11.2 *SEVERE WEATHER IDENTIFICATION USING POLARIMETRIC RADAR AND MACHINE LEARNING TECHNIQUES*

Tulipa Silva *
SIMEPAR - Paraná Meteorological System, Curitiba, Paraná
Cesar Beneti
SIMEPAR - Paraná Meteorological System, Curitiba,Paraná
P. H. Siqueira
Federal University of Paraná, Curitiba, Paraná
M. F. Buzzi
Federal Technological University of Paraná, Curitiba, Paraná
L. Calvetti
Federal University of Pelotas, Pelotas, Rio Grande do Sul

1. INTRODUCTION

This work presents our experience with radar data analysis for single and dual polarimetric systems, using machine learning techniques to explore complex data and obtain an overview and better understanding of the observed phenomena.

In less than 5 years, in Brazil we increased our radar coverage, from 23 single polarization weather radars to 15 more new dual polarization radars, illustrated in Fig. 1 and Fig. 2. Those radars are mainly S-Band, with a concentration in the southern region, an area prone to severe weather, mostly related to Mesoscale Convective Systems. Weather radars in the south of Brazil play an important role in quantitative precipitation estimation and severe weather monitoring and forecasting. The major economic activity in this region is agrobusiness and energy production. This region is also responsible for more than 35% of hydropower energy generation used in the country, directly dependent on precipitation distribution, water availability and severe storms impacts.

Using machine learning techniques as Support Vector Machine (SVM) and Multilayer Perceptron (MLP), we aim to identify convective storms that can become a Severe Weather Event (SWE) within the next 30 minutes. For this study we analysed a series of SWE which occurred in the state of Paraná, Brazil, from January 2015 to July 2016. The input variables of the model are from polarimetric weather radar, specifically: Z, ZDR, KDP and RHOHV and radar products such as height of the maximum reflectivity (HMAX), and azimuthal, vertical and radial shear (Silva, 2017).

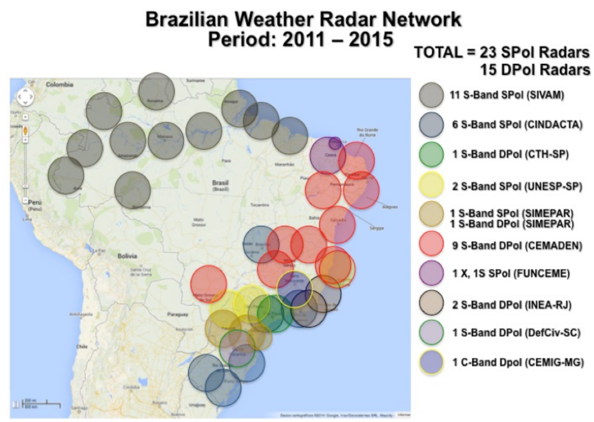

Figure 1: Brazilian weather radar network.

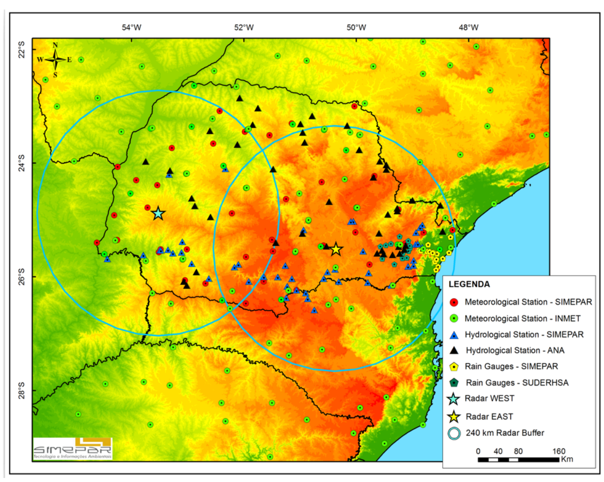

Figure 2: Hydrometeorological System in Parana State, with 2 S-Band weather radars, 1 Single Pol and 1 Dual Pol.

2. SEVERE WEATHER EVENTS AND MACHINE LEARNING

2.1 SEVERE WEATHER EVENTS

In the south of Brazil, from January 2015 to July 2016, according to Defesa Civil (2015, 2016), there

*Corresponding author address:
Tulipa Silva; e-mail: tulipajs@gmail.com

were more than 106 severe weather events caused by convective storms (storms capable of generating windstorm, heavy rain, hail, tornado or atmospheric electrical discharges (AED) (Maddox, 1980).

These SWEs caused great damage, directly affecting more than 39,000 people. A example of such damages can be seen in Figure 3.

Fig. 3 (a) A damage occurred on 12/28/2015. This SWE had a precipitation of 61 mm in 4 hours, causing several floods.

Fig. 3 (b) Damage caused by a tornado on 03/22/2016 (Cândido Rondon city).

Fig. 3 (c) Tornado on 03/22/2016 (Cândido Randon city)

Fig. 3 (d) A funnel cloud occurred on 13 july 2015.

On 07/13/2015, 3 SWEs occurred in an interval of less than 3 hours. Of the 35 cities that suffered from these SWEs (G1PR, 2015), the cities of Matelândia, Ampére and Francisco Beltrão were the most affected, 3564 people faced heavy rains and hailstones reaching the region, leaving at least 19 injured and 50 homes destroyed. Figure 2 shows a funnel cloud that occurred on 07/13/2015, near the region affected by the SWEs.

During the studied period a tornado occurred, a phenomenon of rare occurrence in the region. The tornado registered occurred on 03/22/2016 at 03:15 AM, in the city Captain Leônidas Marques, damaging crops of the region, plucking trees, breaking power cables, leaving street forbidden and at least one person injured (Porta CLICK3: 2016) .

2.2 MACHINE LEARNING

Although the types of SWEs are different from each other, they are all vertical development of a convective storm, so one way to study such phenomena is to use radar data, because it is a data that allows a three-dimensional view of the storm (Fabry, 2015).

Machine Learning (ML) techniques have become popular for working large volumes of data (Marsland, 2015), such as three-dimensional radar data.

ML is an artificial intelligence subarea that studies ways to construct algorithms that enhance one's performance in a specific task using a finite set of data (Mitchell, 997). ML is dedicated to obtaining, through principles of inference, generic conclusions from a set of examples (Lorena, 2007). Two AM techniques used are MLP and SVM, since they have been identified as the best techniques for classification of available standards (Marsland, 2015).

2.2.1 *Multilayer Perceptron Artificial Neural Network*

As defined by Haykin (1999), an Artificial Neural Network (ANN) is a system that is in parallel distributed and massively interconnected by simple processing units called neurons whose modeling is inspired by the functioning of biological neurons. The Multilayer Perceptron (MLP) is the most used artificial neural network.

2.2.2 *Support Vector Machine*

Support Vector Machine (SVM) is an ML technique, of supervised learning, developed by V. Vapnik that seeks to minimize the error in the training set (empirical risk), as well as the error in the test set (generalization risk) (Vapnik , 1974).

According to its mathematical formulation, the SVM corresponds to obtain an optimal hyperplane so that it better segregates the classes (Haykin, 1999).

3. METODOLOGY

The methodology follows the following steps:

1. Polarimetric radar data collection;
2. Pre-processing of data (correction of missing data and resolution, normalization, calculation of shearings, removal of noise and reduction of dimensionality of the input);
3. Inclusion of the classes to the data set;
4. Separation of the test and training set;
5. Training of the MLP and SVM techniques with the polarimetric radar data, generating the M-MLP and M-SVM models respectively;
6. Identification of SWE regions by Lightning Jump and Cloud to ground lightning density;
7. Comparison of the obtained models with each other and with the identification obtained by the use of Lightning Jump and Cloud to ground lightning density, M-ENTLN model;
8. Study and analysis of the results obtained.

To validate the regions where the occurrence of SWE is unknown the M-ENTLN model was obtained by using ENTLN total lightning data, that means a SWE identification by using high total-lightning density or lightning jumps.

4 RESULTS

For the model obtained by the MLP technique, two events were not identified during the entire 30 minute period before the SWE occurrence. For the model using SVM, only one event of heavy rainfall was not identified during the period of study.

Both, M-MLP and M-SVM, models received the same learning set, such that Table 1 shows the percentages of correct identification in the set presented for learning the models.

Table 1: Correct identification by M-MLP and M-SVM in training, test and general data set.

Model	Training set	Test set	General set
M-MLP	78,78 %	90 %	81,40 %
M-SVM	90 %	100 %	93,02 %

As shown in Table 1, the M-SVM model identified 93.02% of the regions of interest, and good classification (90%) and good generalization (100%). Here the region of interest are the regions of the radar data where the occurrence of SWE is known.

Some examples of output models are shown in figures 4, 5, and, 6. Those examples are the same exemplified in Fig. 3(a), Fig. 3(b), Fig. 3(c) and Fig. 3(d).

In figures 4, 5 and 6, the black dots represent the points identified by the respective model, plotted on reflectivity PPI. In these figure, the circle indicates the input regions in the model training where the occurrence of SWE is known. For each SWE, the circled region is the only region of the radar file where the occurrence of SWE is known. The other regions identified by the models (black dots outside of the circle) are regions where the information of the occurrence of SWE is unknown.

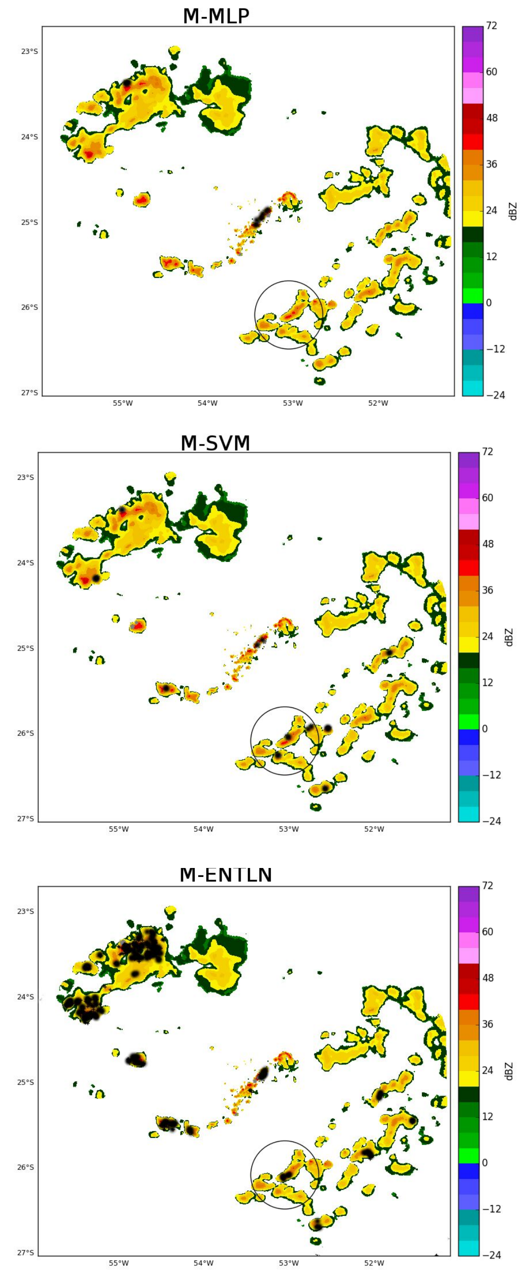

Figure 4: Output from models M-MLP, M-SVM and M-ENTLN on 12/28/2015, 09:15 PM (SWE in Fig 3(a).

The M-ENTLN model, obtained by Cloud-to-ground lightning density ou lightning jump, has 100% of the accuracy in the region of interest, and therefore the identified regions outside of the circle, by M-MLP and M-SVM are compared with identification by atmospheric electric discharges (M-ENTLN) in order to validate the results in unknown SWE (Table 2).

Table 2 indicates that, in general, M-MLP has 68.58% of its identifications (presented in data not previously identified) also identified by M-ENTLN. And, the regions identified by M-ENTLN, 32.02% are not identified by M-SVM.

In the same way, in the total number of regions identified by M-SVM, 67.84% are regions that are also identified by M-ENTLN. And, in the regions identified by M-ENTLN, 23.76% are not identified by M-SVM.

Among those regions where the model M-SVM does not identify and M-ENTLN identifies, 89.58% of them are more than 120 km radar. That is, up to 120 km radar, the agreement of M-SVM with M-ENTLN is 97.52%. For M-MLP this agreement is 82.89%.

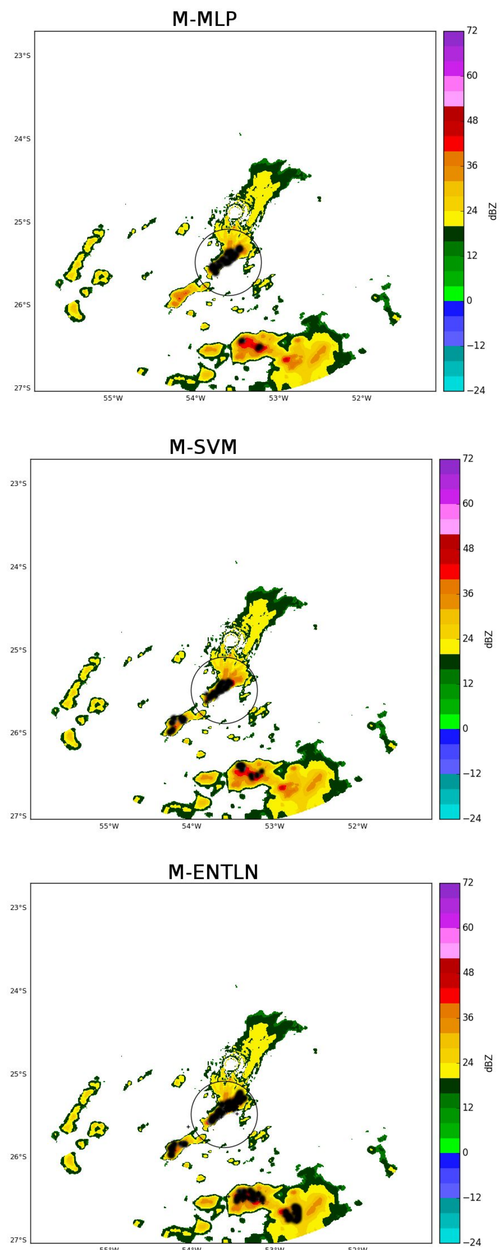

Figure 5: Output from models M-MLP, M-SVM and M-ENTLN on 03/22/2016, related Fig, 2(b) and Fig. 2(c).

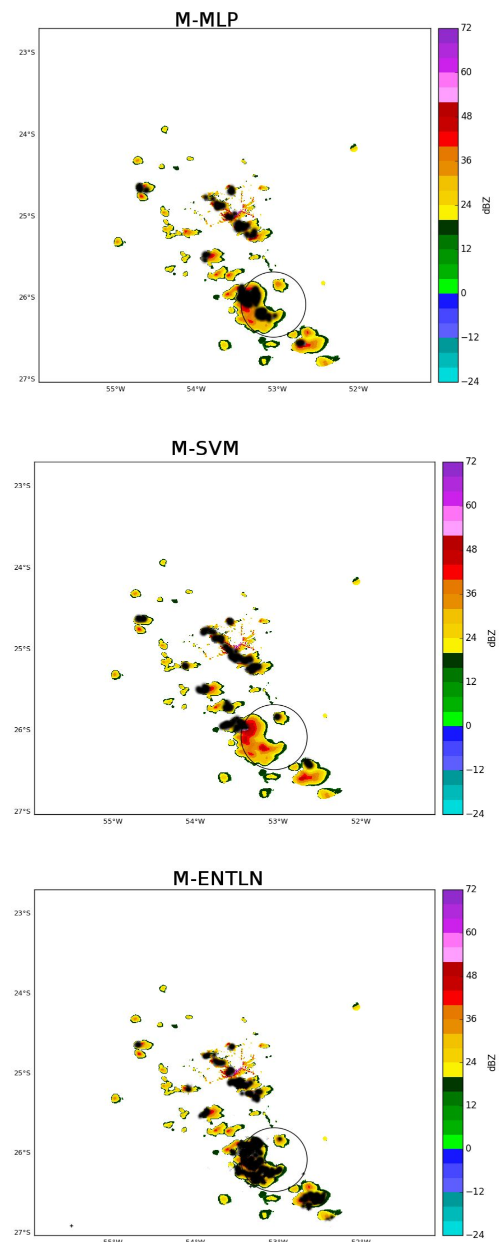

Figure 6: Output from models M-MLP, M-SVM and M-ENTLN on 07/13/2015, related Fig 2(d).

Table 2: Identification by M-MLP and M-SVM models compared to M-ENTLN.

Set	Model	Models and M-ENLTLN	Only M-ENTLN
Training	M-MLP	75 %	32,5 %
	M-SVM	75 %	23,94 %
Test	M-MLP	57,37 %	30,88 %
	M-SVM	55,42 %	23,33 %
General	M-MLP	68,58 %	32,02 %
	M-SVM	67,84 %	23,76 %

4. CONCLUSIONS

The performance of the models may be directly related to the inaccuracy of location information, time and duration of events, but 80% of the events were fully identified during the 30 minutes preceding the SWE. In addition, some regions which were not previously identified as SWE, were indicated by the models as possible occurrence of severe weather. In regions where the models were not applied during training, a comparison was made with total lightning data. This comparison showed good correlation between the identified regions and high values of total lightning data. In this way, the study is efficient as a tool to support the decision, by meteorologists, in the identification and prediction of SWE. In order to improve these results, it is necessary to include new events to train the model, as well as to further investigate the regions with no previous classification of SWE but which were identified as possible severe events.

5. ACKNOWLEDGEMENTS

This research was partially supported by CAPES/Brazil. Earth Networks provided the total lightning data.

6. REFERENCES

Defesa Civil, 2015: Reconhecimentos realizados em 2015. http://www.mi.gov.br/reconhecimentos-em-2015.
.

Defesa Civil, 2016: Reconhecimentos realizados em 2016. http://www.mi.gov.br/reconhecimentos-em-2016.

FABRY, F. 2015: Radar Meteorology Principles and Practice. Cambridge University Press.

G1PR: 2015. Vendaval destelha casas, derruba árvores e postes no oeste do Paraná. http://g1.globo.com/pr/oeste-sudoeste/noticia/2015/01/vendaval-destelha-casas-derruba-arvores-e-postes-no-oeste-do-parana.html

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

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

MADDOX, R. A., 1980: Mesoscale convective complexes. Bull. Am. Meteorol. Soc., v. 61, p. 1374–1387.

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

MITCHELL, T. M. 1997: Machine learning. 2. ed. McGraw-Hill, 1997.

Porta CLICK3, 2016: Temporal causa destruição em Capitão Leônidas Marques. ref:http://www.click3.com.br/site/noticias.php?id=3211&temporal-causa-destruicao-em-capitao-leonidas-marques

RBJ, 2015: Chuva provoca alagamento de ruas em Francisco Beltrão. http://www.rbj.com.br/geral/chuva-provoca-alagamento-de-ruas-em-francisco-beltrao-2438.html

SILVA, T., 2017: Severe Weather Event identification using machine learning techniques on polarimetric radar data.(In Portuguese) Master thesis, Federal University of Paraná.

VAPNIK, V.; CHERVONENKIS, A. 1974: Theory of Pattern Recognition. Akademie Verlang.