



On the Possibility of Using Neural Networks for the Thunderstorm Forecasting

Elena Stankova¹ (✉) , Irina O. Tokareva², and Natalia V. Dyachenko³

¹ Saint-Petersburg State University, 7-9, Universitetskaya nab., St.Petersburg 199034, Russia
e.stankova@spbu.ru

² IBS AppLine, St. Skladochnaya, 3, Building 1, 127018 Moscow, Russia

³ Russian State Hydrometeorological University, 98, Malookhtinsky pr., St.Petersburg 195196, Russia

Abstract. The paper explores the possibility of forecasting such dangerous meteorological phenomena as a thunderstorm by applying five types of neural network to the output data of a hydrodynamic model that simulates dynamic and micro-physical processes in convective clouds. The ideas and the result delivered in [1] are developed and supplemented by the classification error calculations and by consideration of radial basic and probabilistic neural networks. The results show that forecast accuracy of all five networks reaches values of 90%. However, the radial basis function has the advantages of the highest accuracy along with the smallest classification error. Its simple structure and short training time make this type of neuralnetwork the best one in view of accuracy versus productivity relation.

Keywords: Machine learning · Neural networks · Perceptron complex · Radial basic neural network · Probabilistic neural network · Numerical model of convective cloud · Weather forecasting · Thunderstorm forecasting

1 Introduction

Throughout human history, meteorological processes prediction has always been a complex task mainly because the Earth's atmosphere system is very complex and dynamic.

Weather forecasts are calculated based on meteorological data collected by a network of weather stations, radiosondes, radars and satellites around the world. Data is sent to meteorological centres, where it is entered into forecast models for atmospheric conditions calculation. Such models are based on physical laws and work according to extremely complex algorithms.

Precipitation is determined by the physical processes occurring in the cloud, namely the physics of the interaction of water droplets, ice particles and water vapor. Convective clouds are very variable due to the large vertical speeds within the cloud and its environs. It is also difficult to conduct control experiments involving them. All this leads to the fact that the development of the cloud is usually analyzed using computer simulation, which allows us to do this without resorting to expensive field experiments.

As a result of computer simulation of the cloud we get a data set that can be further used for forecasting various dangerous convective phenomena such as thunderstorm, hail and heavy rain.

In recent decades mathematicians and programmers are working hard to improve existing numerical weather forecasting models. Nowadays machine learning methods, especially neural networks are considered to be one of the most promising tool of such improvement [2, 3]. Authors in [2] state that advantages of neural networks are the intrinsic absence of model bias and possible savings of computational resources due to ability of neural network very efficiently calculate forecasts with new data after corresponding training.

The use of machine learning methods allows us to automate the forecasting process, which greatly facilitates data analysis. These methods conduct a series of computational experiments with the aim of analyzing, interpreting and comparing the simulation results with the given behavior of the object under study and, if necessary, subsequently refining their input parameters.

The idea to use neural networks to process output from numerical weather prediction models had been explored in far 1998 year in [4] in order to give more accurate and localized rainfall predictions already.

Prediction of rainfall amounts is very popular application for neural networks usage [5–7]. Thus in [5] researchers in Thailand tried to predict possible flooding dangers by estimating rainfall amounts using feed-forward neural networks. Authors in [6] tried to accurately predict heavy precipitation events ($>25 \text{ mm h}^{-1}$) over Germany using also neural networks.

Tao et al. [8, 9] use deep neural networks for forecasting precipitation amount among meteorological factors and obtained promising results.

Authors in [10] used neural networks for predicting probabilities and quantitative values of precipitation with the help of the Eta atmospheric model and upper air soundings.

Researcher in [11] has investigated how effectively neural networks can perform classification prediction of freezing and gusty events as well as minimum temperature and maximum gust values. Paper [11] contains also the detailed review of neural networks application for solution of meteorological problems.

Neural networks have also been used to predict various weather phenomena (wind speed, barometric pressure, fog [12]) including extreme events, such as tornadoes [13] and typhoons [14, 15].

In [14] a multilayer perceptron is used to predict changes in tropical cyclone intensity in the northwestern Pacific Ocean. The paper [15] uses a generative adversarial network to predict typhoon trajectories. The neural network generates an image showing the future location of the typhoon center and cloud structure using satellite images as an input.

In this paper we continue the studies described in [1, 16–18] and analyze the possibility of the use of neural networks for dangerous convective phenomena forecasting by processing the output data of numerical model of convective cloud [19–22]. The idea is to retrieve the possibility of thunderstorm forecasting from the data of the model, able to simulate only dynamical and microphysical characteristics of convective clouds, but

not electrical characteristics. The ideas and the result delivered in [1] are developed and supplemented by the classification error calculations and by consideration of radial basic and probabilistic neural networks.

2 Initial Data

Research using machine learning methods is based on data, therefore, in order to obtain the best results, it is necessary to use reliable sources of information to obtain data and form their correct structure.

In this work, the data was obtained using the following algorithm:

1. We receive data on the date and place of meteorological phenomena occurrence;
2. We select the data corresponding to the presence of a thunderstorm or the absence of any meteorological phenomena;
3. We obtain data from atmosphere radio sounding for the certain date and place;
4. We convert the radio sounding data to the model input data format;
5. Using the hydrodynamic model, we obtain the integral and spectral characteristics of the cloud;
6. We determine the height and time corresponding to the maximum development and maximum water content of the cloud. The cloud parameters corresponding to these height and time will be used for the thunderstorm forecasting.

Formed data set contains 416 records, where 220 samples correspond to the presence of a thunderstorm and 196 samples to its absence. This data was divided into training and test data sets. The training one contains 333 samples and the test one contains 83. Due to the small amount of data we decided to use test data set for validation.

We also created labels for each sample in the data set. Since there are only two cases, the presence and absence of phenomenon, we could have created one label per sample. But we decided to use two labels per sample, one for each case, mainly because we will need to divide the output variables of the neural network at some point.

3 Data Preprocessing

Solution of machine learning problems require to find an unknown relationship between a known set of objects and a set of answers. In our case the fact of dangerous phenomenon occurrence can be considered as an answer, and the results of numerical modeling, can be considered as an object. Radiosonde sounding data are used as the model input.

Neural networks, like all machine learning algorithms, depend significantly on the quality of the source data. Therefore, before proceeding to the construction of a neural network, we will need prepare the data.

First, we normalize the data using the Standard Scaler method from the Python scikit-learn library, which converts the data to the standard normal distribution.

Then we select the most significant features. To do this, we use the Recursive Feature Elimination method from the scikit-learn library with Random Forest algorithm as an estimator. The method is as follows. The estimator is firstly trained on the initial set

of features, then the least important feature is pruned and the procedure is recursively repeated with smaller and smaller set. Figure 1 shows the resulting graph of the prediction accuracy versus the number of features used. As can be seen from the figure, maximum accuracy is achieved when using 8 features. Their names and their importance are shown in the Fig. 2. Thus, we will use the following features: vapor, aerosol, relative humidity, density, temperature excess (inside cloud), pressure, velocity, temperature (in the environment).

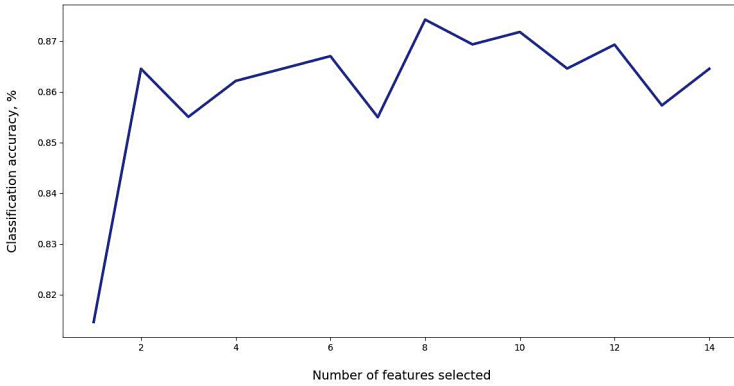


Fig. 1. Graph of prediction accuracy versus the number of features involved

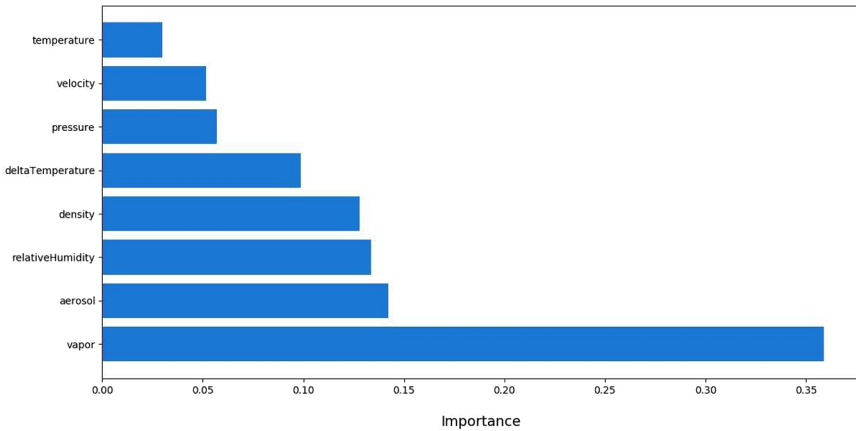


Fig. 2. Selected features and their importance.

4 Classical Multi-layer Perceptron and Perceptron Complexes

The main ideas and results achieved while using classical multilayer perceptron structure (Fig. 3) and perceptron complexes were described in [1]. Some additional explanations

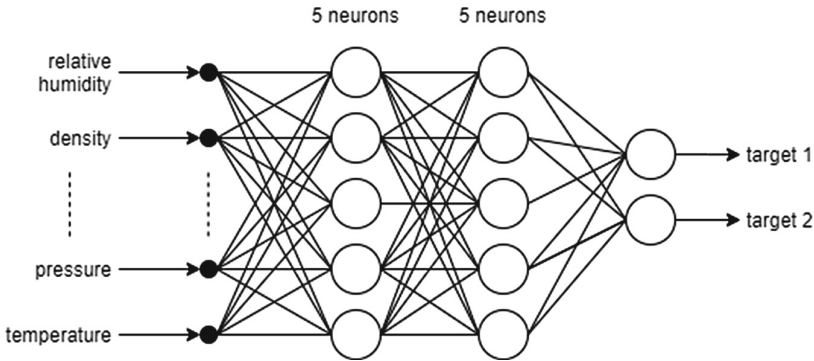


Fig. 3. Classical multi-layer perceptron

for using perceptron complexes and the values of the classification errors can be added to what is said there.

The article [23] mentions that the ratio of the volume of the training data set and the number of trainable network parameters is one of the factors that affect the modeling ability of the perceptron. If this ratio is close to 1, the perceptron will simply remember the training set, and if it is too large, the network will average the data without taking the details into account. In this regard, in most cases, it is recommended that this ratio falls in the range from 2 to 5. In our case, this ratio is.

$$\frac{333}{123} = 2.7$$

which falls into this range. However, our training data set is small and the use of algorithms based on neural networks may be ineffective with small amounts of experimental data [24]. So we decided to use one of the methods that can help to increase the efficiency of our neural network.

The method is described in [23]. It consists in dividing the set of input and output variables into several perceptrons with a simpler structure and then combining them into a single perceptron complex. Figure 4 shows the general structure of such a complex.

The perceptron complex training algorithm is as follows [24]:

1. For each first level perceptron:
 - a. Given the input and output variables of the current perceptron, we construct the training and test data sets for it based on the initial data;
 - b. Perceptron training is executed;
 - c. For all samples of training and test data sets, the values of the perceptron outputs are calculated and stored.

2. For the resulting perceptron:
 - a. Given the input and output variables of the perceptron, we construct the training and test data sets for it based on the initial data and the calculated output values of the first level perceptrons;
 - b. Perceptron training is executed.

Two variants of the perceptron complexes were described in our previous work [1]. Here we can only add that classification errors are equal to 0.081 and 0.078 for the first and the second perceptron complexes consequently.

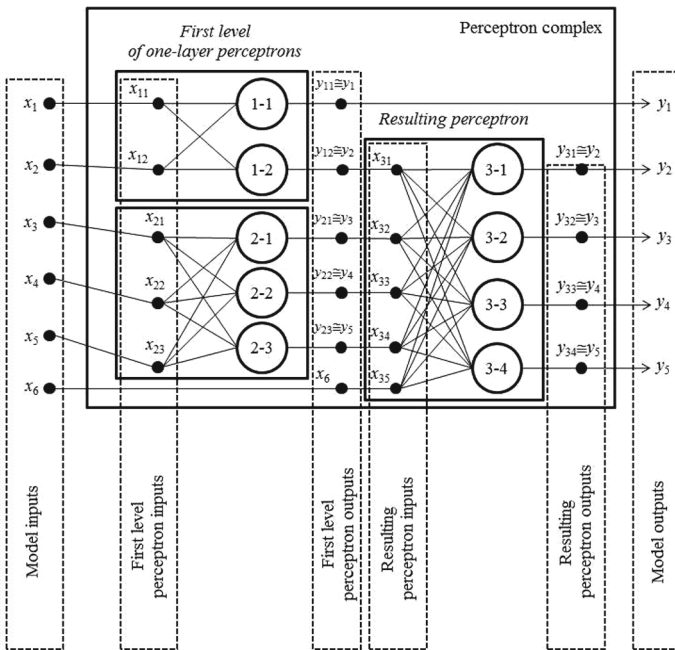


Fig. 4. General structure of the perceptron complex.

5 Radial Basis Function Network

Two types of networks that belong to radial basic networks are considered as they show good results in problems of binary classification. Also, their advantage is a simple structure where there is only one hidden layer.

In the process of training this network, three sets of parameters are determined. We considered several ways to set their initial values and established how many neurons there should be in the hidden layer to get the highest prediction accuracy.

The resulting neural network is shown on Fig. 5. Its accuracy is 91.6%, classification error is 0.069.

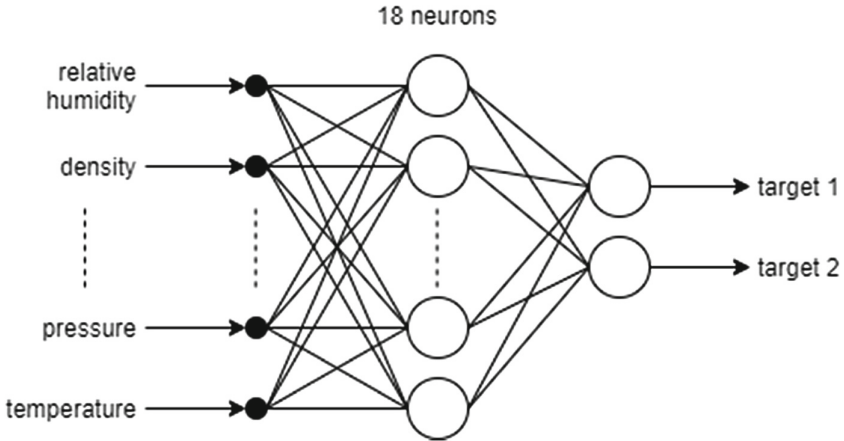


Fig. 5. Radial basis function network

6 Probabilistic Neural Network

A feature of the probabilistic neural network is that the number of neurons in the hidden layer is equal to the number of examples in the training set, that is, the network simply stores the entire training set.

The structure of the network is shown on the Fig. 6. The accuracy is 90.4%, classification error is 0.096.

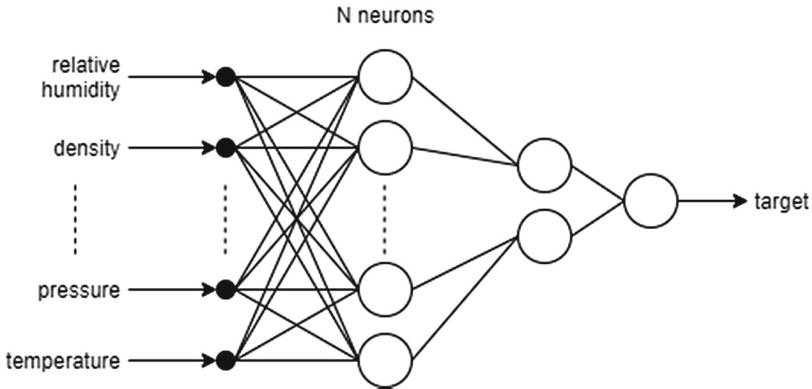


Fig. 6. Probabilistic neural network

7 Classification Accuracies and Classification Errors of the Neural Networks of Different Types

Table 1 presents the values of classification accuracy and classification error of the five types of neural networks considered by the authors. As can be seen from the table the best accuracy is achieved using the second perceptron complex and the radial basis function network.

Table 1. Maximum forecast accuracy and classification error of the five types of neural networks

Neural network	Classification accuracy, %	Classification error
Multi-layer perceptron	89,1	0,082
First perceptron complex	90,4	0,081
Second perceptron complex	91,6	0,078
Radial basis function network	91,6	0,069
Probabilistic neural network	90,4	0,096

8 Conclusions

The work analyzed the possibility of using neural networks to build forecasts of dangerous convective phenomena by the example of a thunderstorm.

The initial data set was obtained using numerical modeling of a convective cloud.

Using machine learning methods at the stage of data analysis and processing of features, the most significant features were identified.

Five networks were considered. The best accuracy was achieved using the second perceptron complex and the radial basis function network. However, the radial basis function network gave the smallest classification error. Also, its advantages over the perceptron complex are simple structure and short training time.

In future we will further explore the possibility of using neural networks for forecasting thunderstorm and other dangerous convective phenomena and specifically our research should be focused on obtaining sufficient number of radiosonde soundings with the corresponding model simulations for training data sets formation.

References

1. Stankova, E.N., Tokareva, I.O., Dyachenko, N.V.: On the effectiveness of using various machine learning methods for forecasting dangerous convective phenomena. In: Gervasi, O., et al. (eds.) ICCSA 2020. LNCS, vol. 12254, pp. 82–93. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58817-5_7
2. Schultz, M.G., et al.: Can deep learning beat numerical weather prediction? *Phil. Trans. R. Soc. A* **379**, 20200097 (2021). <https://doi.org/10.1098/rsta.2020.0097>

3. Scher, S., Messori, G.: Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as a study ground. *Geosci. Model Dev.* **12**, 2797–2809 (2019). <https://doi.org/10.5194/gmd-12-2797-2019>
4. Kugliowski, R.J., Barros, A.P.: Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Weather Forecast.* **13**(4), 1194–1204 (1998)
5. Hung, N.Q., Babel, M.S., Weesakul, S., Tripathi, N.K.: An artificial neural network model for forecasting in Bangkok, Thailand. *Hydrol. Earth Syst. Sci.* **13**(8), 1413–1425 (2009)
6. Unwetterklimatologie: Starkregen. <https://www.dwd.de/DE/leistungen/unwetterklima/starkregen/starkregen.html>. Accessed 30 April 2020
7. Luk, K.C., Ball, J.E., Sharma, A.: An application of artificial neural networks for rainfall forecasting. *Math. Comput. Model.* **33**(6–7), 683–693 (2001). [https://doi.org/10.1016/S0895-7177\(00\)00272-7](https://doi.org/10.1016/S0895-7177(00)00272-7)
8. Tao, Y., Gao, X., Ihler, A., Sorooshian, S.: Deep neural networks for precipitation estimation from remotely sensed information. In: *Proceedings IEEE Congress on Evolutionary Computation*, Vancouver, BC, Canada, pp. 1349–1355. IEEE (2016)
9. Tao, Y., Gao, X., Ihler, A., Sorooshian, S., Hsu, K.: Precipitation identification with bispectral satellite information using deep learning approaches. *J. Hydrometeor.* **18**, 1271–1283 (2017)
10. Hall, T., Brooks, H.E., Doswell, C.A., III.: Precipitation forecasting using a neural network. *Weather Forecast.* **14**(3), 338–345 (1999)
11. Culclasure, Andrew, *Using Neural Networks to Provide Local Weather Forecasts* (2013). *Electronic Theses and Dissertations*. 32. <https://digitalcommons.georgiasouthern.edu/etd/32>
12. Santhanam, T., Subhajini, A.C.: An efficient weather forecasting system using radial basis function neural network. *J. Comput. Sci.* **7**(7), 962–966 (2011)
13. Marzban, C., Stumpf, G.J.: A neural network for tornado prediction based on Doppler radar-derived attributes. *J. Appl. Meteorol.* **35**(5), 617–626 (1996)
14. Baik, J.-J., Paek, J.-S.: A Neural Network Model for predicting typhoon intensity. *J. Meteor. Soc. Japan.* (2000). <https://doi.org/10.2151/jmsj1965.78.6857>
15. Ruettgers, M., Lee, S., Jeon, S., You, D.: Prediction of a typhoon track using a generative adversarial network and satellite images. *Sci. Rep.* **9**, 6057 (2019). <https://doi.org/10.1038/s41598-019-42339-y>
16. Stankova, E.N., Grechko, I.A., Kachalkina, Y.N., Khvatkov, E.V.: Hybrid approach combining model-based method with the technology of machine learning for forecasting of dangerous weather phenomena. In: Gervasi, O., et al. (eds.) *ICCSA 2017. LNCS*, vol. 10408, pp. 495–504. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62404-4_37
17. Stankova, E.N., Balakshiy, A.V., Petrov, D.A., Korkhov, V.V., Shorov, A.V.: OLAP technology and machine learning as the tools for validation of the numerical models of convective clouds. *Int. J. Bus. Intell. Data Min.* **14**(1/2), 254 (2019). <https://doi.org/10.1504/IJBIDM.2019.096793>
18. Stankova, E.N., Khvatkov, E.V.: Using boosted k-nearest neighbour algorithm for numerical forecasting of dangerous convective phenomena. In: Misra, S., et al. (eds.) *ICCSA 2019. LNCS*, vol. 11622, pp. 802–811. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24305-0_61
19. Raba, N.O., Stankova, E.N.: Research of influence of compensating descending flow on cloud's life cycle by means of 1.5-dimensional model with 2 cylinders. In: *Proceedings of MGO*, vol. 559, pp. 192–209 (2009). (in Russian)
20. Raba, N., Stankova, E.: On the possibilities of multi-core processor use for real-time forecast of dangerous convective phenomena. In: Taniar, D., Gervasi, O., Murgante, B., Pardede, E., Apduhan, B.O. (eds.) *ICCSA 2010. LNCS*, vol. 6017, pp. 130–138. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12165-4_11

21. Raba, N.O., Stankova, E.N.: On the problem of numerical modeling of dangerous convective phenomena: possibilities of real-time forecast with the help of multi-core processors. In: Murgante, B., Gervasi, O., Iglesias, A., Tanar, D., Apduhan, B.O. (eds.) ICCSA 2011. LNCS, vol. 6786, pp. 633–642. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-21934-4_51
22. Raba, N.O., Stankova, E.N.: On the effectiveness of using the GPU for numerical solution of stochastic collection equation. In: Murgante, B., et al. (eds.) ICCSA 2013. LNCS, vol. 7975, pp. 248–258. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39640-3_18
23. Dudarov, S.P., Diev, A.N.: Neural network modeling based on perceptron complexes with small training data sets. *Math. Meth. Eng. Technol.* **26**, 114–116 (2013). (in Russian)
24. Dudarov, S.P., Diev, A.N., Fedosova, N.A., Koltsova, E.M.: Simulation of properties of composite materials reinforced by carbon nanotubes using perceptron complexes. *Comput. Res. Model.* **7**(2), 253–262 (2015). <https://doi.org/10.20537/2076-7633-2015-7-2-253-262>