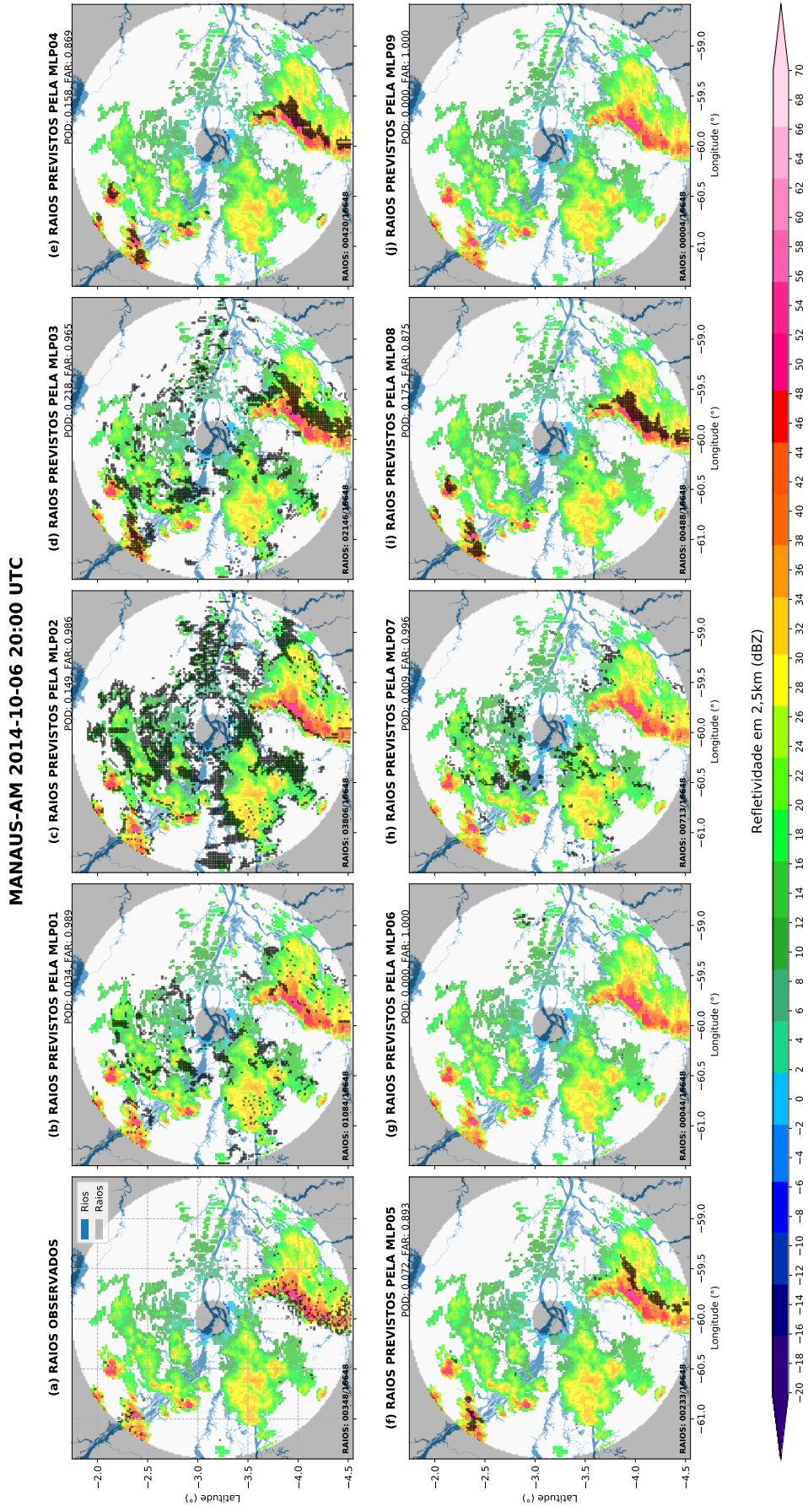


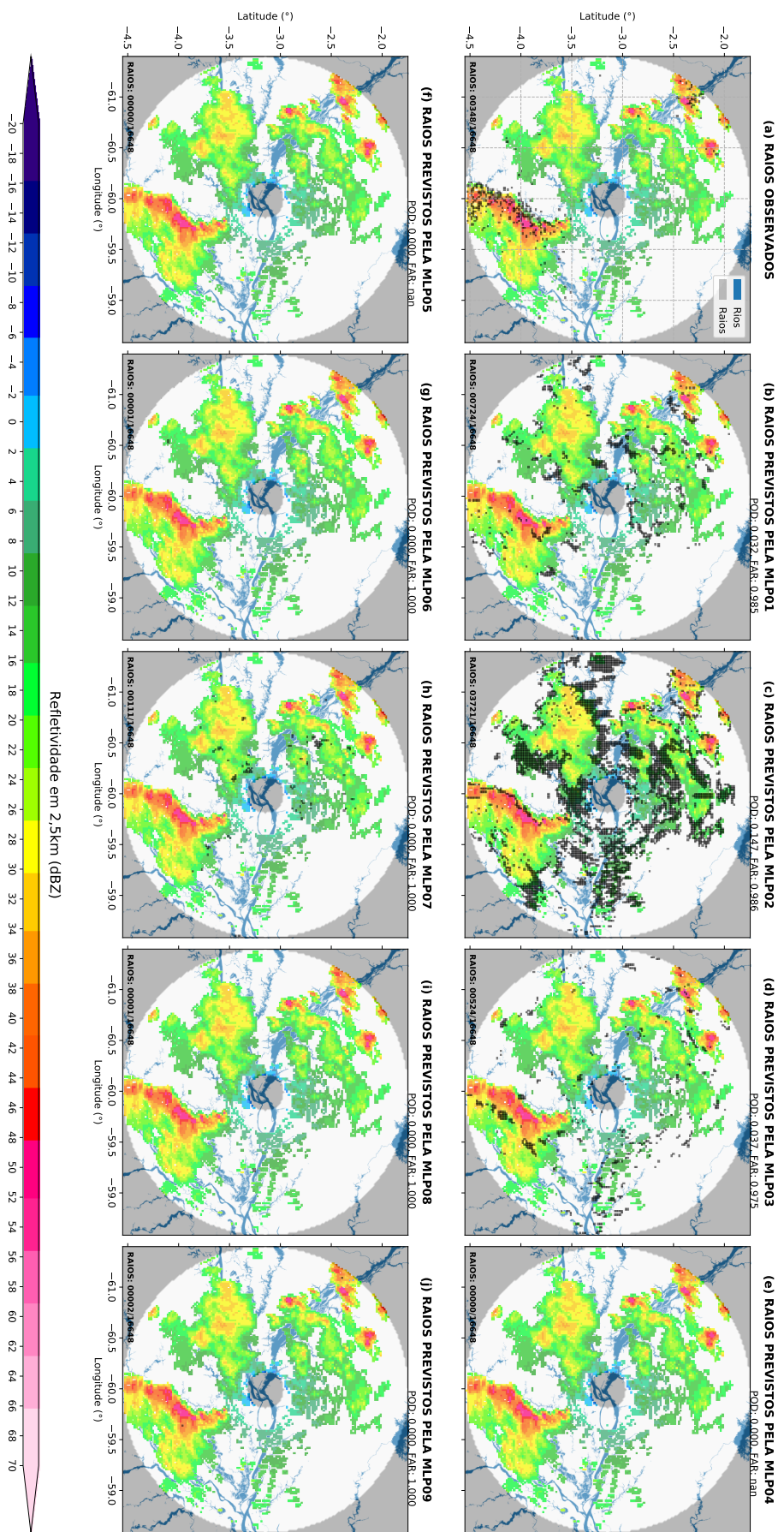
Figura C.2 - Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade >80%.



Fonte: Produção Autor.

Figura C.3 - Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade >90%.

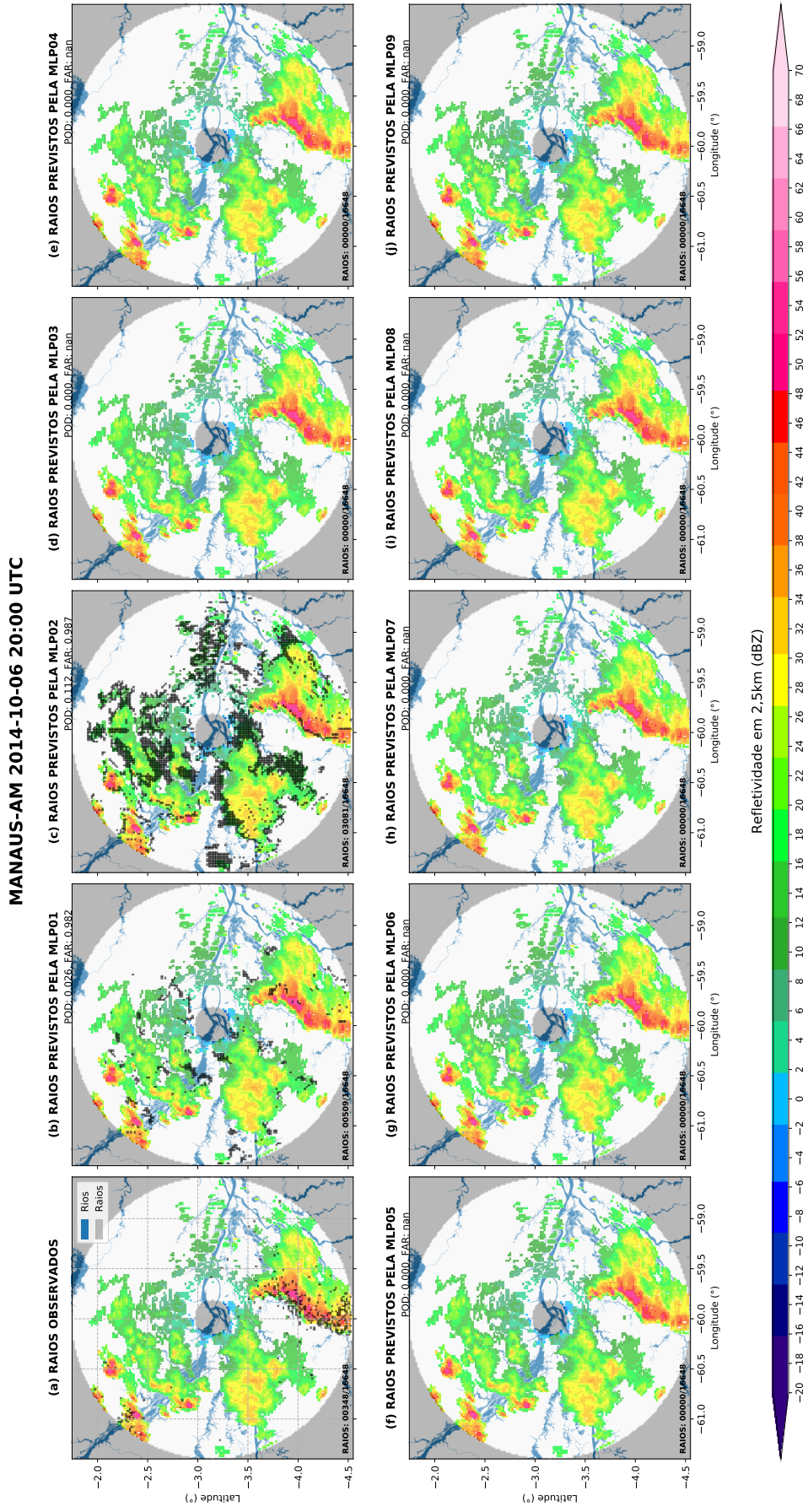
**MANAUS-AM 2014-10-06 20:00 UTC**



Fonte: Produção Autor.



Figura C.4 - Mosaico de previsão com todas as configurações de entrada com a MLP com limiar de probabilidade >98%.



Fonte: Produção Autor.

