

# *Previsão de Condição de Tempo Severo em Sistemas Convectivos de Mesoescala utilizando Support Vector Regression*

Tulipa Silva, Paulo H. Siqueira  
PPGMNE  
Universidade Federal do Paraná  
Curitiba, Brasil  
tulipa@simepar.br

Cesar A. A. Beneti  
Departamento de Monitoramento e Previsão  
SIMEPAR  
Curitiba, Brasil

**Resumo**—Este trabalho apresenta uma aplicação da técnica de Aprendizagem de Máquina conhecida como *Support Vector Regression* (SVR) para a previsão de Condição de Tempo Severo em Sistemas Convectivos de Mesoescala (SCM). Os SCMs estão, em geral, associados a Eventos de Tempo Severo (ETS), causando prejuízos materiais e humanos. No entanto, as condições específicas para que um SCM venha a possuir um ETS ainda não são totalmente conhecidas. A fim de prever um ETS a partir da coleta de dados de sensoriamento remoto de um SCM, utilizamos o SVR com diversas combinações de parâmetros. Os dados utilizados como entrada do modelo, vetor de características, é constituído de dados coletados do radar de dupla polarização pertencente ao Sistema Meteorológico do Paraná (SIMEPAR) e dados de Descargas Elétricas Atmosféricas (DEA) cedidos pela Earth Networks Total Lightning Network (ENTLN). Após o treinamento obteve-se um modelo de regressão com erro de regressão e generalização inferiores a 0,082. Aplicando o modelo aos dados de estudo, foi verificado que dados classificados como prováveis ETS são aqueles em que a saída do modelo é igual ou superior a 0,76.

**Palavras-chave**—Regressão; *Support Vector Regression*; *Now-casting*; Radar Meteorológico; Descargas Elétricas Atmosféricas

## I. INTRODUÇÃO

No período de janeiro de 2015 a julho de 2016 a Defesa Civil efetuou 106 reconhecimentos, no estado do Paraná, causados por tempestades convectivas.

Segundo a Defesa Civil [2] e [3], milhares de pessoas foram diretamente afetadas. Este número seria muito maior se incluirmos prejuízos à economia e agricultura pois, como apresentado em [12], a região do Estado do Paraná tem a agricultura como atividade essencial.

Tempestades Convectivas estão associadas a Sistemas Convectivos de Mesoescala (SCM), os quais são, frequentemente associados vendaval, chuvas intensas, granizo e tornados e descargas elétricas atmosféricas [17]. E sua classificação, detecção e previsão ainda é pouco consolidada. Implicando

na não existência de modelo conceitual que descreva o desenvolvimento do SCM [18] e [1].

Devido ao desenvolvimento vertical dos SCMs, é de grande valia seu estudo com a utilização do radar meteorológico, pois possibilita conhecimento vertical da estrutura da nuvem [10].

Além disso, é possível encontrar SCMs com base nos dados de descargas atmosféricas, fato já explorado por [18], [19], [16] e [20], os quais, utilizam os dados de descargas atmosféricas para estimar, inclusive, a fase de desenvolvimento do SCM.

Com o intuito de estimar a evolução de um SCM em possível Evento de Tempo Severo (ETS), iremos utilizar a técnica de aprendizado de máquina, a *Support Vector Regression* (SVR) aplicada a dados de radar e descargas atmosféricas.

O SVR é uma técnica derivada da *Support Vector Machine* (SVM). O SVM tem obtido resultados satisfatórios na classificação de agrupamentos, como a exemplo [4] que utilizou para classificação entre precipitação convectiva e estratiforme, ou [8] que realizou remoção de Clutter.

## II. CONCEITOS METEOROLÓGICOS

### A. Sistemas Convectivos de Mesoescala

Sistemas Convectivos de Mesoescala (SCM) são sistemas de precipitação associados a fortes campos verticais de vento e altas taxas pluviométricas [1]. Ou ainda, um sistema de nuvem associado à ocorrência da aglomeração das tempestades [17].

Segundo [5], um SCM tem tempo de vida de 6 a 12 horas e produz aproximadamente, uma área de 100 km de precipitação, podendo ou não causar Evento de Tempo Severo (ETS). Cada nuvem de tempestade em um SCM dura cerca de uma hora, e portanto, o SCM pode conter diversas nuvens em estágios diferentes, formando, maturando e ou dissipando.

### B. Radar

O radar (*R*Adio *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 [9].

Em termos de monitoramento e previsão meteorológica, o radar é uma das melhores ferramentas disponíveis, pois oferece monitoramento em tempo real, com boa resolução espacial e possibilitando uma visão tridimensional dos dados [10].

Para tanto, muito do que se conhece da estrutura interna dos sistemas de precipitação se deve ao estudo dos dados de radar [12].

A coleta de dado pelo radar ocorre quando o radar emite (através do transmissor) um feixe de radiação eletromagnética e capta (através do receptor) a energia refletida pelas partículas presentes na atmosfera. Essa energia é amplificada e, conhecido o intervalo de tempo entre emissão e retorno é possível determinar a distância do alvo ao radar [15]. A partir da informação da energia refletida e a distância do radar ao alvo é possível inferir as variáveis de radar: refletividade, velocidade radial e largura espectral [6].

Em radares de dupla polarização, é possível obter variáveis adicionais devido a orientação do sinal eletromagnético emitido e ou captado, horizontal e ou vertical. Esse é o caso das variáveis Refletividade Diferencial (ZDR) e Correlação Copolar (RHOHV).

1) *Refletividade (Z)*: É proporcional a energia refletida para o radar e o quadrado da distância do radar ao alvo. A fim de diminuir o intervalo de valores que esta variável pode assumir, convencionalmente, utiliza-se o valor em sua escala logarítmica (dBZ). A refletividade é dependente do número e tamanho das partículas presentes no volume de coleta de dado, sendo uma das variáveis de maior utilização na previsão meteorológica. Pois, valores elevados de refletividade estão, em geral, associados a chuva intensa [6].

A variável dBZ emitida e recebida na direção horizontal, após tratada pelo controle de qualidade do *software* do radar, é representado por DBZH.

2) *Velocidade Radial (V)*: Utilizando o efeito Doppler, o radar estima a velocidade, em metros por segundo ( $m/s$ ), com que o alvo se afasta ou se aproxima do radar na direção do feixe emitido.

A variável V emitida e recebida horizontalmente é representada por VELH.

3) *Largura Espectral (W)*: Representa o desvio padrão das medidas 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 [6].

Para W emitida e recebida horizontalmente utilizamos a representação WIDTHH.

4) *Refletividade Diferencial (ZDR)*: A Refletividade oriunda da emissão e captação de um sinal horizontal é representada por  $Z_{HH}$ . Analogamente,  $Z_{VV}$  representa a emissão e captação da Refletividade associada a um sinal vertical. Alguns radares podem ter sinal alternado, gerando  $Z_{HV}$  ou  $Z_{VH}$ . [6]

A Refletividade Diferencial (ZDR), é a diferença entre  $Z_{HH}$  e  $Z_{VV}$ , em unidade logarítmica [11].

5) *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 [9].

### C. Descargas Elétricas Atmosféricas

Os SCMs possuem maior quantidade de raios em relação aos demais. A integração de informações de dado de radar e Descargas Elétricas Atmosféricas (DEA) para detecção de SCMs já foram exploradas por [20] e [16], e se mostraram acuradas. Além disso, segundo [19], eventos de tempo severo ocorrem logo após um pico de DEA.

O mecanismo de eletrização das nuvem não é bem conhecida porém, as Descargas Elétricas Atmosféricas (DEA) são resultado do desenvolvimento de intensos centros de cargas na nuvem, que excedida a capacidade de isolamento elétrico do ar, resulta na dissipação dos centros de carga. Os tipos de DEA são classificados de acordo com local de onde se originam e onde terminam [18]. Neste trabalho, as DEA são classificadas em *Intra Cloud* (IC) e *Cloud to Ground* (CG).

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 [20].

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 [20].

### III. Support Vector Regression (SVR)

Aprendizado de Máquina (AM) é uma sub-área da Inteligência Artificial que se dedica a obter, por meio de princípios de inferência, conclusões genéricas a partir de um conjunto de exemplos [13].

Ao conjunto de exemplos sobre o qual a técnica de AM é treinada dá-se o nome de conjunto de treinamento. Em geral, entre 60% e 90% do conjunto de dados. O erro associado ao conjunto de treinamento é dito erro de regressão. Os demais dados farão parte do conjunto de teste [14].

Após seu treinamento, a técnica de AM, baseada no conhecimento extraído do grupo de treinamento, deve classificar novas amostras apresentadas (conjunto de teste), até então desconhecidas pelo modelo, ao padrão mais próximo às suas características. Ou seja, entradas semelhantes produzem saídas semelhantes [14].

Support Vector Regression (SVR) é uma adaptação de uma técnica supervisionada (os dados são previamente classificados) de AM: o Support Vector Machine (SVM). Enquanto o SVM consiste em encontrar um hiperplano separador entre as classes, o SVR busca um hiperplano tal que o conjunto de treinamento fique o mais próximo possível. O hiperplano ótimo é aquele em que todos os dados se situam entre as margens e cujo interior é a função de regressão [21].

Seja  $\{(x_1, y_1), (x_2, y_2), \dots, (x_p, y_p)\}$  um conjunto de treinamento, onde  $x_i \in \mathbb{R}^n$  representa um vetor de características de entrada e  $y_i \in \mathbb{R}$  a saída desejada para a entrada  $x_i$ . Obter

o hiperplano ótimo para o SVR, é equivalente a resolver o problema de otimização (1), apresentado em [22].

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

O vetor  $e$  possui  $p$  coordenadas, todas iguais a 1.

As variáveis  $\xi$  e  $\xi^*$  são ditas variáveis de folga não negativas, que tem como objetivo penalizar dados que se situem fora da margem especificada por  $\epsilon$ , e  $C$  é uma constante positiva de penalização arbitrária. Quanto maior o valor de  $C$  menor tende a ser o erro do modelo, no entanto, grandes valores de  $C$  podem causar *overfitting*, que é o processo de não ser capaz de generalizar. Analogamente, valores muito pequenos de  $C$ , podem causar *underfitting*, ou seja, não ser capaz de classificar [14].

A função  $\phi: \mathcal{R}^p \rightarrow \mathcal{R}^n$ , com  $n > p$ , mapeia os dados em um espaço de características, no intuito de tornar os dados linearmente separáveis no espaço característico. Porém, devido a complexidade em se obter, ou mesmo garantir a existência de  $\phi$ , é possível utilizar as funções *kernel*( $K$ ), que representam produtos internos da forma  $\phi(x_i) \cdot \phi(x_j)$ , sem efetivamente obter  $\phi$  [14], [13], [21].

Sejam  $\alpha$  e  $\alpha^*$  os multiplicadores de Lagrange e  $L(w, b, \xi, \xi^*, \alpha, \alpha^*)$  a função Lagrangeana associada ao problema primal (1) e  $Q$  uma matriz semidefinida positiva da forma  $Q_{ij} = K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ . O problema apresentado em sua forma dual será:

$$\begin{aligned} \min_{\alpha, \alpha^*} \quad & \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \epsilon e^T (\alpha + \alpha^*) - y^T (\alpha - \alpha^*) \\ \text{Sujeito a:} \quad & e^T (\alpha - \alpha^*) = 0, \\ & 0 \leq \alpha_i, \alpha_i^* \leq C, \quad i = 1, \dots, p. \end{aligned} \quad (2)$$

O problema (1) na forma (2) implica, pelas condições de complementariedade de Karush-Kuhn-Tucker [14], que apenas pontos de treinamento sobre as margens ou fora delas  $\epsilon$  e  $-\epsilon$  contribuem no cálculo do vetor de pesos  $w$ , pontos localizados fora das margens representam os erros do modelo e os vetores de suporte são caracterizados pelos vetores sobre as margens. E, a função de decisão no espaço original será uma regressão linear da forma  $\sum_{i=1}^p (\alpha - \alpha^*) K(x_i, x) + \rho$ .

A escolha da função *kernel* depende da complexidade dos dados. Na Tabela III são listados alguns exemplos de função *kernel*.

#### IV. DADOS E METODOLOGIA

O radar a ser utilizado neste trabalho é um radar banda S (frequências de emissão utilizada está entre 2-4 GHz), localizado em Cascavel (Paraná), operante desde 2014. Esse

Tabela I  
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	$K(x_i, x_j) = \gamma((x_i + x_j) + r)^d$
Gaussiano	$K(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$
Sigmoidal	$K(x_i, x_j) = \tanh(\gamma(x_i \cdot x_j) + r)$

Tabela II  
EVENTOS ESTUDADOS

Data	Hora	Tipo de evento	Cidade afetada
12/07/2015	18:00	Chuva Intensas	Cafezal do Sul
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	Tornados	Francisco Beltrão
23/11/2015	13:30	Vendaval	Umuarama
28/12/2015	21:15	Vendaval	Francisco Beltrão
27/02/2016	09:53	Chuva Intensas	Francisco Beltrão
22/03/2016	03:15	Vendaval	C. Leônidas marques

radar é dual polarimétrico e executa uma varredura completa em 7,5 minutos. As variáveis dBZ, V e W, são identificadas nos dados deste radar como DBZH, VELH e WIDTHH, respectivamente.

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 do radar de estudo (240 km distantes da posição do radar) e alguns desses dados não possuem a resolução máxima do radar. Além disso, foi necessário escolher os dados de acordo com a disponibilidade dos dados da ENTLN.

Os dados que se enquadram nas especificações de disponibilidade, localização e resolução são os que se referem aos eventos indicados na Tabela IV.

Para cada evento foi selecionada uma região da tempestade próxima ao evento, e rastreado essa tempestade em até uma hora antes do evento.

Devido as 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 próximo ao radar em detrimento aos mais distantes. A fim de padronizar a resolução dos dados, os mesmos fora interpolados em uma grade cartesiana, com o auxílio do software Pyart [7].

As variáveis de radar utilizadas são DBZH, VELH, WIDTHH, ZDR e RHOHV. Quanto aos dados da ENTLN, são consideradas todas as DEA que ocorrem durante os 7.5 minutos de varredura do radar, contados a partir da hora de início da varredura, e que estejam no range do radar. Os dados de DEA são separados em quantidade por  $km^2$  de IC e CG por varredura.

Assim, cada ponto de grade cartesiana possui valores de DBZH, VELH, WIDTHH, ZDR, RHOHV, IC, CG e a altitude

Tabela III  
ESPECIFICAÇÃO DOS PARÂMETROS  $\gamma$ ,  $C$  e  $d$  PARA APRIMORAMENTO DA REGRESSÃO.

Parâmetro	Máximo	Mínimo	Varição
$\gamma$	10	0	0,01
$C$	100	0	1
$d$	10	2	1

do ponto, o que corresponde, após mapeados no intervalo  $[-1, 1]$ , aos vetores de características de entrada para a SVR.

Para o treinamento foram escolhidos aleatoriamente 75% dos vetores para treinamento e 25% para validação. Esta proporção foi escolhida de forma arbitrária.

Para treinamento da SVR foi utilizado o pacote scikitlearn, do Python [22]. Esse pacote possui vasta gama de métodos de AM implementados e de fácil utilização.

Para efetuar a previsão de curto prazo (30 minutos aproximadamente), também chamada de *nowcasting*, de ETS a partir de SCMs, considerou-se que dados uma hora antes do evento tem saída do modelo igual a 0, dados coletados a menos de 30 minutos da hora aproximada do evento tem saída maior que 0,5 e menor ou igual a 1.

O objetivo em relação a saída do modelo é que seja uma regressão crescente tal que dados semelhantes aos que ocorrem até 30 minutos antes dos ETS estudados possuam sempre valores de saída do SVR maiores que 0,5.

Para obter os melhores parâmetros de entrada da SVR foram utilizados os 4 *kernel* disponíveis no *software* utilizado, variando  $\gamma$ ,  $C$  e  $d$ . Na Tabela IV são apresentados o máximo, mínimo e variação utilizada para os parâmetros. Em relação aos parâmetros  $\epsilon$  e  $r$  os mesmos foram mantidos em 0,1 e 0, respectivamente.

Os ensaios ocorreram em todas combinações de *kernel* e parâmetros, sendo abortado quando excedido tempo limite de processamento, adotado como 1 hora de processamento para cada combinação.

Após o treinamento da técnica, os dados do conjunto de treinamento e teste são apresentados separadamente para verificação dos erros de regressão (associado ao conjunto de treinamento) e generalização (erro associado ao conjunto de teste).

## V. RESULTADOS

Na Tabela V estão listados, em ordem crescente do Erro Quadrático médio (EQm) para o conjunto de treinamento, os 23 menores valores de Erro Quadrático Médio associados ao conjunto de treinamento, obtidos utilizando os parâmetros como descritos na Tabela IV.

Os demais ensaios possuem EQm superior a 0,158325 ou excedem o tempo limite de processamento.

Pela Tabela V vemos que o *kernel* mais adequado ao problema, dentre os estudados, é o *kernel* Gaussiano. Com erro de regressão de 0,079182 e erro de generalização de 0,081396.

Tabela IV  
ERRO QUADRÁTICO MÉDIO (EQM) DA REGRESSÃO EM RELAÇÃO A FUNÇÃO KERNEL E AOS PARÂMETROS  $\gamma$ ,  $C$  e  $d$ .

<i>kernel</i>	$\gamma$	$C$	$d$	EQm (treinamento)	EQm (teste)
Gaussiano	1	100	-	0,079182	0,081396
Gaussiano	0,5	100	-	0,095492	0,096690
Gaussiano	1	1	-	0,104277	0,105202
Polinomial	0,5	10	5	0,107059	0,107773
Gaussiano	0,5	10	-	0,108197	0,109143
Gaussiano	1	0,1	-	0,119461	0,120085
Polinomial	0,5	1	5	0,115429	0,116211
Gaussiano	0,5	1	-	0,120367	0,120962
Polinomial	0,5	0,1	5	0,123579	0,124417
Gaussiano	0,1	100	-	0,125062	0,125489
Polinomial	0,1	100	5	0,127163	0,127983
Polinomial	0,5	10	3	0,130468	0,130677
Gaussiano	0,5	10	-	0,130964	0,131261
Gaussiano	0,1	10	-	0,132271	0,132423
Polinomial	0,1	10	5	0,134843	0,135661
Polinomial	0,1	0,1	3	0,137871	0,138455
Linear	0,01	0,001	-	0,155308	0,155178
Linear	0,1	0,001	-	0,155308	0,155178
Sigmoidal	0,01	0,1	-	0,155329	0,155200
Sigmoidal	0,01	1	-	0,158070	0,158023
Linear	0,01	100	-	0,158308	0,158275
Linear	0,1	0,1	-	0,158361	0,158325

Quando aplicado o modelo SVR (*kernel* = *Gaussiano*,  $\gamma = 1$ ,  $C = 100$ ) ao conjunto de dados, temos que todos os dados coletados a menos de 30 minutos do ETS possuem máximos de regressão superiores a 0,76. As Fig. 1 e 2 exemplificam a saída do modelo e representam dados 30 minutos antes de um ETS e aproximadamente durante o ETS, respectivamente.

Observando as Fig. 1 e 2 é possível notar que nem todos os dados contribuem de fato para a previsão. Existem pontos comuns a todos os casos e que podem estar contribuindo demasiadamente para o erro do modelo.

No caso da Fig. 1, mesmo que os dados não tenham necessariamente o valor de saída do modelo de 0,5, nenhum valor é maior que 0,76. Para a Fig. 1, existem dados com saída entre 0,76 e 1. Esta situação se repete a todos os casos: eventos a mais de 30 minutos do ETS não possui saída superior a 0,76 e dados a menos de 30 minutos do ETS sempre possuem saída entre 0,76 e 1. Os eventos mais próximos dos ETSs sempre tem ao menos um ponto de saída igual a 1.

## VI. CONCLUSÕES

Aceitando dados com saída do modelo SVR (*kernel* = *Gaussiano*,  $\gamma = 1$ ,  $C = 100$ ) igual ou superior a 0,76 em algum ponto como provável ETS, o modelo de regressão obtido apresentou bons resultados e possui boa generalização.

Porém, observando as saídas, como nas Fig. 1 e 2 é possível ver que são poucos os dados representativos para a SVR, se comparados ao volume total de dados, comportamento que se

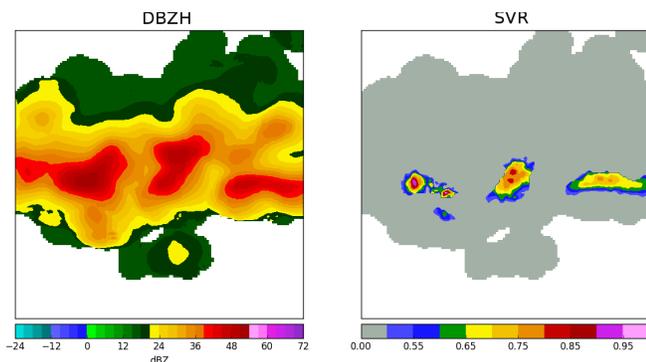


Figura 1. Saída do modelo SVR( $kernel = Gaussiano$ ,  $\gamma = 1$ ,  $C = 100$ ) para 12/07/2015, 17:30.

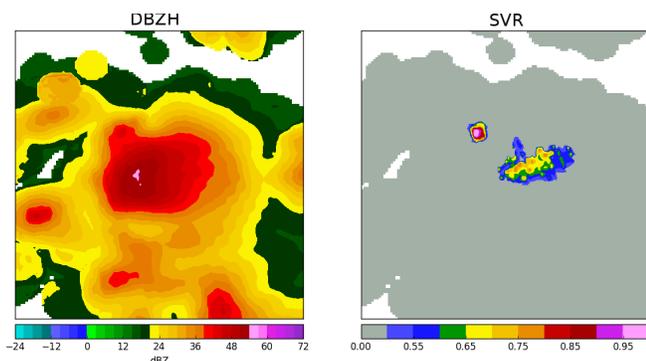


Figura 2. Saída do modelo SVR( $kernel = Gaussiano$ ,  $\gamma = 1$ ,  $C = 100$ ) para 13/07/2015, 21:52.

repete em todos os casos. Os dados pouco representativos e ou atípicos podem aumentar os erros do modelo e ou dificultar o seu treinamento.

Como continuidade deste trabalho, serão utilizadas técnicas de AM que possam extrair os dados menos significativos como pré-processamento. Espera-se, com isso, garantir maior acurácia ao modelo.

## REFERÊNCIAS

- [1] E. N. Anagnostou and C. W. Kummrow, *Stratiform and Convective Classification of Rainfall using SSM/I-GHZ Brightness Temperature Observations*, Journal of Atmospheric and Oceanic Technology, v. 14, p. 570–575. 1997.
- [2] Defesa Civil, *Reconhecimentos realizados em 2015*, Disponível em: <http://www.mi.gov.br/reconhecimentos-em-2015>. Acessado em 17 de fevereiro de 2016.
- [3] Defesa Civil, *Reconhecimentos realizados em 2016*, Disponível em: <http://www.mi.gov.br/reconhecimentos-em-2016>. Acessado em 30 de agosto de 2016.
- [4] E. A. Damian, *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. Paraná, 2011.
- [5] W. Cotton, R. Anthes, *Storm and Cloud Dynamics*, International Geophysics Series, 44 – Academic Press. Inc. 883 pp. 1989.
- [6] F. Fabry, *Radar Meteorology Principles and Practice*, 1. ed [S.I]: Cambridge University Press 2015.
- [7] J. Helmus, S. M. Collis, *The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Pro-*

- gramming Language*, Journal of Open Research Software, 4: e25 <http://dx.doi.org/105334/jors.119>
- [8] M. A. Neto, *Mineração de dados: extração da 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. Paraná, 2008.
- [9] R. E. Rinehart, *Radar for Meteorologists* 4 ed. Columbia: Rinehart Publications, 2004.
- [10] H. Sauvegeot, *Radar Meteorology* 1 ed. Artch House Publishers, 1992.
- [11] C. A. Knight, , Journal of Open Research Software, 4: e25 <http://dx.doi.org/105334/jors.119>
- [12] C. A. A. Beneti, I. Nozu, E. A. Saraiva, *Monitoramento da Precipitação e de Eventos de Tempo Severo com Radar Meteorológico no Estado do Paraná*. Congresso Brasileiro de Meteorologia Brasília, 1998.
- [13] A. C. Lorena, A. C. P. L. F. Carvalho, *Uma introdução às Support Vector Machines*. Revista de Informática Teórica e Aplicada, v. 14, p. 43, 2007.
- [14] S. Haykin, *Neural Networks - A Comprehensive Foundation*. Macmillian College Publishing, inc., 1994.
- [15] C. G. Collier *Applications of Weather Radar Systems: a Guide to Uses of Radar Data in Meteorology and Hydrology*. 2th. ed. [S.I.]: John Wiley & Sons, 1996.
- [16] M. J. Murphy, R. L. Holle *Warnings of cloud-to-ground lightning hazard based on total lightning and radar information*. Second Conf. on Meteorological Applications of Lightning Data, Amer. Meteorol. Soc., Atlanta, GA, paper 2.2 (CD-ROM).
- [17] R. A. Maddox, *Mesoscale convective complexes*. Bull. Am. Meteorol. Soc., v. 61, p.1374-1387, 1980.
- [18] K. C. Lima, *Descargas Eléctricas Atmosféricas em Sistemas Convectivos de Mesoescala no Sul da América do Sul*. Dissertação (mestrado) - Programa de Pós-Graduação em Meteorologia, Universidade Federal do Pelotas. Rio Grande do Sul, 2005.
- [19] C. B. Darden, D. J. Nadler, B. C. Carcione, R. J. Blakeslee, G. T. Satano, D. E. Buechler, *Utilizing total lightning information to diagnose convective trends*. Bull. Am. Meteorol. Soc., 91: 167–175, 2010.
- [20] C. Liu, S. Heckman, *Total Lightning Data and Real-Time Severe Storm Prediction*. WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation. 2012.
- [21] F. Parrella, *Online Support Vector Regression*. Master Science Thesis. Department of Information Science, University of Genoa, Italia, 2007.
- [22] A. J. Smola, *A Tutorial on Support Vector Regression*. Bernhard Schölkopf -Statistics and Computing archive Volume 14 Issue 3, August 2004, p. 199-222