Research and forecasting of the influence of the Earth's electromagnetic field on the accident rate on public roads

A Yu Abalyaev¹, L V Grunskaya*¹ and I A Leshchev¹
¹ Vladimir State University named after A.G. Stoletov and N.G. Stoletov, Gorky street 87, Vladimir, 600000, Russia

*grunsk@vlsu.ru

Abstract. This work is devoted to predicting the average accident rate on public roads. At the physical experimental range of the Department of General and Applied Physics of Vladimir State University (VSU) monitoring of the Earth's electromagnetic fields, meteorological data and radiation background is carried out. Research is being conducted to assess the relationship of road traffic accidents (RTA) in the Vladimir region with geo- and heliophysical characteristics. The use of neural networks made it possible to solve certain problems of predicting road accidents and to take into account the influence of solar activity on the value of the predicted parameter.

1. Introduction

The relevance of this topic is due to the persistent high level of road traffic injuries. Every year in Russia in road traffic accidents (RTA) about 30 thousand people are killed and more than 180 thousand people are injured. Predicting the level of accidents on public roads is an important task, since it allows us to estimate the number of road accidents (RTA), their impact on the economy of the region and the main thing is to estimate the required level of preparedness of emergency services to eliminate the consequences of RTA.

Since 1999, at the physical testing ground of the Department of General and Applied Physics of Vladimir State University, continuous monitoring of electromagnetic fields of the infra-low frequency range, meteorological factors and radiation background has been carried out. Since 2009, together with the Office of Rospotrebnadzor in the Vladimir region, using statistical data on the mortality of the population, as well as on the circulation of children with respiratory diseases for the period from July 2004 to December 2006 in the city of Vladimir, work has been carried out to study the impact of space and geophysical factors on the indicators of public health. In the works of many prominent scientists, it is proved that natural electromagnetic fields of the infra-low frequency range have an impact on human health indicators [1-4].

Separately, we can say about the geomagnetic field, which, by the nature of its variations, is also infra-low-frequency. Geomagnetic disturbances (geomagnetic storms) are a natural risk factor for human health, disrupting the time sequence of information signals used by the body. The geomagnetic field is one of the many environmental factors that affects the regulatory mechanisms of the body at all levels: molecular, intracellular, intercellular, etc. [5-7].

At the first stage of these studies the influence of the electric and geomagnetic fields of the Earth, radiation background on various indicators of human health was studied: broncho-pulmonary diseases, mortality of the population. Besides one of the objectives of the research was to determine the influence
of long-term, periodically and randomly acting causal factors on the long-term dynamics of victims of collisions in road accidents, to determine the seasonality [9, 10]. The research uses databases of the Earth's electromagnetic field and meteorological data based on the monitoring results at the VlSU physical testing ground. The data on the road traffic accidents in Vladimir region were provided by the Medical and Sanitary Unit of the Ministry of Internal Affairs of Russia for the Vladimir Region and the Department of the State Traffic Safety Inspectorate of the Ministry of Internal Affairs of Russia for the Vladimir Region [11, 12].

At the next stage of the research the task was to assess the possibility of predicting the accident rate on the roads of the Vladimir region using neural networks. Predicting the level of accidents on public roads is an important task to assess the number of road traffic accidents (RTA), their impact on the required level of preparedness of emergency services to eliminate the consequences of RTA. To predict the accident rate the data of the number of the accidents from open sources were used [11]. The analysis of the change in the accident rate during the year (according to 2001-2010 data) shows that the accident rate is highly variable in nature [8]. Of course, it is influenced by a number of factors: calendar, seasonal, weather, geographic, social, economic, human factor and others. These features make it difficult to predict directly the number of accidents using the statistical methods.

The use of neural networks (NN) allows not only to solve effectively a wide range of classification and forecasting problems, but also to take into account the influence, including implicit, various additional factors, both known and unknown ones at the time of making the forecast, on the value of the predicted parameter. The first stage of the research was devoted to predicting the average accident rate based on data for 2001-2010 [8]. A recurrent neural network (RNN) consisting of two layers of long short-term memory (LSTM) [13] and an output layer of 7 elements was created, which allows predicting the change in the accident rate for the next seven days outside the current period. The RNN training process consisted of a number of epochs, during each of the both of them the correction of the synapse weights based on the data from the training sample and the control of the training process based on the data from the control sample are carried out. To assess the suitability of using the RNN in the process of predicting the level of accidents on public roads, a study was carried out of the influence on the accuracy of the resulting forecast not only of the number of learning epochs, but also the number of RNN presentations of each sequence from the set of the training sample (repetitions) during each epoch. The values of the objective function of the error were obtained on the training and control samples. It was noted that an increase in the number of training epochs led to a significant decrease in the training error, which, in real conditions, could contribute to an increase in the forecast accuracy.

An example of predicting the average reduced accident rate for an arbitrarily selected sequence included in the control sample for the case of training the RNS for 10 epochs with 300 repetitions in each of them shows that, in general, using RNS it is possible to predict the trend of change in the average accident rate, however, in some details, the forecast is not accurate enough.

Based on the obtained results, it was concluded about the possibility of using RNS to predict changes in the average accident rate with sufficient accuracy. However, in order to improve the accuracy of the forecast it is necessary both to implement the forecast of the maximum and minimum accident rate, and to take into account the additional natural factors which affect the predicted value.

This article presents the results of assessing the correlation of RTA parameters with the characteristics of the natural environment and the results of predicting the RTA level taking into account one of the important geophysical factors, namely the solar activity, expressed in Wolf numbers. The forecast takes into account the total Wolf number. The average accident rate is also predicted. Prediction was performed using a recurrent neural network built and trained using the Tensor Flow 2 library [13].

2. Using recurrent neural networks to predict the accident rate on public roads

2.1. Analysis of the relationship of road traffic accidents with the level of intensity of the Earth's electric field, geomagnetic field, solar activity
The assessment of the correlations between the vertical component of the electric field strength, the Earth's magnetic field, Wolf numbers and the parameters of road accidents was done: the number of road accidents, the number of deaths and injuries as a result of road accidents.

Table 1 shows the sample results for the obtained significant correlation coefficients for the period 2001-2006.

**Table 1.** Significant correlation coefficients with the probability of error (p) for the number of road accidents, deaths, injured in road accidents in the Vladimir region and Wolf numbers for 2001-2006.

| Analyzed processes                                      | Date                         | Correlation coefficient | Probability of error, p |
|---------------------------------------------------------|------------------------------|-------------------------|-------------------------|
| RTA and the Earth electrical field                      | 28.06.2001-24.07.2001        | -0.50                   | ≤0.02                   |
| RTA and the Earth magnetic field                        | 28.06.2001-08.08.2001        | 0.40                    | ≤0.01                   |
| RTA and Wolf Number                                     | 28.06.2001-29.08.2001        | 0.50                    | ≤0.001                  |
| Perished as a result of RTA and Wolf Number             | 19.07.2001-19.09.2001        | -0.50                   | ≤0.001                  |
| Perished as a result of RTA and a geomagnetical field   | 10.07.2002-21.08.2002        | -0.41                   | ≤0.01                   |
| Perished as a result of RTA and Wolf Number             | 11.07.2003-11.08.2003        | -0.46                   | ≤0.01                   |
| Perished as a result of RTA and the Earth electrical field | 28.01.2005-28.02.2005        | -0.574                  | ≤0.001                  |
| Perished as a result of RTA and the Earth magnetic field | 05.04.2005-16.05.2005        | -0.46                   | ≤0.01                   |
| RTA and Wolf Number                                     | 28.01.2005-10.03.2005        | 0.454                   | ≤0.01                   |
| RTA and the Earth electrical field                      | 28.01.2005-28.02.2005        | -0.574                  | ≤0.001                  |
| RTA and Wolf Number                                     | 25.01.2005-25.02.2005        | 0.50                    | ≤0.01                   |
| Perished as a result of RTA and the Earth electrical field | 22.03.2005-02.05.2005        | 0.50                    | ≤0.001                  |
| Perished as a result of RTA and Wolf Number             | 20.05.2005-30.06.2005        | -0.46                   | ≤0.01                   |
| Perished as a result of RTA and the Earth electrical field | 01.04.2005-02.05.2005        | 0.52                    | ≤0.01                   |
| Perished as a result of RTA and Wolf Number             | 01.04.2005-02.05.2005        | 0.50                    | ≤0.01                   |
| Injured men as a result of RTA and Wolf Number          | 05.04.2005-16.05.2005        | -0.50                   | ≤0.02                   |
| Injured people as a result of RTA and the Earth magnetic field | 28.01.2005-28.02.2005        | -0.50                   | ≤0.01                   |
In some parts of the 2001-2006 time series a fairly high correlation was revealed between the vertical component of the electric field strength, the Earth's magnetic field, Wolf numbers and the number of road accidents, the number of victims in the Vladimir region.

For example, Figures 1-2 show the joint time series for the number of accidents and Wolf numbers for different time intervals.

2.2 The prediction of the accident rate level on roads using RNN taking into account the solar activity influence
Recurrent neural networks (RNNs) are a powerful tool for time series analysis and forecasting. Based on the analysis of time series information, they allow forecasting not only one value, but also a sequence of them. The accuracy of such a forecast will depend on the nature of the time series, the duration of the analyzed time interval, the features of the RNN implementation and the forecast period. The use of RNN for predicting the average accident rate based on the results of [8] has shown its promise. However, taking into account only one factor, namely, the average accident rate, leads to insufficient forecast accuracy, since it does not take into account the influence of additional factors.

In the process of learning and working, the RNN makes the best use of information about the events sequential in time. The principle of the RNN operation can be represented in the form of a diagram shown in [14] in Figure 3.

![Figure 3. Diagram of recurrent neural network layer.](image)

This figure shows a repeating unit in a standard RNN, consisting of one layer. In accordance with the given above scheme, the output $h_t$ of each neuron on the recurrent layer is determined by the sum of the products of the input weights and the output value of the previous $h_{t-1}$ time step and the current value of the time series $x_t$, for example, the hyperbolic tangent ($\tanh$), which the weighting function is applied:

$$h_t = \tanh(w_1 h_{t-1} + w_2 x_t + b),$$

where: $w_1, w_2$ are weight coefficients; $b$ is a shift.

The analysis of expression (1) shows that the previous events, except the preceding this one, have an insignificant effect on the current value of the neuron output. And what is more, the influence of the events remote from the current moment turns out to be indirect, that is, it is taken into account only in the output value of the neuron that is directly responsible for the previous event. Such a scheme of RNN operation, although it turns out to be the simplest, is not fully suitable for predicting the values of time series, whose subsequent events are strongly influenced by a number of previous events, sometimes significantly far from the current value. The Long short-term memory (LSTM) RNN, first described by Hochreiter and Schmidhuber [14], lacks this drawback. The principle of its operation is illustrated by the diagram shown in Figure 4.

![Figure 4. Diagram of LSTM-layer of recurrent neural network.](image)
In this diagram, the repeating model in the LSTM network consists of four interacting layers. In accordance with the given above scheme, not only the output of the neuron that processes the previous event is transmitted among the neurons of the RNN layer, but also the complete information about the state of the neurons that process the previous events in the considered time interval Ct-1. The corresponding values can be determined as follows:

\[
\begin{align*}
    f_t &= \sigma(w_{f1}h_{t-1} + w_{f2}x_t + b_f); \\
    i_t &= \sigma(w_{i1}h_{t-1} + w_{i2}x_t + b_i); \\
    o_t &= \sigma(w_{o1}h_{t-1} + w_{o2}x_t + b_o); \\
    C'_t &= \tanh(w_{c1}h_{t-1} + w_{c2}x_t + b_t); \\
    C_t &= f_t C_{t-1} + i_t C'_t; \\
    h_t &= o_t \tanh(C_t),
\end{align*}
\]

where: \(wf1, wf2, wi1, wi2, wo1, wo2, wc1, wc2\) are weight coefficients; \(\sigma\) - is a transfer function (sigmoid); \(C'_t\) - the contribution of the current time step to the change in the state of the layer of LSTM neurons; \(f_t\) - a function showing the contribution of the previous state of the RNN LSTM layer to the current state, the forget gate layer function; \(i_t\) - a function showing the contribution of the current state to the current state of the layer, the input filter (input gate layer); \(o_t\) - a function showing the contribution of the current event to the generated output of a neuron, an output filter (output gate layer).

The analysis of the above expressions shows that for each event in the time series, its significance, the significance of previous events and the impact of both the current state and the previous event on the output of the neuron are determined. As a result, it becomes possible to take into account all previous events from the beginning of the considered time interval and a particularly significant event, to forget all of them or partially. This approach has shown its outlook to predict the average accident rate [1]. The use of LSTM RNS made it possible to achieve a training accuracy of more than 50% (on a training sample) and a forecast accuracy of slightly less than 50% (on a test sample) [8]. However, such results were achieved without considering additional factors that are believed to have an impact on road safety [8].

One of these factors is solar activity, expressed by the Wolf number. The impact of the solar activity on meteosensitive people can have a delayed character, namely, the peak of the impact may occur for several days before or after the change in the level of the solar activity. An indirect confirmation of this thing is the fact that the correlation between the level of the solar activity (total Wolf number) and the average accident rate is less than 0.6. In this regard the use of LSTM RNS for predicting the average accident rate seems to be justified, taking into account the influence of solar activity.

For taking into account the effect of the total Wolf number on the average accident rate, the data on the number of accidents were supplemented with data on the level of solar activity [15]. As in the previous work [8], the array of the information about the accident rate was divided into training (3000 elements) and control (652 elements) samples, and each of them was divided into sequences consisting of 32 elements for the period preceding the predicted one, and 7 control values of the forecast period. The results of RNS training are presented in Table 2.

The analysis and comparison of the data from Table 2 and the data of the previous article [1] shows that the learning error, in the case of taking into account the influence of the total Wolf number, decreases, regardless of the parameters of the learning process. However, the forecast error is not so unambiguous. Thus, it decreases in the case of a smaller number of learning epochs and repetitions in each epoch, and slightly increases in the case of an increase of the number of learning epochs and repetitions in each epoch. In general, it can be noted that taking into account the influence of the total Wolf number on the results of predicting the average accident rate makes it possible to obtain a more accurate forecast than without it.
Table 2. Values of the objective function of the error on the training and control samples.

| Item No | Epoch | Quantity | Iterations | Error of training | Error of forecasting |
|---------|-------|----------|------------|-------------------|----------------------|
| 1       | 10    |          | 200        | 0.4702            | 0.5233               |
| 2       | 30    |          | 200        | 0.4379            | 0.5915               |
| 3       | 10    |          | 400        | 0.4516            | 0.5467               |
| 4       | 30    |          | 400        | 0.4174            | 0.5820               |
| 5       | 20    |          | 300        | 0.4600            | 0.5733               |
| 6       | 10    |          | 300        | 0.4624            | 0.5810               |
| 7       | 30    |          | 300        | 0.4417            | 0.5712               |
| 8       | 20    |          | 200        | 0.4568            | 0.5256               |
| 9       | 20    |          | 400        | 0.4215            | 0.6079               |

Table 3 shows a comparison of the forecasts of the average reduced accident rate for an arbitrarily selected sequence included in the control sample obtained with taking into account (forecast 1) the influence of solar activity and without it (forecast 2).

Table 3. Comparison of forecasts of the average reduced accident rate for an arbitrarily selected sequence included in the control sample, obtained with and without taking into account the effect of solar activity and without it.

| Day | Exact Value | Forecast 1 | Forecast 2 |
|-----|-------------|------------|------------|
| 1   | 1.992715    | 2.458192   | 3.992518   |
| 2   | 0.773658    | 2.066775   | 2.252165   |
| 3   | 0.947809    | 1.419484   | 2.539621   |
| 4   | 0.773658    | 0.99381846 | 0.775908   |
| 5   | 0.077054    | 1.102194   | -1.62354   |
| 6   | -0.4454     | 0.37185985 | -0.67739   |
| 7   | -0.0971     | 0.528004   | -1.66768   |

Comparison of the results of predicting the average accident rate for one of the control samples is shown in Figure 5.

**Figure 5.** The results of predicting the normalized number of accidents using the RNS taking into account solar activity and without it. (Straight Line – Exact value; Dotted Line - forecast based on sun activity; Dot-dash - forecast based on sun activity).

In general, it can be concluded based on the above information that the RNS can be used to predict changes in the average accident rate with sufficient accuracy. The analysis of the data presented in the figure and in the table shows that the forecast without taking into account solar activity gives an overestimated accident rate for the entire forecast period. The most accurate forecast is for 1-3 days, and then the accuracy of the forecast deteriorates.
3. Conclusion
It can be concluded that the RNS can be used to predict changes in the average accident rate with sufficient accuracy. In order to improve the accuracy of the forecast, it is necessary both to implement the forecast of the maximum and minimum accident rate and to take into account additional, natural, factors that affect the predicted value.

It should be noted that taking into account the effect of the total Wolf number on the results of predicting the average accident rate makes it possible to obtain a more accurate forecast than without it.

To improve the forecasting results in further work, it will be necessary to connect additional factors (the Earth's electric field, geomagnetic field, meteorological factors), and it is also necessary to develop a method for additional processing of the experimental data of these factors.

4. References
[1] Chizhevsky A L 1995 *Space Pulse of Life* (Moscow : Misl) p 786
[2] Edited by N V and Krasnopogorskaya N V 1984 *Electromagnetic fields in the biosphere. In two volumes* (Moscow : Nauka) p 700
[3] Pressman A S 1968 *Electromagnetic Fields and Life* (Moscow : Nauka) p 288
[4] Temuryants N A, Vladimirskiy B M and Tishkin O G 1992 *Electromagnetic Signals in the Biological* (Kiev : Nauk. Dumka) p 187
[5] Aghajanyan N A, Oraevsky V N, Makarova I I and Kanonidi H D 2001 *Biomedical effects of geomagnetic disturbances* (Moscow: IZMIRAN) p 136
[6] Temuryants N A 1994 *Biological effect of electromagnetic fields* (Priroda) p 14-20
[7] Adey W R 1988 *Physiological signaling across cell membranes and cooperative influences of extremely low frequency electromagnetic fields* Biological coherence and response to external stimuli (New York : Springer) p 97
[8] Abalyaev A Yu, Grunskaya L V and Leshyov I A 2020 *Using recurrent neural networks for prediction of accident rate level on public roads* (Vladimir : Vladimir State University) p 272-5
[9] Grunskaya L V, Isakevich V V and Isakevich D V 2020 Geophysical and astrophysical processes in Earth’s electric field *J. Phys.: Conf. Ser.* 1604 1-7
[10] Grunskaya L V and Leshchev I A 2020 Evaluation of the influence of solar activity on road and transport accidents in the Vladimir region (Aghajanian’s reading materials of III all-Russia scientific-practical conference with international participation) (Moscow : Peoples' friendship university of Russia) 64-7
[11] Official website of GIBDD MVD of Russia: [http://www.gibdd.ru/section/stat](http://www.gibdd.ru/section/stat)
[12] Regions of Russia. Socio-economic indicators 2010: [http://www.gks.ru/wps/wcm/connect/rosstat/](http://www.gks.ru/wps/wcm/connect/rosstat/)
[13] Poornachandra Sarang 2020 *Artificial Neural Networks with TensorFlow 2: ANN* (Architecture Machine Learning Projects Apress) p 726
[14] Hochreiter S and Schmidhuber J 1997 *Long Short-Term Memory* (Neural Computation, 9) 1735-80
[15] [http://sidc.oma.be/silso/datafiles](http://sidc.oma.be/silso/datafiles)