UDC 004.8
DOI: 10.20535/SRIT.2308-8893.2019.4.03

FORECASTING SO_2 EMISSION OF KILAUEA VOLCANO USING INTELLIGENT METHOD OF DATA ANALYSIS

S.I. ZABIELIN

Abstract. Kilauea is one of the most active and well-known volcanoes in the world and most of our knowledge of volcanism originates from its research. During a long study of volcanoes, many different methods of forecasting their activity were proposed, from the seismological analysis to the statistical analysis of their emissions. However, a comprehensive analysis of data arrays with the help of intelligent methods of data analysis has not been carried out before. Using fuzzy data processing methods, a neural network, volcanic and atmospheric indicators, we forecast emissions SO_2 for a period of one to three months.

Keywords: neural network, volcanology, fuzzy logic, LSTM.

INTRODUCTION

Kilauea is a shield volcano on the island of Hawaii on the southeastern tip of the Hawaiian Archipelago. It is one of the most active and well-known volcanoes in the world and most of our knowledge of volcanism originates from his research. Kilauea often erupted and expanded over a long period of time.

Kilauea is one of the most studied volcanoes in the world. Since the arrival in its neighborhood of the first Christian missionaries in 1823, detailed descriptions of the eruption of the volcano were conducted. In 1912, after the construction of the Hawaiian Observatory of Volcanoes on the caldera Kilauea, ongoing scientific research was carried out [1]. Today, this observatory has become one of the world's leading centers of volcanological research.

During a long study of volcanoes, many different methods of forecasting their activity were proposed, from seismological analysis to statistical analysis of their emissions. However, a comprehensive analysis of data arrays with the help of intelligent methods of data analysis has not been carried out before.

MODEL FOR VOLCANIC ERUPTIONS

Many different methods of volcanology try to predict the behavior of volcanoes using temperature data or seismic models [2], lava motion models, or gas motion models. The proposed model used to predict volcanic activity consists of three components: volcanological, atmospheric, and seismic models (Fig. 1).

During volcanic eruptions, a hot mixture of particles and volcanic gases is typically released into the air. The main parameters of the volcanological model include indicators of carbon dioxide (tons per month), sulfur dioxide (tons per month) and a discrete volcanic explosivity index from 0 to 8, where 0 to

10,000 meters of cubic ejected materials, 8 – more than 1,000 cubic kilometers of ejected materials [3].

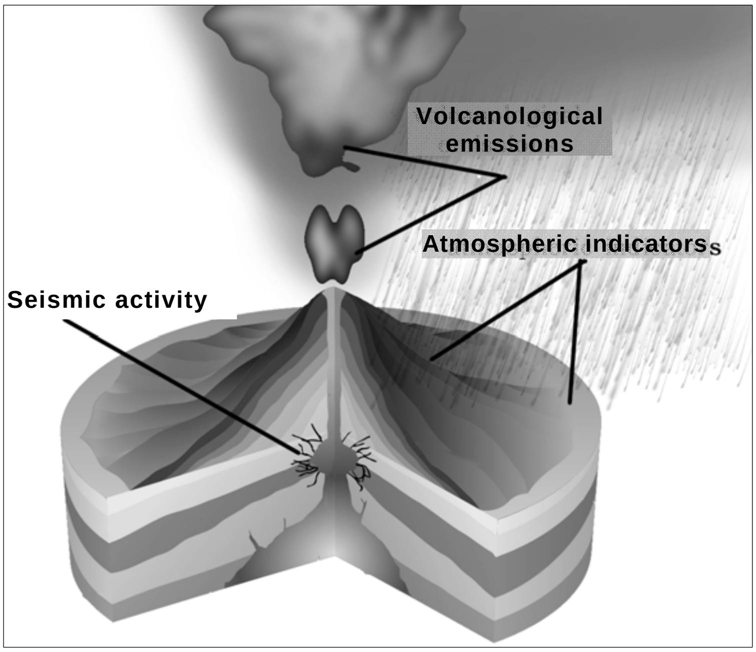

Fig. 1. Volcanic Eruption Model

Emissions of gases and hot particles affect the atmospheric performance of the metrological stations closest to the volcano. The main parameters of the atmospheric model include temperature anomaly (°C/ month), precipitation (mm per month), air pressure, wind speed. Thus, an increase in the temperature anomaly precedes the eruption of a volcano and can provide the necessary data for forecasting [4].

The eruption can also be preceded by seismic activity several hours or days before the eruption. The main parameters of the seismic model is the relative magnitude of soil vibrations [5].

STATEMENT OF THE PROBLEM

Task is to predict the behavior of the volcano using the volcanic eruption model described above and various data collected from volcanic studies over a long period.

Data

The first component is the data. It was taken from a different number of sources and reports [6–9]. Data can be divided into several categories.

Volcanological indicators include:

1. Carbon dioxide CO_2. The gas that is released during the volcanic activity of a volcano.

2. Sulfur dioxide SO_2. A colorless gas with a pungent odor that irritates the skin and mucous membranes of the eyes, nose and throat.

3. Indicator of volcanic activity. An indicator of the strength of a volcanic eruption based on an estimate of the volume of erupted products and the height of the ash column.

These indicators directly show the activity of the volcano. With a sharp increase in emissions, it can be concluded about the onset of volcanic activity.

Atmospheric indicators include:

4. Temperature anomaly (°C /month). Deviation from the reference value or long-term average.

5. Precipitation (mm per month).

6. Air pressure.

7. Wind speed. Allows you to determine the speed of gas propagation from the bowels of the volcano and make the appropriate correction for volcanological indicators.

Seismic indicators include.

Relative magnitude of soil vibrations, calculated by the following formula:

$$M_r = \frac{M}{d^2},$$

where M is the magnitude of the earthquake at the epicenter and d is the distance from the epicenter to the volcano.

METHODS USED

To analyze data, fuzzy logic and neural networks were used.

Echo time series

For higher accuracy of the neural network, the echo time series have been developed. This method is used for time series with high sparseness. Instead of fading immediately, a surge in the value in the time series generates an echo that propagates further along the time series. Each echo term in a time series is described by the following formula:

$$E_i = x_i + kx_{i-1} + k^2 x_{i-2} + \ldots + k^n x_{i-n},$$

where E_i is the echo value of the time series, x_i is the value of the input time series at time i, k is the attenuation coefficient $(k < 1)$, n is the attenuation limit, which determines the limit of the influence of previous values on the current value.

This method was used in the current problem for the relative soil vibration magnitudes, since earthquakes are rare and sporadic.

Fuzzification

To use fuzzy data it was fuzzified. Fuzzification is the process of changing a real scalar value into a fuzzy value. This is achieved with the different types of fuzzifiers (membership functions) [10].

Fuzzy data were presented in the form of triangular fuzzy numbers, described using three numbers — $\{a,b,c\}$, where b is the mode of the fuzzy number, a and c are the degrees of fuzziness of the fuzzy number. They were calculated from initial data using the following formulas:

$$b_i = x_i\,,\ a_i = x_i\left(\frac{N-1}{N}\right),\ c_i = x_i\left(\frac{N+1}{N}\right),$$

where x_i is the initial clear data at time *i*, *N* is the number of input variables, for this problem $N = 8$.

Thus, using the above-described method of fuzzification, we can obtain fuzzy numbers, which, with an increase in the number of input parameters, tend to real scalar data. This consideration is justified by the assumption that the forecast will be less fuzzy using all the indicators of the data. In other words if we had the opportunity to obtain the location and velocity vector of all molecules within a radius of 10 km from the volcano, then the forecast could be given with undeniable accuracy.

Neural networks

To solve this problem, it was decided to use recurrent neural networks, namely long-term short-term memory (LSTM). Like most recurrent neural networks, an LSTM network is universal in the sense that, with a sufficient number of network elements, it can perform any calculation that a regular computer is capable of, which requires an appropriate weight matrix that can be considered as a program. Unlike traditional recurrent neural networks, the LSTM network is well suited for training on the problems of classification, processing and forecasting time series in cases where important events are separated by time lags with an indefinite duration and boundaries. The relative immunity to the duration of time gaps gives LSTM an advantage over alternative recurrent neural networks, hidden Markov models, and other training methods for sequences in various fields of application [11].

RESULTS AGGREGATION PROCESS

To predict volcanic activity, we used the process of aggregation of results, presented in Fig. 2.

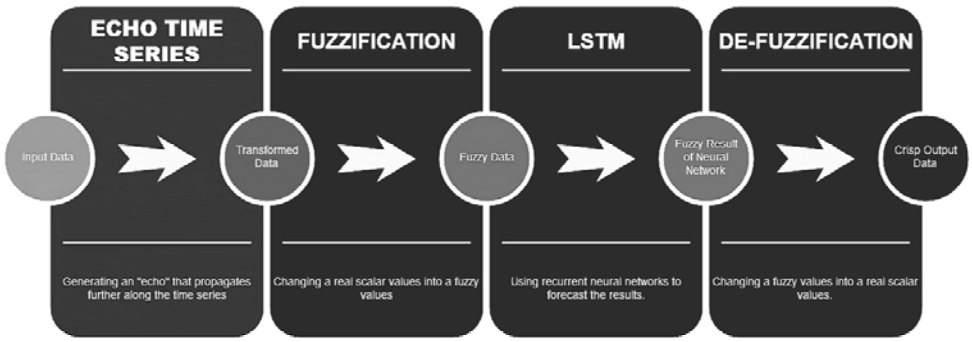

Fig. 2. Schematic process of aggregating results

The input data that was described earlier goes through several stages of data conversion, fuzzification, forecasting and de-fuzzification. Each part of the process receives input from the previous stage and, after appropriate transformations, transfers it to the next stage.

Below in Fig. 3, 4 and Table, you can see an example of the aggregation process with intermediate values obtained at various stages of its operation.

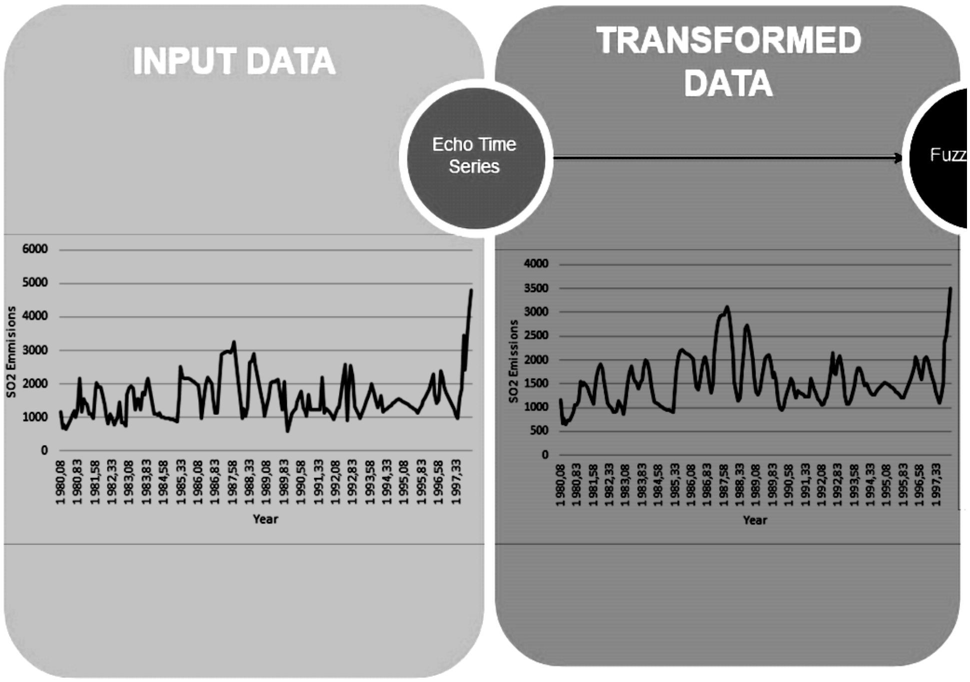

Fig. 3. Example of an aggregation process for SO_2 emissions

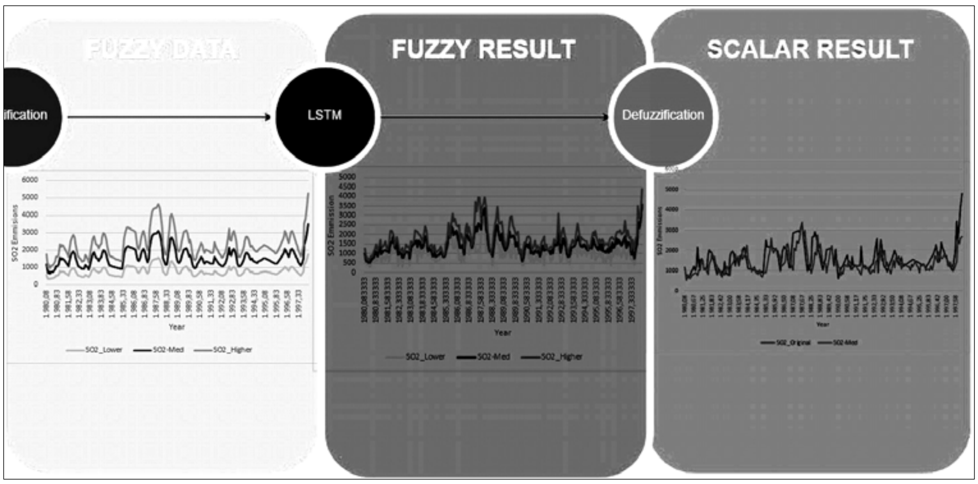

Fig. 4. Example SO_2 Emissions Aggregation Process

An example of forecasting

Date	Input SO_2	Fuzzy Low	Fuzzy High	Forecasted Low	Forecasted High	Scalar Result	MPE
1983,00	1670,20	835,10	2505,30	1632,10	1765,20	1698,65	0,04
1983,08	1872,33	936,17	2808,50	1193,17	2347,33	1770,25	0,25
1983,17	1944,51	972,26	2916,77	1415,26	2753,51	2084,38	0,24
1983,25	1864,81	932,40	2797,21	1091,40	2090,81	1591,11	0,24
1983,33	1233,60	616,80	1850,41	816,80	1422,60	1119,70	0,21
1983,42	1560,11	780,05	2340,16	1189,05	2202,11	1695,58	0,23
1983,50	1552,24	776,12	2328,35	444,12	1856,24	1150,18	0,38
1983,58	1229,47	614,74	1844,21	966,74	1140,47	1053,60	0,08
1983,67	1740,36	870,18	2610,53	1026,18	2327,36	1676,77	0,28
1983,75	1680,17	840,08	2520,25	1015,08	2161,17	1588,13	0,27
1983,83	2066,06	1033,03	3099,09	1411,03	1964,06	1687,55	0,14
1983,92	2157,86	1078,93	3236,79	1065,93	2583,86	1824,90	0,29

FORECASTING RESULTS

Prediction was carried out for emissions of SO_2 for a period of one to three months.

The predicted indicators are a triangular fuzzy number, where each of the parameters is calculated by a separate neural network. Each of the three neural networks receives real scalar data, the first is the lower boundary of the triangular number, the second is the mode, and the third is the upper boundary.

The result of the work of these networks and their various training became predicted fuzzy number. Cross-validation of data was performed when dividing the data into a training and verification sample. Data was taken from 1980 to 1997. The results are given below.

As can be seen when forecasting at 1 month (fig. 5–6), good forecasting results were obtained with an original SO_2 level that stayed in the boundaries of the predicted values.

CONCLUSIONS

In this paper, we presented a method for predicting SO_2 emissions using a large number of diverse input data, phasing methods, recurrent neural networks, and time series echoes. High accuracy indicators were obtained when forecasting from one to three months in advance.

This method can be used to detect volcanic activity at an early stage of its initiation and to prevent the catastrophic consequences that volcanic activity can bring with it.

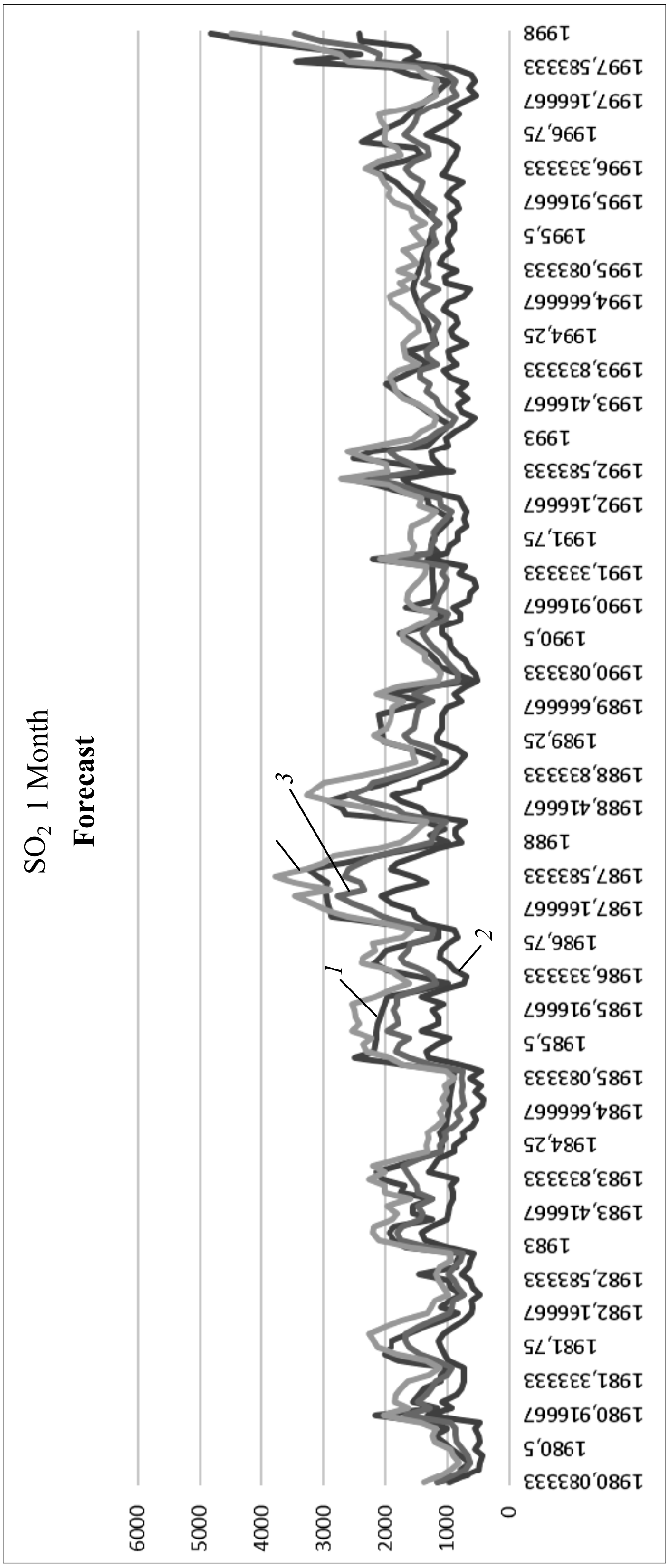

Fig. 5. Forecasting emissions for 1 month: *1* — SO_2 (Original); *2* — SO_2 (Low); *3* — SO_2 (Med); *4* — SO_2 (High)

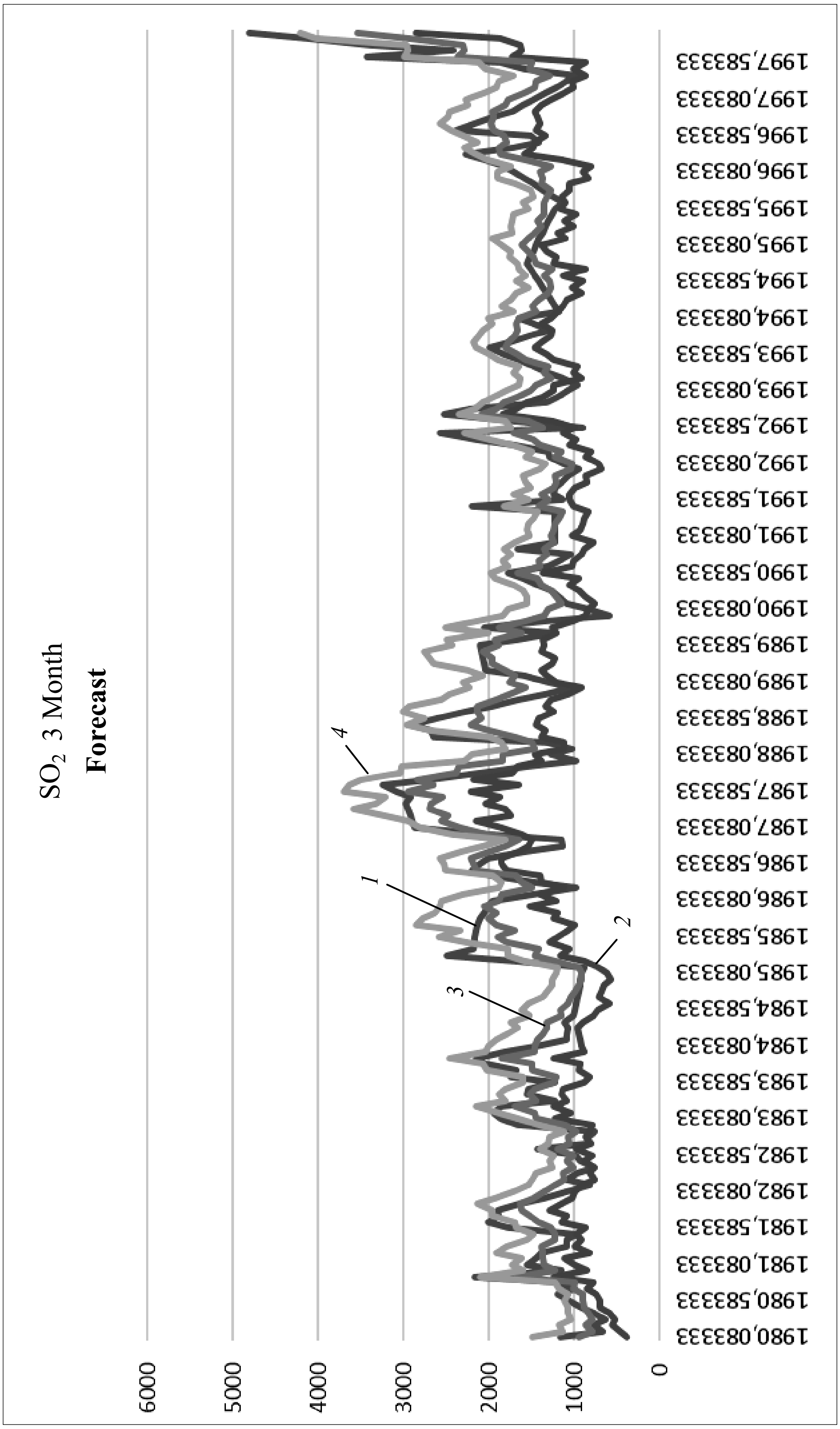

Fig. 6. Forecasting emissions for 3 month: *1* — SO_2 (Original); *2* — SO_2 (Low); *3* — SO_2 (Med); *4* — SO_2 (High)

CONCLUSIONS

In this paper, we presented a method for predicting SO_2 emissions using a large number of diverse input data, phasing methods, recurrent neural networks, and time series echoes. High accuracy indicators were obtained when forecasting from one to three months in advance.

This method can be used to detect volcanic activity at an early stage of its initiation and to prevent the catastrophic consequences that volcanic activity can bring with it.

REFERENCES

1. *Hawaii* Volcanoes National Park (N.P.), natural resources management plan. — Hawaii: Department of the Interior, 1974. — P. 32.
2. *Shimozuru D.* A seismological approach to the prediction of volcanic eruptions / D. Shimozuru // The Surveillance and Prediction of Volcanic Activity. — Paris, 1971. — P. 19–45.
3. *Aiello G.* Volcanoes: geological and geophysical setting, theoretical aspects and numerical modeling, applications to industry and their impact on the human health / G. Aiello. — London: IntechOpen, 2018. — 285 p.
4. *Gvishiani A.D.* Artificial intelligence and dynamic systems for geophysical applications / A.D. Gvishiani, J.O. Dubois. — Berlin: Springer, 2011. — P. 239–283.
5. *Zobin V.M.* Introduction to volcanic seismology / M. V. Zobin. — Amsterdam: Elsevier, 2017. — P. 29–43.
6. *Elias T.* Sulfur dioxide emission rates of Kilauea Volcano, Hawaii, 1979–1997. Menlo Park, CA: U.S. / T. Elias // Geological Survey. — 1998.
7. *Carey R.* Hawaiian volcanoes: from source to surface / R. Carey. — Washington, D.C: American Geophysical Union, 2015. — P. 393–404.
8. *Poland M.P.* Characteristics of Hawaiian volcanoes / M.P. Poland, T.J. Takahashi, C.M. Landowski. — Reston, Virginia: U.S. Department of the Interior, U.S. Geological Survey, 2014. — 429 p.
9. *Helz R.T.* Whole-rock analyses of core samples from the 1988 drilling of Kilauea Iki lava lake, Hawaii / R.T. Helz, J.E. Taggart // Open-File Report. — 2010. — doi: 10.3133/ofr20101093
10. *Bhargava A.K.* Fuzzy set theory fuzzy logic and their applications / A.K. Bhargava // Ram Nagar, New Delhi: S CHAND & CO LTD, 2013. — P. 315–348.
11. *Goodfellow I.* Deep learning. Cambridge / I. Goodfellow, Y. Bengio, A. Courville. — MA: MIT Press, 2017. — P. 408–412.

Received 25.10.2019

From the Editorial Board: the article corresponds completely to submitted manuscript.