Application of GIS forecasting in forestry based on a neural network

N Yagotinceva¹, E Istomin¹, O Kolbina¹, T Safonova¹ and A Kochnev²

¹Department of applied information science, Russian state hydrometeorological university, Voronezhskaya Street 79, St. Petersburg 192007, Russian Federation
²ITMO University, Kronverksky Avenue 49, letter A, St. Petersburg, 197101, Russian Federation

*Corresponding e-mail: yagotintceva@yandex.ru

Abstract. This article deals with the application of modern information technology for planning activities and operations in forestry. A GIS of meteorological forecasting using a neural network is presented as a planning tool. The structure of GIS is proposed. The stochastic method for training a neural network that suggests the most likely outcome of an event based on a previous sample is described. The results described in the article show the possibility of application of the selected method of neural network training and data use in geoinformation systems of decision-making in forestry.

1. Introduction

The forest and all its components are in close relationship with the environmental conditions. As the most important physical environmental parameters, meteorological factors have a great influence on all aspects of forest life: determining the possibility of forest production, its diversity, its productivity, the course of all life processes in the forest, as well as the conditions and methods of the forest economic activity. Therefore, the meteorological measurements are an important part of forestry, silvicultural, botanical and other studies. Meteorological data are widely used in the practical activities of forestry specialists in conducting forestry activities, design and survey work and in drafting a variety of projects.

Already today, subdivisions of Roshydromet provide forestry organisations with meteorological information. Such provision is aimed not only at minimising the consequences of forest fires, but also makes it possible to protect forest crops and planting materials from frost and other similar calamities.

It has been shown that the planning and implementation of forestry activities largely depend on weather conditions. Consequently, meteorological data and its forecasting have an impact on forest performance. Improving the efficiency of forecasting meteorological conditions and forming a GIS to support decision-making based on weather forecasts is an urgent task.

Existing methods of weather forecasting can be divided into two types: synoptic and hydrodynamic. Both of these methods have been used successfully in forecasting weather conditions with and without the use of information technology [1].

The hydrodynamic method of forecasting involves displaying information without visual representation, while the synoptic method is based on the construction of geographical maps with overlapping thematic layers.
In the synoptic method, the GIS must directly display a graphical representation of the different natural phenomena on a map and provide a user with tools to assist decision-making. This method requires a trained staff and enormous financial resources to purchase the necessary sensors to cover the entire area where weather conditions are monitored and short-term forecasts are made. Also, this method provides a more convenient tool for displaying atmospheric precipitation on the ground [2].

In the hydrodynamic method huge computational work has to take place for more accurate forecasting than in the synoptic method. Inherent in this method is the availability of large amounts of data due to the need to determine probabilistic changes in weather conditions. Neural networks can be used to process large amounts of statistical data in the operation of such a system [3].

Historical data are not used to their full potential in large weather forecasting systems. Mathematical models of weather forecasting, laid in the work of modern GIS use statistical data [4], which in turn affects the accuracy of the forecast. Thus, using the hydrodynamic method to build a mathematical model based on historical data and synoptic method to display the results in the visual representation gives a synthesis that increases the percentage of successful weather forecasts [5].

The automation of neural network training will serve to further regulate and improve GIS performance.

2. Methods and Materials

2.1. Structure of a GIS using neural networks in meteorological forecasting

Any modern geoinformation system consists of at least three layers: data, processing and demonstration of this processing. A geo-information system for decision support based on meteorological forecasts structurally consists of 6 blocks connected with each other (figure 1).

![Figure 1. Decision support GIS structure diagram using neural networks.](image)

All forecasting involves loss of data and time to make the forecast, but the use of a neural network will optimize and improve the efficiency of forecasting meteorological conditions. In order to implement a GIS decision support system based on weather forecasts, the optimal forecasting method and training method for the neural network must be selected.
2.2. Neural network training methodology in hydrodynamic prediction methods

Prediction methods in neural networks are roughly divided into three groups depending on the type of problem to be solved:

- The deterministic method.
- The method of expert judgement.
- Stochastic method [7].

A more detailed comparison of these methods is presented in table 1.

| The method          | Expert assessments                                      | Stochastic                        | Deterministic                      |
|---------------------|--------------------------------------------------------|-----------------------------------|------------------------------------|
| Advantages          | Accurate results (expressed in more than just numbers) | Probability results               | Exact data (expressed numerically) |
|                     | Consideration from different perspectives              | Scattered of the correct results  | Easiest way to display the result   |
|                     |                                                        | Processing speed                  | Processing speed                   |
| Disadvantages       | Cost                                                   | Decreasing probability over time  | Exact data (assumes exact hits on a set of values) |
|                     | Processing speed                                       |                                    |                                    |

Based on this table, the stochastic method will be used to solve the prediction problems in the GIS under consideration. It will suggest the most probable outcome of an event based on the previous sample and the chance that it will occur. Another advantage of this method is that it can be used not only to train a new neural network, but also for an already trained one, while excluding local minima.

To train the neural network considered by GIS, it is necessary to make a pseudo-random change in the value of the weights, fixing only those combinations that lead to improvement. Training the network using the stochastic method boils down to a set of the following sequential steps:

1. an array of weather data values \( x \);
2. generate a randomly generated weight;
3. perform correction of the generated weight by a random number;
4. calculate the resulting outputs;
5. compare the resulting outputs with the available statistical sample and calculate the value of the difference between them;
6. obtain the target function;
7. if the result of step 5 decreases the target function, save it, otherwise return to the original weight value;
8. repeat steps 1 to 6 until the network is sufficiently trained [8].

Some array of weather data values \( x \) with size \( m \) is fed into the system. After the primary processing according to some given condition inside the primary blocks (circles), the new values \( w \), which are already weight values, are fed into the secondary processing. Then a new value \( k \) is obtained, as well as \( w \) going into processing. Thereafter a completely new value \( y \), which is the resultant value, is already obtained [9]. In this scheme, if we consider it as a stochastic method, \( y \) is the most likely outcome and \( \min(w, k) \) and \( \max(w, k) \) are permissible limits of the value deviation.

### 3. Results and Discussion

In order to perform forecasting of meteorological data, an application program with a trained neural network was implemented, which fully meets the requirements of the "Data Processor" block (ure 1). After the processing of the input data by this program, the output is:

- Probable weather;
- Probability spread.

To work with the database, the library "Microsoft.Office.Interop.Excel" was connected, which allows to work with the extensions ".xlsx" and ".xls". An empty knowledge base was also created to
be automatically populated with the resultant indicators from the neural network. In this knowledge base, the first line is responsible for storing some values, such as:

- date of last update;
- forecast deviation (Eps);
- success.

The date of the last update is needed for later use of the program so that the data in the program is not repetitive. Forecast deviation is a variable responsible for the success of the forecast. If the probable temperature is -3 and the actual temperature is -2.7, the deviation is 0.3 and the forecast is considered successful. The success rate is a statistical variable, calculated using the formula:

\[
\text{Success rate} = \frac{\text{Number of successful predictions}}{\text{Total number of forecasts}}
\]

Figure 2 shows the knowledge base after processing by the application program. A sample of 250 values was taken for testing, from 1.11.2018 to 1.12.2018.

![Figure 2. The processed knowledge base.](image)

The initial deviation was taken as number 2. If the prediction was successful, this number was divided by 2. If it was not successful, it was increased by 2x or 0.1, depending on the condition. The data processor is the main link for running the software application.

Figure 3 shows two curves that use columns 'B' and 'C' as curve points. The orange line shows the temperature predicted by this software application, the blue line shows the actual temperature from the weather diary. The graph shows that the predicted temperature does not deviate much from the real temperature line.
Figure 3. Ratio of probable and actual temperature.

Figure 4 shows a graph of the variation of the temperature deviation from the probable forecast. It can be seen that the deviation is quite sharp in this model. It can be adjusted more smoothly, but only if the sample is significantly larger than the 250 values presented in this paper. In this case, however, it is noticeable that the deviation at the end of processing varies from 0.1 to 1.1 degrees.

Figure 4. Deviation of the probable temperature from the actual temperature.

Figure 5 shows a prediction success rate plot, which tends to 55%, with a final deviation of 0.3. This is a reasonably good result for the weather forecasting model used, given the small amount of input data.

Figure 5. Rate graphon prediction success.
Figure 5 shows the real temperature and the simulation values. The tops of the graph plotted from the simulated values are higher than the real temperature values, indicating that if the real temperature values change drastically and the input parameters to the neural network are insufficient, the latter gives significant deviations.

It can be seen from the graphs presented that the success of the forecasts will increase and the deviations will decrease with increasing input data. This is not a problem, because weather diaries consist of terabytes of information.

4. Conclusion
From the results presented, it can be concluded that it is possible to apply neural networks to predict meteorological conditions and use the data in geographic information systems for decision-making in forestry.

Such a GIS is a hardware and software man-machine composition that provides collection, processing, display and distribution of spatial and geographical data [10]. The model provides integration of data and knowledge about the territory for resolving the scientific and applied tasks related to the inventory, analysis, modelling, forecasting, environmental management and territorial organization of society.

References
[1] Vagizov M R, Istomin E P, Dobrovolskyi A A, Zhernova A P and Yagotintceva N V 2020 Technological aspects of the development of the automated method of air-photo interpretation of forest stands. *IOP Conference Series: Earth and Environmental Science* vol 574 (1)
[2] Bykov F L 2020 Statistical correction of COSMO weather forecasts using neural networks *Meteorology and hydrology* vol 3 pp 5-20
[3] Yaremenko I A 2020 Method for the distributed processing of meteorological information based on artificial neural network technology *Proceedings of the Mozhaisky Military Space Academy* vol S674pp 263-270
[4] Istomin E, Martyn I, Petrov Y, Stepanov S, Sidorenko A. 2021 Development of a mathematical model of wind waves in the area of the proposed construction of a hydraulic structure. *IOP Conference Series: Earth and Environmental Science* vol 723 (5)
[5] Istomin E, Petrov Y, Stepanov S, Kolbina O and Sidorenko A 2019 About technology of risk management in forestry *IOP Conference Series: Earth and Environmental Science* vol 316(1)
[6] Mityushkin Yu I, Mokin B I and Rotstein A P 2002 *Soft Computing: identification of regularities by fuzzy knowledge bases*. Vinnytsia, p 145
[7] Vagizov M R, Ustyugov V A and Kvochkin D O 2017 Determination of the forest inventory indicators according to the photographs of the unmanned aerial vehicles *Ecology, Environment and Conservation* vol 23(1) pp 582-586
[8] Panfilov P N 2001 Introduction to neural networks *Modern trading* vol 2 pp 12-17
[9] Galimyanov F A 2020 Comparative analysis of back propagation algorithms for neural network training *Scientific and Technical Bulletin of the Volga Region* vol 2pp 69-72
[10] Istomin E, Khramov I, Sokolov A, Yagotintseva N 2021 Database of estimated parameters in the near-shore zone of the northern seas *Database registration certificate* Application No2020622792.