ISSN 0027-1349, Moscow University Physics Bulletin, 2024, Vol. 79, Suppl. 2, pp. S807–S817. © Allerton Press, Inc., 2024.

Diagnostics of Geoinduced Currents in High Latitude Power Systems Using Machine Learning Methods

A. V. Vorobev[1,2*] and G. R. Vorobeva[2,3]

[1] *Geophysical Center of the Russian Academy of Scinces, Moscow, Russia*
[2] *Ufa University of Science and Technology, Ufa, Russia*
[3] *Ufa State Petroleum Technological University, Ufa, Russia*
Received October 1, 2024; revised October 10, 2024; accepted October 20, 2024

Abstract—It is known, that geoinduced currents (GICs) take place in spatially distributed current-carrying technical systems (main pipelines, power transmission lines and telegraph lines, railway infrastructure facilities, etc.) due to geomagnetic variations (GMVs), which rate of change in high-latitude regions is often about several hundred nT/min. Flowing through the grounded windings of power transformers of system-forming electrical circuits, extreme GICs are capable of transferring their magnetic systems into saturation mode, which, in turn, can cause a failure of the corresponding electrical systems. However, due to little knowledge of the mechanisms of the emergence and development of GIC, as well as the fragmentation and heterogeneity of the available empirical data, the problem of their predicting and diagnostics today is associated with many uncertainties and remains practically unsolved. The research based on machine learning methods examines approaches to diagnostics of gas and electric power level in the electric network "Severnyi Transit." In this case, both geomagnetic data recorded by magnetic stations in the subregion (Kola Peninsula, Russia) and natural (visible) indicators of extreme geomagnetic activity are used as input parameters. Using an annual sample of more than 35 000 records as an example, it was shown that the approach to GICs diagnostics, based on multiple linear regression, provides a root mean square error (RMSE) of $\sim$0.122 A^2. The use of an artificial neural network with the ReLU activation function can slightly improve the diagnostic accuracy (RMS $\sim$ 0.119 A^2). However, the interpretability and theoretical significance of the model is significantly reduced. The application, in turn, of the Bayesian classifier to the data of optical observations of auroras showed that the posterior probability of the fact that in the north the GIC level at the Vykhodnoy station during auroras will exceed 2 A is 5.78%, while the probability of exceeding this value during auroras in the zenith and south are 10.04 and 14.93%, respectively. In the absence of auroras, the model indicates that the probability of achieving a GIC of a similar level does not exceed 0.26%, and the probability of exceeding 3 A is practically zero.

Keywords: geoinduced currents, geomagnetic variations, magnetic storms, machine learning, statistical analysis

DOI: 10.3103/S0027134924702278

1. INTRODUCTION

The highest risks of reducing the level of technospheric safety are associated with the effects of space weather on high-latitude infrastructure facilities (failures of short-wave radio communication systems and railway automation, the occurrence of additional errors in magnetic inclinometers, failures of power equipment systems, decreased durability of main pipelines due to an increase in the rate of their corrosion, etc.) are detected in the auroral oval area, which is a belt of intense luminosity created by the intrusion of electrons into the atmosphere from near-Earth space [1–4].

For example, the magnetic storm on March 13, 1989 caused the failure of power transformers and a cascade shutdown (blackout) of power transmission lines for more than 9 h in the province of Quebec (Canada) [5]. In the unified power system in the north-west of Russia in November 2001, due to geomagnetic activity (GMA), there was twice a unilateral shutdown of the Olenegorsk–Monchegorsk overhead power line (330 kV) from the Olenegorsk substation,

*E-mail: geomagnet@list.ru

as a result of which consumers with a total capacity of more than 70 MW were disconnected [6, 7]. In October 2003, a similar reason led to a power outage for 20–50 min in the power system in Malmo in southern Sweden, as well as to a "false activation" of a relay at the Olenegorsk substation at the very beginning of a magnetic storm [8]. From the report of the Zurich Insurance Group it follows that in the United States alone, as a result of electrical equipment failures during periods of magnetic storms from 2005 to 2015, insurance payments exceeded \$1.9 billion [9]. Also the research [10–12] note that current pickups after almost every strong magnetic storm are the cause of synchronous anomalies in the operation of signal automation of the northern branches of the Oktyabrskaya and Northern railways. The problem is aggravated by the fact that during periods of extreme GMA, due to the shift of the auroral oval to lower latitudes, these risks also become relevant for power systems operating at midlatitudes. Described in the research [13–15] the relationship between geomagnetic variations (GMV) and the level of GIC allows the diagnostics of current interference in the presence of appropriate sources of geomagnetic data and with quality depending on their parameters and number. For example, according to [13], the 15-minute averaged GIC level at the Vykhodnoy (VKH) station can be estimated from expression (1) with a root-mean-square error of $\sim$0.16 A^2:

$$|J_{\mathrm{VKH}}| = w_0 + w_1 \left|\frac{dB}{dt}\right|, \qquad (1)$$

where w_0 and w_1 are the regression coefficients.

Figure 1 shows an example of variations of magnetic field (Y-component), field variability ($|dY/dt|$), and GIC during the period from 2015-03-16 00 : 00 (UT) to 2015-03-18 23 : 59 (UT). This time interval includes one of the powerful magnetic storms during recent years—St. Patrick's Day Geomagnetic Storm (March 17, 2015). Data are averaged over 15 min intervals.

In addition, there are a series of research [16, 17], which are concerned with the analysis of similar risks caused by the impact of space weather factors on the energy systems of Australia. However, the accuracy of the proposed methods is practically comparable to the accuracy of the GIC assessment by calculating the first derivative of the horizontal component of geomagnetic variations recorded by the nearest magnetic station, as defined in expression (1).

Despite the well calculated accuracy of the method, the limits of its applicability, within which the presented dependence remains linear, remain not entirely clear. In addition, the approach is practically unsuitable for regions that do not have dense coverage with reliable sources of geomagnetic data, for example, the Taimyr Peninsula, the Gydan Peninsula, the northern regions of the Republic of Sakha (Yakutia), etc.

Also, a certain idea of the current GIC level can be obtained by relying on satellite data on solar wind parameters recorded by the ACE and DSCOVR devices in quasiperiodic Lissajous orbits near the first Lagrange point (L1). However, due to various reasons (for example, lack of satellite data, stable Internet connection, etc.), the effectiveness of the approach, despite all its potential advantages, may be lower than expected.

To summarize, we can conclude that the existing monitoring of space weather parameters and geomagnetic field variations in the Arctic is limited to only a small group of satellites and several dozen magnetic stations located mainly in the United States, Canada, northern and central Europe. It is obvious that the situation practically excludes the possibility of promptly and reliably diagnostics of the GIC level for most of the Arctic zone of the Russian Federation (AZRF), where in fact the only available indicator of the state of space weather is still auroras.

2. SOURCE DATA AND THEIR PREPROCESSING

In the research, the Lovozero Observatory (LOZ), which is part of the Polar Geophysical Institute (PGI) and is practically the only station on the territory of the Russian Federation (Fig. 2), continuously and for a long time conducting observations and recordings, is used as the main source of data on the presence/absence of auroras, magnetic field variations and other geophysical effects in high latitudes caused by processes in the magnetosphere, ionosphere and atmosphere of the Earth. Data on auroras in the vicinity of the LOZ observatory were analyzed for more than 10-year period (from October 10, 2011 to December 31, 2021), corresponding to the highest quality results of synchronous observations of the sky and the GIC level in the subregion limited to 67.97° N, 35.02° E (Lovozero village, Murmansk region, Russia) and 68.83° N, 33.08° E (Vykhodnoy transformer substation (VKH), Murmansk region, Russia).

Thus, since 2009, the results of auroras optical observations have been openly published by PGI in the form of quarterly sets of ascaplots (Fig. 3) [18], available at: http://pgia.ru/lang/ru/archive_pgi. However, as experience shows, the data presentation format that has been established since the 1970s is practically unacceptable in tasks of intellectual analysis of large volumes of this kind of information [19]. In this regard, based on specially developed algorithms [20],

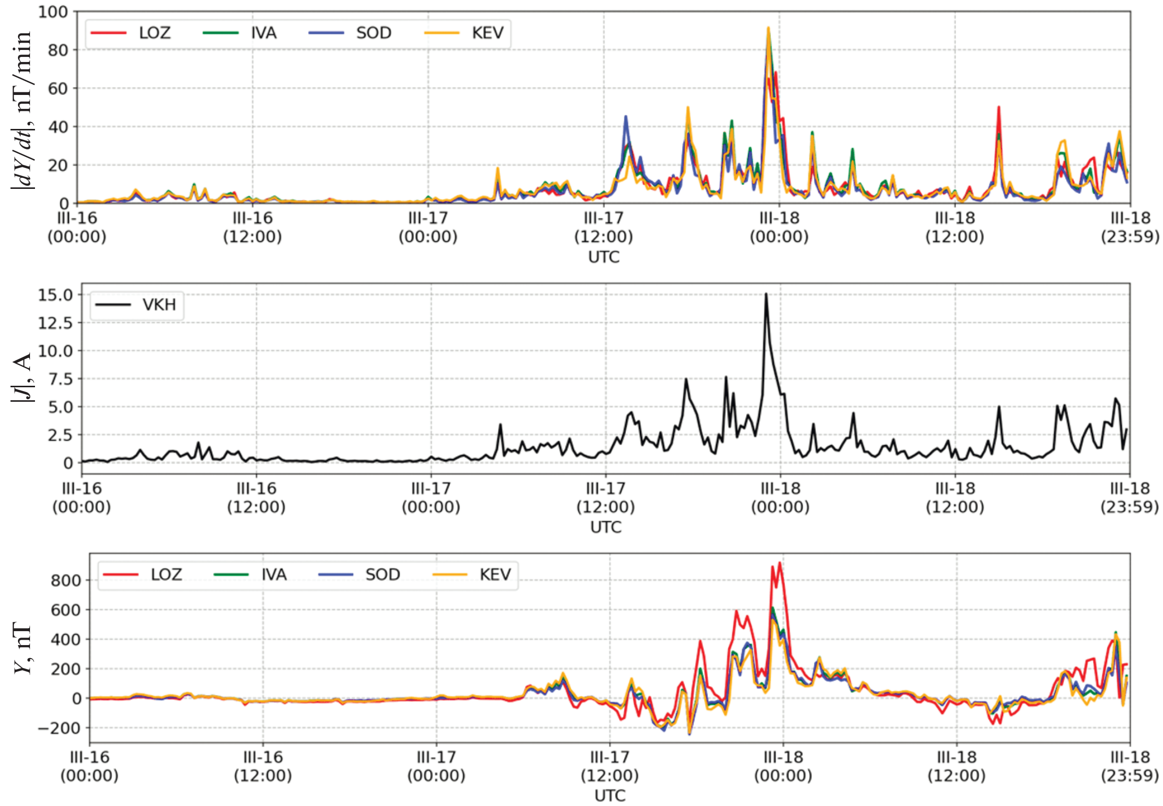

Fig. 1. A fragment of the time series (averaged over 15-min intervals) of the initial data, comprising the magnetic storm on March 17, 2015. In the time series of the Y component of the geomagnetic variations, the baseline has been excluded.

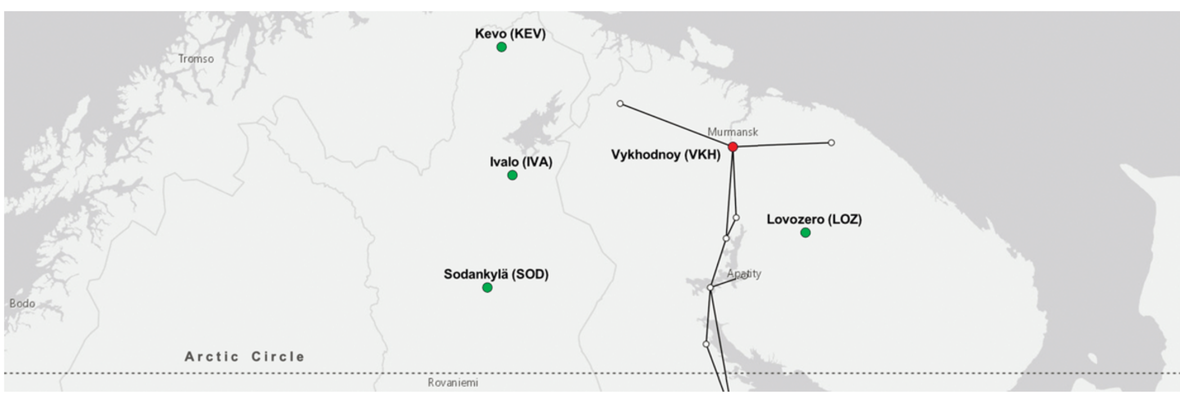

Fig. 2. Geography of data sources: magnetometers (green circles) and GIC recording station (red circle). Black solid lines denote the 330/440 kV power lines.

the original ascaplots were converted into their corresponding spreadsheets, which, in turn, were synchronized with the GIC values recorded at the VKH station. This became possible largely due to the fact that in 2011, the Kola Science Centre of the Russian Academy of Sciences (KSC RAS), together with PGI and with the assistance of the Federal Grid Company of the Unified Energy System (FGC UES), created a regional system for monitoring currents in transformer neutrals, which has accumulated a significant amount of information on the impact of GMA on the main electrical network with a length of over 800 km [21]. As a result, in 2022, a database of GIC measurements in autotransformer neutrals was published at three substations ("Vykhodnoy," "Louhi," "Kondopoga") of the 330 kV main electric network "Severny Transit" for the period 2011–2022 (certificate on state registration of database no. 2022623220 "Geoinduced currents in the main electrical network "Severny Transit," http://gic.en51.ru) [22].

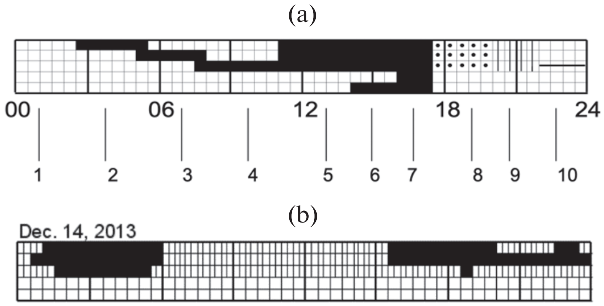

Fig. 3. Format for presenting data in the form of an ascaplot: 1—no aurora is observed; 2—aurora is in the northern region; 3—aurora is at the zenith; 4—aurora is in the south; 5—aurora is at the zenith, northern and southern regions; 6—moderate aurora at the zenith; in addition, aurora presents in the northern and southern regions; 7—strong aurora at the zenith; in addition, aurora presents in the northern and southern regions; 8—partly cloudy; 9—overcast; 10—registration was not carried out (a); example of an ascaplot from the LOZ observatory for December 14, 2013 [23] (b).

Thus, as a result of digitizing 1921 ascaplots for 2011–2021, we have 92 208 episodes of 30-min synchronous observations of the sky in the vicinity of the LOZ observatory and the GIC level at the VKH station, presented in the form of a Table 1:

$$J_{\mathrm{VKH}_n} = \frac{1}{N} \sum_{m=n}^{n+\Delta t_1/\Delta t_2} |J'_{\mathrm{VKH}}|_m, \qquad (2)$$

where Δt_1 is sampling step of optical observations of auroras (ascaplots) $\Delta t_1 = 30$ min, Δt_2 is sampling step for GIC data $\Delta t_2 = 0.5$ s, J'_{VKH} is original GIC data published by PGI.

Figure 4 represents, as an example, a time diagram of the synchronous registration of GIC at the VKH substation and aurora by the LOZ observatory for December 14, 2013. As follows from the figure, the periods of presence/absence of aurora correspond to the time intervals of significant GIC variations occurrence. Moreover, the existence of auroras in the sky southern part correlates with the appearance of extreme GIC values.

The research is also based on the data from magnetic stations of the SuperMag project [https:// supermag.jhuapl.edu/mag], which, in addition to collecting and storing geomagnetic data, also implements some procedures for their preprocessing, for example, excluding the daily component of GMF variations, the annual trend and constant displacements [24]. Prepared data further in the research are denoted by the symbol Δ (for example, ΔZ_{LOZ}) and are considered relative to the local magnetic coordinate system NEZ proposed by SuperMag [24].

3. MACHINE LEARNING METHODS FOR GIC DIAGNOSTICS

To define the methods for extreme GIC diagnostics it is necessary to describe statistical relationships betweeen GIC and various geophysical parameters. The main reson for this is that the type and character of the statistical distribution, in addition to the homogeneity of the general samples, can indicate the physical mechanisms responsible for the appearance of one or another kind of disturbance. For example, the summing up effect of many random weak independent impacts a normal distribution is formed. In a closed system, the energy of its components is distributed according to the exponential law or the Laplace law. Also a random multiplicative set of several parameters leads to a lognormal distribution, etc.

The results of the statistical analysis of the GIC values from VKH station $|J|$ and the geomagnetic field variability $|dY/dt|$ are shown in Fig. 5. The calculated probability density function (PDF), the survival function (SF) and the empirical survival function (ESF) reveal that the distribution of both parameters is best described by the lognormal distribution. Starting from ∼99.3 percentile a heavy tail can be traced, corresponding to generalized Pareto distribution law (Fig. 5a).

Analysis of the intensity of the geomagnetic variations of Y-component from neighboring to VKH magnetic stations KEV, SOD, and LOZ (Table 1) found a distribution of variability values $|dY/dt|$ similar to that shown in Fig. 5. It indicates that fluctuations of both GIC amplitude and local geomagnetic field variability are determined mainly by rare intense deviations, probably due to substorm activity and geomagnetic pulsations in the Pc5-Pi3 range [25]. At the same time, there are a number of researches that find it is convective periods that generate the extremes rather than pulsations [26]. This situation may indicate the need for additional research in this area.

The analysis of the distribution of the regional IE-index characterizing the magnetic disturbance in the auroral zone [27] demonstrates the structural similarity of the statistics. However, the distribution (not shown) predominantly corresponds to the exponential law with exponential tails.

The results of assessing the Pearson correlation coefficient r between the GIC values and the corresponding geomagnetic data are repreresnted in Table 2. To reduce the influence of the stochastic component on the result, at this stage the time series are averaged over 15-min intervals taking into account that the duration of a substorm and the time interval between two following substorms are both usually at least 30 min.

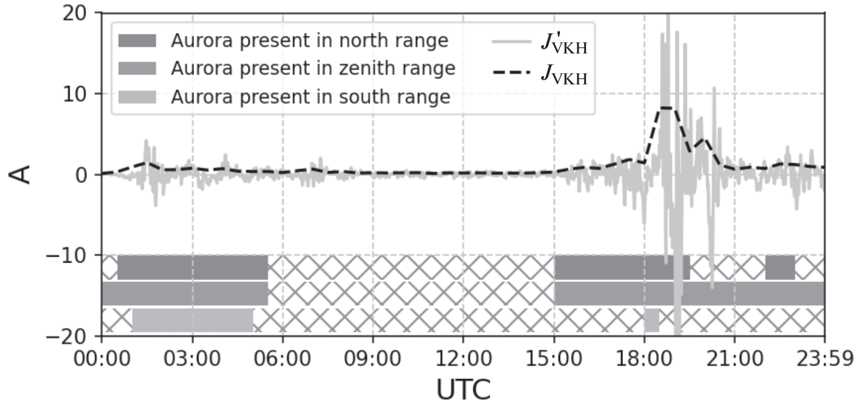

Fig. 4. Comparison of the GIC level at the VKH station and the auroral observation area in the vicinity of the LOZ observatory for December 14, 2013.

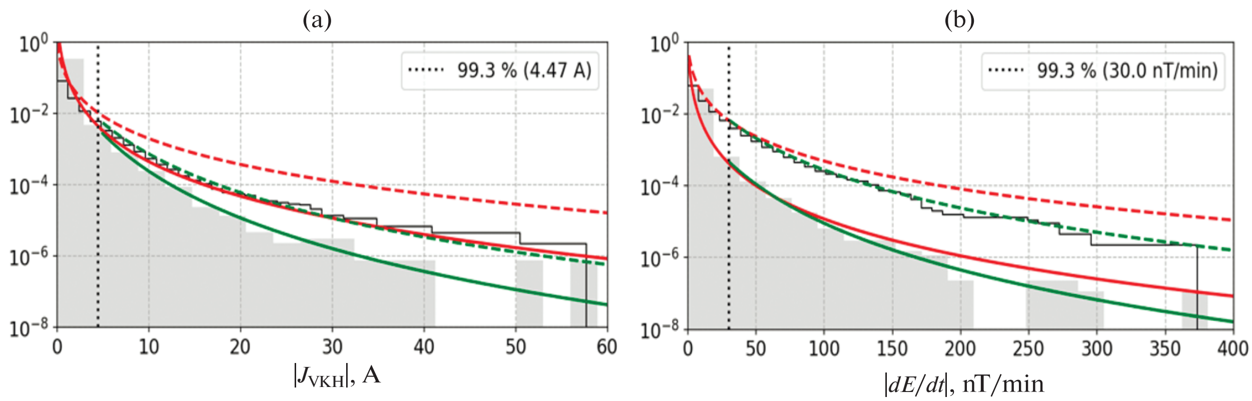

Fig. 5. Statistics of the distribution of the GIC magnitudes $|J|$ at VKH station (a) and the rate of change of Y-component of geomagnetic variations recorded at IVA station (b) for 2015 ($\Delta t = 1$ min). Red solid and dashed lines correspond to PDF and SF of the lognormal distribution, green solid and dashed lines describe the tail of the distribution and the PDF and SF of the generalized Pareto distribution; the black solid line corresponds to ESF.

The values of the Pearson correlation coefficient (Table 2) correspond to the ranking of the statistical test for the compliance of heavy tails to the generalized Pareto law results. It eliminates the cases of the false correlation, and is well explained by the spatial (km) and latitudinal (° N) separation of magnetic stations from the VKH station (Table 2). The statistics obtained indicates the consistency of the results with physical laws and geostatistical principles.

Next, the estimation of the coefficient of determination $R2$ showed that the diagnostics of the GIC values $|J_{\mathrm{VKH}}|$ based on the geomagnetic field variability from nearby magnetic stations is reasonable. Now, here, we can apply the methods based on multiple linear regression and an artificial neural network (ANN). The regression relationship can be defined as follows:

$$f(x,\beta) = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k = \sum_{j=1}^{k} \beta_j x_j = x^T \beta^T, \quad (3)$$

where $x^T = (x_1, x_2, \ldots, x_k)$ is a vector of regressors; $\beta^T = (\beta_1, \beta_2, \ldots, \beta_k)$ is a vector column of coefficients; and k is a number of model input variables.

Also to increase the stability of the model, the calculation of the regression coefficients values in (3) is performed by the elastic net method, which is a hybrid method of LASSO and ridge regression [28]:

$$\hat{\beta}_{\mathrm{ElasticNet}} = \arg\min_{\beta} \left(\sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{k} (\alpha \beta_j^2 + (1-\alpha)|\beta_j|) \right), \quad (4)$$

where λ is a regularization parameter, and $\lambda \sum_{j=1}^{k} (\alpha \beta_j^2 + (1-\alpha)|\beta_j|)$ is a model complexity penalty.

The analysis of the model's feature objects quality by relief scoring method [29] revealed that the input

Table 1. Fragment of data from synchronous observation of aurora and GIC after preprocessing

No.	UTC	J_{VKH_n}, A	Aurora in the north	Aurora at the zenith	Aurora in the south
...	...	...	...	...	...
12191	2013-12-14 18:00	1.415	1	1	2
12192	2013-12-14 18:30	8.226	1	1	1
12193	2013-12-14 19:00	8.179	1	1	2
12194	2013-12-14 19:30	2.878	1	1	2
...	...	...	...	...	...

J_{VKH_n}—GIC value determined in accordance with expression (2); 0—no aurora; 1—auroras are present; 2—cloudy.

variable $|dY_{SOD}/dt|$ has the least quality at a sufficiently high correlation coefficient with the objective function (Table 3). This indicates multicollinearity of the regressors and the necessity to exclude $|dY/dt|$ from SOD from analysis.

Finally, we have:

$$|J_{\mathrm{VKH}}| = \beta_0 + \beta_1\left(\frac{dY_{\mathrm{LOZ}}}{dt}\right) + \beta_2\left(\frac{dY_{\mathrm{IVA}}}{dt}\right) + \beta_3\left(\frac{dY_{\mathrm{KEV}}}{dt}\right) + \beta_4 IE, \qquad (5)$$

where $\beta_0 = 0.1$; $\beta_1 = 90.56 \times 10^{-3}$; $\beta_2 = 32.25 \times 10^{-3}$; $\beta_3 = 32.36 \times 10^{-3}$; $\beta_4 = 0.37 \times 10^{-3}$.

The result of GIC diagnostic based on expression (6) is shown in Fig. 6. The regression coefficients for equation (6) were obtained for the year 2015 and tested on events from the same year. In the same plot, the results of the ANN with a similar set of input features, 20 neurons in the hidden layer, and the ReLU activation function are presented. For iterative updating of ANN weights, the Adam optimization algorithm has been used, which is an extension of the stochastic gradient descent method. Table 3 shows the evaluation results of both approaches.

Table 2. Correlation of the values $|J_{\mathrm{VKH}}|$ with values of geomagnetic variations (data averaged over 15 min intervals)

	$\left\|\frac{dY_{\mathrm{LOZ}}}{dt}\right\|$	$\left\|\frac{dY_{\mathrm{IVA}}}{dt}\right\|$	$\left\|\frac{dY_{\mathrm{SOD}}}{dt}\right\|$	$\left\|\frac{dY_{\mathrm{KEV}}}{dt}\right\|$	IE-index
r	0.882	0.878	0.847	0.841	0.772

When assessing the correlation coefficient between two time series, if the ith value was absent in the first time series, then it was excluded along with the ith value of the second time series and vice versa.

Table 3. Correlation of the values $|J_{\mathrm{VKH}}|$ with values of geomagnetic variations (data averaged over 15 min intervals)

Approach	Metrics			
	R^2	MSE	RMSE	MAE
Regression	0.807	0.122	0.349	0.168
ANN	0.812	0.119	0.345	0.156

MSE is a mean squared error, RMSE is a root mean square error, MAE is a mean absolute error.

Actually, the research is based on data for 2015, which corresponds to the maximum of the 24th solar cycle. However, in fact, 28 032 (out of 35 040) observation fragments are used to build regression models, which corresponds to 80% of the total number of 15-min observation fragments and represents a training sample. The validation metrics of the model were taken based on the remaining 7008 (out of 35 040) observation fragments, which represent a test sample.

As follows from Fig. 6 and Table 3, the diagnostics of the GIC by means of the ANN because of the quantity of neurons in the hidden layer is able to provide just a slightly better result as compared to the regression model (5). However, the ANN requires much more computational resources (computer time) and is less interpretable.

As is known, the type and nature of the distribution of a random variable are largely determined by the physical mechanisms of the process being under research. It is advisable to take into account the presence of heavy tails of the distribution, indicating that the dispersion of a random variable is determined primarily by rare intense rather than frequent small deviations. Experiment results demonstrated that GIC values have statistical relationships with aurora observations too.

As follows from Fig. 7, the nature of the J_{VKH} values distribution during simultaneous observation of auroras in the sky best corresponds to the lognormal law (6) [30]. This is confirmed by the results of the Kolmogorov–Smirnov test [31], consistent with the

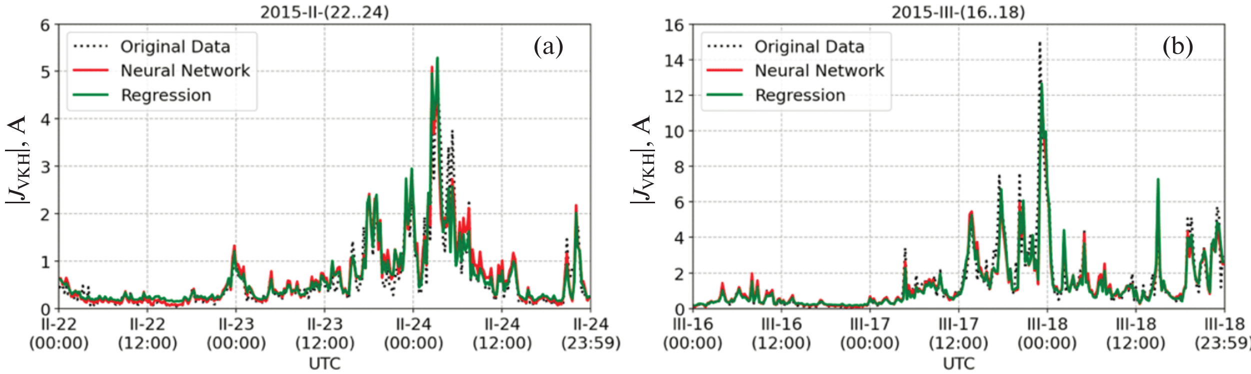

Fig. 6. Estimation of the GIC level $|J_{\rm VKH}|$ using the regression and ANN models for the magnetic storm of February 24, 2015 (Kp = 5) and March 17, 2015 (Kp = 8).

previously obtained results [13, 14], and also does not contradict the research results published by PGI [18]

$$PDF(x, s) = \frac{1}{sx\sqrt{2\pi}} \exp\left(-\frac{\log(x)^2}{2s^2}\right), \quad (6)$$

where s is a shape parameter.

Analysis of the GIC values distribution, presented in Fig. 7, showed that the most probable level of $J_{\rm VKH}$ when observing auroras in the north, zenith and south is 0.08, 0.23, and 0.68 A, respectively (Fig. 7b), which is explained by the expansion of the auroral oval during periods of strong geomagnetic activity and provides the ability to determine the level of currents induced in high-latitude power lines as a function of the area where aurora occur. Thus, as follows from Figs. 7 and 8, when observing auroras in the north, the probability that the average half-hour GIC level will exceed, for example, 2 A is $\sim 6\%$, while when observing auroras in the zenith and in the south, the probability that the GIC will exceed a similar level is $\sim 10\%$ and $\sim 15\%$, respectively (Fig. 7b). The probability that $J_{\rm VKH}$ will exceed 10 A during the period of auroras in the south is 0.15%, versus 0.06% and 0.04% when observing auroras at the zenith and in the north, respectively.

These same physical mechanisms are responsible for the formation of the geometry of the tail distribution associated with the frequency of extreme GIC occurrence (Fig. 8). For example, the statistics of GIC values during auroras in the south (Fig. 7c) has minimal values of asymmetry and kurtosis, which characterizes the thickest tail, and therefore the maximum frequency of occurrence of extreme GICs during these periods.

Also, during periods of weak diffuse auroras observations in the north, or their absence, the GIC statistics are characterized by the highest asymmetry and kurtosis, which indicates that the $J_{\rm VKH}$ values are most concentrated in the lower range and have the least uncertainty (Fig. 8a).

Analysis of Figs. 7 and 8 also indicates that the occurrence of extreme GIC practically determines the

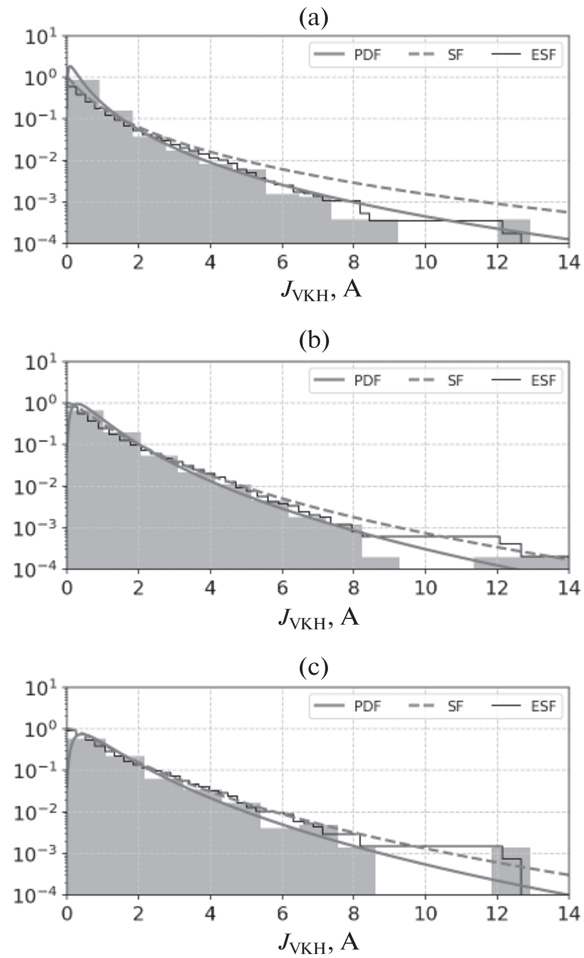

Fig. 7. GIC statistics when observing auroras in the north (a), at the zenith (b), and in the south (c). The gray solid and dotted lines correspond to the probability density function (PDF) and survival function (SF) of the lognormal distribution, respectively. Black solid line—empirical survival function (ESF).

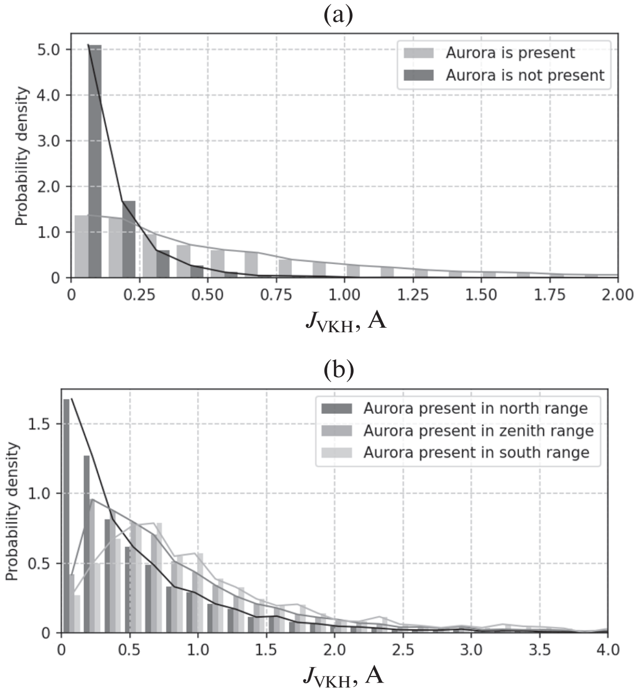

Fig. 8. Histogram of the probability density distribution of GIC values in the presence/absence of auroras (a) and when they are differentiated by areas of the sky (b). The width of the histogram intervals in this case is determined according to the rule: $h_n = 3.49sn - 1/3$, where n is the sample size, s is the standard deviation and corresponds to $\sim$0.15 A.

presence of auroras, but the observation of auroras does not guarantee the occurrence of extreme GIC values, i.e., it is not a sufficient condition for their appearance.

Correlation analysis of J_{VKH} values with the region where auroras occur, in turn, also reveals a clear connection between current pickups in high-latitude power systems and the region where auroras are observed. Moreover, during periods of observation of auroras at the zenith, the Spearman rank correlation coefficient is determined at a level of $\sim$0.7, which is 2 times higher than when auroras occur in the north or south. This result may indicate that GIC is nonlinearly related to the level of GMA and strongly depends on the location of the object of influence relative to the boundaries of the auroral oval.

4. RESULTS OF GIC DIAGNOSTICS WITH PROPOSED METHODS

Let's consider a basic approach to diagnosing the level of GIC (Fig. 9) based on observational data of auroras based on Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{7}$$

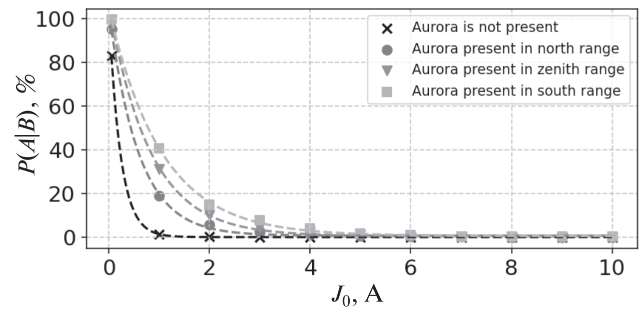

Fig. 9. A posteriori probability of exceeding the GIC at the VKH station of the J_0 level during simultaneous observation of auroras in different areas of the sky. Markers indicate calculated (empirical) values; the dotted line is the approximation of empirical values by expression (7).

where $P(A)$ is a prior probability of the hypothesis A or prior distribution; $P(A|B)$ is a probability of the hypothesis A upon the occurrence of an event B (posterior probability); $P(B|A)$ is a probability of an event B occurring if the hypothesis A is true; $P(B)$ is a the total probability of the event B occurrence, determined in accordance with the expression (8):

$$P(B) = \sum_{i=1}^{N} P(B|A_i)P(A_i), \tag{8}$$

where probabilities under the sum sign are known or can be estimated experimentally.

Then, in the context of the problem being solved, we have:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\sim A)P(\sim A)}, \tag{9}$$

where $P(A|B)$ is the probability that when observing auroras in a given area $J_{\mathrm{VKH}} \geq J_0$, where $J_0 =$ const is some given GIC value; $P(B|A)$ is the probability that when observing auroras in a given area $J_{\mathrm{VKH}} \geq J_0$; $P(A)$ and $P(\sim A)$ are the probabilities that $J_{\mathrm{VKH}} \geq J_0$ and $J_{\mathrm{VKH}} < J_0$, respectively; $P(B|\sim A)$ is the probability that when observing auroras in a given area $J_{\mathrm{VKH}} < J_0$.

Thus, the posterior probability that when auroras are observed in the north, the GIC level at station VKH will exceed 2 A is 5.78%, while the probability of exceeding this value when auroras exist in the zenith and in the south is 10.04 and 14.93%, respectively (Fig. 9). In the absence of auroras, the probability of J_{VKH} reaching a similar level does not exceed 0.26%, and the probability of exceeding 3 A is already practically zero.

As follows from Fig. 9, the dependence of the probability of exceeding the GIC level J_0 has an

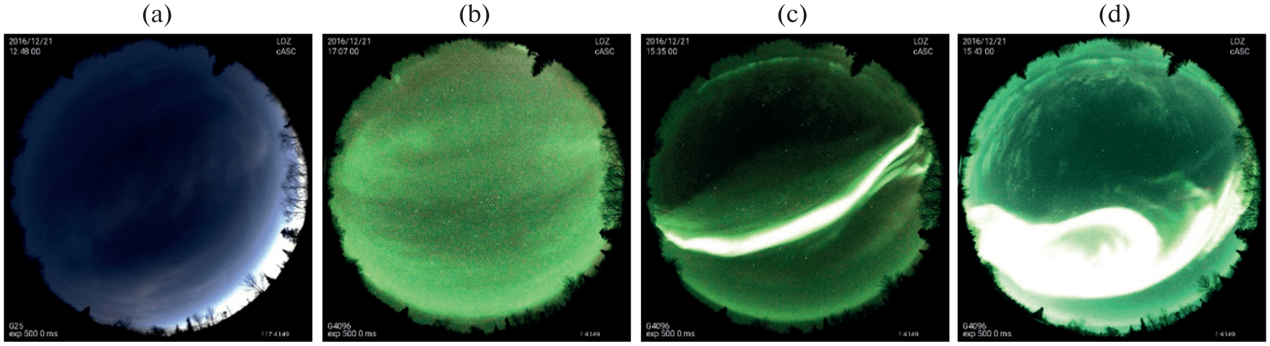

Fig. 10. State of the sky recorded by the All-sky camera of the LOZ observatory at different times of the day on December 21, 2016: (a) no auroras; (b) diffuse auroras; (c) "arc" type aurora; (d) "vortex" type aurora.

exponential character, depends on the area of manifestation (relative to the object of influence) of the auroras, and can be approximated quite well (with a discrepancy not exceeding 10^{-8} of the measured value) by an expression:

$$P(A|B) \approx P(J_0) = a\exp(bJ_0) + c, \quad (10)$$

where $a = 102.87$ for cases where there is no aurora, $a = 102.68,\ 104.69,\ 103.60$ for cases of observing auroras in the north at the zenith and in the south, respectively; similarly $b = -4.34,\ -1.69,\ -1.21,\ -0.95$ and $c = 0.04,\ 0.68,\ 0.53,\ 0.62$ for cases of absence of auroras and their observation in the north, zenith and south, respectively.

5. DISCUSSION

The results obtained seem to indicate that the presence of auroras is a necessary but not sufficient condition for the existence of extreme GICs. At the same time, the relationship between the area of observation of auroras and the value of induced currents in high-latitude power systems of a given subregion is clearly visible, which, under certain conditions, can be used as a natural indicator characterizing the probable level of GIC.

Preliminary studies also reveal a correlation between the level of GIC and the intensity of the auroras and the structure of the auroras (Fig. 10). Thus, for example, according to the data of synchronous registration of the state of the sky and GIC as of December 21, 2016, for periods without auroras (12 : 48 UT), the average minute level of GIC was 0.1 A, for diffuse auroras (17 : 07 UT)—0.7 A, and for intense auroras of the "arc" (15 : 35 UT) and "vortex" (15 : 43 UT) types—1.34 and 13.06 A, respectively.

Thus, it is natural to assume that prompt identification of the intensity of the glow and the type of auroras (for example, "diffusion," "arc" or "vortex") can significantly increase the efficiency of using natural indicators of the space weather state to assess the effects of its impact on high-latitude infrastructure objects.

In conclusion, it should be noted that since the research is predominantly based on statistical methods, the numerical values of some results obtained here are estimated and in the case of other experimental data may vary, but the relationships between them will remain the same.

6. CONCLUSIONS

The research demonstrated that the best result of the GIC diagnostics from ground-based magnetometer data is provided by regression methods or ANN using the variability of the Y-component of the geomagnetic field (dY/dt). Verification tests showed that the ANN-based approaches provide a slightly higher diagnostic accuracy ($\text{MSE} = 0.119\ A^2$) compared to the regression methods ($\text{MSE} = 0.122\ A^2$). However, the ANN methods are less interpretable and require more computational power when implemented. The MSE of the obtained relationship for the GIC diagnostics is $\sim$7.5 times lower than the MSE of similar expressions obtained previously in [13]. This advancement has been achieved thanks to a detailed analysis and careful selection of feature objects, comprising statistical analysis of heavy tails, pairwise correlation analysis, and assessing the quality of the regression model's feature objects, etc.

However, despite the fact that the highest risks of reducing the level of technospheric safety associated with the effects of space weather on power systems are identified in the auroral oval area, existing monitoring systems designed for diagnostic of extreme GIC in the power systems of the Russian Arctic are practically ineffective. As a result, the only universally available indicator of the state of space weather remains auroras, the analysis of the properties of which [30–32] can reduce the general level of situational ignorance about the probable level of current interference.

Thus, having analyzed 1921 ascaplots over a more than 10-year observation period, including 92 208 episodes of 30-min observations of the sky in the vicinity of the LOZ station, it was shown that the most probable level of GIC at the VKH station when recording auroras in the north, zenith, and south is 0.08, 0.23, and 0.68 A, respectively. In this case, the posterior probability that during auroras in the north J_{VKH} will exceed 2 A is 5.78%, while the probability of exceeding this value during auroras in the zenith and in the south is 10.04 and 14.93%, respectively. In the absence of auroras, the probability of J_{VKH} reaching a similar level does not exceed 0.26%, and the probability of exceeding 3 A is practically zero.

It is also shown that the probability of GIC exceeding a certain value is exponential and clearly depends on the area where the auroras occur. It is possible to increase the overall efficiency of the proposed approach by additionally identifying the intensity of the glow and the shape of the observed auroras. At the same time, a natural limitation of the applicability of the proposed approach is that ground-based registration of auroras in the night sky at high latitudes is possible only for up to 7 months a year, subject to favorable meteorological conditions.

ACKNOWLEDGMENTS

The authors express their gratitude to the Polar Geophysical Institute (PGI) for providing data on the observation of auroras by the Lovozero observatory, as well as PGI and the Center for Physical and Technical Problems of Northern Energy, Kola Science Centre of the Russian Academy of Sciences, for data on geoinduced currents recorded at the Vykhodnoy station.

FUNDING

This work was supported by the Russian Science Foundation (project no. 21-77-30010).

CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

REFERENCES

1. O. N. Sokolova, Ya. A. Sakharov, S. S. Gritsutenko, and N. V. Korovkin, Izv. Ross. Akad. Nauk, Energ., No. 5, 33 (2019). https://doi.org/10.1134/S0002331019050145
2. N. G. Ptitsyna, M. I. Tyasto, V. V. Kasinsky, and N. N. Lyakhov, Solnechno-Zemnaya Fizika, No. 12-2, 360 (2008).
3. A. D. Gvishiani and R. Yu. Lukyanova, in *Fundamental Basis of Innovative Technologies in the Oil and Gas Industry. Materials of the All-Russian Scientific Conference Dedicated to the 30th Anniversary of the Institute of Oil and Gas Engineering of the Russian Academy of Sciences* (2017), p. 46.
4. V. Pilipenko, A. Chernikov, A. Soloviev, N. Yagova, Ya. Saharov, D. Kudin, D. Kostarev, O. Kozyreva, A. Vorobev, and A. Belov, Russ. J. Earth Sci. **7** (3), 72 (2021). https://doi.org/10.2205/2023es000824
5. R. Kataoka and Ch. Ngwira, Prog. Earth Planet. Sci. **3** (1), 23 (2016). https://doi.org/10.1186/s40645-016-0101-x
6. V. I. Pulyaev and Yu. V. Usachev, Energetik, No. 7, 18 (2002).
7. G. A. Danilov, Yu. M. Denchik, M. N. Ivanov, G. V. Sitnikov, V. P. Gorelov, and V. G. Sal'nikov, *Improving the Quality of Operation of Power Lines: A Monograph* (Direkt-Media, Moscow, 2015).
8. W. Radasky, Z. Emin, R. Adams, et al., "CIGRE TB 780: Understanding of geomagnetic storm environment for high voltage power grids," Technical Report (2019).
9. R. W. Dobbins and K. Schriiver, Electrical claims and space weather measuring the visible effects of an invisible force June 2015. https://centerforsecuritypolicy.org/.
10. L. M. Zeleny and A. A. Petrukovich, Priroda, No. 9, 31 (2015).
11. Kh. D. Kanonidi, V. N. Oraevsky, A. V. Belov, et al., in *Problems of Forecasting Emergency Situations: Collection of Materials from a Scientific and Practical Conference* (2002), p. 41.
12. V. Pilipenko, A. A. Chernikov, A. Soloviev, N. Yagova, Ya. Saharov, D. Kudin, D. Kostarev, O. Kozyreva, A. Vorobev, and A. Belov, Russ. J. Earth Sci. **23**, ES2008 (2023). https://doi.org/10.2205/2023es000824
13. A. V. Vorobev, V. A. Pilipenko, Ya. A. Sakharov, and V. N. Selivanov, Solnechno-Zemnaya Fizika **5** (1), 48 (2019). https://doi.org/10.12737/szf-51201905
14. A. Vorobev, A. Soloviev, V. Pilipenko, G. Vorobeva, and Ya. Sakharov, Appl. Sci. **12** (3), 1522 (2022). https://doi.org/10.3390/app12031522
15. A. V. Vorobev, V. A. Pilipenko, Ya. A. Sakharov, and V. N. Selivanov, in *Problems of Geocosmos—2018*, Ed. by T. Yanovskaya, A. Kosterov, N. Bobrov, A. Divin, A. Saraev, and N. Zolotova, Springer Proceedings in Earth and Environmental Sciences (Springer,

Cham, 2020), p. 39.
https://doi.org/10.1007/978-3-030-21788-4_5
16. R. A. Marshall, E. A. Smith, M. J. Francis, C. L. Waters, and M. D. Sciffer, Space Weather **9**, 10004 (2011).
https://doi.org/10.1029/2011sw000685
17. R. A. Marshall, C. L. Waters, and M. D. Sciffer, Space Weather **8**, S05002 (2010).
https://doi.org/10.1029/2009sw000553
18. O. I. Yagodkina, V. G. Vorobyov, and E. S. Shekunova, Trudy Kol'skogo Nauchnogo Tsentra Rossiiskoi Akademii Nauk, Geliogeofizika, No. 10, 8 (2019).
https://doi.org/10.25702/KSC.2307-5252.2019.10.8
19. A. Vorobev, A. Soloviev, V. Pilipenko, G. Vorobeva, A. Gainetdinova, A. Lapin, V. Belahovskiy, and A. Roldugin, Solnechno-Zemnaya Fizika **9** (2), 26 (2023).
https://doi.org/10.12737/szf-92202303
20. A. V. Vorobev, A. N. Lapin, and G. R. Vorobeva, Informatika i Avtomatizatsiya **22**, 1177 (2023).
https://doi.org/10.15622/ia.22.5.8
21. M. B. Barannik, A. N. Danilin, Yu. V. Kat'kalov, V. V. Kolobov, Ya. A. Sakharov, and V. N. Selivanov, Instrum. Exp. Tech. **55**, 110 (2012).
https://doi.org/10.1134/s0020441211060121
22. V. N. Selivanov, T. V. Aksenovich, V. A. Bilin, V. V. Kolobov, and Ya. A. Saharov, Solnechno-Zemnaya Fizika **9** (3), 100 (2023).
https://doi.org/10.12737/szf-93202311
23. PGI Geophysical data. January, February, March 2013, Ed. by V. Vorob'ev (Polyarnyi Geofizicheskii Institut, Kol'skii Nauchnyi Tsentr Rossiiskoi Akademii Nauk, Apatity, Murmansk oblast, 2013).
24. J. W. Gjerloev, J. Geophys. Res.: Space Phys. **117**, A09213 (2012).
https://doi.org/10.1029/2012ja017683
25. N. G. Kleimenova and O. V. Kozyreva, Geomagn. Aeron. (Engl. Transl.) **49**, 1199 (2005).
https://doi.org/10.1134/s0016793209080325
26. M. P. Freeman, C. Forsyth, and I. J. Rae, Space Weather **17**, 827 (2019).
https://doi.org/10.1029/2018sw002148
27. S. V. Apatenkov, V. A. Pilipenko, E. I. Gordeev, A. Viljanen, L. Juusola, V. B. Belakhovsky, Ya. A. Sakharov, and V. N. Selivanov, Geophys. Res. Lett. **47**, e2019GL086677 (2018).
https://doi.org/10.1029/2019gl086677
28. H. Zou and T. Hastie, Journal of the Royal Statistical Society Series B: Statistical Methodology **67**, 301 (2005).
https://doi.org/10.1111/j.1467-9868.2005.00503.x
29. I. Kononenko, E. Šimec, and M. Robnik-Šikonja, Appl. Intell. **7**, 39 (1997).
https://doi.org/10.1023/A:1008280620621
30. E. Limpert, W. A. Stahel, and M. Abbt, BioScience **51**, 341 (2001).
https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2
31. D. S. Dimitrova, V. K. Kaishev, and S. Tan, J. Stat. Software **95** (10), 1 (2020).
https://doi.org/10.18637/jss.v095.i10
32. A. V. Vorobev, V. A. Pilipenko, R. I. Krasnoperov, G. R. Vorobeva, and D. A. Lorentzen, Russ. J. Earth Sci. **20**, ES6001 (2020).
https://doi.org/10.2205/2020es000721

AI tools may have been used in the translation or editing of this article.