IDENTIFICATION OF SEVERE WEATHER EVENT WITH METEOROLOGICAL RADAR AND MULTILAYER PERCEPTRON

Tulipa G. Silva
tulipa.silva.simepar@gmail.com
Federal University of Paraná
Av. Cel. Francisco H. dos Santos, 210, 81530-900, Curitiba, PR, Brazil

Paulo H. Siqueira
paulohs@ufpr.br
Federal University of Paraná
Av. Cel. Francisco H. dos Santos, 210, 81530-900, Curitiba, PR, Brazil

Cesar Beneti
cesar.beneti@simepar.br
SIMEPAR - Parana Meteorological System
Av. Cel. Francisco H. dos Santos, 210, 81530-900, Curitiba, PR, Brazil

Maiko Buzzi
maikobuzzi@hotmail.com
Federal Technological University of Paraná
Av. Sete de Setembro, 3165, 80230-901, Curitiba, PR, Brazil

Leonardo Calvetti
lcalvetti@gmail.com
Federal University of Pelotas
Av. Cel. Gomes Carneiro, 1, 96010-610, Pelotas, RS, Brazil

Abstract. *This paper 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. An application of the Multilayer Perceptron technique (MLP) is applied for the Severe Weather Event (SWE) forecast using weather radar as input data. After training, were obtained models that can support the decision about SWE alerts in the state of Paraná. The results indicate a detection of 81.40% for the studied SWEs and a 78.06% agreement for cases identified by a lightning detection network. This preliminary study showed that, for events up to 157 km radar, it is enough to evaluate the model that has as input the CAPPI radar product. This result can facilitate the forecast, because it does not require the volumetric input of the data in most cases of SWE detection.*

CILAMCE 2017
Proceedings of the XXXVIII Iberian Latin-American Congress on Computational Methods in Engineering
R.H. Lopez, L.F.F. Miguel, P.O. Farias (Editor), ABMEC, Florianópolis, SC, Brazil, November 5-8, 2017

Keywords: *Multilayer Perceptron, Severe weather event, Polarimetric radar*

1 INTRODUCTION

From January 2015 to July 2016, Civil Defense conducted 106 records of severe weather events in the state of Paraná, and at least 39753 people were directly affected by strong wind, heavy rain, hail and tornadoes (Defesa Civil, 2015,2016).

For nowcasting, radar is one of the best tools available, it offers real-time monitoring, with good resolution and enables three-dimensional view of the data (Fabry, 2015). In addition, it is possible to analyze a Severe Weather Event (SWE) based on the data of atmospheric electric discharges, a fact already explored by Lima (2005), Darden (2010), Murphy (2006) and Liu (2012).

Due to the large volume of data, machine learning techniques have been used in meteorology, both for classification as shown by Damian (2011) and Neto (2008) and for forecasting as shown by Anochi (2015). In order to identify the occurrence of SWEs, the Multilayer Perceptron (MLP) machine learning technique will be applied to the radar data, to obtain identification models that estimate if a storm contain a SWE in very short time (30 minutes or less). The results are compared to atmospheric electric discharges data to better evaluate the model.

2 METEOROLOGICAL CONCEPTS

2.1 Severe Weather Event (SWE)

A Severe Weather Event (SWE) is a storm capable of generating strong wind, heavy rain, hail, tornadoes and lightning (Maddox, 1980). These SWE usually occur in a convective environment of vertical development, due to the generation of currents within the storm cells (Fabry, 2015). These SWE are also associated with large social and economic impacts (Doswell, 2001).

In fact, from January 2015 to June 2016, more than 100 SWEs has occured in Paraná state, causing social and economic impacts, affecting 39753 people directly. Among the events reported by the Civil Defense, the different SWEs presented in Table 1 were selected for this study.

2.2 Dual Polarization Weather Radar

Radar is one of the best tools available for monitoring and forecasting as it provides real-time monitoring with good spatial resolution and allows a three-dimensional view of the data (Sauvageot, 1992), (Rinehart, 2004).

In a simplified form, the radar consists of a transmitter, a receiver, an antenna and a decoding and processing system. Radar is a remote sensing tool, that is, through the electromagnetic radiation, collects data without coming into direct contact with the study target (Novo, 1992).

Data collection occurs when the radar transmits a beam of electromagnetic radiation and captures the energy reflected by the particles present in the atmosphere. This energy is amplified and, given the time interval between emission and echo return, it is possible to determine the distance from the target to the radar.

Table 1: SWEs used in the study

Date	Time	SWE type	Affected city
01/02/2015	12:00 MN	windstorm	Foz do Iguaçu
07/13/2015	7:23 PM	windstorm	Matelândia
07/13/2015	8:15 PM	windstorm	Ampére
07/13/2015	9:53 PM	windstorm/heavy rain	Francisco Beltrão
07/14/2015	10:15 AM	tornado	Francisco Beltrão
11/23/2015	1:30 PM	windstorm	Umuarama
11/26/2015	11:46 PM	heavy rain	Francisco Beltrão
12/28/2015	9:15 PM	windstorm	Francisco Beltrão
02/27/2016	9:53 AM	heavy rain	Francisco Beltrão
03/22/2016	3:15 AM	tornado	Capitão Leônidas Marques

In conventional radar, with single polarization, the emitted and received electromagnetic signal has only one orientaton (vertical or horizontal), with horizontal orientation being more common. Thus, a simple polarizing radar is able to obtain information from the targets only in one orientation. The dual polarization radar has the propagation of electromagnetic signal with orientations both horizontally and vertically, being able to obtain information in the two orientations (Fabry, 2015). The polarimetric variables provide additional information about the shape and even size of the target (Kumjian, 2013), (Fabry, 2015).

The data collected by the radar are in spherical coordinates, where the position of each data is described in terms of (θ, ϕ, r), where θ is the angle of antenna elevation, ϕ is the angle of clockwise rotation of the antenna relative to the geographic North, called the azimuth angle, and r is the distance from the radar to target. The data collected from the radar for a full scan and fixed elevation is called PPI (Plan Position Indicator). The collection of successive PPIs of the radar is a volumetric scanning. A two dimensional representation of the volume data is obtained by a selection of ranges at several PPIs, called CAPPI (Constant Altitude Plan Position Indicator). Therefore, a CAPPI product (Fig. 1) is the information at a certain altitude, so that all elevations can contribute to the generated product (Fabry, 2015).

The following variables, obtained from radar volume data were used in this study:

Reflectivity (Z). The measure of a target's efficiency in intercepting and returning the energy emitted by the radar is called reflectivity. Thus, reflectivity is dependent on the sizes, formats, quantity and dielectric properties of the targets. Reflectivity is one of the most widely used variables in weather forecasting. High reflectivity values (up to 45 dBZ) are generally associated with heavy rainfall, if these values are between 5 and 12 km of altitude, they are associated with severe storm (Fabry, 2015).

Radial velocity (V). The radial velocity is how fast the detected target is moving away or approaching the radar. Using the Doppler effect, the radar estimates the velocity at which the target moves away or stay closer to the radar in the direction of the beam. By convention, negative values represent particles comming toward the radar and positive values, particles moving

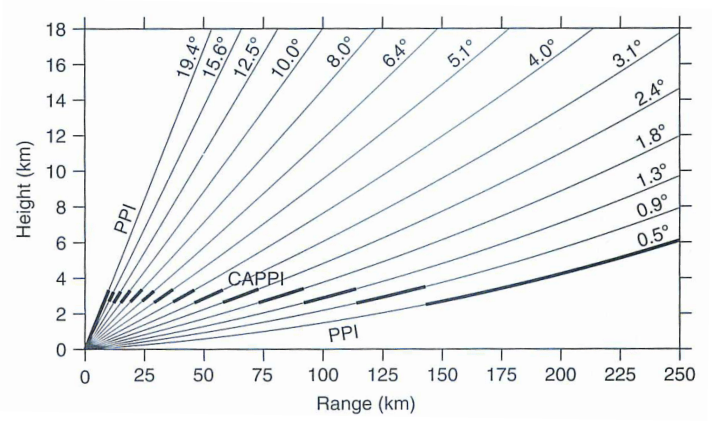


Figure 1: Example of PPI and CAPPI. In this case, a CAPPI is obtained by a selection of ranges at several PPIs to create a 2D image of the volume scan. Figure source: Fabry (2015).

away from radar (Fabry, 2015).

Differential Reflectivity (ZDR). Differential Reflectivity (ZDR) is the difference between Z_{HH} (reflectivity with signal emitted and captured horizontally) and Z_{VV} (reflectivity with signal emitted and captured vertically). The variable ZDR is widely used to differentiate rain, hail, snow, and even non-meteorological targets (Kumjian, 2013).

Copolar Correlation (RHOHV). The Copolar Correlation (RHOHV) has no unit of measure and represents the correlation between the horizontal and vertical polarized Z signals at a given point in space (Rinehart, 2004). The values of $RHOHV$ vary between 0 and 1.0 and can be seen as a measure of the similarity between horizontal and vertical polarization signals (Kumjian, 2013). This variable describes physical characteristics of the target, being closer to 1.0 the more uniform the target, since the vertical and horizontal signals tend to be the same.

Phase Differential (PHIDP). The phase differential (PHIDP), measured in degrees, is the difference between the phase emitted and received horizontally and the phase emitted and received vertically in a pulse (Fabry, 2015). The variable $PHIDP$ has radial variations, since cumulative changes in the phase difference for the complete pulse journey. Thus, $PHIDP$ is not rarely replaced by its spatial derivative, KDP (Fabry, 2015).

Specific Phase Differential (KDP). The Specific Phase Differential (KDP) variable, is the spatial derivative of PHIDP (Fabry, 2015). KDP is an excellent estimator for precipitation, as presented by Ruzanski and Chandrasekar (2012), values above 2°/km indicate a significant amount of liquid or oblate water.

Due to vertical development, convective storms are associated with shear, especially at low and medium levels (Fabry, 2015). Traditionally, the shear is obtained by the difference of the Doppler velocity between two points, divided by the distance between these two points, being possible to describe 3 shear fields: vertical, radial and azimuthal (Newman et al., 2013), as described below .

The ***azimuthal shear (AZS)*** is the difference of velocity between 2 bins at the same distance r from radar, at the same elevation θ and consecutive azimuths ϕ_i and ϕ_{i+1} and divided by the distance d between these two bins.

The ***radial shear (RS)*** is the difference in velocity between consecutive 2 bins r_i and r_{i+1},

at the same elevation θ and azimuth ϕ, divided by the distance d between these two bins.

The ***vertical shear (VS)*** is the velocity difference between 2 bins with the same azimuth ϕ and radar distance r, but in consecutive elevations θ_i andθ_{i+1}, divided by the distance d between these two bins.

Fig. 2 shows the area of the radar used in this study, a polarimetric weather radar located in the west of Paraná state.

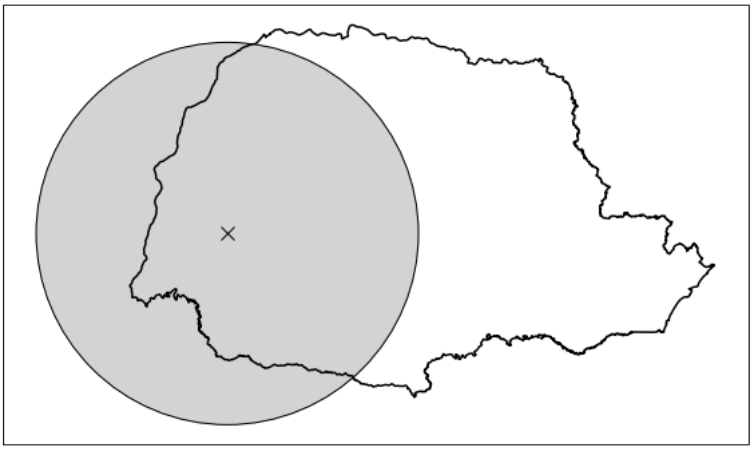


Figure 2: Example of the range covered by polarimetric radar from SIMEPAR over Paraná state. The circle centering position is where is the radar.

2.3 Atmospheric electric discharge

Atmospheric electric discharges are the result of the development of intense electric charge centers in the cloud, which exceeded the electrical insulation capacity of air resulting in the dissipation of the electric charge centers. Atmospheric electric discharges can indicate the severity of a storm by its position and amount. As presented by Holle, et. al. (1994), lightning tends to increase and concentrate in the convective regions during the storm maturity of a convective storm and decay in the phase of storm dissipation.

The types of atmospheric electrical discharges are classified according to where they originate and where they arrive (Lima, 2005): Intra Cloud (IC) and Cloud to Ground (CG). The increase in total density of atmospheric electrical discharges (Total Lightning-TL), sum of the densities of CG and CI, is a good indicator of SWE. This phenomenon is called Lightning Jump (LJ), and it happens minutes before the occurrence of a SWE (Schultz et al, 2011).

Earth Networks Total Lightning Network (ENTLN) is the world's largest global network for detecting IC and CG types, with high sensor density and operating in real time (Liu and Heckman, 2012). In this work, IC and CG data provided by ENTLN are used to calculate CG density and LJ from total lightning (CG added to IC).

3 MULTILAYER PERCEPTRON (MLP)

As defined by Haykin (1999), an artificial neural network is a parallel distributed system massively interconnected by simple processing units called neurons whose modeling is inspired by the functioning of biological neurons.

The figure 3 exemplifies the structure of a neuron and its major components: input signals x, sinaptic weights w, summing junction, bias b, activation function $f(.)$, and output y. The k-index indicates that it is the k-th neuron of the artificial neural network.

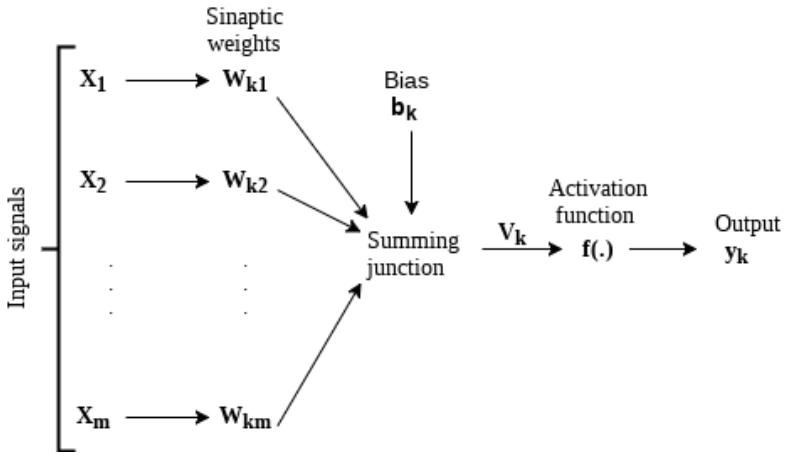


Figure 3: Representation of an artificial neuron.

The synaptic weights indicate the weight that each component of the input have in the summing junction, representing the connection among neurons in the network. The summing junction is a linear combiner, it is responsible for the sum of the input signals,considerate by the weight of each synapse. The activation function generates the output of the neuron from the value of the sum function and the associated bias, also ensuring that the output is in a finite amplitude range. The bias is a linear term to be added to the summing junction, it represents a factor with external explanation.

Mathematically, given an input $x \in \Re^n$, such that $x = (x_1, x_2, ..., x_n)$ and has n features, a weight vector $w_k \in \Re^n$, with $w = (w_{k1}, w_{k2}, ..., w_{kn})$. Then, the output of this neuron is given by

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

such that

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

is the summing junction, $b_k \in \Re$ is the bias and $f(.)$ is the activation function.

In general, for output $\in [0, 1]$, the most used activation function is the sigmoidal, presented by Eq. 1 (Marsland, 2015).

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

The learning of an artificial neural network is stored in the synaptic weights that are modified (updated) at each learning iteration, until one or more pre-determined stop criteria are satisfied (Haykin, 1999).

The perceptron neural network is a supervised network (uses the desired output to improve its learning) that uses one or more independent neurons for processing, capable of solving linearly separable problems. The Multilayer Perceptron (MLP) network is an enhancement of

perceptron, including intermediate layers of neurons, making possible its application to non-linearly separable problems, widely used for its simplicity and easy implementation (Marskin, 1999).

The Fig. 4 exemplifies the architecture of an MLP.

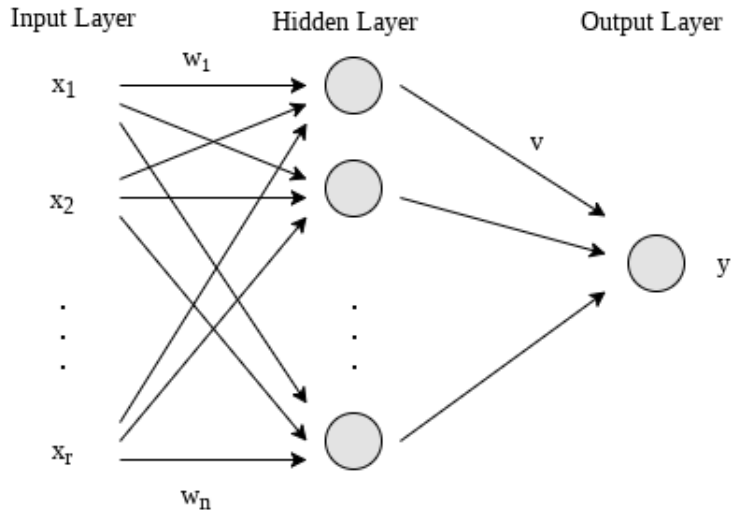


Figure 4: Example of the architecture of a MLP.

The MLP exemplified in Fig. 4 has 3 layers: the input layer, output layer and an intermediate layer, also known as layer of hidden neurons or simply hidden layer. This MLP has input with r attributes, n hidden neurons, and 1 output.

Thus, in this MLP (Fig. 4), given an input $x = (x_1, x_2, ..., x_r)$ then

$$z_j = f(z_j^* + \theta_{aj}) = f(w_j^T x + \theta_{aj}), \tag{2}$$

where $f(.)$ is a activation differentiable function, $w_j^T x$ is the scalar product between the weight vector w_j and the input vector x. The θ_{aj} component is the linear term (bias) related to j-th hidden neuron position .

In that case, the output y has the form:

$$y = f(y^*) = f(v^T z + \theta_b), \tag{3}$$

where θ_b is the linear term related to output, $v^T z$ is the scalar product between the vector of hidden layer weights v and z vector whose components are obtained by Eq.2 and $f(.)$ is a activation function.

One way to update the synaptic weights during MLP learning is to use the backpropagation algorithm.

The backpropagation algorithm is a down-gradient-based method for calculating the derivatives of the error function in relation to the synaptic weights in order to find a set of weights to minimize error through repeated exposure of the instances (Bishop, 1995). According to Haykin, (1999) and Marsland (2015), it is the main algorithm used in the training of supervised neural networks.

Let E be the error associated with output y, v_j the component of the vector of weights v

and d the desired output for y. By the chain rule

$$-\frac{\partial E}{\partial v_j} = \frac{\partial E}{\partial y^*}\frac{\partial y^*}{\partial v_j} = \frac{\partial E}{\partial y^*} z_j.$$

The output error is the difference between the desired output and the desired output. Thus, in order to proceed in the opposite direction of the error, set

$$\frac{\partial E}{\partial y} = -(d - y).$$

Then, the weight adjustment for v_j is given by Eq. 4. In Eq. 4 η is the learning rate. According to Bishop (1995), η is a positive and small value. Under these conditions, the weight vector converges to the point at which the error is minimized. However, if η is too small, learning becomes slow, and if η is large value, learning may diverge (Bishop, 1995).

$$\Delta v_j = \eta(d - y) f'(y^*) z_j \tag{4}$$

Suppose $f(.)$ Is the sigmoidal function as the activation function in all neurons of the output layer. Thus,

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

implies in Eq. 5.

$$f'(y^*) = \frac{e^{-y^*}}{(1 + e^{-y^*})^2} \tag{5}$$

Substituting 5 into 4 and simplifying, the term for updating the weights of the output layer is given by

$$\Delta v_j = -y(1 - y)(d - y) z_j. \tag{6}$$

That is, v_j is updated by

$$v_j = v_j + \Delta v_j. \tag{7}$$

Similarly, with a sigmoidal activation function, the update term for the weights w_{ij} is

$$\Delta w_{ij} = \eta \sum_k (d_k - y_k) y_k (1 - y_k) v_{jk} z_j (1 - z_j) x_i, \tag{8}$$

and updated w_{ij} is

$$w_{ij} = w_{ij} + \Delta w_{ij}. \tag{9}$$

The learning of a neural network supervised through backpropagation can be summarized by the following steps:

1. Present a training instance, to the artificial neural network;
2. Determine the output y by the artificial neural network for the presented input;
3. Calculate the error in the network output comparing to the desired output d;
4. Update the weights of the neurons of the outer layers to reduce the error;
5. Backpropagate the error to the neurons from the innermost layers, towards the input layer;
6. Adjust the weights of the innermost neurons;

7. Repeat the previous steps until the stop criterion is satisfied.

The main optimal stopping criterion for MLP is the minimum error curve of the validation set. The validation set is extracted from the test set that has its error computed at each iteration of the learning process. If the error in the validation set goes through to a minimum it means that the network is losing its generalization capacity (Marsland, 2015).

4 METHODOLOGY

The methodology applied in this study can be described by the following steps:

1. Polarimetric radar data collection;
2. Pre-processing of data (correction of missing data and resolution, normalization, shear calculation, removal of noise and reduction of dimensionality of the input) and inclusion of the classes to the data set;
3. Separation of the test and training sets;
4. Training of the MLP technique, generating the models M-MLP3D and M-MLP2D;
5. Identification of SWE regions using atmospheric electric discharges data, and generation of the M-ENTLN model;
6. Comparison of the models obtained by MLP with each other and with M-ENTLN and study and analysis of the results.

Each volumetric data (radar file), has points to each 500 m from radar up to 240 km, for between 10 to 17 elevations and 360 azimuths. The CAPPI product (Fig 1) has points to each 500 m from radar up to 240 km.

As pre-processing data was used correction of missing data and resolution, normalization, calculation of shears, noise removal, reduction of dimensionality of the input and transform the radar file data to tridimensional volume in cartesian coordinates (grid).

After the radar data interpolation into a cartesian grid was performed, the points where Z were less than 30dBZ were excluded because they are not representative in a convective storm (Maddox, 1980).

A point in a convective storm that "the information of the occurrence of SWE in this storm is unknown" is classified as 0 (desired output) and a point that "the information of the occurrence of SWE in this storm is known" is classified as 1 (desired output).

The M-MLP3D model is obtained by the training MLP technique over the volumetric radar data, and M-MLP2D model is obtained training the MLP technique with CAPPI radar product.

Each radar file has only one region where the information of the SWE occurrence is known. Among the 43 files (with a complete volumetric scan) available, 33 were used for the training of both techniques (M-MLP3D and M-MLP2D) and 10 to verify the generalization (test set).

The input data set consists of vectors containing the normalized values of Z, ZDR, RHOHV, KDP, HMAX, AZS , RS, and and VS for each grid point and for each radar file, in that order. Where HMAX is the altitude of the point of greatest reflectivity per column.

Both M-MLP3D and M-MLP2D uses one hidden layer and 43 hidden neurons and sigmoidal activation function. More than 43 neurons and other activation function did not show improvement to the models. For MLP training, the free software Scikitlearn, from Python (Smola, 2004) was used. This package has a wide range of machine learning techniques implemented and easy to use (Marsland, 2015).

For the selection of the test set files, an event should be contained integrally in the set. The files related to the randomly selected SWE were those for the event of 07/13/2015.

The M-ENTLN model was obtained by the identification of CG density and LJ by region. If there was a high CG density or there was LJ, probability there was a SWE there. This model is an indirect way to estimate a SWE occurrence and compare with the other methods.

5 RESULTS AND DISCUSSIONS

MLP training generated two identification models, M-MLP3D using volumetric radar scan and M-MLP2D that uses the CAPPI product at 3 km, while model M-ENTLN was obtained only using atmospheric electric discharges.

Although the input (and output) in the models were point-by-point in the grid, we evaluated regions within the grid, because the SWE is a phenomenon without a precise location within the storm.

Fig. 5 shows an example of the models obtained by MLP M-MLP3D and M-MLP2D and also M-ENTLN. In this case, M-MLP3D identified a region where the other models did not. In fact, M-MLP2D identifies less regions than M-MLP3D, but in most of cases, those regions are unidentified by M-ENTLN, suggesting that those regions can be misidentified by M-MLP3D. The identification by M-MLP2D is the same as M-MLP3D to the previously identified regions as presented by Table 2. That is, there is no loss of information for the regions in which the occurrence of SWE is known. Generally M-MLP2D indetify 21.93 less % regions than M-MLP2D. In the test set this difference in lower: 10.11 % less.

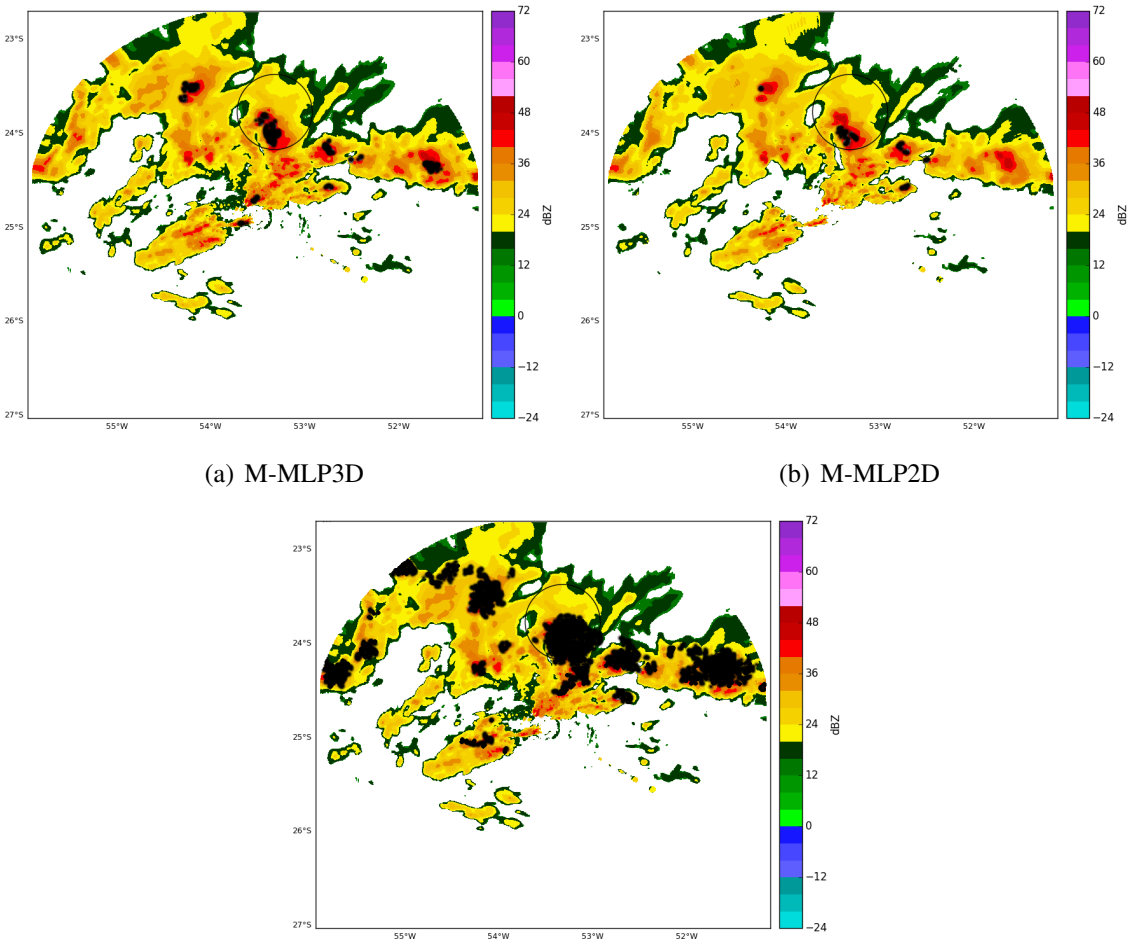
In Table 2, the accuracy of the training set shows the performance of model, while the accuracy of test set shows the generalization capacity of models and the general set shows a complete view about the models.

Table 2: Accuracy for the M-MLP3D and M-MLP2D models.

Model	Training set	Test set	General set
M-MLP3D	78.78 %	90.00 %	81.40 %
M-MLP2D	78.78 %	90.00 %	81.40 %

In Fig. 5, the region inside the circle was an example of pre-identified region with SWE occurrence. This region was classified as 1, but not all internal points were classified by the models. This is because SWE did not occur in all the area, and these points could be misclassified.

In the training of the M-MLP3D and M-MLP2D models, the regions in which the occurrence of SWE are known (exemplified by the circle in the Fig. 5) refer to 10 specific events.



(a) M-MLP3D

(b) M-MLP2D

(c) M-ENTLN

Figure 5: Output obtained by (a) M-MLP3D, (b) M-MLP2D and (c) M-ENTLN for 11/23/2015. The images represent the points identified by the respective models (in black) on the reflectivity attribute image. The circled region shows where SWE occurred within 30 minutes.

However, it can be seen by the Fig. 5 that there are other regions identified by the models, but whose occurrence information of SWE is unknown.

All regions that do not correspond to the events studied do not have previous information about the occurrence of SWE. However, this information can be obtained indirectly when comparing the regions identified with M-ENTLN. In addition, the M-ENTLN model obtained 100 % accuracy in the identification of the known SWEs, and therefore it is the most accurate model in the study data. Because of that, the M-ENTLN model is used to study not pre-classified regions.

The occurrence of SWE is possible without occurrence of LJ, but whenever there is high density CG it can be said that it is a SWE of the electric storm type. Therefore, classification by M-ENTLN may still contain poorly classified data, in a small scale.

In this way, in order to validate the identification of M-MLP3D and M-MLP2D in regions where information on the occurrence of SWE is unknown and also to verify if the regions disregarded by M-MLP2D are relevant, the M-ENTLN model will be used.

Table 3 presents the percentage of regions identified by MLP models where M-ENTLN

confirms the SWE and the percentage of regions in which the MLP models identify SWE but M-ENTLN does not confirm. In this table R1, R2, R3, R4, R4 and R5 represent:

- **R1**: Regions identified by M-MLP3D and M-ENTLN with 240 km radius from radar;
- **R2**: Regions identified by M-MLP2D and M-ENTLN with 240 km radius from radar;
- **R3**: Regions identified by M-ENTLN and not identified by M-MLP3D with 240 km radius from radar;
- **R4**: Regions identified by M-ENTLN and not identified by M-MLP2D with 240 km radius from radar;
- **R5**: Regions identified by M-ENTLN and not identified by M-MLP3D with 157 km radius from radar;
- **R6**: Regions identified by M-ENTLN and not identified by M-MLP2D with 157 km radius from radar.

Table 3: Comparison between identification of SWE by M-ENTLN, M-MLP3D and M-MLP2D

Set	R1	R2	R3	R4	R5	R6
Training	75.00%	89.89%	32.50%	37.5%	16.19%	11.07%
Test	57.31%	62.12%	30.88%	33.78%	31.52%	15.20%
General	68.58%	78.06%	32.02%	35.98%	17.11%	13.72%

From the results presented in Table 3, it is possible to say that the M-MLP2D model has a better correlation with the regions identified by M-ENTLN up to a radius of 157 km radar. It fails to identify, in the test set, 15.20 % regions that M-ENTLN identifies, while M-MLP3D does not identify 31.52 % on the same radius.

Based on the Table 3, it is possible to affirm that the regions not identified by M-MLP2D are not relevant to the identification of up to 157 km of radar, but if there is a need to study beyond 157 km, M-MLP3D has a higher rating.

Both M-MLP3D and M-MLP2D models presented good decision support tools, since they have a correct identification in 81.40 % of the cases studied. In the test set, the correct identification is 90 %, showing the generalization capacity of the models.

The M-ENTLN model, for the regions where the occurrence of SWE is known, has an accuracy of 100 % and therefore, it was used to validate the models M-MLP3D and M-MLP2D. However, M-MLP3D has more identified regions that are not identified by M-ENTLN when compared to M-MLP2D.

If the entire range of the radar is considered, M-MLP2D identified more regions than M-MLP3D in which M-ENTLN agrees, but more regions in which M-ENTLN identifies and M-MLP2D does not identify.

When considered up to 157 km from radar, the number of regions in which M-ENTLN identifies and M-MLP2D does not identify becomes smaller than the error of M-MLP3D. In other words, M-MLP2D identifies less regions than M-MLP3D, but identifies more regions that

agree with M-ENTLN.

Thus M-MLP2D is a good tool for the identification of SWE, up to 157 km radar, with better identification than MLP3D.

Since from 157 km (for 0.5 degree elevation) the CAPPI data are the same as the PPI and may not represent the storm at the chosen altitude (they are above 3 km, as illustrated by Fig. 1), it was expected that regions over 157 km from radar could be misclassified. Thus, for data up to 157 km from radar, it is adequate to use only M-MLP2D for SWE identification, decreasing the input volume and consequently processing time.

ACKNOWLEDGEMENTS

This research was funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Paraná Meteorological System (SIMEPAR). The atmospheric electrical discharges data was kindly provided by Earth Networks Total Lightning Network (ENTLN).

REFERENCES

Anochi, J., 2015. *Previsão climática de precipitação por redes neurais autoconfiguradas*. Instituto Nacional de Pesquisas Espaciais/Cachoeira.

Bishop, C., 1995. *Neural Networks for Pattern Recognition.* Clarendon Press.

Damian, E., 2011. *Duas metodologias aplicadas à classificação de precipitação convectiva e estratiforme com radar meteorológico: SVM e K-means*. Master thesis, Universidade Federal do Paraná/Curitiba.

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

DeSoto, C.,1945. Radar Techniques - Primer Principles. *QST*, vol XXIX, pp. 1-20.

Doswell, C., 2001. *Severe Convective Storms*. American Meteorological Society, Boston.

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

Gatlin, P. & Goodman, S.,, 2010. A Total Lightning Trending Algorithm to Identify Severe Thunderstorms. *Journal of atmospheric and oceanic technology*, vol. 27, pp. 744-755.

Harrington, P., 2010. *Machine Learning in Action*. Manning Publications, ISBN 9781617290183.

Haykin, S., 1999. *Neural Networks: A Comprehensive Foundation*. Pearson Education, ed. 2.

Holle, R., Watson, A., Lopez, R., MacGorman, D. & Ortiz, R., 1994. The life Cycle of Lightning an Severe Weather in 3-4 June 1985 PRE-STORM Mesoscale Convective System. *Monthly Weather Review*, vol. 122, pp. 1798-1808

Kumjian, M., 2013. Principles and Applications of Dual-Polarization Weather Radar. Part I: Description of the Polarimetric Radar Variables. *J. Operational Meteorology*, pp. 226-242, `http://nwafiles.nwas.org/jom/articles/2013/2013-JOM19/2013-JOM19.pdf`.

Kumjian, M., 2013. Principles and Applications of Dual-Polarization Weather Radar. Part II: Warm- and Cold-Season Applications. *J. Operational Meteorology*, pp. 243-264.

Lima, K., 2005. *Descargas Elétricas Atmosféricas em Sistemas Convectivos de Mesoescala no Sul da América do Sul*. Master thesis, Universidade Federal de Pelotas/Pelotas.

Maddox, R., 1980. Mesoscale convective complexes. *Bull. Am. Meteorol. Soc.*, vol. 61, pp. 1374-1387.

Marsland, S., 2015. *Machine Learning: An Algorithmic Perspective*. CRC, ed. 2.

Mitchell, T., 1997. *Machine learning.* McGraw-Hill, New York.

Murphy, M. & Holle, R., 2006. Warnings of cloud-to-ground lightning hazard based on total lightning and radar information. *Second Conf. on Meteorological Applications of Lightning Data.* Atlanta, GA.

Neto, M., 2008. *Mineração visual de dados: extração do conhecimento a partir das técnicas de visualização da informação e mineração de dados*. Master thesis, Universidade Federal do Paraná/Curitiba.

Newman, J., Lakshmanan, V., Heinselman, P., Richman, M. & Smith, T., 2013. Range-Correcting Azimuthal Shear in Doppler Radar Data. *Weather and Forecasting*, vol. 28, pp. 194-211.

Novo, E., 1992. *Sensoriamento Remoto: princípios e aplicações.* Edgard Blücher.

Rinehart, R., 2004.*Radar for Meteorologists*. Columbia, ed. 4.

Ruzanski, E. & Chandrasekar, V., 2012. Nowcasting Rainfall Fields Derived from Specific Differential Phase. *Journal of applied Meteorology and Climatology*, vol 51, pp. 1950-1959.

Sauvageot, H., 1992. *Radar Meteorology*. Artech House Publishers, ed. 1.

Schultz, C., Petersen, W., Carey, L., 2011. Lightning and Severe Weather: A Comparison between Total and Cloud-to-Ground Lightning Trends. *American Meteorological Society*, vol. 26, pp.744-755.

Smola, A., 2004. A Tutorial on Support Vector Regression. *Bernhard Schölkopf-Statistics and Computing archive*, vol. 14, n. 3, pp. 199-222.