A comparative study of the output skill scores of a forecast cloud-to-ground lightning system based on a neural network and different mesoscale models

Wendell Farias
Federal University of Santa Catarina

Julio Pablo Reyes Fernandez
National Institute for Space Research, Brazil

GROUND'2008
&
3rd LPE

International Conference on Grounding and Earthing
&
3rd International Conference on
Lightning Physics and Effects
Florianopolis - Brazil November, 2008

A COMPARATIVE STUDY OF THE OUTPUT SKILL SCORES OF A CLOUD-TO-GROUND LIGHTNING FORECAST SYSTEM BASED ON A NEURAL NETWORK AND DIFFERENT MESOSCALE MODELS

Zepka, G. S. Pinto Jr., O. Farias, W. R. G. Fernandez, J. P. R.

INPE, National Institute for Space Research, S. J. Campos, SP, 12201-970, Brazil

Abstract - To assure that an artificial neural network (NN) has a lightning forecast character with real practical applicability, at first it seems reasonable to choose input variables really representative of the phenomenon which confer results with errors inside acceptable limits. The purpose of this work is to compare the output skill scores of a lightning forecast neural system when it is fed with meteorological parameters simulated from two different mesoscale models: Eta and WRF (Weather Research and Forecasting). In both cases, accumulated values of cloud-to-ground (CG) lightning data provided by Brazilian Lightning Detection Network (BrasilDat) were used as input to NN as well. All input variables were selected for a specific time in the morning in order to predict (the NN output) the CG lightning activity that would occur in the afternoon of the same day. The forecasting output is presented in terms of a CG lightning activity index as: low, medium and high electrical activity. Based on the results of this comparative analysis, no significant differences were found in the output skill scores even though the WRF numerical outputs were integrated in time and the same did not happen with Eta model whose outputs were analysis fields.

1 - INTRODUCTION

Due to the increasing dependence of society by efficient technologies able to forecast or even to detect severe weather, studies about the adequate meteorological conditions to the occurrence of cloud-to-ground (CG) lightning have been lately intensified. Thunderstorms do not form just anywhere. Thunderstorms need a favorable environment in which to develop. This favorable environment, in its most elemental analysis, requires just two ingredients to come together. One is a potentially unstable air mass, and the other is a lifting mechanism to start the release of the instability [1]. There are no specific amounts of each component that ensures thunderstorm development. In fact, a low quantity of one may be compensated by an abundance of the other. That is the challenge of thunderstorm forecast.

The success of numerical modeling combined with the rapid progress of computational power has led to the mesoscale models improvement in the operational environment of severe weather forecasting. To take advantage as much as possible of the benefits offered by these systems requires mastering techniques to process and analyze the excessive amount of information from the numerical outputs [2]. Applications of a mesoscale numerical model can provide an effective method to research the cause of the thunderstorm and consequently the lightning formation. Based on meteorological satellite data, it is known that most thunderstorms in Brazil are associated with local convection conditions, fronts and large mesoscale convective systems [3].

Neural network (NN) is an interesting way of approaching the solution of intelligence problems, simulating artificially the human brain operations [4]. Compared to other statistical tools, an NN stands out because of it does not require any prior knowledge of a solution, trying to recognize data patterns and regularities. However, this technique in lightning forecast has not been commonly used due to the complexity of atmospheric processes and, until recently, to a representative climatology of the lightning activity. At the moment, only a few forecast lightning systems using artificial intelligence as basic tool are available.

To associate mesoscale models and neural networks represents an effort of applying an artificial intelligence technique to predict the storm occurrence with lightning. The fundamental idea is to feed an NN with some meteorological variables simulated in numerical models getting as output the level of the storm electrical activity in a next future. Figure 1 presents this goal. In this case, the NN can be seen as a transfer function, relating some inputs with one output. Zepka et al. [5] showed preliminary results of a CG lightning forecast system using as NN inputs CG lightning data and analysis fields of meteorological parameters obtained with the Eta model. The proposal of the present work is to compare the output skill scores of this lightning forecast neural system when it is fed with meteorological parameters simulated from two different mesoscale models: Eta and WRF (Weather Research and Forecasting). Accumulated values of CG lightning data were used as input to NN as well. Once the NN has learned the atmospheric dynamics, the lightning activity is predicted for a few hours later in terms of a severity index defined as: low, medium and high lightning activity.

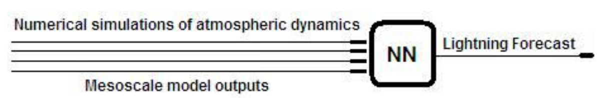

Figure 1 - Basic diagram of the lightning forecast neural system.

2 - LIGHTNING FORECAST NEURAL SYSTEM

A neural network can be defined as a processing structure capable of implementation in electronic devices, composed by interconnected units called artificial neurons which present a specific behavior of input/output determined by its transfer function, the interconnections with other adjacent units and possibly by external inputs ([4], [6]). Although it is possible to design an NN from the role it should play, combining the effects of all individual neurons, an NN usually adapts itself to achieve the desired functionality from one or more learning strategies which will act near by NN configurable parameters [6]. So the benefit to use NN in conjunction with an engineering issue is that it is only necessary to focus on input and output, and a problem can be solved without aware of exact principle of cause and effect [5].

The NN proposed architecture to the lightning forecast system is a backpropagation, multilayer, feedforward and fully connected network. It was used a backpropagation with momentum as training rule and an axon as activation function with genetic algorithm ([4], [7]). Figure 2 illustrates the MultiLayer Perceptron (MLP) adopted in this study. The MLP network trained with the backpropagation algorithm has been the NN model most often used in problems of patterns classification. Its main characteristic is the ability of generalization and universal approximation of functions ([4], [6]).

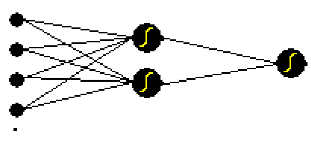

Figure 2 - MLP topology composed by the input layer, only one intermediate layer with two hidden neurons and the output layer.

3 - LIGHTNING FORECAST NEURAL SYSTEM INPUTS

The input set is composed by hourly number of lightning flashes from 00 to 09 local time (LT) and meteorological parameters obtained from the Eta and WRF mesoscale models at 12 UT (10 LT because of the summer time in Brazil). In order to the NN to have a forecast character with real practical applicability these possible meteorological predictor variables were picked and chosen based on the necessary conditions for the storm formation, i. e., atmospheric thermal profile, humidity and upward movement.

CG lightning data from December 2005 and January 2006 were provided by the Brazilian Lightning Detection Network (BrasilDat) for the *Companhia Paulista de Força e Luz* (CPFL Energy) area which extends from 21^{o} to 24^{o} S latitude and 50.5^{o} to 45.5^{o} W longitude across the state of São Paulo (Figure 3). More details about the BrasilDat can be obtained in Pinto Jr. et al. ([8], [9]).

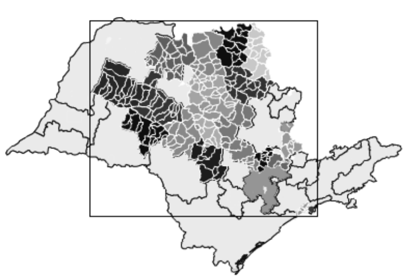

Figure 3 - Map of the São Paulo state (in the Southeast region of Brazil). Inside the black rectangle is located the CPFL area.

Analysis fields at 12 UT of Convective Available Potential Energy (CAPE) index, Best Lifted Index (BLI) and divergence from the Eta model with horizontal resolution of 20 km were chosen as meteorological NN inputs. These variables were extracted for the CPFL area from Eta domain covering most of the South America continent and adjacent oceans. The initial condition was taken from analysis of the NCEP (National Centers for Environmental Prediction) and the lateral boundary conditions were taken from the CPTEC (Center for Weather Forecasting and Climate Studies)/ COLA (Center for Ocean-Land-Atmosphere Studies) Global Model forecasts and updated every 6 hours. Model details are given in Black [10].

Similar meteorological parameters were selected from WRF model as NN inputs with the intent to compare the NN performance and to analyze which mesoscale model better represents the atmosphere during the lightning occurrence. It was used forecasts at 12 UT of Convective Available Potential Energy (CAPE) index, Lifted Index (LI) and divergence to feed the forecast neural system instead of analysis fields. The WRF model is run on a 20 km resolution grid over the CPFL area. The initial and boundary data were taken from the 00 UT NCEP Global Forecast System (GFS) analysis. The detailed description of WRF is presented in Wang et al. [11] and Skamarock et al. [12].

4 - COMPARATIVE ANALYSIS

A running average was applied to CG lightning data and to the meteorological variables before becoming NN inputs. The number of CG lightning was converted in an index, called CG Lightning Activity Index (LAI), in order to minimize random variations. This index was classified as follows: 0 to low, 1 to medium and 2 to high CG lightning activity.

The NN was trained in a dynamic way. The whole month of December 2005 was used in order to predict the first day of January 2006. The January predicted days were replacing the December trained days so as to maintain always 31 days in training set. So, January 2006 was the forecast month. The trained NN has foreseen the CG lightning behavior in the afternoon (15-18 LT) of the same inputs days according to the scale provided by the index cited above.

Figure 4 shows the result of the lightning forecast neural system using CG lightning data and analysis fields from Eta model as inputs. The output skill score achieved was 81%. The black curve represents the real quantity of CG lightning which occurred from 15 to 18 LT in each January 2006 days converted to LAI index. The blue curve is the lightning forecast neural system outputs. Its values are decimal fractions. Thus, in order to compare the real and predicted outputs it was necessary to make an approximation of the NN outputs to the nearest LAI index value. This approach is represented by the orange curve in the graphic.

Figure 5 presents the result of the lightning forecast neural system using now CG lightning data and meteorological forecasts from WRF model as inputs. Comparing the real and predicted output curves note that the output skill score is 84%.

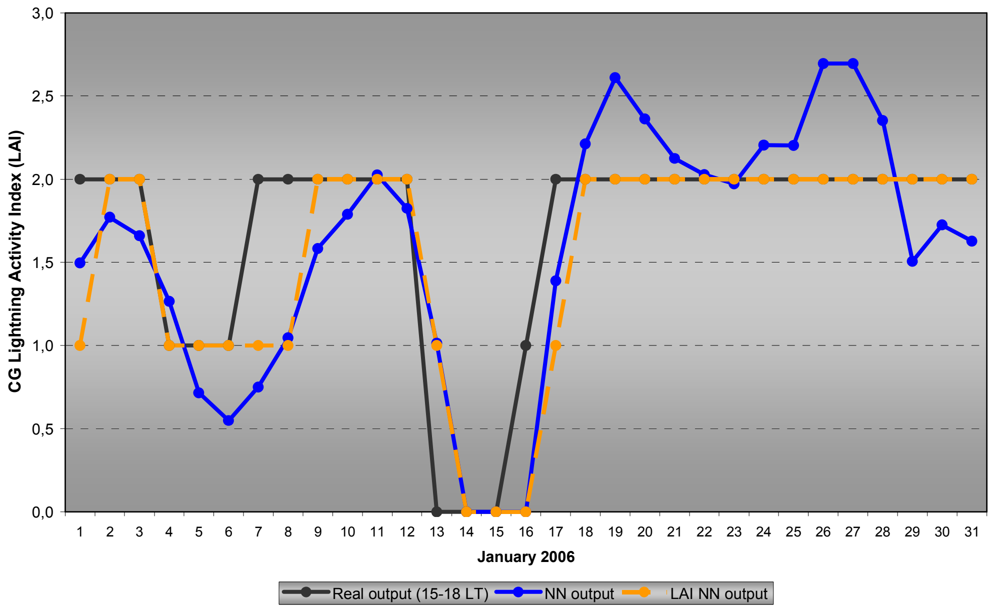

Figure 4 - CG lightning forecast in terms of LAI index using lightning data and meteorological variables CAPE, BLI and divergence from Eta model as NN inputs.

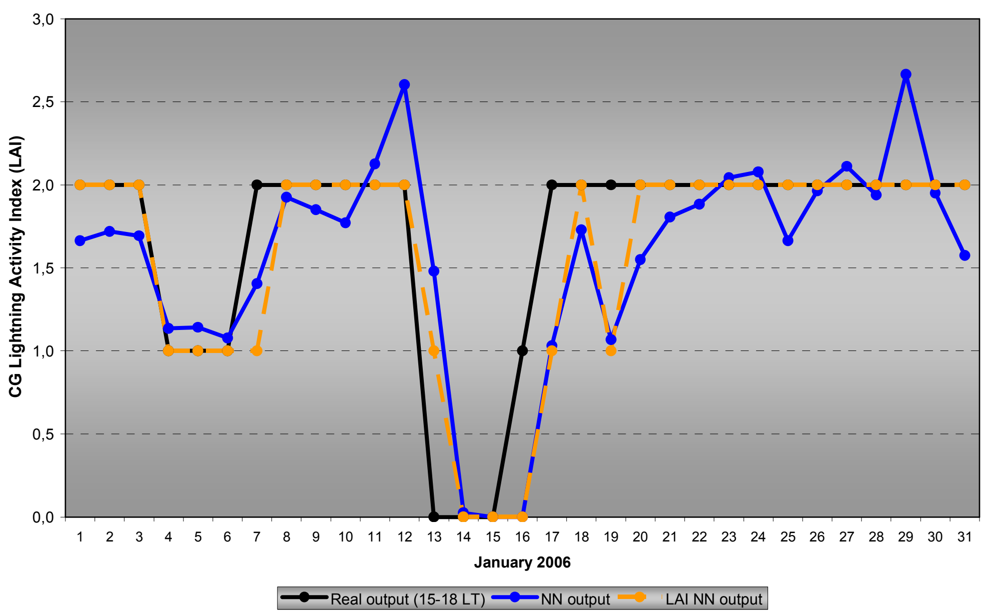

Figure 5 - CG lightning forecast in terms of LAI index using lightning data and meteorological variables CAPE, LI and divergence from WRF model as NN inputs.

5 - CONCLUSIONS

The integrated use of thorough knowledge about the responsible mechanisms by weather systems formation in different scales and results of prognostic, diagnostic and statistic models compose the basis for an efficient short term forecast system. The study presented in this work allows one to conclude that an NN can be an important tool for a lightning forecast system when it was fed with CG lightning data and meteorological parameters from mesoscale models. No significant differences were found in the output skill scores when the meteorological parameters are obtained from the Eta and WRF models, even though the WRF numerical outputs were integrated in time and the Eta numerical outputs were analysis fields. The results although satisfactory promising should be corroborate through more analyses in the future.

Independent of the mesoscale model considered, it was verified that the chosen NN input variables well represent the phenomenon, or either, they are correlated with the electrical storm. It is important to say that the good performance of the forecast system depends on the precision and resolution of the mesoscale model and the lightning detection system for the considered area.

6 - REFERENCES

[1] McCann, D. W. (1992) A Neural Network Short-Term Forecast of Significant Thunderstorms. Forecasting Techniques, 525-534.
[2] Cortinas Jr., J. V. and Stensrud, D. J. (1995) The Importance of Understanding Mesoscale Model Parameterization Schemes for Weather Forecasting. Weather and Forecasting 10, 716-740.
[3] Pinto, I. R. C. A. and Pinto Jr., O. (2003) Cloud-to-ground lightning distribution in Brazil. Journal of Atmospheric and Solar-Terrestrial Physics 65, 733-737.
[4] Fausett, L. (1994) Fundamentals of neural networks: Architecture, Algorithms and Applications, pp. 461, Prentice Hall, New Jersey.
[5] Zepka, G. S., Pinto Jr., O., Farias, W. R. G., Carretero, M. A. and Carneiro, J. C. (2008) A forecast cloud-to-ground lightning system based on a neural network – Preliminary results. In: International Lightning Meteorology Conference (ILMC), Vaisala Inc., Tucson, Arizona.
[6] Haykin, S. (1994) Neural Networks: A Comprehensive Foundation, pp. 696, Prentice Hall, New Jersey.
[7] Beale, R. and Jackson, T. (1990) Neural Computing: An Introduction, pp. 264, IOP Publishing Ltda, Bristol, UK.
[8] Pinto Jr., O., Naccarato, K. P., Saba, M. M. F., Pinto, I. R. C. A., Garcia, S. A. M. and Cazeta-Filho, A. (2006) Recent upgrades to the Brazilian integrated lightning detection network. In: International Lightning Detection Conference (ILDC), Vaisala Inc., Tucson, Arizona.
[9] Pinto Jr., O., Pinto, I. R. C. A. and Naccarato, K. P. (2007) Maximum cloud-to-ground lightning flash densities observed by lightning location systems in the tropical region: A review. Atmospheric Research 84, 189-200.
[10] Black, T. L. (1994) The New NMC Mesoscale Eta Model: Description and Forecast Examples. Weather and Forecasting 9, 265-278.
[11] Wang, W., Barker, D., Bruyere, C., Dudhia, J. and Michalakes, J. (2004) WRF version 2 modeling system User's guide.
[12] Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W. and Powers, J. G. (2005) A description of the advanced research WRF Version 2.

Main author
Name: Gisele dos Santos Zepka
Address: Avenida dos Astronautas, 1758 – São José dos Campos, SP, Brazil – 12227-010
Phone: ++55 (12) 3945-6841
E-mail: gzepka@dge.inpe.br