

UNIVERSIDADE FEDERAL DO PARANÁ

TULIPA GABRIELA GUILHERMINA JUVENAL DA SILVA

IDENTIFICAÇÃO DE EVENTO DE TEMPO SEVERO UTILIZANDO TÉCNICAS
DE APRENDIZAGEM DE MÁQUINA EM DADOS DE RADAR
POLARIMÉTRICO

CURITIBA
2017

TULIPA GABRIELA GUILHERMINA JUVENAL DA SILVA

IDENTIFICAÇÃO DE EVENTO DE TEMPO SEVERO UTILIZANDO TÉCNICAS
DE APRENDIZAGEM DE MÁQUINA EM DADOS DE RADAR
POLARIMÉTRICO

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, do Departamento de Matemática, Setor de Ciências Exatas e do Departamento de Construção Civil, Setor de Tecnologia, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Métodos Numéricos em Engenharia.

Orientador: Prof^o. Dr. Paulo Henrique Siqueira

Coorientador: Dr. Cesar Augustus Assis Beneti

CURITIBA

2017

S586i

Silva, Tulipa Gabriela Guilhermina Juvenal da
Identificação de evento de tempo severo utilizando técnicas de
aprendizagem de máquina em dados de radar polarimétrico / Tulipa
Gabriela Guilhermina Juvenal da Silva. – Curitiba, 2017.
78f. : il. [algumas color.] ; 30 cm.

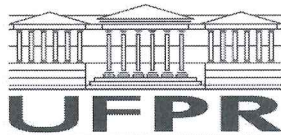
Dissertação (mestrado) - Universidade Federal do Paraná, Setores
Ciências Exatas e de Tecnologia, Programa de Pós-Graduação em Métodos
Numéricos em Engenharia, 2017.

Orientador: Paulo Henrique Siqueira -- Coorientador: Cesar Augustus
Assis Beneti.

Bibliografia: p. 73-78.

1. Radar meteorológico. 2. Descargas elétricas atmosféricas. I.
Universidade Federal do Paraná. II. Siqueira, Paulo Henrique. III. Beneti,
Cesar Augustus Assis. IV. Título.

CDD: 621.3848



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **TULIPA GABRIELA GUILHERMINA JUVENAL** intitulada: **IDENTIFICAÇÃO DE EVENTO DE TEMPO SEVERO UTILIZANDO TÉCNICAS DE APRENDIZAGEM DE MÁQUINA EM DADOS DE RADAR POLARIMÉTRICO**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 28 de Junho de 2017.

PAULO HENRIQUE STQUEIRA

Presidente da Banca Examinadora (UFPR)

LEONARDO CALVETTI

Avaliador Externo (UFPEL)

DEISE MARIA BERTHOLDI COSTA

Avaliador Interno (UFPR)

AGRADECIMENTOS

À minha família, que apoiou e esperou pacientemente pelas promessas (nunca cumpridas) de ir vê-los mais vezes ou de ficar por mais tempo. A família a qual me refiro não é aquela em que nasci, mas sim aqueles que me acolheram: família Siewert (Oma e Opa), família Hertel (Beatriz, Roni, Renan e Larissa), Célia Belanda (mãe Célia), Rafael Klosowski e Bianca B. Maske. Uma fase se encerra, outra começa, e tenho certeza de que eles estarão apoiando novamente.

À todos os meus amigos que sempre me apoiaram nos estudos e nas horas de dificuldades. Em particular, aos amigos do SIMEPAR, que muito me ensinaram, a Taninha, ao Rômulo Leite e a Patrícia S. Pacheco.

Ao Prof. Dr. Paulo Siqueira, ao Dr. Cesar Beneti e ao Prof. Dr. Maiko Buzzi que me orientaram neste trabalho. Mais que a conclusão do trabalho, devo a estes três a oportunidade, a confiança, muito aprendizado e apoio, inclusive financeiro.

Aos demais professores e alunos do PPGMNE, ao secretário Jair dos Anjos e demais funcionários do programa, por todo apoio e por terem feito com que tudo funcionasse da melhor maneira possível.

Agradeço, também, à CAPES e ao SIMEPAR pelo apoio financeiro.

Obrigada a todos!

RESUMO

Análises e estudos meteorológicos permitem realizar previsões sobre o tempo e antecipar eventos severos como chuva intensa, tempestades elétricas e tornados. Radares meteorológicos têm a característica de possibilitar a previsão de eventos severos à curtíssimo prazo, sendo possível identificar e monitorar tempestades severas. Este trabalho apresenta uma aplicação das técnicas de Aprendizagem de Máquina *Support Vector Machine* (SVM) e *Multilayer Perceptron* (MLP) para a previsão de Evento de Tempo Severo (ETS). O vetor de características, utilizado para informar os dados de entrada do modelo, é constituído de dados coletados do radar de dupla polarização pertencente ao Sistema Meteorológico do Paraná (SIMEPAR). Os dados de Descargas Elétricas Atmosféricas (DEA) são utilizados para análise dos resultados obtidos, cedidos pela Earth Networks Total Lightning Network (ENTLN). Após o treinamento obteve-se modelos que podem servir de apoio à decisão em alertas de ETS no estado do Paraná. O modelo obtido por SVM se mostrou superior ao obtido por MLP, identificando 93,02% dos ETSs estudados e com 67,84% do total de regiões identificadas confirmadas por DEA, se tornando uma possível ferramenta de apoio à decisão.

Palavras-chave: Radar Polarimétrico; *Nowcasting*; *Support Vector Machine*; *Multilayer Perceptron*; Descargas Elétricas Atmosféricas; Aprendizagem de Máquina.

ABSTRACT

Analyzes and meteorological studies make possible to forecast weather and anticipate severe events such as heavy rain, electrical storms and tornadoes. Meteorological radars have the characteristic of making possible to predict severe events in the short term, allows to identify and monitor severe storms. This paper presents an application of the Machine Learning techniques for the Severe Weather Event (ETS) prediction. These techniques are: Support Vector Machine (SVM) and Multilayer Perceptron (MLP). The data used as input of the model, characteristic vector, consists of data collected from the double polarization radar belonging to the Paraná Meteorological System (SIMEPAR). The data of Atmospheric Electric Discharges (DEA) are used to analyze the results obtained, provided by Earth Networks Total Lightning Network (ENTLN). After the training, were obtained models that can serve as support for the decision on ETS alerts in the state of Paraná. The model obtained by SVM was superior to that obtained by MLP, identifying 93,02 % of the ETSs studied and 67.84 % of the total number of identified regions confirmed by DEA, becoming a possible decision support tool.

Key-words: Polarimetric Radar; Nowcasting; Support Vector Machine; Multilayer Perceptron; Atmospheric Electric Discharges; Machine Learning.

LISTA DE FIGURAS

FIGURA 1 – REPRESENTAÇÃO DA TUPLA DE COORDENADAS (θ, ϕ, R) , DOS DADOS COLETADOS PELO RADAR.	20
FIGURA 2 – ESQUEMA DE UMA VARREDURA VOLUMÉTRICA DE RADAR. .	20
FIGURA 3 – VISUALIZAÇÃO DOS MOMENTOS DOPPLER.	21
FIGURA 4 – COMPARAÇÃO ENTRE RADARES DE DUPLA POLARIZAÇÃO E POLARIZAÇÃO SIMPLES NA DETECÇÃO DE PARTÍCULAS.	24
FIGURA 5 – VISUALIZAÇÃO DAS VARIÁVEIS POLARIMÉTRICAS.	25
FIGURA 6 – REPRESENTAÇÃO DE UM NEURÔNIO ARTIFICIAL.	35
FIGURA 7 – EXEMPLO DE ARQUITETURA DE UMA MLP.	39
FIGURA 8 – HIPERPLANO ÓTIMO DE SEPARAÇÃO E MARGEM δ	44
FIGURA 9 – RADAR POLARIMÉTRICO DE CASCAVEL.	50
FIGURA 10– DADOS DE 14/07/2015 ÀS 9:52 UTC EM GRADE CARTESIANA, ONDE Z, ZDR, RHOHV, KDP, HMAX, AZS, RS, E VS REPRESENTAM REFLETIVIDADE, REFLETIVIDADE DIFERENCIAL, CORRELAÇÃO COPOLAR, DIFERENCIAL DE FASE ESPECÍFICA, ALTITUDE DO MÁXIMO VALOR DE REFLETIVIDADE POR COLUNA, CISALHAMENTO AZIMUTHAL, CISALHAMENTO RADIAL E CISALHAMENTO VERTICAL.	57
FIGURA 11– CORRELAÇÃO ENTRE OS ATRIBUTOS DE ENTRADA. AS POSIÇÕES DA DIAGONAL APRESENTAM OS ATRIBUTOS CUJA CORRELAÇÃO ESTÁ NA LINHA E COLUNA CORRESPONDENTE ACIMA DA DIAGONAL E O GRÁFICO DE DISPERSÃO E A RESPECTIVA CURVA DE REGRESSÃO ESTÃO NA LINHA E COLUNA CORRESPONDENTE ABAIXO DA DIAGONAL.	60
FIGURA 12– SAÍDA OBTIDA COM A UTILIZAÇÃO DOS MODELOS M-MLP E S-SVM RESPECTIVAMENTE. AS IMAGENS REPRESENTAM OS PONTOS IDENTIFICADOS PELOS RESPECTIVOS MODELOS (EM PRETO)	

SOBRE A IMAGEM DO ATRIBUTO REFLETIVIDADE. A REGIÃO CIRCULADA MOSTRA ONDE OCORREU ETS DENTRO DE 30 MINUTOS.	61
FIGURA 13– TEMPESTADE DE 26/11/2015, QUE OCASIONOU UM ETS ÀS 23:45, INTERNO AO CÍRCULO DAS FIGURAS. ESTE ETS NÃO FOI IDENTIFICADO PELOS MODELOS M-MLP E M-SVM.	62
FIGURA 14– BOXPLOT DOS VALORES DE HMAX QUE RESULTAM EM SAÍDA POSITIVA PELOS MODELOS M-MLP E M-SVM.	66
FIGURA 15– BOXPLOT DOS VALORES DE ALTITUDE DOS PONTOS QUE RESULTAM EM SAÍDA POSITIVA PELOS MODELOS M-MLP E M-SVM.	67
FIGURA 16– BOXPLOT DA QUANTIDADE DE PONTOS POR COLUNA IDENTIFICADOS PELOS MODELOS.	68

LISTA DE SIGLAS

AM	Aprendizagem de Máquina
DEA	Descargas Elétricas Atmosféricas
ENTLN	Earth Networks Total Lightning Network
ETS	Evento de Tempo Severo
MLP	Multilayer Perceptron
RNA	Rede Neural Artificial
SIMEPAR	Sistema Meteorológico do Paraná
SVM	Support Vector Machine

LISTA DE SÍMBOLOS

h	Altitude do <i>bin</i>
$HMAX$	Altitude do máximo valor de refletividade por coluna
θ	Ângulo de azimute
ϕ	Ângulo de elevação da antena
AZS	Cisalhamento azimutal
RS	Cisalhamento radial
VS	Cisalhamento vertical
$RHOHV$	Correlação copolar
KDP	Diferencial de fase específica
$PHIDP$	Diferencial de fase
r	Distância da partícula ao radar
W	Largura espectral
LJ	Lightning Jump
ZDR	Refletividade diferencial
Z	Refletividade representada pela escala logarítmica
V	Velocidade Doppler

SUMÁRIO

1	INTRODUÇÃO	12
1.1	MOTIVAÇÃO	13
1.2	OBJETIVOS	13
1.2.1	Objetivo geral	13
1.2.2	Objetivos específicos	13
1.2.3	Limitações do trabalho	13
1.3	ESTRUTURA DO TRABALHO	14
2	CONCEITOS METEOROLÓGICOS	15
2.1	EVENTO DE TEMPO SEVERO (ETS)	15
2.2	RADAR METEOROLÓGICO	18
2.2.1	Momentos Doppler	20
2.2.1.1	Refletividade (Z)	21
2.2.1.2	Velocidade Radial (V)	22
2.2.1.3	Largura Espectral (W)	22
2.2.2	Variáveis Polarimétricas	23
2.2.2.1	Refletividade Diferencial (ZDR)	25
2.2.2.2	Correlação Copolar (RHOHV)	26
2.2.2.3	Diferencial de Fase (PHIDP)	26
2.2.2.4	Diferencial de Fase Específica (KDP)	27
2.2.3	Cisalhamento	27
2.3	DESCARGAS ELÉTRICAS ATMOSFÉRICAS (DEA)	28
3	APRENDIZAGEM DE MÁQUINA (AM)	31
3.1	FUNDAMENTOS DE APRENDIZAGEM DE MÁQUINA	31

3.2 CLASSIFICAÇÃO DE PADRÕES	33
3.3 REDES NEURAIS ARTIFICIAIS (RNA)	35
3.3.1 Algoritmo Backpropagation	37
3.3.2 Multilayer Perceptron (MLP)	38
3.4 SUPPORT VECTOR MACHINE (SVM)	42
4 MATERIAIS E MÉTODOS	49
4.1 MATERIAIS	49
4.1.1 Especificações do Radar	49
4.1.2 Especificações dos dados de Descargas Elétricas Atmosféricas	50
4.1.3 Especificações dos Eventos de Tempo Severo	50
4.2 MÉTODOS	51
4.2.1 Pré-processamento	52
4.2.2 Inclusão das classes	53
4.2.3 Treinamento	54
4.2.4 Análise	56
5 RESULTADOS	58
5.1 PRÉ-PROCESSAMENTO	58
5.2 MODELOS M-MLP E M-SVM	59
5.3 COMPARAÇÃO ENTRE M-MLP E M-SVM AOS DADOS DE DESCARGAS ELÉTRICAS ATMOSFÉRICAS	64
5.4 ESTUDO DAS ENTRADAS POSITIVAS	65
6 CONCLUSÕES	70
6.1 SUGESTÕES PARA TRABALHOS FUTUROS	72
REFERÊNCIAS	73

1 INTRODUÇÃO

No período de janeiro de 2015 a julho de 2016 a Defesa Civil efetuou 106 registros eventos de tempo severo, no estado do Paraná, causados por tempestades convectivas. Segundo a Defesa Civil, ao menos 39753 pessoas foram diretamente afetadas por estas tempestades (DEFESA CIVIL, 2015, 2016). Os eventos registrados são do tipo vendaval, chuvas intensas, granizo ou tornados.

Para o monitoramento e previsão meteorológica, o radar é uma das melhores ferramentas disponíveis, pois oferece monitoramento em tempo real, com boa resolução espacial e possibilita uma visão tridimensional dos dados.

Além disso, é possível analisar um Evento de Tempo Severo (ETS) com base nos dados de Descargas Elétricas Atmosféricas (DEA), fato já explorado por Lima (2005), Darden et al. (2010), Murphy e Holle (2006) e Liu e Heckman (2012), os quais utilizam estes dados para estimar, inclusive, a fase de desenvolvimento da tempestade.

As técnicas de Aprendizagem de Máquina (AM) têm sido utilizadas na meteorologia, quer seja para classificação como mostram Damian (2011) e Neto (2008), ou para previsão como mostra o trabalho de Anochi (2015).

Com o intuito de identificar a ocorrência de ETSs, são aplicadas neste trabalho duas técnicas de AM, *Support Vector Machine* (SVM) e *Multilayer Perceptron* (MLP), aos dados de radar polarimétrico para obter um modelo de predição que estima se uma tempestade convectiva pode conter um ETS em muito curto prazo (30 minutos ou menos). Os resultados obtidos são também comparados aos dados de DEA, para melhor avaliação dos modelos obtidos.

1.1 MOTIVAÇÃO

A convecção é um processo rápido, muitas das células convectivas duram cerca de uma hora ou menos (FABRY, 2015). Como os dados de radar meteorológico e DEA fornecem informações em tempo real, convém estudar a convecção com base nestas ferramentas. Além disso, este estudo possibilita verificar características comuns a ETS em dados de radar. Assim, estudar ETS em dados de radar aplicando técnicas de AM possibilita, por meio da identificação de padrões nos dados, a previsão de muito curto prazo de ETS .

1.2 OBJETIVOS

1.2.1 OBJETIVO GERAL

Utilizar técnicas aprendizagem de máquina para classificar tempestades convectivas em potenciais eventos de tempo severo em até 30 minutos com dados de radar polarimétrico.

1.2.2 OBJETIVOS ESPECÍFICOS

- Mostrar aplicações de AM associadas à meteorologia;
- Estudar DEA e dados de radar para melhor compreensão dos resultados obtidos e do conjunto de dados sobre o qual se deseja gerar um modelo;
- Gerar dois modelos de classificação de ETS (utilizando MLP e SVM);
- Confrontar os modelos obtidos com a identificação de ETS por dados de DEA, a fim de verificar as extensões dos modelos.

1.2.3 LIMITAÇÕES DO TRABALHO

A seguir são listadas algumas limitações do trabalho:

- Este trabalho se restringe ao estudo de ETSs ocorridos na região do Paraná e captados de forma volumétrica pelo radar do SIMEPAR, localizado em Cascavel;
- Os dados estudados representam de 30 minutos antes da ocorrência do ETS, não sendo estudado o que ocorre após o ETS nem antes dos 30 minutos de antecedência;
- A localização e hora utilizada como ocorrência de ETS são as mesmas obtidas pelos relatórios do Defesa Civil do Paraná;
- A pré-classificação dos dados se baseia apenas nas informações de Defesa Civil, sendo um dos objetivos do trabalho confrontar tais informações;
- São utilizadas duas técnicas escolhidas com base no tipo de problema e dos dados disponíveis.

1.3 ESTRUTURA DO TRABALHO

No capítulo 1, apresenta-se uma breve introdução sobre o problema estudado, os dados e método de análise dos resultados. Além disso, apresentou-se os objetivos deste trabalho, bem como algumas de suas limitações.

O capítulo 2 é dedicado a desenvolver o problema e justificar sua solução. Nele também descrevem-se alguns trabalhos que resolvem o problema proposto, são ainda apresentadas algumas justificativas sobre a utilização dos dados.

No capítulo 3, são apresentados os conceitos associados a AM, e suas técnicas MLP e SVM que são usadas no trabalho. No decorrer do capítulo são esclarecidos os motivos de algumas escolhas de metodologia do trabalho.

No capítulo 4, são apresentados os materiais e métodos. Neste capítulo as limitações são discutidas.

Os resultados são discutidos no capítulo 5, juntamente com sua apresentação.

Por fim, no capítulo 6 são apresentadas as conclusões do trabalho.

2 CONCEITOS METEOROLÓGICOS

Prever os eventos de tempo severo é uma necessidade para mitigar prejuízos tanto humanos quanto econômicos. Os modelos de previsão de eventos de tempo severo operacionais são baseados em diversos tipos de dados meteorológicos e técnicas de previsão. Neste capítulo são apresentados alguns conceitos meteorológicos necessários a este trabalho, tais como definições de evento de tempo severo, radar e suas variáveis, e descargas elétricas atmosféricas.

2.1 EVENTO DE TEMPO SEVERO (ETS)

Uma tempestade convectiva severa é uma tempestade capaz de gerar vendaval, chuva intensa, granizo, tornados e ou descargas elétricas atmosféricas (MADDOX, 1980).

Cada um dos eventos do tipo vendaval, chuvas intensas, granizo, tornados e tempestades elétricas, decorrentes de tempestade convectiva é chamado Evento de Tempo Severo (ETS) (MADDOX, 1980) .

Os tornados são uma violenta coluna rotativa de ar em movimento, e ocorrem em condições de tempestade severa (DOSWELL, 2001). Vendavais são ventos de alta velocidade, também provenientes de uma tempestade severa. Chuva intensa trata-se de uma forte precipitação em um curto espaço de tempo.

Embora os tipos de ETSs sejam diferentes entre si, todos são decorrentes do desenvolvimento vertical de uma tempestade convectiva, devido a geração de correntes de ar para cima e para baixo nestas células de tempestade (FABRY, 2015).

Os limiares para velocidade do vento, tamanho do granizo, volume de precipitação e outros fenômenos associados a ETSs são baseados na região de estudo.

DOSWELL (2001), em uma revisão sobre tempestades convectivas, discute a arbitrariedade dos limiares para se definir ETS, pois parâmetros de corte podem incluir ou excluir eventos muito semelhantes, sendo adequado que os limiares sejam baseados no conceito de que um ETS é caracterizado por ser capaz de gerar fenômenos de impacto social e econômico.

De fato, no período de janeiro de 2015 a junho de 2016, foram relatados, pela Defesa Civil, mais de 100 ETSs ocorridos na região do estado do Paraná, causando impactos sociais e econômicos. A seguir são apresentados alguns destes eventos que afetaram aproximadamente 39753 pessoas de forma direta (DEFESA CIVIL, 2015, 2016).

Em 02/01/2015 ocorreu um forte vendaval na região de Foz do Iguaçu, destelhando casas, derrubando postes e árvores e causando alagamentos, noticiado pelos portais de notícias (G1PR, 2015b), (PORTAL DA CIDADE, 2015). Ainda que nenhum morador tenha deixado a residência em decorrência do alagamento (PORTAL DA CIDADE, 2015), 26286 pessoas foram afetadas por este evento (DEFESA CIVIL, 2015).

No dia 13/07/2015 ocorreram 3 ETSs em um intervalo inferior a 3 horas. Das 35 cidades que sofreram com estes ETSs (G1PR, 2015a), as cidades de Matelândia, Ampére e Francisco Beltrão foram as mais prejudicadas, 3564 pessoas enfrentaram vendavais, chuvas intensas e granizo atingindo a região.

Em 14/07/2015 e 22/03/2016 ocorreram tornados em Francisco Beltrão e Capitão Leônidas Marques, respectivamente, ambos confirmados pela Defesa Civil. No tornado de 14/07/2015 muitas residências foram destruídas e pelos menos 50 pessoas ficaram feridas (Agência de notícias do Paraná, 2015). O tornado que ocorreu em 22/03/2016 afetou diretamente 120 pessoas e as rajadas de vento passaram de 100 km/h.

O município de Francisco Beltrão foi o mais afetado em todo o período (janeiro de 2015 a junho de 2016), sofrendo com chuvas intensas e vendaval, nas datas de 26/11/2015, 28/12/2015 e 27/02/2016, afetando um total de 30600 pessoas nas 3

datas (DEFESA CIVIL, 2015, 2016).

Trabalhos como em Campos e Eichholz (2010), Damian (2011), têm sido publicados com o intuito de prever os ETS ou mesmo classificar regiões de tempestade convectiva propícias à ocorrência de ETS. Damian (2011) utiliza *kmeans* aplicado a dados de radar para classificar chuva convectiva, Campos e Eichholz (2010) estudam características físicas em um modelo teórico. A seguir apresentam-se outros trabalhos relacionados a ETS, as entradas e as técnicas aplicadas.

MERCER et al. (2009) utilizam o modelo *Weather Research and Forecast* como entrada para as técnicas *Support Vector Machine* (SVM) e regressão logística para a identificação de tornados com 24 horas de antecedência. De acordo com este estudo, o SVM foi utilizado a fim de diminuir os falsos alarmes obtidos anteriormente com a utilização de redes neurais artificiais.

O trabalho de KITZMILLER, McGOVERN e SAFLE (1995) utiliza dados de radar Doppler, tais como refletividade e velocidade Doppler como entrada do algoritmo estatístico *severe weather potential*, a fim de indicar a probabilidade da ocorrência de em ETS dentro de 20 minutos, comparando os resultados à medida de valor integrado de líquido.

JOE et al. (2004) apresentam uma comparação entre 4 algoritmos operacionais para previsão de ETS: *National Severe Storm Laboratory Warning Decision Support System*, *Meteorological Service of Canada Canadian Radar Decision Support*, *National Center for Atmospheric Research Thunderstorm Initiation, Tracking, Analysis, and Nowcasting* e *Bureau of Meteorology Research Centre polarized C-band polarimetric algoritmo*. Estes algoritmos utilizam técnicas estatísticas, técnicas de aprendizagem de máquina e técnicas de processamento de imagem aplicadas a dados de radar, raios, dados de superfície, entre outros dados meteorológicos disponíveis a fim de detectar, classificar e acompanhar células de tempestades convectivas. Este estudo apresentou resultados distintos para cada uma das 3 configurações de radar usadas para teste. Isto se deve a estrutura de cada algoritmo, bem como os tipos de dados

utilizados como entrada para cada um deles e principalmente da qualidade dos dados (JOE et al., 2004). Segundo JOE et al. (2004), muitos são os algoritmos de previsão de ETS operacionais no mundo, sendo os 4 citados escolhidos por conveniência.

Um exemplo de outro algoritmo operacional de previsão de ETS é apresentado pelo estudo de GAGNE, McGOVERN e BROTZGE (2009), o qual aplica dados de radar às técnicas J48 e *random forests*, obtendo performance semelhante às demais técnicas de aprendizagem de máquina já utilizadas na classificação de ETS. O estudo conclui que é possível automatizar a classificação de ETS e sugere que, após o treinamento de uma técnica de aprendizagem de máquina, o modelo que se obtém é tal que identifica características comuns a um ETS, sendo suficiente, a partir de então, estudar apenas tempestades com as características identificadas pelos métodos de aprendizado de máquina.

Neste trabalho são estudados os eventos severos denominados como vendaval, tornado e chuva intensa, citados apenas como ETS. Embora os três eventos sejam diferentes entre si, todos são decorrentes de tempestades convectivas, sendo resultado de seu desenvolvimento vertical, e portanto, presume-se que possam ser estudados de forma conjunta.

2.2 RADAR METEOROLÓGICO

O radar (*RA*dio *D*etection *A*nd *R*anging), foi desenvolvido inicialmente para fins militares. Porém, é mais conhecido pelas suas aplicações no controle de tráfego e monitoramento de eventos meteorológicos (RINEHART, 2004).

Para o monitoramento e previsão meteorológica, o radar é uma das melhores ferramentas disponíveis, pois oferece monitoramento em tempo real, com boa resolução espacial e possibilita uma visão tridimensional dos dados (SAUVAGEOT, 1992). Muito do que se conhece da estrutura interna dos sistemas de precipitação se deve ao estudo dos dados de radar (BENETI, 2012).

De uma forma simplificada, o radar é constituído por um transmissor, um receptor, uma antena e um sistema de decodificação e processamento. O radar é uma ferramenta de sensoriamento remoto, ou seja, por meio da radiação eletromagnética, coleta dados sem entrar em contato direto com o alvo de estudo (LANDGREBE, 1978), (LILLESAND; KIEFFER, 1987), (NOVO, 1992).

A coleta de dados ocorre quando o radar transmite um feixe de radiação eletromagnética e capta a energia refletida pelas partículas presentes na atmosfera. Essa energia é amplificada e, conhecido o intervalo de tempo Δt entre emissão e retorno do eco, é possível determinar a distância do alvo ao radar. Essa distância r pode ser calculada através da Equação 1, onde c_m é a velocidade de propagação das ondas eletromagnéticas no meio (COLLIER, 1996).

$$r = \frac{c_m \Delta t}{2} \quad (1)$$

Os dados coletados pelo radar são obtidos em coordenadas esféricas, onde a posição de cada dado é descrita em termos de (θ, ϕ, r) , onde θ é o ângulo de elevação da antena, ϕ é o ângulo de rotação, em sentido horário, da antena em relação ao Norte geográfico, chamado de ângulo de azimute, e r a distância do dado ao radar. A Figura 1 apresenta a posição de um alvo em termos das componentes θ , ϕ e r .

Após o retorno do sinal refletido, o software do radar separa o sinal refletido em intervalos igualmente espaçados na direção radial. O número de separações deste intervalo é chamado de número de bins, cada um destes bins recebe o valor da média ponderada dos alvos detectados entre um bin e o próximo.

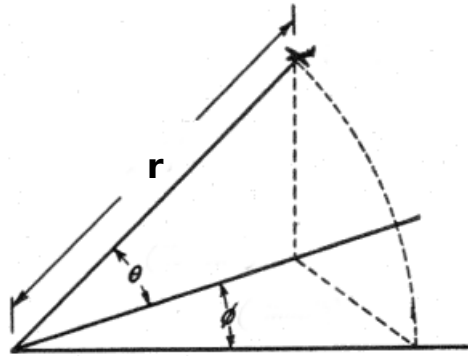


FIGURA 1: Representação da tupla de coordenadas (θ, ϕ, r) , dos dados coletados pelo radar.

FONTE: DeSoto (1945) adaptado.

A coleta de sucessivos giros da antena do radar a ângulos fixos, é conhecida como varredura volumétrica (FABRY, 2015), pois representa o dado em sucessivos cones, como exemplificado pela Figura 2

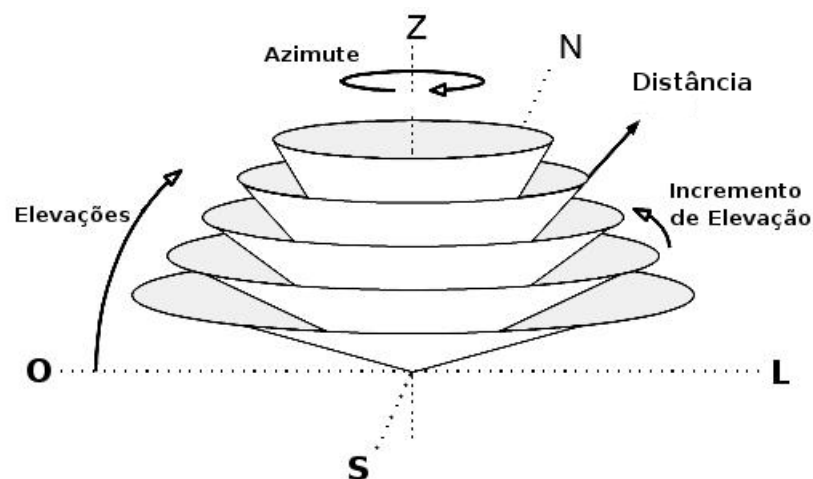


FIGURA 2: Esquema de uma varredura volumétrica de radar.

FONTE: SELUZNIAK (2016).

2.2.1 MOMENTOS DOPPLER

Radares do tipo Doppler operam comparando a frequência do sinal transmitido com o sinal recebido, a fim de calcular a velocidade da partícula captada pelo sinal (SILVEIRA, 1998). Para um radar meteorológico de tipo Doppler, são medidos três momentos: Refletividade (Z), Velocidade Radial (V) e Largura Espectral (W) (FABRY, 2015). Na Figura 3 estão ilustrados os momentos Doppler.

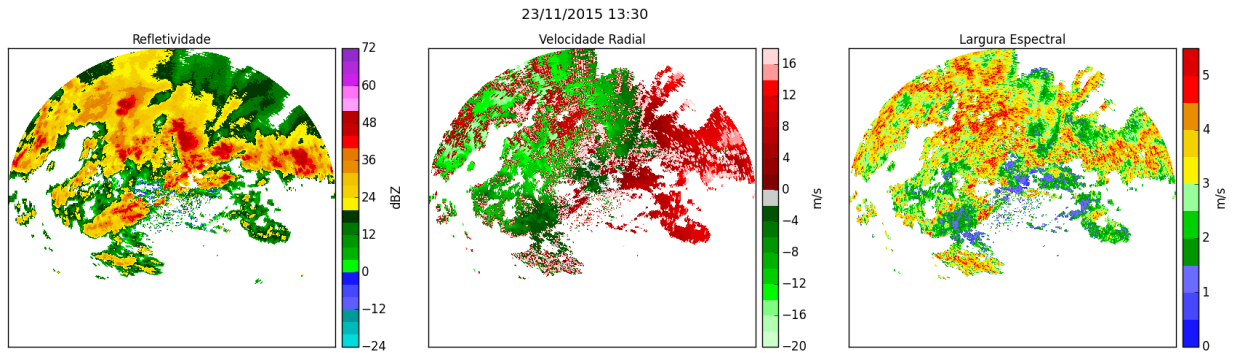


FIGURA 3: Visualização dos momentos Doppler.

FONTE: O autor (2017).

2.2.1.1 REFLETIVIDADE (Z)

A medida de eficiência de um alvo em interceptar e retornar a energia emitida pelo radar é chamada de refletividade. Dessa forma, refletividade é dependente dos tamanhos, formatos, quantidade e propriedades dielétricas dos alvos, relação expressa pela Equação 2, onde c_r é chamada constante do radar, definida por características do radar e suas configurações, e p_r é a energia refletida para o radar (FABRY, 2015).

$$z = c_r r^2 p_r. \quad (2)$$

A refletividade é expressa pela unidade mm^6/m^3 . Porém, os valores que a refletividade pode assumir com essa unidade variam entre 0 até 36000, sendo convencional fazer a compressão desses valores, adotando a escala logarítmica. Nesse caso, a refletividade pode ser escrita como:

$$Z = 10 \log_{10} \left(\frac{z}{1mm^6/m^3} \right). \quad (3)$$

Assim, na nova escala, a refletividade é denotada pela unidade dBZ e possui valores típicos entre -24 dBZ e 72 dBZ (FABRY, 2015).

A refletividade é uma das variáveis de maior utilização na previsão meteorológica.

Valores elevados de refletividade estão, em geral, associados à chuva intensa e altos valores de refletividade, 45 dBZ ou mais, entre 5 e 12 km de altitude podem estar associados a tempestade severa (FABRY, 2015).

2.2.1.2 VELOCIDADE RADIAL (V)

Utilizando o efeito Doppler, o radar estima a velocidade com que o alvo se afasta ou se aproxima do radar na direção do feixe emitido (FABRY, 2015).

A velocidade é a derivada da distância r da partícula com relação ao tempo t , sendo descrita como:

$$V = \frac{dr}{dt}. \quad (4)$$

A velocidade radial é a velocidade com que o alvo detectado está se distanciando ou se aproximando do radar. Por convenção, valores negativos representam partículas se aproximando do radar e valores positivos, partículas se afastando do radar (FABRY, 2015).

2.2.1.3 LARGURA ESPECTRAL (W)

A largura espectral (W) representa o desvio padrão das medidas de velocidade obtidas. Em geral, está associado a turbulência no volume e sua unidade de medida é expressa em metros por segundo (m/s). Altos valores de largura espectral podem indicar convecção (FABRY, 2015).

Seja V_i a velocidade individual de cada elemento na amostra de N elementos e V_{ave} a velocidade média da amostra. A largura espectral, é definida pela Equação 5

$$W = \frac{\sum(V_i - V_{ave})^2}{N - 1}. \quad (5)$$

2.2.2 VARIÁVEIS POLARIMÉTRICAS

No radar convencional, de polarização simples, o sinal eletromagnético emitido e recebido possui apenas uma orientação (vertical ou horizontal), sendo capaz de obter informações dos alvos em apenas uma orientação. O radar polarimétrico dual ou radar de dupla polarização, tem a propagação de sinal eletromagnético com orientações tanto horizontal como vertical, sendo capaz de obter informações nas duas orientações (FABRY, 2015).

Alguns estudos para a previsão meteorológica utilizando as variáveis polarimétricas são listados por Kumjian (2013b), pois as variáveis polarimétricas fornecem informações adicionais sobre o alvo.

Quando o sinal é emitido e interpretado pelo radar na mesma orientação dizemos que este radar possui copolarização. As variáveis coletadas desta forma recebem os índices HH ou VV , como é o caso de Z_{HH} , que representa a refletividade emitida e interpretada horizontalmente (FABRY, 2015).

A polarização cruzada se refere a interpretar o sinal refletido apenas na orientação contrária à emitida. As variáveis coletadas em configuração de polarização cruzada recebem os índices HV ou VH , indicando as orientações de emissão e captação. Por exemplo, o índice da variável Z_{HV} indica que Z foi emitida com orientação horizontal e interpretada na orientação vertical (FABRY, 2015).

As partículas na atmosfera, em geral, não são esféricas implicando que, quando comparados os dados coletados nas possíveis combinações de orientações, é possível obter informações sobre a forma e até mesmo tamanho do alvo.

Granizo, neve e chuva são alguns exemplos de alvos meteorológicos cuja identificação é beneficiada pelo radar polarimétrico. A Figura 4, apresenta uma comparação entre as perspectivas de um radar convencional e um radar de polarimétrico para estes hidrometeoros. Por exemplo, neve e granizo são semelhantes em sua dimensão horizontal (linha vermelha), sendo distinguíveis por sua dimensão vertical (linha azul),

assim como chuva e neve são distinguíveis na dimensão vertical mas semelhantes na dimensão horizontal. Assim, quando os sinais são combinados, os alvos antes muito semelhantes em sua detecção, são distinguíveis.

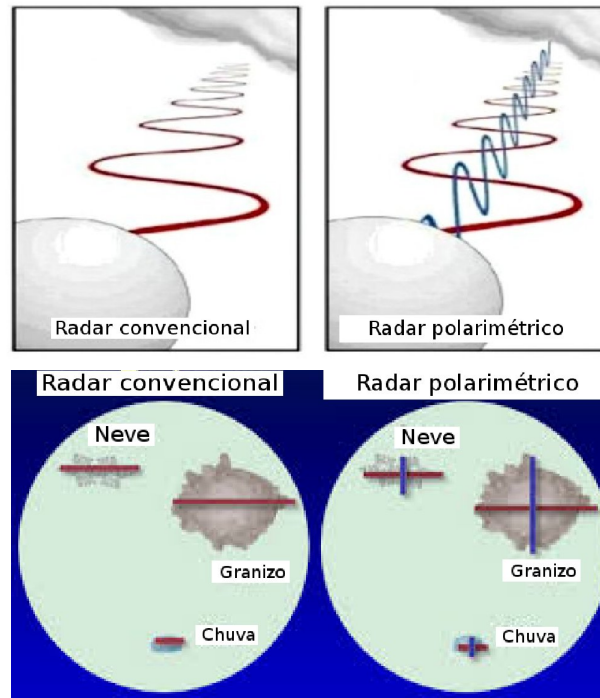


FIGURA 4: Comparação entre radares de dupla polarização e polarização simples na detecção de partículas.

FONTE: MetEd (2014) adaptado.

Devido às combinações de orientações possíveis, além dos momentos Doppler, o radar de dupla polarização gera novas variáveis tais como ZDR , $RHOHV$, $PHIDP$ e KDP , denominadas variáveis polarimétricas. Na Figura 5 estão exemplificadas as variáveis polarimétricas.

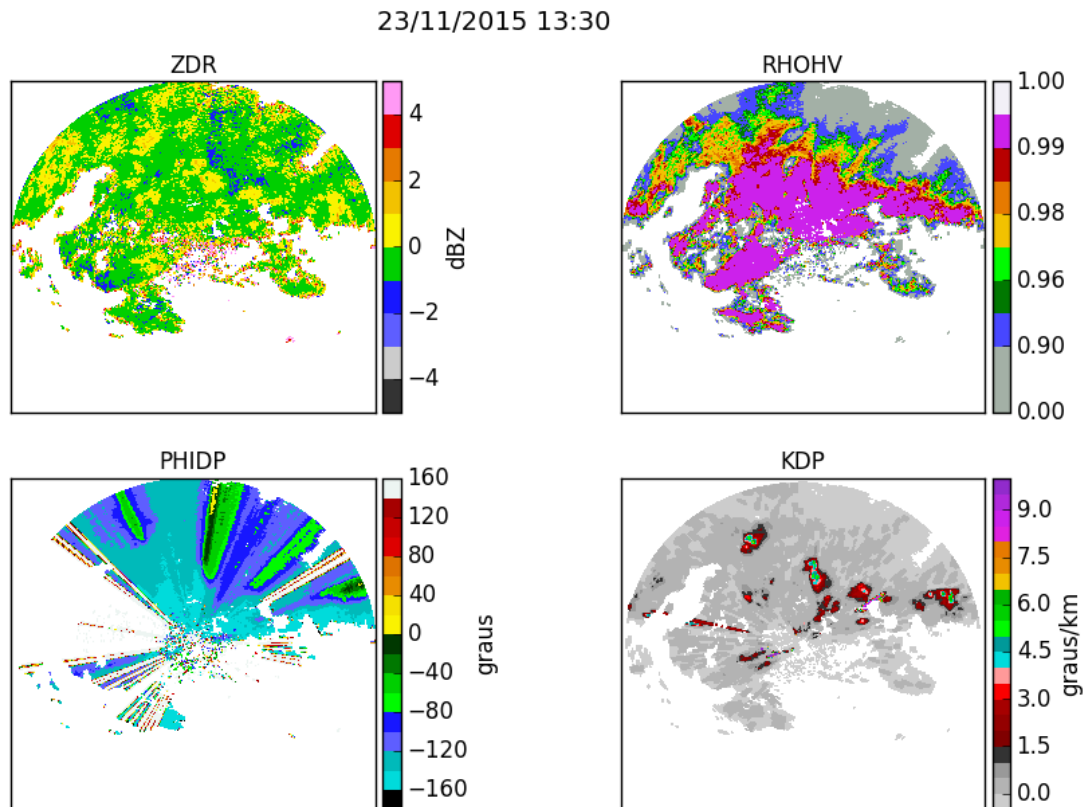


FIGURA 5: Visualização das variáveis polarimétricas.

FONTE: O autor (2017).

2.2.2.1 REFLETIVIDADE DIFERENCIAL (ZDR)

A Refletividade Diferencial (ZDR), é a diferença entre Z_{HH} e Z_{VV} , ou seja,

$$Z_{DR} = Z_{HH} - Z_{VV}. \quad (6)$$

Quando as partículas são quase esféricas, ZDR é próximo de 0 dB, no entanto, gotas de chuva são achatadas verticalmente (formato oblato) e quanto maiores e mais pesadas forem as gotas, mais oblato se tornam (FABRY, 2015). Valores típicos de ZDR variam de -4 dB a 7 dB, com valores acima de zero, em geral, associados a chuva (FABRY, 2015).

Portanto, ZDR contém informações sobre a forma das partículas, orientação e

índice de refração, tornando-se útil para diferenciar o que representa chuva, granizo, neve, e até mesmo alvos não meteorológicos (KUMJIAN, 2013a).

2.2.2.2 CORRELAÇÃO COPOLAR (RHOHV)

A Correlação Copolar (RHOHV), não possui unidade de medida e representa a correlação entre os sinais horizontais e verticais polarizados de Z, em um determinado ponto no espaço (RINEHART, 2004), cuja equação é definida por:

$$RHOHV = \frac{\langle S_{VV}S_{HH}^* \rangle}{\langle |S_{HH}|^2 \rangle^{1/2} \langle |S_{VV}|^2 \rangle^{1/2}}, \quad (7)$$

onde S representa a intensidade do sinal e S* é o conjugado complexo da intensidade do sinal, os índices representam a forma de emissão e de captação do sinal. No caso de S_{VV} , o sinal foi emitido e recebido verticalmente (FABRY, 2015).

Os valores de $RHOHV$ variam entre 0 e 1,0 e podem ser vistos como medida da similaridade entre os sinais horizontal e vertical de polarização (KUMJIAN, 2013a). Esta variável descreve características físicas do alvo, sendo mais próximo de 1,0 quanto mais uniforme for o alvo, pois os sinais vertical e horizontal tendem a serem iguais.

Por ser calculado por meio da correlação dos sinais, o RHOHV possui independência em relação à concentração de hidrometeoros na atmosfera e a efeitos de propagação, possuindo apenas dependência em relação às formas do hidrometeoros, suas inclinações e as misturas dos mesmos (BRINGI; CHANDRASEKAR, 2001).

2.2.2.3 DIFERENCIAL DE FASE (PHIDP)

Seja Φ a fase em que o pulso foi emitido e recebido, o diferencial de fase (PHIDP), medido em graus, representa a diferença entre as fases Φ_{HH} e Φ_{VV} para o trajeto completo do pulso (FABRY, 2015), sendo descrito por:

$$PHIDP = \Phi_{HH} - \Phi_{VV} \quad (8)$$

A variável $PHIDP$ está exemplificada na Figura 3, onde é possível notar que esta

variável possui variações radiais, pois apresenta mudanças cumulativas na diferença de fase para a viagem completa do pulso.

A *PHIDP* não raramente é substituída pela sua derivada espacial, a *KDP* (FABRY, 2015).

2.2.2.4 DIFERENCIAL DE FASE ESPECÍFICA (KDP)

A variável Diferencial de Fase Específica (KDP) é a derivada espacial de PHIDP (Equação 9) (FABRY, 2015).

$$KDP = \frac{1}{2} \frac{dPHIDP(r)}{dr} = \frac{PHIDP(r_2) - PHIDP(r_1)}{2(r_2 - r_1)} \quad (9)$$

Na Equação 9, r_1 e r_2 são dois bins consecutivos e a multiplicação por 2 expressa a dependência do sinal emitido e retornado. KDP tem como unidade de medida $^{\circ}/\text{km}$ e é um excelente estimador para precipitação, como apresentado por Ruzanski e Chandrasekar (2012), com valores acima de $2^{\circ}/\text{km}$ indicando quantidade significativa de água líquida e ou em formato oblato.

2.2.3 CISALHAMENTO

Devido ao desenvolvimento vertical, as tempestades convectivas estão associadas a cisalhamento, especialmente a baixo e médio níveis (FABRY, 2015).

Tradicionalmente, o cisalhamento é obtido pela diferença da velocidade Doppler entre dois pontos, dividido pela distância entre estes dois pontos, sendo possível descrever 3 campos de cisalhamento: vertical, radial e azimutal Newman et al. (2013), conforme descritos a seguir.

O cisalhamento azimutal (AZS) é a diferença da velocidade entre 2 bins à mesma distância de radar r , na mesma elevação θ e em azimutes consecutivos ϕ_i e ϕ_{i+1} e dividido pela distância d entre estes dois bins, como resumido pela Equação 10.

$$AZS = \frac{V_{(\theta, \phi_i, r)} - V_{(\theta, \phi_{i+1}, r)}}{dV_{(\theta, \phi_i, r), V_{(\theta, \phi_{i+1}, r)}}, \quad (10)$$

onde $V_{(\theta, \phi_i, r)}$ e $V_{(\theta, \phi_{i+1}, r)}$ representam, respectivamente, a velocidade Doppler nas coordenadas esféricas (θ, ϕ_i, r) e (θ, ϕ_{i+1}, r) .

O cisalhamento radial (RS) é a diferença da velocidade entre 2 bins consecutivos r_i e r_{i+1} , na mesma elevação θ e azimute ϕ e dividido pela distância d entre estes dois bins, como apresentado na Equação 11.

$$RS = \frac{V_{(\theta, \phi, r_i)} - V_{(\theta, \phi, r_{i+1})}}{dV_{(\theta, \phi, r_i), V_{(\theta, \phi, r_{i+1})}}, \quad (11)$$

onde $V_{(\theta, \phi, r_i)}$ e $V_{(\theta, \phi, r_{i+1})}$ representam, respectivamente, a velocidade Doppler nas coordenadas esféricas (θ, ϕ, r_i) e (θ, ϕ, r_{i+1}) .

O cisalhamento vertical (VS) é a diferença da velocidade entre 2 bins com o mesmo azimute ϕ e distância do radar r , porém em elevações consecutivas θ_i e θ_{i+1} e dividido pela distância d entre estes dois bins, relação expressa pela Equação 12.

$$VS = \frac{V_{(\theta_i, \phi, r)} - V_{(\theta_{i+1}, \phi, r)}}{dV_{(\theta_i, \phi, r), V_{(\theta_{i+1}, \phi, r)}}, \quad (12)$$

onde $V_{(\theta_i, \phi, r)}$ e $V_{(\theta_{i+1}, \phi, r)}$ representam, respectivamente, a velocidade Doppler nas coordenadas esféricas (θ_i, ϕ, r) e (θ_{i+1}, ϕ, r) .

As equações 10, 11 e 12 são campos escalares, que representam as variações de velocidade azimutal, radial e verticalmente. Essas variações podem indicar rotação, ou mesmo que existe diferença significativa entre a velocidade da tempestade em baixa, média e alta altitude.

2.3 DESCARGAS ELÉTRICAS ATMOSFÉRICAS (DEA)

As Descargas Elétricas Atmosféricas (DEA) são resultado do desenvolvimento de intensos centros de cargas na nuvem, que excedeu a capacidade de isolamento elé-

trico do ar resultando na dissipação dos centros de carga. Os tipos de DEA são classificados de acordo com local de onde se originam e onde terminam (LIMA, 2005).

As DEA são classificadas em Nuvem-Nuvem (*Intra Cloud* - IC) e Nuvem-Solo (*Cloud to Ground* - CG).

Os dados de DEA podem indicar a severidade de uma tempestade por sua posição e quantidade. Conforme apresentado por Holle et al. (1994), os raios tendem a se concentrar nas regiões convectivas e durante a maturidade de uma tempestade convectiva a quantidade de descargas do tipo CGp (tipo CG com polaridade positiva) se torna maior que a quantidade de descargas do tipo CGn (tipo CG com polaridade negativa), para depois decair novamente na fase de dissipação da tempestade.

O crescimento na densidade total de DEA (*total lightning* - TL), soma das densidades de CG e IC, também é um indicador de ETS possuindo resultados superiores às técnicas de detecção por CG (SCHULTZ; PETERSEN; CAREY, 2011). Este fenômeno é chamado de *lightning jump* (LJ), e ocorre minutos antes da ocorrência de um ETS (SCHULTZ; PETERSEN; CAREY, 2011).

De acordo com GATLIN e GOODMAN (2010), o LJ é um bom indicador de potencial ETS ainda que, em seu estudo, 10% dos ETS não foram precedidos pelo fenômeno, bem como 37% dos LJ foram alarme falso.

SCHULTZ et al. (2017) examinou 39 tempestades com dados de radar polarimétrico e densidade total de DEA a fim de estudar as diferenças entre tempestades que ocorre LJ e as que não ocorre LJ. As tempestades severas em que não se observou o LJ foram aquelas em que se observa correntes de ar que sobem em menor velocidade, ou de forma mista.

A densidade de descargas do tipo CG pode indicar tempestades elétricas, como mostrado por Liu e Heckman (2010), mas não necessariamente indica outros tipos de ETS, permanecendo importante o estudo do LJ quer seja da densidade total de raios (TL) quanto das descargas do tipo IC.

Neste capítulo foram mostradas a definição de ETS e uma breve revisão da pesquisa e previsão de ETSs, conceitos relacionados ao radar polarimétrico e a utilização de DEA para previsão de ETS. No próximo capítulo, serão apresentados conceitos de aprendizagem de máquina e as técnicas que serão utilizadas neste trabalho.

3 APRENDIZAGEM DE MÁQUINA (AM)

Aprendizado de Máquina (AM) é uma subárea de inteligência artificial que estuda formas de construir algoritmos que aprimoram o próprio desempenho em uma tarefa específica, utilizando-se de um conjunto finito de dados (MITCHELL, 1997). A AM se dedica a obter, por meio de princípios de inferência, conclusões genéricas a partir de um conjunto de exemplos (LORENA; CARVALHO, 2007).

Os algoritmos de AM podem ser usados para classificação binária e multiclasse, extração de características, regressão, criação de regras de associação, reconhecimento de padrão, compreensão e ou detecção de valores atípicos, sendo usada desde diagnósticos médicos a ajustes de funções (ALPAYDIN, 2010).

3.1 FUNDAMENTOS DE APRENDIZAGEM DE MÁQUINA

Em um determinado conjunto de dados, cada valor, característica ou aspecto que o representa é considerado um atributo. Uma instância é formada por valores de atributos referentes aos dados estudados. Assim, uma instância x , também denominada como exemplo, padrão ou dado, é uma tupla de valores de n atributos da forma (x_1, x_2, \dots, x_n) que descreve o objeto de interesse (HARRINGTON, 2010).

Uma vez que o aprendizado se baseia em experiência, a qualidade dos dados influencia diretamente no aprendizado, bem como na qualidade do modelo (HARRINGTON, 2010). Além do pré-processamento nos dados, por exemplo diminuindo a dimensão do espaço de entrada e retirada de dados discrepantes, é comum mapear cada característica da entrada no intervalo $[-1, 1]$ (MARSLAND, 2015). Além disso, ao se aplicar um técnica de AM a um conjunto de dados tem-se, por hipótese, que os atributos do conjunto de dados são suficientes para explicar o fenômeno de interesse

e que existem padrões o bastante para o aprendizado da técnica (representatividade estatística) (HARRINGTON, 2010).

Em geral, o conjunto de instâncias é dividido, pelo menos, entre conjunto de teste e conjunto de treinamento. O conjunto de treinamento é o conjunto de instâncias que são apresentados ao algoritmo de AM durante o processo de aprendizado, em geral, entre 60% e 90% do conjunto. Após o processo de aprendizado (geração do modelo) é apresentado o conjunto de teste verificando se o modelo continua a ser válido para novas instâncias, ou seja, se possui capacidade de generalização (HARRINGTON, 2010). Neste sentido, entradas semelhantes devem produzir saídas semelhantes (HAYKIN, 1999).

Quando um algoritmo possui baixa capacidade de generalização, pode haver um superajuste aos dados (*overfitting*), ou seja, a hipótese se especializou nos dados do conjunto de treinamento, perdendo sua capacidade de inferir sobre novas instâncias. Por outro lado, se o algoritmo de aprendizagem de máquina possui, após o aprendizado, uma baixa taxa de acerto mesmo no subconjunto de treinamento, significa que o mesmo possui subajustamento (*underfitting*) o que significa que o aprendizado não foi eficiente (MITCHELL, 1997).

O algoritmo utilizado para aprender a partir das instâncias apresentadas é chamado indutor. A resposta do indutor ao ser apresentado a uma instância é chamada saída. Além disso, após realizado o aprendizado, o indutor treinado pode ser considerado um modelo, pois para qualquer instância do objeto de interesse será possível induzir uma saída (BRUCE, 2001).

Os algoritmos de AM podem ser classificados de acordo com seu tipo de aprendizagem em supervisionado, não supervisionado, semi-supervisionado ou por reforço, de acordo com os dados disponíveis para a realização do processo de indução (ALPAYDIN, 2010).

No aprendizado supervisionado, também conhecido como aprendizagem com professor, as instâncias possuem como uma de suas características a saída desejada.

Esta saída pode ser binária, categórica, real ou um vetor. Assim, durante o processo de treinamento o indutor utiliza a informação da saída desejada para verificar seu desempenho e aprimorar sua indução (ALPAYDIN, 2010).

No aprendizado não supervisionado, as instâncias são constituídas apenas pelos atributos de entrada, ou seja, a saída desejada é desconhecida. Em geral, este tipo de aprendizado retorna padrões de agrupamento em conjunto de dados (ALPAYDIN, 2010).

O aprendizado semi-supervisionado é um misto do aprendizado supervisionado e não supervisionado. Muito utilizado para conjuntos de dados em que se possui poucos dados rotulados, porém um amplo conjunto de dados e pelo menos um exemplo de cada rótulo (BRUCE, 2001).

Na aprendizagem por reforço, o sistema de aprendizado deve aprender a escolher as ações interagindo com o ambiente (SUTTON; BARTO, 1998). Neste caso, o treinamento ocorre reforçando as saídas que melhorem o desempenho do algoritmo e ou penalizando as saídas menos vantajosas (ALPAYDIN, 2010).

3.2 CLASSIFICAÇÃO DE PADRÕES

O problema de classificação de padrão consiste em, a partir de uma entrada x , decidir a qual das N classes possíveis x pertence, tal que cada entrada pertença a apenas uma classe (MARSLAND, 2015).

A classificação de padrões está presente em diversas áreas como meteorologia, biologia, energia, *software*, danos estruturais, reconhecimento de imagem e medicina. Estes problemas podem ser resolvidos de diversas formas tais como método exato (determinístico), estatística, aprendizagem de máquina e métodos híbridos (unem uma ou mais técnicas na solução) (MARSLAND, 2015), (RAMAKRISHNAN, 2016).

Devido ao crescimento do volume de dados disponíveis nas últimas décadas, métodos de AM tem se tornado cada vez mais utilizados na área de reconhecimento e

classificação de padrões (MARSLAND, 2015).

Um algoritmo de aprendizagem de máquina aplicado em tarefas de classificação de padrões deve ser capaz de: aprender a partir de um conjunto de treinamento e criar regras capazes de relacionar os valores dos atributos precursores de uma instância do conjunto de treinamento aos valores de seu atributo classe (MICHALSKI; CARBONELL; MITCHELL, 1986), tal que a classe é o atributo que descreve o fenômeno de interesse (HARRINGTON, 2010).

A técnica support vector machine (SVM) é indicada como uma das melhores técnicas de classificação de padrões, pois quase sempre apresenta excelentes resultados, embora seja uma técnica muito custosa para grandes conjuntos de dados e sensível a dados ruidosos e mal classificados (MARSLAND, 2015).

Trafalis, Adrianto e Richman (2007) utilizam SVM para previsão de tornados apresentando uma discussão sobre a dificuldade em se classificar os dados para a inserção na técnica. Gavrishchaka e Ganguli (2001) utilizam SVM para obter conhecimento a partir de dados meteorológicos multidimensionais comparando a técnica com o uso de RNA como sendo de mesma performance ou superior. Radhika e Shashi (2009) utilizam SVM para previsão de temperatura. Damian (2011) e Neto (2008) utilizam SVM para classificação de chuva convectiva e estratiforme e detecção de alvos não meteorológicos em dados de radar respectivamente, mostrando a versatilidade da SVM.

As redes neurais artificiais são técnicas de AM que podem ser usadas tanto para resolução de problemas de previsão quanto de classificação, sendo mais usadas em previsão. Por exemplo, os estudos de Litta, Idicula e Mohanty (2013), Narvekar e Fargose (2015), Maqsood e Abraham (2004), Hung, Babel e Tripathi (2009) e HALL, BROOKS e DOSWELL (1998) propõem RNAs para a resolução de problemas de previsão e LACORTE (2011) apresenta a utilização de RNA para a previsão de tornados, um problema de classificação.

As redes neurais são muito utilizadas por serem capazes de aprender mesmo que os dados sejam ruidosos e ou inconsistentes, bastando, para isso, que disponham de

um conjunto de teste grande o suficiente (MARSLAND, 2015). A seguir são detalhadas as RNAs e a técnica SVM.

3.3 REDES NEURAIS ARTIFICIAIS (RNA)

Como definido por Haykin (1999), uma Rede Neural Artificial (RNA) é um sistema paralelamente distribuído e maciçamente interligado por unidades simples de processamento denominadas neurônios, cuja modelagem é inspirada no funcionamento dos neurônios biológicos.

A Figura 6 apresenta a estrutura de um neurônio e seus principais componentes: sinais de entrada x , pesos sinápticos w , combinador linear *Soma*, bias b_k , função de ativação $f(\cdot)$, e saída y .

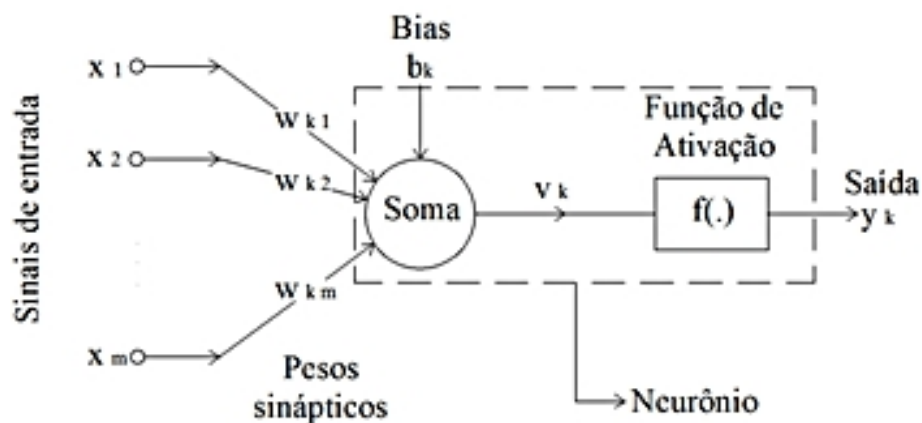


FIGURA 6: Representação de um neurônio artificial.

FONTE: Araújo, Neto e Filho (2015).

Os pesos sinápticos indicam o peso que cada componente da entrada terá no combinador linear, representando a força de ligação entre os neurônios da rede. O combinador linear é responsável pela soma dos sinais de entrada, ponderados pelo peso de cada sinapse. A função de ativação gera a saída do neurônio a partir do valor da função soma e a bias associada, assegurando também que a saída esteja em um intervalo de amplitude finita. O bias é um termo linear a ser acrescentado ao combinador linear, representa um fator com explicação externa.

Matematicamente, dada uma entrada $x \in \mathfrak{R}^n$, tal que $x = (x_1, x_2, \dots, x_n)$ e possui n atributos, um vetor de pesos $w_k \in \mathfrak{R}^n$, onde $w = (w_{k1}, w_{k2}, \dots, w_{kn})$. Então, a saída deste neurônio (ou resposta) é dada por

$$y_k = f(v_k) = f(u_k(x) + b_k),$$

onde

$$u_k(x) = \sum_{i=1}^n w_{ki}x_i$$

é o combinador linear, $b_k \in \mathfrak{R}$ é o bias e $f(\cdot)$ é a função de ativação. O índice k indica que se trata do k -ésimo neurônio, uma RNA pode ser construída por um ou mais neurônios.

Alguns exemplos de função de ativação são: degrau (Equação 13), linear (Equação 14), sigmóide (Equação 15) ou tangente hiperbólica (Equação 16). Outras funções podem ser utilizadas, desde que sejam diferenciáveis, pois a diferenciabilidade é necessária para o processo de minimização do erro. Em geral, para saída $\in [0, 1]$, a função de ativação mais usada é a sigmoideal (MARSLAND, 2015).

$$f(v_k) = \begin{cases} 0, & \text{se } v_k \geq 1, \\ 0, & \text{se } v_k < 0. \end{cases} \quad (13)$$

$$f(v_k) = v_k \quad (14)$$

$$f(v_k) = \frac{1}{1 + e^{-v_k}} \quad (15)$$

$$f(v_k) = \frac{1 - e^{v_k}}{1 + e^{v_k}}, \quad (16)$$

onde $v_k = u(x) + b_k = \sum_{i=1}^n w_{ki}x_{ki} + b_k$.

Alguns exemplos de RNAs citadas por Marsland (2015) são as seguintes:

- *Perceptron*: rede supervisionada que utiliza um ou mais neurônios independentes para o processamento, capaz de resolver problemas linearmente separáveis.
- *Multilayer perceptron (MLP)*: é um melhoramento da *Perceptron*, incluindo camadas intermediárias de neurônios, tornando possível sua aplicação a problemas não linearmente separáveis.
- *Self organization maps*: rede não supervisionada, utilizada para reconhecimento de padrões, agrupamento de dados e ou extração de características.
- Elman: RNA recorrente, utiliza a própria saída como parte da entrada, amplamente utilizada para previsão de séries temporais.
- RNA de base radial (RBF): método semi-supervisionado, muito utilizada para aproximação de funções.

O aprendizado de uma RNA é armazenado nos pesos sinápticos que são modificados (atualizados) a cada iteração de aprendizagem, até que um ou mais critérios de parada pré-determinados sejam satisfeitos. Parar ao atingir erro inferior a um determinado valor, parar ao atingir uma determinada quantidade de iterações, parar ao atingir determinado uso de memória ou parar ao se obter o mínimo de erro em um conjunto de validação, são alguns critérios de parada para a fase de aprendizagem de uma técnica de AM (HAYKIN, 1999).

3.3.1 ALGORITMO BACKPROPAGATION

O algoritmo *Backpropagation* é um método baseado no gradiente descendente para calcular as derivadas da função de erro em relação aos pesos sinápticos, a fim de encontrar um conjunto de pesos que permita minimizar o erro através da exposição repetida das instâncias (BISHOP, 1995). Segundo Haykin (1999) e Marsland (2015), trata-se do principal algoritmo utilizado no treinamento de RNAs supervisionadas.

O aprendizado de uma RNA por meio do *Backpropagation* pode ser organizado pelos seguintes passos:

1. Apresentar uma instância de treinamento à rede;
2. Determinar a saída y da rede para a entrada apresentada;
3. Calcular o erro na saída da rede comparando com a saída desejada d ;
4. Atualizar os pesos dos neurônios mais externos para diminuir o erro;
5. Retropropagar o erro para os neurônios mais internos;
6. Ajustar os pesos dos neurônios mais internos;
7. Repetir os passos anteriores até que o critério de parada seja satisfeito.

Os passos 1 a 3, representam a fase *forward* e os passos 4 a 6, a fase *backward* do algoritmo. Cada uma das componentes w_{ki} do vetor de pesos sinápticos w_k , é atualizada na t -ésima iteração por:

$$w_{ki}(t+1) = w_{ki}(t) - \eta \frac{\partial E(t)}{\partial w_{ki}} \quad (17)$$

Na Equação 17, o termo η é um valor arbitrariamente fixado chamado taxa de aprendizagem, $E(t)$ é o valor do erro e $-\frac{\partial E(t)}{\partial w_{ki}}$ é a direção para qual ocorre o maior decréscimo do erro (gradiente descendente).

3.3.2 MULTILAYER PERCEPTRON (MLP)

A RNA conhecida como *Multilayer Perceptron* (MLP) é uma RNA de aprendizado supervisionado, que possui uma ou mais camadas de neurônios entre a camada de entrada e de saída, muito utilizada por sua simplicidade e fácil implementação (MARS-LAND, 2015). A Figura 7 exemplifica a arquitetura de uma MLP.

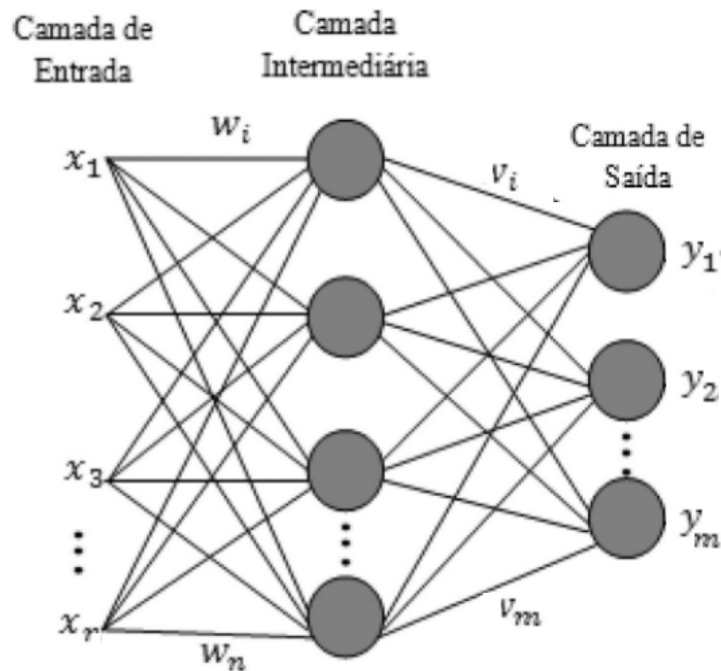


FIGURA 7: Exemplo de arquitetura de uma MLP.

FONTE: Coutinho, Silva e Delgado (2016) adaptado.

A MLP exemplificada na Figura 7 possui 3 camadas: a camada de entrada, de saída e uma camada intermediária, também conhecida como camada de neurônios escondidos ou simplesmente camada escondida, com n neurônios escondidos, r atributos de entrada e m saídas.

Por arquitetura de uma RNA entende-se a estrutura da rede, incluindo padrão de ligação entre os neurônios, algoritmo de aprendizado, tipo de aprendizado, forma e quantidade de entradas, formato e quantidade de saídas e funções de ativação. Nesta arquitetura (Figura 7), dada uma entrada $x = (x_1, x_2, \dots, x_r)$ vem que

$$z_j = f(z_j^* + \theta_{aj}) = f(w_j^T x + \theta_{aj}), \quad (18)$$

em que $f(\cdot)$ é uma função de ativação diferenciável, $w_j^T x$ é o produto escalar entre o vetor de pesos w_j e a entrada x e θ_{aj} é o termo linear (bias) relacionado a j -ésima posição.

Assim, para cada k-ésima saída y_k tem-se que

$$y_k = f(y_k^*) = f(v_k^T z + \theta_{bk}), \quad (19)$$

onde θ_{bk} é a k-ésima posição do vetor de bias relacionado a camada de saída, $v_k^T z$ o produto interno entre o vetor de pesos v_k e o vetor z cujos termos são obtidos pela Equação 18 e $f(\cdot)$ uma função de ativação.

Uma forma de atualizar os pesos sinápticos durante o aprendizado da MLP é utilizar o algoritmo backpropagation (seção 3.3.1). Assim, os passos 1 a 3 do *backpropagation* são descritos pelas equações 18 e 19.

O passo 4 se refere à atualização dos pesos da camada mais externa, ou seja de v_1, v_2, \dots, v_m no exemplo apresentado pela Figura 7.

Sejam E_k o erro associado a saída y_k , v_{jk} a componente k do vetor de pesos v_j e d_k a saída desejada para y_k . Pela regra da cadeia

$$-\frac{\partial E_k}{\partial v_{jk}} = \frac{\partial E_k}{\partial y_k^*} \frac{\partial y_k^*}{\partial v_{jk}} = \frac{\partial E_k}{\partial y_k^*} z_j.$$

O erro da saída é a diferença entre a saída desejada e a saída obtida. Assim, para seguir na direção contrária ao erro, defini-se

$$\frac{\partial E_k}{\partial y_k} = -(d_k - y_k).$$

Então, o ajuste de pesos para v_{jk} é dado por

$$\Delta v_{jk} = \eta (d_k - y_k) f'(y_k^*) z_j. \quad (20)$$

Suponha $f(\cdot)$ a função sigmoideal como função de ativação em todos os neurônios da camada de saída. Assim,

$$f(y_k^*) = \frac{1}{1 + e^{-y_k^*}}$$

implica em

$$f'(y_k^*) = \frac{e^{-y_k^*}}{(1 + e^{-y_k^*})^2}. \quad (21)$$

Substituindo 21 em 20 e simplificando o termo de atualização dos pesos da camada de saída é dado por

$$\Delta v_{jk} = -y_k(1 - y_k)(d_k - y_k)z_j. \quad (22)$$

Assim, v_{jk} atualizado é da forma

$$v_{jk} = v_{jk} + \Delta v_{jk}. \quad (23)$$

Analogamente, obtemos que o termo de atualização para os pesos w_{ij} da camada escondida, com função de ativação também sigmoidal, é dado por

$$\Delta w_{ij} = \eta \sum_k (d_k - y_k) y_k (1 - y_k) v_{jk} z_j (1 - z_j) x_i, \quad (24)$$

e w_{ij} atualizado dado por

$$w_{ij} = w_{ij} + \Delta w_{ij}. \quad (25)$$

A seguir apresenta-se os passos de treinamento de uma MLP Backpropagation.

1. Inicialize os pesos das conexões, dos bias e da taxa de aprendizado;
2. Calcule as saídas na camada escondida (Equação 18) e na camada de saída (Equação 19);
3. Calcule as correções das conexões da camada de saída (Equação 23);
4. Calcule as correções das conexões da camada intermediária (Equação 25);
5. Caso o critério de parada não tenha sido satisfeito repita os passos de 2 a 5.

Segundo BISHOP (1995) a taxa de aprendizado η é um valor positivo e pequeno. Nestas condições, o vetor de pesos converge para o ponto em que o erro é minimi-

zado. No entanto, se η for muito pequeno o aprendizado se torna lento e se η for grande o aprendizado pode divergir.

As funções de ativação, bem como o número de neurônios e camadas escondidas dependem do conjunto de dados estudado. Quanto maior o número de neurônios e camadas escondidas maior será a complexidade da rede, podendo causar até mesmo superajustamento (BISHOP, 1995). Na prática, não é necessário utilizar mais do que uma camada escondida (teorema universal da aproximação) sendo possível obter um modelo complexo incluindo neurônios o suficiente na camada escondida (BISHOP, 1995), (MARSLAND, 2015). Além disto, qualquer função suave pode ser aproximada por uma combinação linear de sigmóides localizadas (MARSLAND, 2015).

O principal critério de parada ideal para a MLP é o mínimo da curva do erro do conjunto de validação. O conjunto de validação é extraído do conjunto de teste que tem seu erro calculado a cada iteração do processo de aprendizagem. Caso o erro no conjunto de validação passe por um mínimo significa que a rede está perdendo sua capacidade de generalização e continuar o treinamento pode implicar em superajustamento (MARSLAND, 2015).

Outros critérios de parada para a MLP são: parar após um número de iterações; parar após uso determinado de CPU; parar após uma função de erro ser menor que um determinado valor ou parar quando a função de erro não ter decréscimo (BISHOP, 1995). Estes critérios de parada são usados como apoio ao critério principal, garantindo, por exemplo, que o treinamento ocorra em tempo finito (MARSLAND, 2015).

3.4 SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine (SVM) ou máquina de vetores de suporte, é uma técnica de AM, de aprendizado supervisionado, desenvolvida por V. Vapnik que busca minimizar o erro no conjunto de treinamento (risco empírico), bem como o erro no conjunto de teste (risco na generalização)(VAPNIK; CHERVONENKIS, 1974).

Da mesma forma que para as RNAs, o objetivo da técnica consiste em obter equilíbrio entre esses erros, evitando o excesso de ajuste (*overfitting*) em prol da capacidade de generalização (HAYKIN, 1999).

Neste caso, o problema de *overfitting* pode ser caracterizado como a memorização das peculiaridades e ruídos dos dados ao invés de detectar características gerais que permitem a classificação de padrão (SMOLA et al., 2000).

De acordo com sua formulação matemática, a SVM corresponde a obter um hiperplano ótimo tal que melhor segregue as classes, ou seja, encontrar um hiperplano com a máxima distância euclidiana do conjunto de treinamento (HAYKIN, 1999).

Considera-se o conjunto de treinamento finito tal que a k -ésima instância é da forma (x_k, y_k) , $x_k \in \mathfrak{R}^n$, $y_k \in \{+1, -1\}$, onde x_k é o padrão de entrada k e y_k é a correspondente resposta desejada. A forma geral da função de decisão linear é dada por

$$D(x) = w^T x + b, \quad (26)$$

onde $w \in \mathfrak{R}^n$ é um vetor de pesos e $b \in \mathfrak{R}$ é o termo independente (bias).

A estratégia de decisão é dada por:

$$\begin{cases} D(x) > 0 & \text{para } y_k = +1 \\ D(x) < 0 & \text{para } y_k = -1, \end{cases}$$

ou seja, as classes são separadas pelo hiperplano $D(x)$ (SEMOLINI, 2002).

A função de decisão linear (Equação 26) é conhecida como SVM de margens fixas, não permitindo que existam dados não classificados. Além disso, o hiperplano resultante não é único, podendo ter infinitas inclinações (MARSLAND, 2015).

Adicionando uma margem $\delta > 0$, admitindo que $y_k = +1$ se $D(x) \geq \delta$ e $y_k = -1$ se $D(x) \leq -\delta$ e redimensionando w e b adequadamente tem-se uma nova estratégia de

decisão:

$$\begin{cases} w^T x + b \geq +1 & \text{para } y_k = +1 \\ w^T x + b \leq -1 & \text{para } y_k = -1, \end{cases}$$

onde, as classes são separadas pelo hiperplano $w^T x + b$, porém com uma margem flexível (SEMOLINI, 2002).

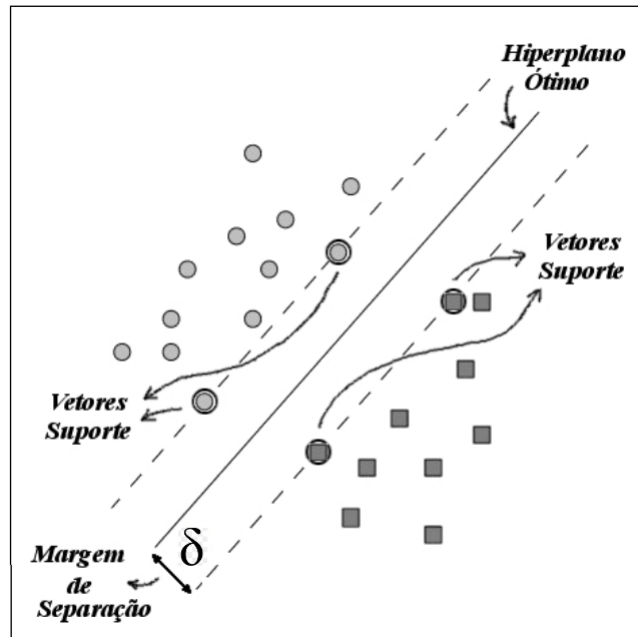


FIGURA 8: Hiperplano ótimo de separação e margem δ .

FONTE: Semolini (2002) adaptado.

A Figura 8 ilustra um hiperplano ótimo e as margens de separação $\delta > 0$. Seja o hiperplano ótimo a superfície de decisão tal que maximiza δ , então este hiperplano é único e os padrões de entrada que se situam sobre os planos mais extremos da margem de separação são denominados vetores suporte (SEMOLINI, 2002).

No caso de conjuntos não linearmente separáveis é acrescentada uma variável ξ_k (variável de folga) a cada padrão de entrada x_k , tal que a função de penalidade,

$$f(\xi) = \sum_k \xi_k,$$

representa a soma dos erros de classificação (HAYKIN, 1999). Em outras palavras, as folgas ξ_k relaxam o problema permitindo que uma determinada quantidade padrões

não seja devidamente classificada.

Obter o hiperplano ótimo da SVM corresponde a obter um hiperplano que maximize a distância δ . Ou seja, minimização do erro de treinamento através da solução do problema (HAYKIN, 1999):

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^p \xi_i \\ \text{Sujeito a:} \quad & y_i(w^T x_i + b) \geq 1 + \xi_i, \\ & \xi_i \geq 0, \quad \forall i = 1, \dots, p. \end{aligned} \tag{27}$$

No problema 27 o termo $C \in \mathfrak{R}_+$ é denominado constante de regularização, a qual determina a importância maior ou menor às folgas, possibilitando que o modelo seja mais ou menos sensível a pontos com padrão não identificado, ou mesmo mal classificados (HAYKIN, 1999).

O problema de minimização 27 é denominado problema primal do tipo quadrático. A resolução deste problema não é trivial, e uma das possíveis formas de resolvê-lo é transformá-lo em um problema mais simples de resolver, através de sua formulação dual. Além de apresentar restrições mais simples, a forma dual do problema 27 possibilita a apresentação do problema de otimização em termos de produtos internos (SMOLA et al., 2000).

O problema primal (Equação 27) tem a função objetivo convexa e os pontos que satisfazem as restrições formam um conjunto convexo que possui um único mínimo global. Ou seja, satisfaz as hipóteses para ser transformado em um problema dual utilizando multiplicadores de Lagrange (HAYKIN, 1999).

Seja

$$L(w, b, \xi, \alpha) = \frac{1}{2} w^T w + C \sum_i \xi_i - \sum_i (\alpha_i (y_i (w^T x_i + b) + 1 + \xi_i))$$

a função lagrangeana associada ao primal 27.

As derivadas parciais de $L(w, b, \xi, \alpha)$ são dadas por:

$$\begin{cases} \frac{\partial L(w, b, \xi, a)}{\partial w} = w - \sum_i y_i \alpha_i x_i \\ \frac{\partial L(w, b, \xi, a)}{\partial b} = \sum_i y_i \alpha_i \\ \frac{\partial L(w, b, \xi, a)}{\partial \xi} = -\sum_i \alpha_i + C \end{cases} .$$

Igualando as derivadas parciais de $L(w, b, \xi, \alpha)$ a zero, segue que

$$\begin{cases} w - \sum_i y_i \alpha_i x_i = 0 \Rightarrow w = \sum_i y_i \alpha_i x_i \\ \sum_i y_i \alpha_i = 0 \Rightarrow \sum_i y_i \alpha_i = 0 \\ -\sum_i \alpha_i + C = 0 \Rightarrow \sum_i \alpha_i = C \end{cases} . \quad (28)$$

Substituindo 28 na função lagrangeana $L(w, b, \xi, \alpha)$ tem-se que

$$\begin{aligned} L(w, b, \xi, \alpha) &= \frac{1}{2} w^T w + C \sum_i \xi_i - \sum_i (\alpha_i (y_i (w^T x_i + b) + 1 + \xi_i)) \\ &= \frac{1}{2} \sum_i y_i \alpha_i x_i \sum_i y_i \alpha_i x_i + \sum_i \alpha_i \sum_i \xi_i - \sum_i (\alpha_i (y_i (\sum_i y_i \alpha_i x_i x_i + b) + 1 + \xi_i)) \\ &= \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j x_i x_j - \sum_i \alpha_i \sum_i \xi_i - \sum_i \alpha_i + \sum_i \alpha_i \sum_i \xi_i - \sum_i \alpha_i y_i (\sum_i y_i \alpha_i x_i x_i + b) \\ &= \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j x_i x_j - \sum_i \alpha_i. \end{aligned} \quad (29)$$

Substituindo 29 e suas condições em 27 obtem-se o problema dual:

$$\begin{aligned} \max_{\alpha} \quad & \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j (x_i x_j) \\ \text{Sujeito a:} \quad & \sum_i y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i, j = 1, \dots, p. \end{aligned} \quad (30)$$

A formulação apresentada em 30 é denominada forma dual de 27.

Sejam $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_N^*)$ a solução para 30 e w, d e ξ solução para 27. Da teoria

de otimização (condições de Kühn-Tucker), segue que

$$\alpha_i^*(y_i(w^T x + b) - 1 + \xi_i) = 0, \quad \forall \quad i = 1, 2, \dots, N,$$

ou seja, os vetores de suporte são os únicos a contribuírem para a determinação do plano separador, pois α^* é diferente de zero apenas se α^* pertence aos planos $D(x) + \delta$ ou $D(x) - \delta$. Desta forma, o parâmetro w apresentado na superfície de decisão 26 pode ser obtido pela primeira equação de 28 e b é dado por (SMOLA et al., 2000):

$$b = \frac{1}{N} \sum_{\alpha_i^* \in \alpha^*} \frac{1}{y_i} - w^T \alpha_i^*.$$

As folgas e margem flexibilizam o SVM, porém ainda é necessário que o conjunto seja linearmente separável, o que não representa a maioria dos problemas de classificação (MARSLAND, 2015). Uma forma de tornar conjuntos não linearmente separáveis em linearmente separáveis é mapeando as entradas em um espaço de características de alta dimensão. Assim, é possível obter superfícies de decisão não lineares no espaço de entrada, separados por um hiperplano no espaço de características (SMOLA et al., 2000).

Seja $\Phi : \mathfrak{R}^m \rightarrow \mathfrak{R}^n$ uma função não linear em \mathfrak{R}^m , simétrica, que satisfaz a desigualdade triangular e com produto interno definido em \mathfrak{R}^m , com $n > m$. Uma função kernel K é tal que, para todo $x_i, x_j \in \mathfrak{R}^m$ tem-se $K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j)$ (SMOLA et al., 2000).

A formulação que considera as funções *kernel* para o problema 30 é dada por:

$$\begin{aligned} \max_{\alpha} \quad & \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} y_i y_j \alpha_i \alpha_j (\Phi x_i^T \Phi x_j) \\ \text{Sujeito a:} \quad & \sum_i y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i, j = 1, \dots, p. \end{aligned} \tag{31}$$

Assim como a função de ativação na MLP, a escolha da função *kernel* depende da complexidade dos dados. Na Tabela 1 são listados alguns exemplos de função *kernel*, onde γ , d e r são parâmetros reais (SMOLA et al., 2000).

TABELA 1: Exemplos de função *kernel*

<i>kernel</i>	Função $K(x_i, x_j)$
Linear	$K(x_i, x_j) = x_i \cdot x_j$
Polinomial	$K(x_i, x_j) = \gamma((x_i + x_j) + r)^d$
Gaussiana	$K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$
Sigmoidal	$K(x_i, x_j) = \tanh(\gamma(x_i \cdot x_j) + r)$

A escolha dos parâmetros para a função *kernel*, bem como a constante de penalização C são críticos na obtenção de um bom classificador e variam de acordo com o problema e tipo de dados (VAPNIK; CHERVONENKIS, 1974). Uma forma de se obter tais parâmetros, quando não se possui conhecimento prévio dos mesmos, é a busca em grade. A busca em grade consiste em executar uma busca exaustiva da combinação de valores para uma lista de parâmetros, a fim de obter a maior acurácia possível (LIEPERT, 2003).

Embora o SVM seja, originalmente, definido para classificação de padrões, o SVM pode ser usado também para classificação multiclases ou regressão, por meio de algumas extensões (adaptações) (MARSLAND, 2015).

Neste capítulo algumas técnicas foram definidas para resolução de problemas de classificação de padrões, a seguir são apresentados os materiais e métodos relacionados ao estudo.

4 MATERIAIS E MÉTODOS

Nos capítulos anteriores foram apresentadas a natureza dos dados e as técnicas utilizadas neste trabalho. Neste capítulo, serão detalhadas especificações dos dados e metodologias utilizadas.

4.1 MATERIAIS

4.1.1 ESPECIFICAÇÕES DO RADAR

Neste trabalho, foram utilizados os dados coletados pelo radar de Cascavel (Figura 9), um radar polarimétrico, banda S (frequência de emissão entre 2-4 GHz), operado desde 2013, pelo Sistema Meteorológico do Paraná (SIMEPAR). Esse radar possui alcance para monitorar grande parte do Estado do Paraná, operando com um raio de 480 km e uma elevação ou 240 km e entre 10 a 17 elevações. Para este radar, a cada sete minutos e meio, é gerado um arquivo com os momentos coletados, em formato NetCDF CF/Radial (RAL.UCAR.EDU, 2017). Além dos momentos doppler e variáveis polarimétricas. O radar possui outros produtos, como por exemplo estimativa de precipitação.

Os momentos Z e V são coletados por este radar com emissão e recepção na mesma direção. Ou seja, temos Z_{HH} , Z_{VV} , V_{HH} e V_{VV} .

Para este trabalho optou-se por utilizar Z_{HH} e V_{HH} que são mencionados como Z e V respectivamente. A não utilização de Z_{VV} e V_{VV} pode ser complementada com a utilização das variáveis polarimétricas, que expressam de forma resumida as informações relevantes em relação às diferentes formas de emissão e captação de pulso.



FIGURA 9: RADAR POLARIMÉTRICO DE CASCAVEL.

FONTE: JORNAL DO OESTE (2015).

4.1.2 ESPECIFICAÇÕES DOS DADOS DE DESCARGAS ELÉTRICAS ATMOSFÉRICAS

A Earth Networks Total Lightning Network (ENTLN) é a mais ampla rede global de detecção de descargas atmosféricas dos tipos IC e CG, com alta densidade de sensores e operando em tempo real (LIU; HECKMAN, 2012).

Os sensores da ENTLN possuem frequência de detecção variando entre 1Hz a 12Hz, o que permite ao sensor detectar DEA com pulsos fracos, principalmente do tipo IC. Toda a informação retornada ao sensor é utilizada na detecção e localização, o que permite acurada localização e detecção de falsos pulsos (LIU; HECKMAN, 2012).

Os dados de DEA disponibilizados pela ENTLN são dos tipos *Intra-Cloud* positivo (ICp), *Intra-Cloud negativo* (ICn), *Cloud-to-Ground positivo* (CGp) e *Cloud-to-Ground negativo* (CGn).

4.1.3 ESPECIFICAÇÕES DOS EVENTOS DE TEMPO SEVERO

Dentre os mais de 100 eventos relatados pela Defesa Civil, no período de janeiro de 2015 a junho de 2016, apenas 27 ocorreram no range volumétrico do radar de estudo (um raio de 240 km distante da posição do radar) e alguns desses dados possuem apenas uma elevação (não são do tipo volumétrico). Resultando em 10

ETSs diferentes que se enquadram nas especificações de disponibilidade, localização e resolução, indicados na Tabela 2 (CIVIL;2015,2016).

TABELA 2: Eventos estudados

Data	Hora	Tipo de evento	Cidade afetada
02/01/2015	00:00	Vendaval	Foz do Iguaçu
13/07/2015	19:23	Vendaval	Matelândia
13/07/2015	20:15	Vendaval	Ampére
13/07/2015	21:53	Vendaval/Chuva Intensa	Francisco Beltrão
14/07/2015	10:15	Tornado	Francisco Beltrão
23/11/2015	13:30	Vendaval	Umuarama
26/11/2015	23:46	Chuva intensa	Francisco Beltrão
28/12/2015	21:15	Vendaval	Francisco Beltrão
27/02/2016	09:53	Chuva Intensa	Francisco Beltrão
22/03/2016	03:15	Tornado	Capitão Leônidas Marques

Para a finalidade de estudar a previsão de ETS a curtíssimo prazo são considerados apenas dados coletados em até 30 minutos antes da ocorrência do ETS. Assim, cada ETS possui entre 3 e 5 arquivos com dados de radar correspondentes a instantes de tempo diferentes, com diferença de 7,5 ou 15 minutos. Cada arquivo possui a coleta volumétrica do dado de radar, iniciando no instante de tempo que descreve, totalizando 43 arquivos com os dados de radar.

4.2 MÉTODOS

A metodologia aplicada pode ser descrita pelos seguintes passos:

1. Coleta dos dados;
2. Pré-processamento dos dados (correção de dados faltantes e resolução, normalização, cálculo dos cisalhamentos, retirada de ruído e redução da dimensionalidade da entrada);

3. Inclusão das classes ao conjunto de dados;
4. Separação do conjunto de teste e treinamento;
5. Treinamento das técnicas MLP e SVM;
6. Identificação de regiões de ETS por LJ e densidade de CG;
7. Comparação dos modelos obtidos entre si e com a identificação obtida pela utilização de LJ e densidade de CG;
8. Estudo e análise dos resultados obtidos.

4.2.1 PRÉ-PROCESSAMENTO

Devido às condições de coleta de dados do radar, a resolução espacial dos dados não é uniforme (a coleta de dados ocorre em coordenadas esféricas). Esta não-uniformidade implica em maior resolução para os dados próximos ao radar em detrimento aos mais distantes e é possível que alguns dados estejam faltando. A fim de padronizar a resolução dos dados e tratar valores faltantes, os mesmos foram interpolados em uma grade cartesiana, com o auxílio do software Pyart (HELMUS et al., 2013).

Para a interpolação em grade cartesiana via PyArt, optou-se por uma função de interpolação exponencial, pois o tempo de processamento foi inferior em 10 vezes às demais funções disponíveis (interpolação polinomial).

A interpolação em grade cartesiana, no entanto, inviabiliza a utilização quantitativa dos momentos Doppler V e W , sendo substituídos pelo módulo dos campos de cisalhamento (seção 2.2.3), a fim de reconstruir características importantes do campo de velocidade.

Após a interpolação, excluiu-se os pontos onde Z é inferior a 30dBZ por se tratarem de dados que não são representativos em uma tempestade convectiva Maddox (1980).

Desta forma, cada ponto da grade cartesiana, a ser considerado, possui valores de Z, ZDR, RHOHV, KDP, HMAX, AZS, RS, VS e altitude.

A Figura 10 mostra um exemplo das entradas Z, ZDR, RHOHV, KDP, HMAX, AZS, RS e VS em grade cartesiana, em altitude de 3000 metros, após serem interpolados. Nesta figura, os círculos mostram a região próxima a um dos ETSs estudados.

Para levar em consideração o crescimento vertical da tempestade, incluiu-se ainda a cada ponto de grade o parâmetro HMAX, que representa a altitude do ponto de maior refletividade por coluna (mesmo par de latitude e longitude).

Após a normalização dos dados, estudou-se a correlação simples entre os atributos a fim de diminuir a dimensionalidade da entrada.

4.2.2 INCLUSÃO DAS CLASSES

A Defesa Civil, bem como noticiários indicam a ocorrência de ETS em região aproximada, em geral apenas num município, não sendo possível obter informação exata sobre a localização e hora da ocorrência do ETS. Logo, os dados apresentados na Tabela 2 são aproximados.

O objetivo de “identificar regiões de ocorrência de ETS dentro de 30 minutos”, implica que a região de interesse é a que possui o atributo “ocorreu ETS dentro de 30 minutos”. Assim, podemos estudar o problema de identificação de ETS como um problema de classificação binária, ou seja, indica-se com 1 as regiões de interesse e zero as que não são de interesse. Neste sentido, convém indicar como região que não é de interesse onde “é desconhecida a informação da ocorrência de ETS dentro de 30 minutos”.

A fim de separar os dados de estudo, com base na localização aproximada dos ETSs, tomou-se como região de interesse toda a região compreendida por um círculo de centro nas coordenadas geográficas centrais dos municípios afetados e raio de 0,5 grau, com o valor de 0,5 obtido empiricamente. Outras metodologias foram utilizadas

como, por exemplo, o acompanhamento do núcleo da tempestade, no entanto, dentre as metodologias aplicadas, a identificação por meio deste círculo obteve os melhores resultados, embora não seja capaz de garantir a não inclusão de dados mal classificados.

Desta forma, todos os dados internos a este círculo são considerados como positivos no que se referem ao ETS indicado (possuem a condição “ocorreu ETS dentro de 30 minutos”) e recebem valor de saída esperada igual a 1. Os demais dados são considerados negativos (“é desconhecida a informação da ocorrência de ETS dentro de 30 minutos”) e recebem valor de saída esperada igual a 0.

É importante ressaltar que a classificação utilizada indica apenas que é conhecida uma região onde ocorreu ao menos um ETS, o que pode não implicar na não ocorrência nas demais regiões do mesmo volume de radar. Assim, é esperada uma taxa razoável de entradas mal classificadas.

Inicialmente, a proporção de dados na classe 1 era de menos de 10%, a fim de aumentar a representatividade destes dados esta proporção foi aumentada, escolhendo menos dados entre aqueles classificados como 0.

4.2.3 TREINAMENTO

Após o pré-processamento dos dados, as técnicas SVM e MLP (apresentadas nas seções 3.4 e 3.3.2) foram treinadas.

Dos 43 arquivos disponíveis na forma descrita em 4.1.3, foram utilizados 33 para o treinamento das técnicas e 10 para validação (conjunto de teste). Para os 33 arquivos do conjunto de treinamento, respeitou-se a proporção de aproximadamente 25% de dados classificados como positivos e os demais 75% classificados como negativos. Foram testadas as proporções de 50%, 60%, 70%, 75%, 80%, 85% e 90% de dados classificados como 0, sendo obtido empiricamente que a proporção 75% apresenta melhores resultados para o problema e modelos utilizados.

Para a escolha dos arquivos de teste sorteou-se um evento que deveria estar contido integralmente ao conjunto de teste. Os arquivos referentes ao evento de 13/07/2015 às 20:15 pertencem integralmente ao conjunto de teste, e os 6 demais arquivos foram escolhidos por sorteio sem importar a qual evento estão relacionados.

O conjunto de entrada consiste de vetores que contém os valores normalizados de Z, ZDR, RHOHV, KDP, HMAX, AZS, RS e VS para cada ponto de grade para cada arquivo e nesta ordem.

Para treinamento da MLP e da SVM foi utilizado o *software* livre Scikitlearn, do Python (SMOLA, 2004). Esse pacote possui vasta gama de técnicas de AM implementadas e de fácil utilização (MARSLAND, 2015).

A MLP foi treinada iniciando-se com 1 neurônio na camada escondida e acrescentando 1 a cada novo treinamento. Após a inserção de 50 neurônios, admitiu-se como modelo, a MLP de menor erro. O número máximo de 50 neurônios foi obtido de forma empírica, pois não houve diminuição do erro após a inserção de 43 neurônios.

A otimização de parâmetros para a SVM utilizou a otimização em grade para as 4 funções *kernel* disponíveis no *software* Scikitlearn, apresentadas na Tabela 1. Na Tabela 3 são apresentados o máximo, mínimo e a variação utilizada para os parâmetros γ , C e d na otimização em grade. Em relação aos parâmetros ε e r os mesmos foram mantidos em 0,1 e 0, respectivamente. Os intervalos e variações para a otimização em grade de parâmetros do SVM também foram obtidos de forma empírica.

TABELA 3: Especificação dos parâmetros γ , C e d para aprimoramento da regressão.

Parâmetro	Máximo	Mínimo	Variação
γ	100	0	0,01
C	100	0	1
d	10	2	1

Ao final do treinamento, aceitou-se como melhor modelo em cada uma das técnicas

cas aquele que apresentou menor erro quadrático médio em relação aos demais.

4.2.4 ANÁLISE

Após a obtenção dos modelos denotados por M-MLP e M-SVM, os dados são classificados segundo os modelos para estudo de métricas e análise comparativa com dados de raios.

Como este estudo tem como objetivo identificar uma região de ocorrência ou não de ETS, as métricas tradicionais não se aplicam aos modelos obtidos, pois estas métricas se baseiam em análise de forma pontual. Assim, são analisados os resultados por regiões identificadas. Para indicar falhas e acertos dos modelos uma comparação foi feita com tempestades onde ocorrem LJ ou alta densidade de CG durante os 30 minutos de estudo. Estas identificações são comparadas às identificações apresentadas pelos modelos M-MLP e M-SVM.

Além disso, um dos objetivos do trabalho é verificar a existência de padrões claros nos dados classificados positivamente. Portanto, faz-se ainda um breve estudo das entradas classificadas como positivas em cada um dos modelos.

Neste capítulo, foram apresentados os materiais e métodos que são utilizados neste trabalho. No próximo capítulo são apresentados os resultados obtidos.

14/07/2015 09:52

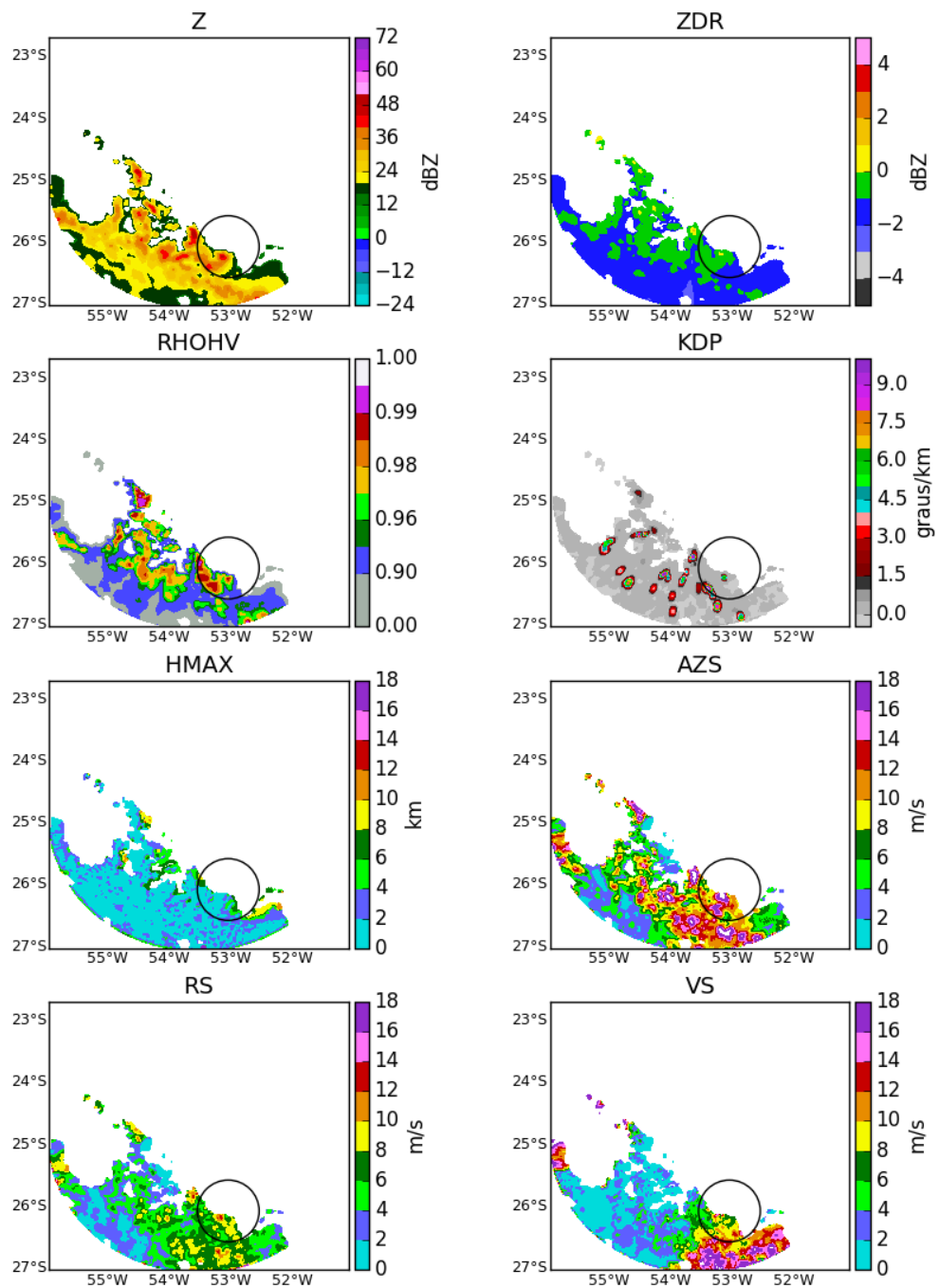


FIGURA 10: Dados de 14/07/2015 às 9:52 UTC em grade cartesiana, onde Z, ZDR, RHOHV, KDP, HMAX, AZS, RS, e VS representam refletividade, refletividade diferencial, correlação copolar, diferencial de fase específica, altitude do máximo valor de refletividade por coluna, cisalhamento azimuthal, cisalhamento radial e cisalhamento vertical.

FONTE: O autor (2017).

5 RESULTADOS

No capítulo 4, foi apresentada a metodologia adotada neste trabalho. Neste capítulo são apresentados e discutidos os resultados obtidos pela aplicação da metodologia apresentada.

5.1 PRÉ-PROCESSAMENTO

A qualidade dos dados usados para o treinamento de uma técnica de AM causa grande influência na qualidade do modelo obtido. A fim de excluir padrões desnecessários ou discrepantes, os padrões que aparecem exclusivamente classificados como negativos foram excluídos do conjunto de dados. Esta exclusão foi realizada por meio dos máximos e mínimos valores de cada atributo, sem deixar de considerar as características físicas que o dado representa.

Por exemplo, nos dados classificados como positivos, o atributo KDP nunca é inferior a $-3,2 \text{ }^\circ/\text{km}$ e Z nunca é superior a 65 dBZ. Ou seja, entradas que contém KDP inferior a $-3,2 \text{ }^\circ/\text{km}$ ou Z superior a 65dBZ são sérias candidatas a serem excluídas do conjunto de dados. No entanto, altos valores de Z indicam altas taxas de precipitação e valores negativos de KDP, em geral, não indicam alvos meteorológicos (seção 2.2.2). Portanto, instâncias com valores de KDP inferiores a $-3,2 \text{ }^\circ/\text{km}$ são excluídos dos conjuntos de dados, mas não são excluídos padrões com valores de Z superior a 65 dBZ, devido às características físicas que este dado representa.

Após retirar os dados discrepantes e aqueles que não são de interesse houve uma diminuição de 1% na quantidade de dados.

Nas técnicas de AM, quando uma característica é combinação linear de outra, a técnica pode atribuir excessivo valor a esta característica em detrimento das outras.

Quando se detecta que uma característica é linearmente dependente de outra é suficiente considerar apenas uma delas, neste caso, dizemos que ocorreu redução na dimensão de entrada.

A fim de verificar quais características são linearmente dependentes utilizou-se o estudo de correlação simples. A Figura 11 ilustra o resultado deste estudo, por meio dos valores de correlação e os gráficos de dispersão.

Na Figura 11, a maior correlação (0,83) ocorre entre as atributos RS e AZS e é inferior a 0,95, não sendo aconselhável reduzir a dimensionalidade da entrada. Além disso, é possível notar que o atributo que, em geral, possui as maiores correlações é Z, indicando que pode ser um dos atributos que mais contribui para o modelo.

5.2 MODELOS M-MLP E M-SVM

Como descrito no capítulo 4, o treinamento das técnicas ocorreu com a apresentação dos dados de 33 arquivos do radar utilizado, com aproximadamente 25% dos dados positivos. A Figura 12 apresenta um exemplo de saída dos modelos obtidos por MLP (M-MLP) e por SVM (M-SVM).

Na Figura 12 são apresentadas as saídas para M-MLP e M-SVM aplicadas aos dados de radar de 28/12/2015 às 21:15, aproximadamente enquanto ocorria um ETS do tipo vendaval na região de Francisco Beltrão (região interna ao círculo). Neste exemplo tem-se que M-MLP não identifica a região de interesse (interna ao círculo), e que ambos os modelos também identificam regiões inicialmente classificadas como não ocorrência de ETS.

Além disso, mesmo quando a região de ocorrência de ETS é identificada, como ocorre pela identificação de M-SVM na Figura 12, não necessariamente todos os pontos pré-classificados como positivos possuem saída positiva para o modelo.

Desta forma, as regiões de interesse em que ocorre ao menos um ponto interno identificado pelo modelo como positivo, a identificação é tomada como correta, pois

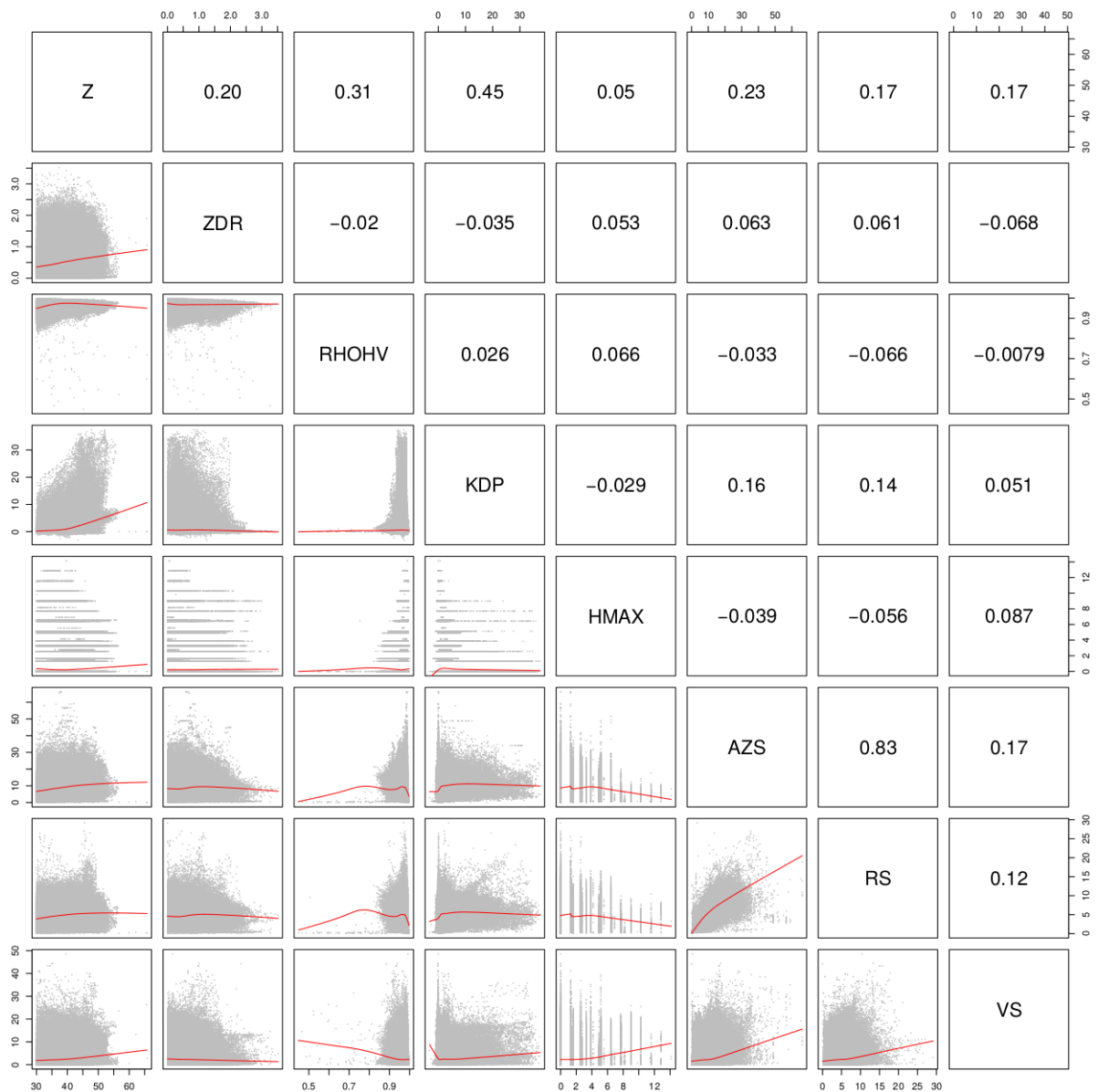


FIGURA 11: Correlação entre os atributos de entrada. As posições da diagonal apresentam os atributos cuja correlação está na linha e coluna correspondente acima da diagonal e o gráfico de dispersão e a respectiva curva de regressão estão na linha e coluna correspondente abaixo da diagonal.

FONTE: O autor (2017).

identificou a a região.

Considerando o conjunto de treinamento para M-MLP, 7 dos 33 arquivos não tiveram a região do ETS identificada e esses arquivos correspondem a apenas dois eventos: 26/11/2015 às 23:46 e 28/12/2015 às 21:15.

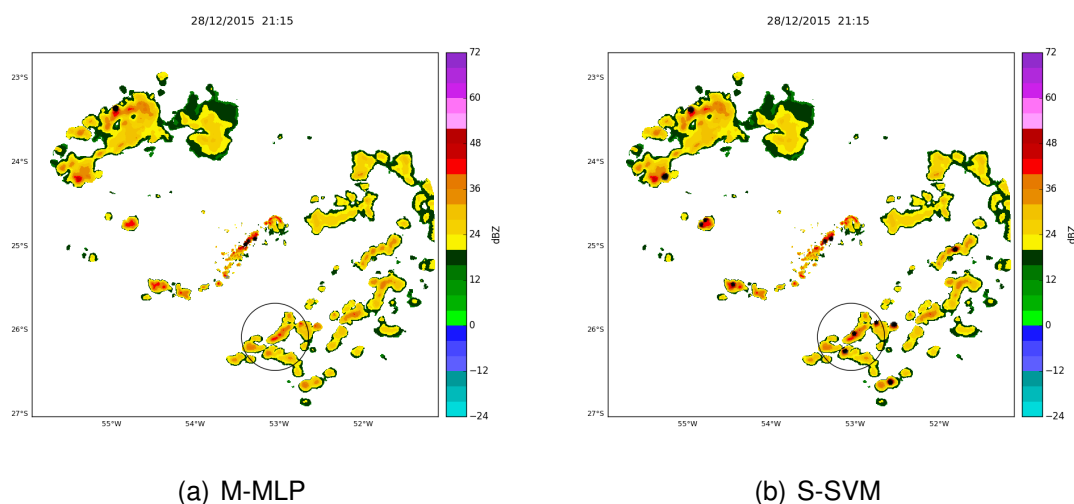


FIGURA 12: Saída obtida com a utilização dos modelos M-MLP e S-SVM respectivamente. As imagens representam os pontos identificados pelos respectivos modelos (em preto) sobre a imagem do atributo refletividade. A região circulada mostra onde ocorreu ETS dentro de 30 minutos.

FONTE: O autor (2017).

Para o modelo M-SVR, 3 dos 33 arquivos não tiveram a região do ETS identificada, esses 3 arquivos correspondem a apenas um evento: 26/11/2015 às 23:46.

A Tabela 4 mostra que, para as regiões de interesse, M-SVM possui identificação superior a M-MLP, com acerto de 93,02%.

TABELA 4: Proporções de acerto para cada um dos modelos

Modelo	Conjunto de Treinamento	Conjunto de Teste	Conjunto Geral
M-MLP	78,78 %	90,00 %	81,40 %
M-SVM	90,00 %	100,00 %	93,02 %

O único ETS que não foi identificado por M-SVM também não foi identificado por M-MLP. Este ETS, ocorrido em 28/12/2015 às 21:15, está relacionado a única tempestade estudada que amadureceu em menos de 30 minutos e que possui valores de Z abaixo de 45 dbZ. A Figura 13 mostra a evolução desta tempestade, durante os 30 minutos antes da ocorrência do ETS, identificado como do tipo "chuva intensa". Embora o ETS de 26/11/2015 às 23:45 não tenha sido totalmente identificado pelos modelos M-MLP e M-SVM, ambos os modelos identificam a região deste evento em pelo me-

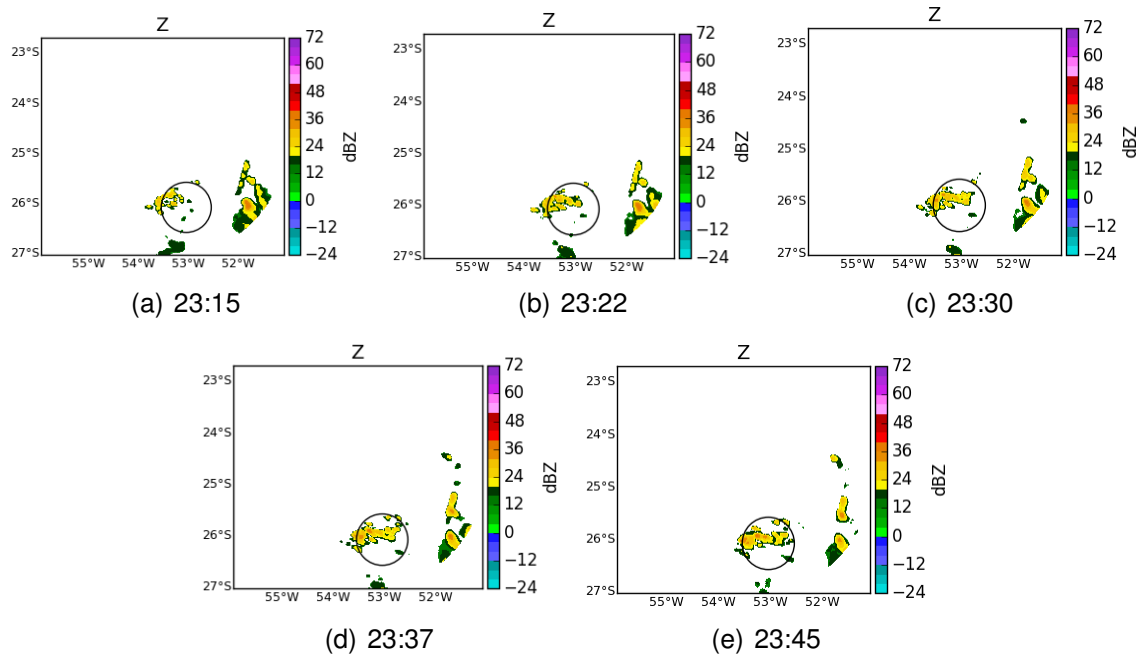


FIGURA 13: Tempestade de 26/11/2015, que ocasionou um ETS às 23:45, interno ao círculo das figuras. Este ETS não foi identificado pelos modelos M-MLP e M-SVM.

FONTE: O autor (2017).

nos um arquivo de radar entre 30 e 22 minutos antes do ETS, mostrando que o evento possui o padrão identificado pelos modelos, mas não por todo o seu desenvolvimento, nem durante a sua ocorrência.

Para M-MLP e M-SVM, o ETS de 26/11/2015 às 23:46 se trata de um padrão atípico, pois possui baixos valores de Z (comparado com os demais casos), teve rápido amadurecimento e é do tipo “chuva intensa” (arquivos do tipo “chuva intensa” representa apenas 3 dos 10 ETS estudados). Uma forma de melhorar os desempenhos de M-MLP e M-SVM neste tipo de dado seria incluir outros dados semelhantes ao conjunto de treinamento.

Em termos de identificação ponto a ponto, a Tabela 5, apresenta a quantidade de pontos inicialmente classificados positivamente, mas cuja saída dos modelos é negativa (falsos negativos). Nesta tabela, nota-se que M-MLP possui mais falsos negativos que M-SVM e que, em geral, pelo menos 13,18% dos dados da região de interesse não são identificados.

TABELA 5: Porcentagem de dados internos às regiões de interesse estudadas que são identificados pelos modelos

Modelo	Conjunto de Treinamento	Conjunto de Teste	Conjunto Geral
M-MLP	14,78 %	06,14 %	13,18 %
M-SVM	32,14 %	06,72 %	27,46 %

A porcentagem de falsos negativos pode estar diretamente relacionada com a escolha da forma de pré-classificação (tomar todos os pontos da região de interesse como positivos), pois é possível que nem todos os pontos classificados como positivos representem o ETS.

Além disto, regiões em que se desconhece a existência de ETS dentro dos próximos 30 minutos foram apresentadas às técnicas tendo 0 como valor desejado de saída. O que não necessariamente implica que as regiões não representem provável ETS, ou seja, não deveriam ser identificadas. A Tabela 6 apresenta a porcentagem de dados que foram identificados pelos modelos, mesmo tendo sido classificados como negativos no pré-processamento.

TABELA 6: Porcentagem de dados externos às regiões de interesse estudadas que são identificados pelos modelos

Modelo	Conjunto de Treinamento	Conjunto de Teste	Conjunto Geral
M-MLP	13,75 %	19,18%	14,97 %
M-SVM	10,95 %	09,82 %	10,70 %

A Tabela 6 indica que, dos dados em que se desconhece a informação de ocorrência de ETS, pelo menos 10,70% (em M-MLP) são identificados pelos modelos, podendo ou não representarem um ETS.

Combinando os resultados contidos nas tabelas 4, 5 e 6 é possível concluir que embora o modelo M-MLP identifique mais pontos com o padrão 1, este modelo não necessariamente identifica mais regiões de interesse. Neste sentido, pode-se dizer que o M-SVM é superior na identificação dos ETSs pré-classificados, pois identifica

mais regiões de interesse conhecidas.

5.3 COMPARAÇÃO ENTRE M-MLP E M-SVM AOS DADOS DE DESCARGAS ELÉTRICAS ATMOSFÉRICAS

Todas regiões que não correspondem aos eventos estudados não possuem informação prévia sobre a ocorrência de ETS. Assim, não é possível garantir que em tais regiões não ocorreram ETS, sendo um dos objetivos deste trabalho extrapolar os resultados da região de interesse, analisando também regiões identificadas com falsos positivos.

Como apresentado na seção 2.3, os dados de DEA podem indicar a ocorrência de ETS. Em particular pelo uso do LJ (*lightning jump*) e pela comparação com os dados de CG. Embora não seja possível obter a informação da ocorrência de ETS de forma direta, esta informação pode ser obtida de forma indireta ao se comparar as regiões identificadas com DEA.

Desta forma, todas as tempestades (convectivas ou não) presentes nos 43 dados de radar tiveram os dados de DEA estudados por LJ e CG, tal que LJ foi analisado apenas em regiões em que não ocorre alta densidade de CG.

Conforme discutido na seção 2.3, a ocorrência de ETS é possível sem que ocorra LJ, mas sempre que ocorre alta densidade CG pode-se dizer que se trata de um ETS do tipo tempestade elétrica. Portanto, a classificação por DEA ainda pode conter dados mal classificados, embora em pequena escala. Ou seja, nada pode ser afirmado sobre regiões identificadas pelos modelos mas que não atendem as condições de DEA para ser considerado ETS.

A Tabela 7 apresenta a porcentagem de regiões identificadas cujos dados de DEA confirmam o ETS (Modelo detectou) e a porcentagem de regiões em que os dados de DEA correspondem a ETS e os modelos não indicam a região (Modelo não detectou).

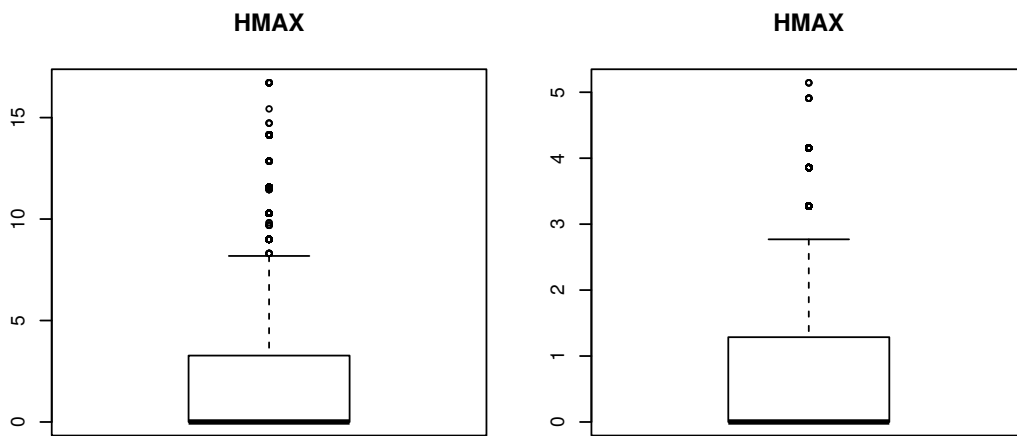
TABELA 7: Comparação entre identificação de ETS por DEA e os modelos M-SVR e M-MLP

Conjunto	Modelo	Modelo detectou	Modelo não detectou
Treinamento	M-MLP	75,00 %	32,50 %
	M-SVM	75,00 %	23,94 %
Teste	M-MLP	57,31 %	30,88 %
	M-SVM	55,42 %	23,33 %
Geral	M-MLP	68,58 %	32,02 %
	M-SVM	67,84 %	23,76 %

Portanto, embora o M-MLP tenha 68,58% de suas identificações endossadas pelos dados de DEA, o modelo deixa de identificar 32,02% das regiões identificadas por DEA. Com o uso de M-SVM, 67,84% das regiões identificadas concordam com os dados de DEA, ou seja, 33,16% das regiões identificadas por M-SVM possuem indícios de estarem indevidamente identificadas. Analogamente, 23,76% estão indevidamente não identificadas.

5.4 ESTUDO DAS ENTRADAS POSITIVAS

A fim de observar relações entre as entradas dos modelos que resultam em identificação positiva foram estudados os boxplots das entradas que resultam em saída positiva para cada um dos modelos: M-MLP e M-SVM. Neste estudo, apenas a entrada HMAX apresentou um padrão visível. O boxplot dos valores do atributo HMAX, que produz saída positiva para cada um dos modelos pode ser visto na Figura 14.

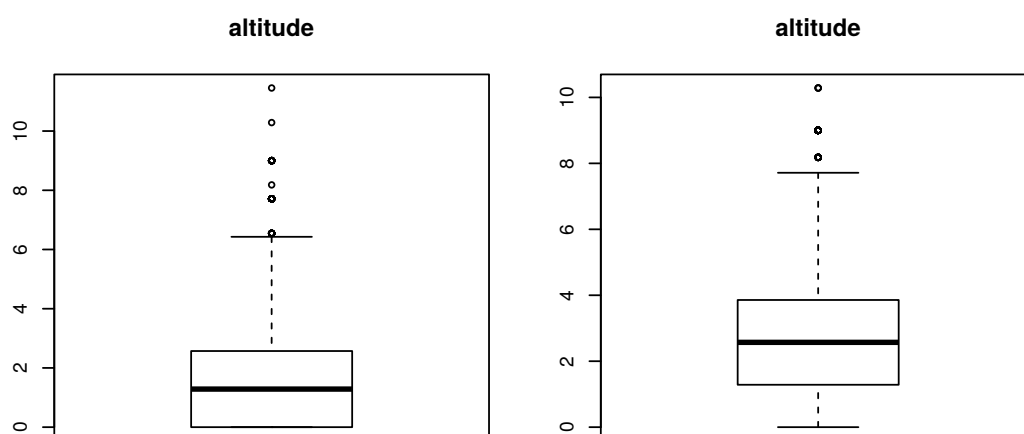


(a) Boxplot dos valores de HMAX que resultam em saída positiva pelo modelo M-MLP. (b) Boxplot dos valores de HMAX que resultam em saída positiva pelo modelo M-SVM.

FIGURA 14: Boxplot dos valores de HMAX que resultam em saída positiva pelos modelos M-MLP e M-SVM.

FONTE: O autor (2017).

A Figura 14 mostra que em 75 % dos dados classificados como positivos pelo modelo M-MLP e M-SVM, têm valor de HMAX (altitude do máximo valor de refletividade) inferior ou igual a 3,13 km e 1,24 km, respectivamente. Este valor de HMAX pode indicar que as características que indicam a ocorrência de ETS estão em baixas ou médias altitudes, fato confirmado pela Figura 15.



(a) Boxplot dos valores de altitude dos pontos que resultam em saída positiva pelo modelo M-MLP. (b) Boxplot dos valores de altitude dos pontos que resultam em saída positiva pelo modelo M-SVM.

FIGURA 15: Boxplot dos valores de altitude dos pontos que resultam em saída positiva pelos modelos M-MLP e M-SVM.

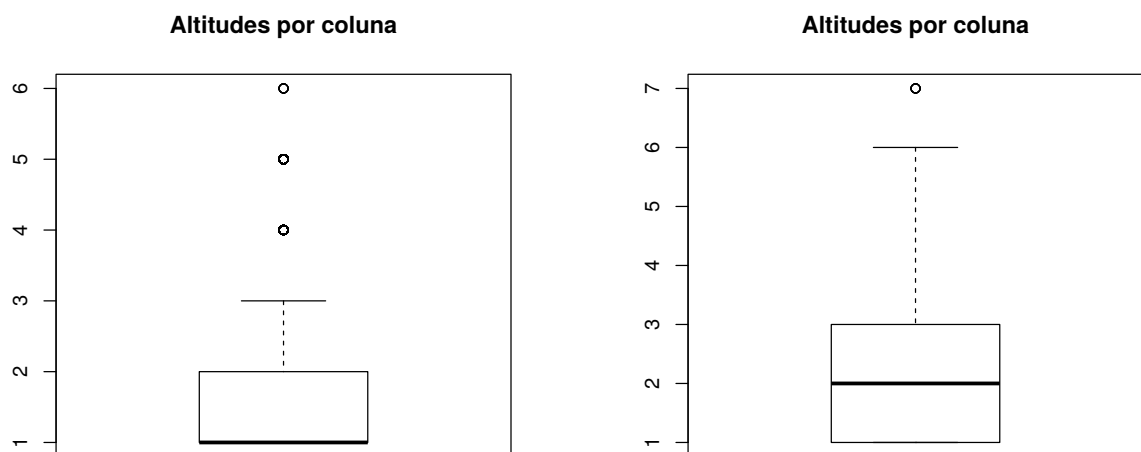
FONTE: O autor (2017).

A Figura 15, apresenta o boxplot da altitudes das entradas que produzem saídas positivas e possuem mediana (50 % dos casos) 2,73 km e 1,76 km para M-SVM e M-MLP respectivamente. Isto significa que as características detectáveis como padrão para estes modelos se encontram, muitas vezes, em baixas altitudes.

Para o desenvolvimento dos modelos foram utilizados os dados em todas as altitudes disponíveis, a fim de que não ocorresse grande perda de informação. Os boxplots relacionados às altitudes (Figura 15) para cada um dos modelos sugerem que existe um subgrupo de altitudes que são suficientes, em 75 % dos casos) para o modelo, não sendo necessário o estudo (previsão) de todas as altitudes disponíveis. Portanto, seria suficiente para 75% dos casos o estudo até 2,85 km (M-MLP) e 4 km (M-SVM) de altitude. Embora a escolha de apenas uma altitude possa não ser suficiente para detectar os padrões, é possível obter um subgrupo de altitudes, diminuindo, por exemplo, o custo computacional de previsão, ou ainda o volume de armazenamento dos dados.

De fato, 89,58 % dos eventos, identificados por DEA e que não foram identificados por M-SVM, estão a mais de 120 km de distância do radar. Para o M-MLP, a proporção fica em 46,56 %, podendo, este percentual de possível mal classificação ser devido ao fato de que a 120 km do radar, a altitude do feixe do mesmo já se encontra a mais de 2 km e, como já foi discutido, para os modelos M-MLP e M-SVM os dados mais representativos se encontram a baixas altitudes.

Em relação à quantidade de pontos por coluna que são identificados pelos modelos (Figura 16), em 75% dos pares de coordenadas, o M-MLP identifica no máximo dois pontos por coluna e o M-SVM no máximo 3. Isto não implica que dados de toda coluna não influenciem na classificação, mas sim que para os modelos obtidos o padrão detectado está presente em não mais que 3 altitudes.



(a) Boxplot da quantidade de pontos identificados para M-MLP por coluna.

(b) Boxplot da quantidade de pontos identificados para M-SVM por coluna.

FIGURA 16: Boxplot da quantidade de pontos por coluna identificados pelos modelos.

FONTE: O autor (2017).

Com os resultados apresentados podemos afirmar que o modelo M-MLP se mostrou inferior ao modelo M-SVM, pois identifica um evento a menos dentre os estudados, tem índice de identificação inferior se comparada a identificação por DEA e

identifica menos pontos por altitude se comparado por M-SVM. No entanto, o modelo M-SVM ainda não é suficiente, pois deixa de identificar ao menos 23,76% (Tabela 7) dos eventos que foram identificados por DEA e 6,92% dos dados estudados, bem como possui 33,16% (Tabela 4) de prováveis falsos positivos (Tabela 7).

A porcentagem de acerto dos modelos mostra, com a devidas ressalvas, que os modelos são capazes de auxiliar a previsão de ETS, conforme os objetivos deste trabalho, pois mostra que se pode identificar um padrão comum às variáveis de radar na ocorrência de ETS. Porém, estes modelos ainda não são capazes de serem autônomos, sendo não mais que uma contribuição de apoio à decisão, ou ainda um ponto de partida para novo estudo de identificação.

A imprecisão da pré-classificação dos dados, bem como o fato de ser o primeiro estudo com estas características implicou em métricas não muito boas para os modelos, que são sensíveis a dados mal classificados. No entanto, considerando a porcentagem estimada de dados mal classificados (em torno da quantidade de falsos negativos e falsos positivos) a obtenção de padrão por parte do modelo é aceitável.

6 CONCLUSÕES

Este trabalho teve como objetivo estudar e utilizar técnicas aprendizagem de máquina para classificar tempestades convectivas ETS dentro de 30 minutos, utilizando dados de radar polarimétrico.

Inicialmente realizou-se um estudo sobre AM e radar para a formulação da metodologia. A metodologia precisou ser adaptada diversas vezes devido à imprecisão da pré-classificação dos dados, conforme mostrado na seção 4.1.

Assim, os dados foram estudados e as técnicas SVM e MLP escolhidas devido ao bom desempenho em problemas semelhantes. Após a aplicação das técnicas foram obtidos dois modelos de predição: M-MLP e M-SVM. Nenhum dos modelos é capaz de fazer predição automática, porém se apresentaram excelentes métodos de apoio à previsão.

Além disso, os modelos reforçaram que a ocorrência de um ETS está relacionado a mais de um ponto por coluna, bem como que os padrões de convecção se apresentam em baixas e médias altitudes, reforçando que as técnicas de AM são capazes de reproduzir características físicas do sistema em que se aplica. Analogamente, a melhor identificação próximo ao radar corrobora com características físicas do problema e qualidade do sinal de radar.

Inicialmente, desejou-se utilizar os dados de DEA para pré-classificação, porém optou-se por utilizá-los para validação. Com as soluções apresentadas neste trabalho obteve-se resultados muito bons, inclusive na identificação de tempestades elétricas, que não correspondiam a nenhum ETS do conjunto de teste.

Vale ressaltar que, ao se escolher as técnicas de AM, levou-se muito em consideração que os dados de estudos eram poucos (apenas 10 eventos e não igualmente

espaçados temporalmente) e que provavelmente muitos dos dados estavam mal classificados. Logo, a técnica MLP poderia não ser bem treinada e a SVM não responder bem. No entanto, o que se observou foi M-SVM com identificação dos ETSs estudados acima de 90 % e concordância com DEA de 67,84 % em geral. Como 89,58 % dos dados em que M-SVM identifica sem a concordância com DEA estão a mais de 120 km do radar, pode-se dizer que a taxa de acerto de M-SVM até 120 km do radar é muito boa.

As técnicas de AM necessitam de grande volume de dados a fim de que seu aprendizado seja garantidamente eficiente. Em termos de volume de dados, foram utilizados 43 arquivos com grades de 12.902.400 pontos, no entanto que representavam apenas 10 eventos. Assim, para garantir a melhoria dos modelos sugere-se a inclusão de novos eventos, bem como mais dados do tipo "chuva intensa" (com o pior desempenho de classificação pelos modelos).

Os bons resultados na classificação, reforçam a robustez da técnicas, em como as mesmas foram capaz de aprender e gerar um modelo razoável (ou muito bom até 120 km do radar, como M-SVM) mesmo com a possibilidade de conterem dados mal classificados. Esta é outra contribuição do trabalho, mostrar que a SVM pode obter resultados satisfatórios mesmo com dados mal classificados.

A pré-classificação dos dados, como foi realizada, embora tenha permitido bons resultados, pode estar diretamente relacionada com os erros do modelo e, portanto, sugere-se obter outras formas de pré-classificar os dados a fim de garantir que as entradas estejam corretamente classificadas.

Uma das hipóteses do uso de AM é de que a entrada é suficiente para explicar o fenômeno, porém é possível que agregar informação com novas entradas como, por exemplo, temperatura possa tornar o modelo mais eficiente. No entanto, é necessário estudar as novas entradas, uma vez que aumentar a dimensão de características não necessariamente implica em se obter um modelo melhor.

Desta forma, pode-se dizer que os objetivos foram cumpridos, embora haja muito a

ser aprimorado, como descrito em 6.1. Pois, o modelo já pode ser utilizado como apoio a decisão com razoável probabilidade de acerto, ainda que seja possível melhorar o modelo por meio de novas técnicas, classificações, entre outras possibilidades, sendo possível que o modelo tornar-se até mesmo autônomo por meio de alguma destas melhorias.

6.1 SUGESTÕES PARA TRABALHOS FUTUROS

Embora este trabalho tenha cumprido seus objetivos, é possível expandí-los a fim de melhorar a identificação de ETS. Para dar continuidade a este trabalho sugere-se:

- Inclusão de novos dados do tipo “chuva intensa”;
- Inclusão de dados de ETS com rápido amadurecimento;
- Buscar outras formas de pré-classificar os dados;
- Incluir outros atributos como por exemplo temperatura, no intuito de agregar informação;
- Aplicar métodos semi-supervisionados incluindo regiões em que se tenha certeza da não ocorrência de ETS;
- Aplicar aprendizagem profunda (*deep learning*);
- Aplicar outras técnicas que minimizem a influência dos dados mal classificados ou com outras abordagens, como por exemplo algoritmo genético;
- Aplicar as técnicas e metodologias a outros problemas como por exemplo a incidência de raios.

REFERÊNCIAS

Agência de notícias do Paraná. **Simepar confirma que tornado atingiu Sudoeste e alerta para mais temporais**. 2015. <http://www.aen.pr.gov.br/modules/noticias/article.php?storyid=84934>. Acessado em 22/03/2017.

ALPAYDIN, E. **Introduction to Machine Learning**. Second edition. [S.l.]: Massachusetts Institute of Technology, 2010.

ANOCHI, J. A. **Previsão climática de precipitação por redes neurais autoconfiguradas**. Tese (Doutorado) — Pós-Graduação em Computação Aplicada, 2015.

ARAÚJO, C. B. C. de; NETO, S. A. D.; FILHO, F. de A. S. Streamflow forecasting for the dam or Ós/ce for hydrometeorological data using perceptrons. **Rev. bras. meteorol.**, v. 30, n. 1, p. 1793–8201, 2015. Acessado em 22/05/2017. Disponível em: <<http://dx.doi.org/10.1590/0102-778620140048>>.

BENETI, C. A. A. **Caracterização hidrodinâmica e elétrica de sistemas convectivos de mesoescala**. Tese (Doutorado em Meteorologia) — Departamento de Ciências Atmosféricas, Universidade de São Paulo, 2012.

BISHOP, C. M. **Neural Networks for Pattern Recognition**. [S.l.]: CLARENDON PRESS, 1995.

BRINGI, V. N.; CHANDRASEKAR, V. **Polarimetric Doppler Weather Radar: Principles and Applications**. 1. ed. [S.l.]: Cambridge University Press, 2001.

BRUCE, R. A baysean approach to semi-supervised learning. **NLPRS**, p. 57–64, 2001.

CAMPOS, C. R. J. de; EICHHOLZ, C. w. Características físicas dos sistemas convectivos de mesoescala que afetaram o rio grande do sul no período de 2004 a 2008. **Revista Brasileira de Geofísica**, 2010. ISSN 0102-261X. Disponível em: <<http://www.scielo.br/pdf/rbg/v29n2/a09v29n2.pdf>>.

COLLIER, C. **Applications of Weather Radar Systems: a Guide to Uses of Radar Data in Meteorology and Hydrology**. 2. ed. [S.l.]: John Wiley & Sons, 1996.

COUTINHO, E. a R.; SILVA, R. M.; DELGADO, A. R. S. Using computational intelligence technique for the meteorological data prediction. **Revista Brasileira de Meteorologia**, v. 31, n. 1, 2016.

DAMIAN, E. A. **Duas metodologias aplicadas à classificação de precipitação convectiva e estratiforme com radar meteorológico: SVM e K-means**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, 2011.

DARDEN, C. B. et al. Utilizing total lightning information to diagnose convective trends. **Bull. Am. Meteorol. Soc.**, p. 167–175, Fevereiro 2010. DOI:10.1175/2009BAMS2808.1. Disponível em: <<http://journals.ametsoc.org/doi/pdf/10.1175/2009BAMS2808.1>>.

DEFESA CIVIL. **Reconhecimentos realizados em 2015**. 2015. Acessado em 17 fevereiro de 2016. Disponível em: <<http://www.mi.gov.br/reconhecimentos-em-2015>>.

DEFESA CIVIL. **Reconhecimentos realizados em 2016**. 2016. Acessado em 30 de agosto de 2016. Disponível em: <<http://www.mi.gov.br/reconhecimentos-em-2016>>.

DESOTO, C. B. Radar techniques - primer principles. **QST**, XXIX, n. 4, p. 1–20, abril 1945. Acesso em: 05 de junho de 2017. Disponível em: <<http://www.rfcafe.com/references/qst/radar-techniques-primer-principles-apr-1945-qst.htm>>.

DOSWELL, C. A. **Severe Convective Storms**. Boston: American Meteorological Society, 2001.

FABRY, F. **Radar Meteorology Principles and Practice**. [S.l.]: Cambridge University Press, 2015.

G1PR. **Chuva causa estragos em 35 cidades do PR e atinge mais de 12 mil pessoas**. 2015. <http://g1.globo.com/pr/parana/noticia/2015/07/chuva-causa-estragos-em-35-cidades-do-pr-e-atinge-mais-de-12-mil-pessoas.html>. Acessado em 22/03/2017.

G1PR. **Vendaval destelha casas, derruba árvores e postes no oeste do Paraná**. 2015. <http://g1.globo.com/pr/oeste-sudoeste/noticia/2015/01/vendaval-destelha-casas-derruba-arvores-e-postes-no-oeste-do-parana.html>. Acessado em 22/03/2017.

GAGNE, D. J.; MCGOVERN, A.; BROTZGE, J. Classification of convective areas using decision trees. **JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY**, p. 1341–1353, 2009.

GATLIN, P. N.; GOODMAN, S. J. A total lightning trending algorithm to identify severe thunderstorms. **JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY**, v. 27, p. 744–755, janeiro 2010.

GAVRISHCHAKA, V. V.; GANGULI, S. B. Support vector machine as an efficient tool for high-dimensional data processing: Application to substorm forecasting. **JOURNAL OF GEOPHYSICAL RESEARCH**, v. 106, n. A12, p. 29911–29914, 2001.

HALL, T.; BROOKS, H. E.; DOSWELL, C. A. Precipitation forecasting using a neural network. **WEATHER AND FORECASTING**, v. 14, p. 338–345, 1998.

HARRINGTON, P. **Machine Learning in Action**. [S.l.]: Manning Publications, 2010. ISBN 9781617290183.

HAYKIN, S. O. **Neural Networks: A Comprehensive Foundation**. 2. ed. [S.l.]: Pearson Education, 1999.

HELMUS, J. et al. The python-arm radar toolkit (py-art), an open source package for weather radar. In: **36th Conference on Radar Meteorology, Breckenridge, CO, USA**. [S.l.: s.n.], 2013.

HOLLE, R. L. et al. The life cycle of lightning an severe weather in 3-4 june 1985 pre-storm mesoscale convective system. **Monthly Weather Review**, v. 122, p. 1798–1808, Agosto 1994.

HUNG, N. Q.; BABEL, S. W. M. S.; TRIPATHI, N. K. An artificial neural network model for rainfall forecasting in bangkok, thailand. **Hydrol. Earth Syst. Sci**, v. 13, p. 1413–1425, 2009. Acessado em 22/05/2017. Disponível em: <<http://www.hydrol-earth-syst-sci.net/13/1413/2009/>>.

JOE, P. et al. The s2k severe weather detection algorithms and their performance. **American Meteorological Society**, v. 19, p. 44–63, 2004.

JORNAL DO OESTE. **Jornal do Oeste**. 2015. Acesso em: 31 de Julho de 2015. Disponível em: <<http://www.jornaldooeste.com.br/regiao/2014/04/novo-radar-meteorologico-comeca-a-operar-em-cascavel/890672/>>.

KITZMILLER, D. H.; MCGOVERN, W. E.; SAFLE, R. E. The wsr-88d severe weather potential algorithm. **American Meteorological Society**, v. 10, p. 414–159, março 1995.

KUMJIAN, M. R. Principles and applications of dual-polarization weather radar. part i: Description of the polarimetric radar variables. **J. Operational Meteor**, n. 19, p. 226–242, 2013. Disponível em: <<http://nwafiles.nwas.org/jom/articles/2013/2013-JOM19/2013-JOM19.pdf>>.

KUMJIAN, M. R. Principles and applications of dual-polarization weather radar. part ii: Warm- and cold-season applications. **J. Operational Meteor**, n. 20, p. 243–264, 2013. Disponível em: <<http://nwafiles.nwas.org/jom/articles/2013/2013-JOM20/2013-JOM20.pdf>>.

LACORTE, S. G. **A neural network approach to tornado forecasting In north alabama and southern middle Tennessee**. Tese (Doutorado) — The University of Alabama in Huntsville, 2011.

LANDGREBE, D. The quantitative approach: Concept and rationale. In: _____. **Remote Sensing: The Quantitative Approach**. [S.l.]: McGRAW-HILL, 1978. p. 1–20.

LIEPERT, M. Topological fields chunking for german with svm's: optimizing svm parameters with ga's. **INTERNATIONAL CONFERENCE ON RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING (RANLP)**, 2003.

LILLESAND, T. M.; KIEFFER, R. W. **Remote Sensing and Image Interpretation**. 2. ed. [S.l.]: John Wiley & Sons, 1987. 721 p.

LIMA, K. C. **Descargas Eléctricas Atmosféricas em Sistemas Convectivos de Me-soescala no Sul da América do Sul**. Dissertação (Mestrado) — Programa de Pós-Graduação em Meteorologia, Universidade Federal do Pelotas, 2005.

LITTA, A. J.; IDICULA, S. M.; MOHANTY, U. C. Artificial neural network model in prediction of meteorological parameters during premonsoon thunderstorms. **International Journal of Atmospheric Sciences**, v. 2013, n. Article ID 525383, 2013. Acessado em 22/05/2017. Disponível em: <<https://www.hindawi.com/journals/ijas/2013/525383/>>.

LIU, C.; HECKMAN, S. The application of total lightning detection and cell tracking for severe weather prediction. **Conference: Proc. Int. Conf. on Ground. and Earth**, 2010.

LIU, C.; HECKMAN, S. **Total Lightning Data and Real-Time Severe Storm Prediction**. 2012.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma introdução as support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, p. 43, 2007.

MADDOX, R. A. Mesoscale convective complexes. **Bull. Am. Meteorol. Soc.**, v. 61, p. 1374–1387, 1980. Disponível em: <[http://journals.ametsoc.org/doi/pdf/10.1175/1520-0477\(1980\)061%3C1374%3AMCC%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/pdf/10.1175/1520-0477(1980)061%3C1374%3AMCC%3E2.0.CO%3B2)>.

MAQSOOD, I.; ABRAHAM, M. R. K. A. An ensemble of neural networks for weather forecasting. **Neural Comput and Applic**, p. 112–122, 2004. DOI 10.1007/s00521-004-0413-4.

MARSLAND, S. **Machine Learning: An Algorithmic Perspective**. 2. ed. [S.I.]: CRC, 2015.

MERCER, A. E. et al. Objective classification of tornadic and nontornadic severe weather outbreaks. **American Meteorological Society**, v. 19, p. 4355–4368, dezembro 2009.

METED. **Weather radar fundamentals**. 2014. Acesso em: 27 de Fevereiro de 2014. Disponível em: <http://www.meted.ucar.edu/radar/basic_wxradar>.

MICHALSKI, S. R.; CARBONELL, G. J.; MITCHELL, M. T. **Artificial Intelligence Approach Volume II**. [S.I.]: Morgan Kaufmann, 1986.

MITCHELL, T. M. **Machine learning**. 2. ed. [S.I.]: McGraw-Hill, 1997. 721 p.

MURPHY, M. J.; HOLLE, R. L. **Warnings of cloud-to-ground lightning hazard based on total lightning and radar information**. 2006.

NARVEKAR, M.; FARGOSE, P. Daily weather forecasting using artificial neural network. **International Journal of Computer Applications**, v. 121, n. 22, p. 0975 – 8887, 2015.

NETO, M. A. S. **Mineração visual de dados: extração do conhecimento a partir das técnicas de visualização da informação e mineração de dados**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, 2008.

NEWMAN, J. F. et al. Range-correcting azimuthal shear in doppler radar data. **Weather and Forecasting**, v. 28, p. 194–211, 2013.

NOVO, E. M. L. M. **Sensoriamento Remoto: princípios e aplicações**. [S.l.]: Edgard Blücher, 1992. 308 p.

PORTAL DA CIDADE. **Temporal atinge Foz do Iguaçu e causa prejuízo aos moradores**. 2015. <http://foz.portaldacidade.com/noticias/5678-temporal-atinge-foz-do-iguacu-e-causa-prejuizo-aos-moradores>. Acessado em 22/03/2017.

RADHIKA, Y.; SHASHI, M. Atmospheric temperature prediction using support vector machines. **International Journal of Computer Theory and Engineering**, v. 1, n. 1, p. 1793–8201, 2009.

RAL.UCAR.EDU. **Radial Radar Data Formats**. 2017. Acesso em: 05 de junho de 2017. Disponível em: <<https://ral.ucar.edu/projects/titan/docs>>.

RAMAKRISHNAN, S. **Pattern Recognition Analysis and Applications**. [S.l.]: ExLi4EvA, 2016.

RINEHART, R. E. **Radar for Meteorologists**. 4. ed. Columbia: Rinehart Publications, 2004.

RUZANSKI, E.; CHANDRASEKAR, V. Nowcasting rainfall fields derived from specific differential phase. **Journal of applied Meteorology and Climatology**, v. 51, p. 1950–1959, 2012.

SAUVAGEOT, H. **Radar Meteorology**. 1. ed. [S.l.]: Artech House Publishers, 1992.

SCHULTZ, C. J. et al. Kinematic and microphysical significance of lightning jumps versus nonjump increases in total flash rate. **American Meteorological Society**, v. 32, p. 275–288, fevereiro 2017.

SCHULTZ, C. J.; PETERSEN, W. A.; CAREY, L. D. Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. **American Meteorological Society**, v. 26, p. 744–755, OCTOBER 2011.

SELUZNIAK, R. H. **Controle de qualidade em dados de radar polarimétrico com apoio de visualizador interativo de dados**. Dissertação (Mestrado) — Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, 2016.

SEMOLINI, R. **Support Vector Machines, Inferência Transdutiva e o Problema de Classificação**. Dissertação (Mestrado) — Pós-graduação da Faculdade de Engenharia Elétrica UNICAMP, 2002.

SILVEIRA, R. **Recognition of clutter in weather radars using polarization diversity information and artificial neural networks**. Tese (Doutorado) — University of Essex, October 1998.

SMOLA, A. J. A tutorial on support vector regression. **Bernhard Schölkopf-Statistics and Computing archive**, v. 14, n. 3, p. 199–222, August 2004.

SMOLA, A. J. et al. **Advances in large margin classifiers**. [S.l.]: MIT Press, 2000.

SUTTON, R. S.; BARTO, A. G. **Introduction to Reinforcement Learning**. 1st. ed.. ed. [S.I.]: MIT Press, 1998.

TRAFALIS, T. B.; ADRIANTO, I.; RICHMAN, M. B. Active learning with support vector machines for tornado prediction. **International Conference on Computational Science**, v. 14, p. 1130–1137, 2007. DOI: 10.1007/978-3-540-72584-8_148.

VAPNIK, V.; CHERVONENKIS, A. **Theory of Pattern Recognition**. [S.I.]: Akademie Verlag, 1974.