DETECTION OF GEOMAGNETIC STORM SUDDEN COMMENCEMENTS WITH THE USE OF NEURAL NETWORK ARCHITECTURES

TARAS VOLOSHIN[1], KONSTANTIN ZAYTSEV[1*]

[1]National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 31 Kashirskoe

Avenue, Moscow, 115409, Russia

E-mail: kszajtsev@gmail.com

ABSTRACT

The purpose of the study is to examine various options to address the task of detecting the starting stage of geomagnetic storms, storm sudden commencement (SSC or SC), based on measurements of the Earth's magnetic field collected by INTERMAGNET observatories. These observatories are located in different regions of the world, allowing the full range of geomagnetic observations to be processed. Through a comprehensive analysis involving time series and machine learning techniques, including both statistical and neural network models, we developed models that integrate scalar and vector data to enhance detection accuracy. Discontinuities on the time scale in the measurements of individual observatories have been registered. In addition to the time series of magnetic field measurements, sudden commencement was detected using such scalar values as the change of the level of induction components and change of rhythm. Various methods of modeling and analyzing time series have been proposed, including statistical and machine-learning methods. To use vector and scalar indicators at the same time, the model was built with two streams of information. Various models were built using the data of both single and multiple laboratories. In the latter case, data from different sources were combined by the methods of *hard voting* and *soft voting*. A quantitative assessment of the results delivered by the models was carried out using accuracy, recall, and precision metrics.

Keywords: *Geomagnetic Storms, Sudden Commencement, Time Series, Machine Learning, Neural Networks.*

1. INTRODUCTION

Today, all spheres of human life are subjected to digital transformation and the introduction of artificial intelligence. At the most recent Bosch Connected World 2024 conference in late February of this year, Elon Musk presented information describing the growth of AI research and utilization as a unique phenomenon, increasing tenfold every 6 months. This concerns not only the economy and the social sphere but many problems in physics and technology. One prominent example of the digitalization of petabyte data and its successful analysis is the task of searching for the Higgs boson at the Large Hadron Collider using machine learning [1]. Another important problem in solar-terrestrial physics is the detection and prediction of geomagnetic storms, particularly their sudden commencement (SC) or only a slight subsequent increase in the activity of the Earth's magnetic field called a sudden impulse (SI). The detection of SC will allow the prediction of magnetic storms, which

majorly affect the communication and navigation of spaceships, and foresee the emergence of eddy currents and even the destruction of energy systems, which in turn can cause humanitarian catastrophes.

Conceptually, SC represents an unexpected and multiple sharp increase in the northern (X) magnetic component registered at INTERMAGNET observatories' monitoring stations. In most cases, this phenomenon is a precursor of electromagnetic storms detected, for example, via the minimum disturbance storm time (Dst) index. Importantly, the emphasis here is placed more on changes in the rhythm of magnetic activity than on the amplitude of possible future magnetic storms. Final registers of SC incidents are formed through visual analysis of magnetograms obtained from five specific low-latitude observation stations. The identification of SC is a topical fundamental scientific problem in geophysics.

Solar-terrestrial physics phenomena that can be utilized to detect and predict geomagnetic storms

have become objects of research around the world [2]. Analysis of recent research publications gives evidence of the increased interest in addressing this issue by various methods.

A.W. Smith et al. [3] proposed making probabilistic forecasts of SC based on data on interplanetary shocks using machine learning. Utilizing four machine learning models (logistic regression, naive Bayesian and Gaussian processes, and random forest), researchers investigate the probability of interplanetary shocks observed in point L1 being associated with SC. The most powerful predictor was found to be the range of the interplanetary magnetic field. In an earlier study, M. Shinohara et al. [4] attempted to detect SC in real-time with a special automated system using magnetometer data.

A research group from the Haystack Observatory of the Massachusetts Institute of Technology and the High-Altitude Observatory of the National Center for Atmospheric Research proposed the creation of the Transputer Integrated Diode Array Spectrometer (TIDAS) 3D system to present high-quality three-dimensional regional ionospheric information to solve the tasks of analyzing the Earth's geomagnetic conditions. Data for the implementation of this project were compiled from various types of ground and space ionospheric observations. These include measurements of total electron content (TEC) by the Global Navigation Satellite System (GNSS), satellite radio observations, and measurements by the Millstone Hill ionospheric radar at Haystack [5].

Researchers from France suggested increasing the prediction horizon of the geomagnetic activity index Kp by several days using a new experimental model, which is based on machine learning and uses images obtained with an Atmospheric Imaging Assembly device on board the Solar Dynamics Observatory spaceship [6].

C. Wang et al. [7] described a method of spectral whitening to predict non-repeating geomagnetic storms. The method is applied to 229 events caused by coronal mass ejections (CME), including 166 events with Kp ≥ 5 and 63 events with Kp < 5 during the 23rd and 24th solar cycles. The study has analyzed a total of 166 geomagnetic storm events and 129 of them were found to be predicted accurately.

Researchers at Cleveland State University [8] put forward the idea of employing a hybrid quantum-classical neural network (HQCNN), which uses quantum computing principles to simulate space weather phenomena. The proposed HQCNN offers 99.9% precise detection of cosmic weather phenomena and thus provides early warnings to mitigate the potential impact on space systems. The study further demonstrates the great potential of early detection of cosmic weather events.

A team of experts [9] has developed a new computer program AI DAGGER, which analyzes data on solar winds (defined as the stream of charged particles emitted by the Sun) gathered by spacecraft and predicts the places on Earth where the upcoming solar storm will strike with a 30-minute advance notice. Preliminary testing of the precision of the DAGGER model during two geomagnetic storms in August 2015 showed the first results to be satisfactory, as the model successfully predicted the two storms.

R. Syiemlieh and E. Saikia [10] examined the effect of cosmic rays on the formation of structural clouds. As solar activity intensifies, coronal mass ejection increases, which leads to a rise in temperature in the Earth's atmosphere. This temperature shift relates to changes in cloud formation and the distribution of precipitation and, therefore, to climate change. This urges researchers to analyze ground-based cloud images for reliable processing of background information or patterns. Geometric study of cloud properties by means of multifractal analysis (MFA) is preferable over standard statistical tools in developing an improved platform to predict future geomagnetic storms.

An article by A. Mourato et al. is devoted to the study of sunspots – magnetic disturbances in the photosphere, which are marked by their dark appearance on the solar disk and are directly connected with phenomena that contribute to these intense storms [11]. This study, combining observations of the Sun and computer vision, utilizes deep learning algorithms U-Net and Mask R-CNN for automatic detection of sunspots (semantic segmentation).

Thus, the purpose of the study was to examine diverse methodologies for detecting the initial phase of geomagnetic storms, known as storm sudden commencement (SSC or SC), utilizing magnetic field measurements gathered by the global network of INTERMAGNET observatories.

2 MATERIALS AND METHODS

2.1 Input data

The input data for our study were gathered from two open resources: the public database of the international system INTERMAGNET [12] and the database of SCs of geomagnetic storms provided by

the Ebre station (EBR; de l'Ebre, Spain). INTERMAGNET data consists of discrete, with a 1-minute step, observations of three coordinates and magnitude of the geomagnetic field. The magnitude is commonly assessed by a separate device – a scalar magnetometer [13].

Data on SC were obtained from a source of the International Service of Geomagnetic Indices (ISGI) [14], which offers data for the period since 1869. The data is presented in an archive of several text files, each containing observations over one year. Timestamps are entered into this roster of SC/SI events based on combined observations of five low-latitude observatories. Figure 1 presents a fragment of data for 2022.

```
DATE        TIME         DOY  MDUR  MAMP   CODES        TYPE   OBSERVATORIES        DURATION              AMPLITUDE                  |
2022-01-08 18:01:00.000 008  5.6   12.3   2 1 1 1 1    SSC    API SJG GUI ABG KNY  5  5  6  6  6   12.2    9.1   10.3   15.8   14.2
2022-01-16 19:09:00.000 016  3.0    7.2   1 1 1 2 1    SI     API SJG GUI ABG KNY  4  2  3  3  3    8.0    4.6    8.2    8.9    6.3
2022-01-21 13:38:00.000 021  4.4   10.5   2 2 3 3 2    SSC    HON SJG GUI ABG KNY  5  6  3  4  4    6.7    7.7    9.5   15.0   13.6
2022-02-01 22:20:00.000 032  4.6   18.1   3 3 3 3 3    SSC    HON SJG TAM ABG KNY  5  4  4  5  5   21.1   12.7   23.7   17.6   15.2
2022-02-09 21:00:00.000 040  6.2    8.1   3 1 2 2 1    SSC    HON SJG TAM ABG KNY  6  6  6  6  7    9.9    6.0    8.9    8.2    7.6
2022-02-11 10:29:00.000 042  7.6   30.8   2 1 2 2 3    SSC    HON SJG TAM ABG KNY  8  7  8  7  8   24.6   15.2   41.1   37.3   35.7
2022-03-13 10:47:00.000 072  4.0   18.4   2 1 1 2 3    SSC    HON SJG GUI ABG KNY  4  4  5  3  4   17.9    7.8   17.0   21.7   27.4
2022-03-31 02:24:00.000 090  3.4   32.2   2 2 3 3 3    SSC    HON KOU GUI ABG KNY  3  4  4  3  3   29.1   29.9   29.7   31.1   41.1
2022-04-12 11:22:00.000 102  4.0    9.5   2 1 2 1 2    SSC    HON KOU GUI ABG KNY  3  5  5  3  4    6.7   12.5   10.2    7.7   10.6
2022-06-06 10:32:00.000 157  6.2   22.1   2 2 2 3 3    SSC    HON KOU GUI ABG KNY  6  8  7  5  5   18.5   23.5   17.2   22.0   29.2
2022-06-15 04:34:00.000 166  3.6   23.3   3 3 3 2 3    SSC    HON KOU GUI ABG KNY  3  4  4  4  3   16.4   26.3   24.6   28.4   20.7
2022-07-07 06:56:00.000 188  6.8   16.4   2 2 2 2 3    SSC    HON SJG GUI ABG KNY  7  7  7  7  6   14.9   12.7   15.6   20.0   18.7
2022-07-11 17:00:00.000 192  5.8   13.3   1 1 1 2 3    SI     HON SJG GUI ABG KNY  5  5  4  8  7   10.0    6.5    9.6   19.2   21.2
2022-07-18 21:19:00.000 199  4.6    9.2   2 1 2 2 1    SI     HON SJG GUI ABG KNY  4  4  5  6  4    6.1    6.8   12.4   13.1    7.5
2022-07-23 02:57:00.000 204  2.8   28.7   3 3 3 3 3    SSC    HON SJG GUI ABG KNY  2  3  3  2  4   22.7   31.3   36.7   21.3   31.7
2022-08-17 02:58:00.000 229  4.4   22.0   1 1 1 2 3    SSC    HON SJG GUI ABG KNY  4  4  4  5  5   17.8   18.7   20.4   25.9   27.2
2022-08-19 17:31:00.000 231  3.4   18.3   2 2 1 3 2    SSC    API SJG GUI ABG KNY  3  3  3  4  4    8.0   14.9   18.5   27.3   22.7
2022-08-29 03:41:00.000 241  4.0   19.6   2 1 1 2 2    SSC    API SJG GUI ABG KNY  3  4  4  5  4   11.8   15.2   15.0   27.5   28.3
2022-09-30 05:16:00.000 273  3.8   13.6   2 2 2 2 2    SSC    HON KOU TAM GAN KNY  3  4  4  4  4    8.2   14.2   15.5   17.7   12.3
```

Figure 1: Data on the sudden commencement of geomagnetic storms for 2022

The structure of data in archival text files varies slightly from year to year. There are both common fields for all the years and different fields for different periods of observation.

Common fields for all files are:
• DATE – date,
• TIME – time,
• DOY – day of the year,
• MDUR – mean duration of the event (in minutes),
• MAMP – mean amplitude (in nT),
• CODES – qualifying codes from a group of 5 observatories (3 or 2 – the event can be unmistakably identified as SC, 1 – the event observed in this record can be an SC but requires confirmation by other observations, 0 – the event is not recognized as SC),
• TYPES – type of event (SSC or SI).

Data from after 1968 include the indication of observatories and those collected after 2004 – the duration and amplitude of the event for each observatory.

Figure 2 presents graphs of components X, Y, and Z of magnetic induction with a red dashed line indicating SSC, at the moment of which we can observe a fluctuation, especially noticeable in the X-component, followed by an increase in activity.
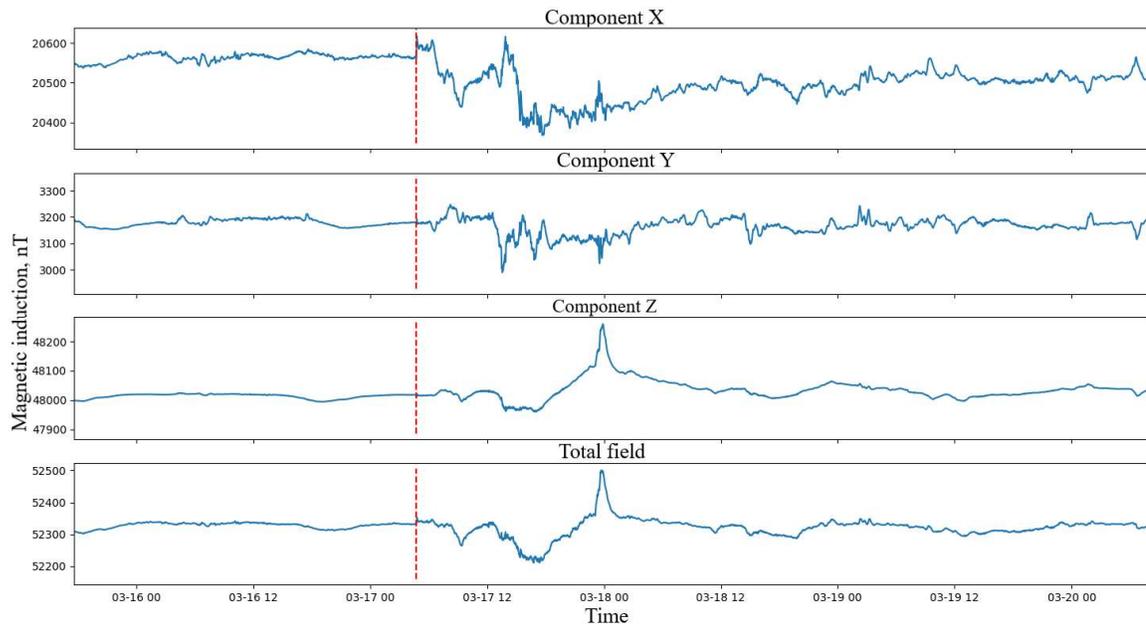
Figure 2: A fragment of magnetic field data from the Boulder observatory observations for the period from March 16, 2015 to March 20, 2015 with a mark indicating a SC of a geomagnetic storm

2.2 Problem statement and SC detection methods

Let L be one of the time scales for measuring the Earth's magnetic field induction (see Figure 2), divided into small equal windows or time intervals l_i, $i=1, 2,\ldots,n,$ $l_i \in L$, in a way so that each window either does not have an SC event (l_{False}) or has only one SC event (l_{True}). The objective is to build a model to solve the task of binary classification, i.e., determining the type of the time window (l_{True}/l_{False}) of induction values or mapping f: $L \to \{l_{True}/l_{False}\}$.

This task is addressed by machine learning methods, which allow calculating the probability of an element belonging to a specific mark.

There are a variety of algorithms to model and analyze time series. Apart from classical statistical methods, such as autoregression, machine learning algorithms have been increasingly used in the past years. Due to the specifics of analyzing time series of values, it is most convenient to apply convolutional and recurrent neural networks. For example, J. et Chen et al. [15] used convolutional networks to predict solar flares from cosmic observation data and P. Wintoft and M. Wik [16] compared different recurrent network architectures to predict the Dst index.

Another approach to this problem can be not to model the series themselves but to calculate their characteristics and generate new features and train models based on them. In the case of SCs of magnetic storms, these indicators can be parameters describing the variation and amplitude of values in the series over the considered period. In this study, we tested a combined approach, under which the model accepted as input both the interval of the time series and additional scalar characteristics.

Data analysis

Prior to solving the set problem, we examined the time series data by INTERMAGNET based on which the classification models were intended to be built. This process included statistical analysis methods. Since many prediction models appropriate for this task are intended for describing stationary series, the data at hand were tested with the augmented Dickey-Fuller test (ADF test) [17]. It was established that the series was not strictly stationary. In some cases, clear trends in data could even be observed visually on graphs. Figure 3 presents a graph of mean hourly values of the full strength of the geomagnetic field according to data from the Borok observatory over the period from 2005 to 2022. The figure clearly shows a gradual rise in the series values. Table 1 provides the results of the ADF test for the same time series.
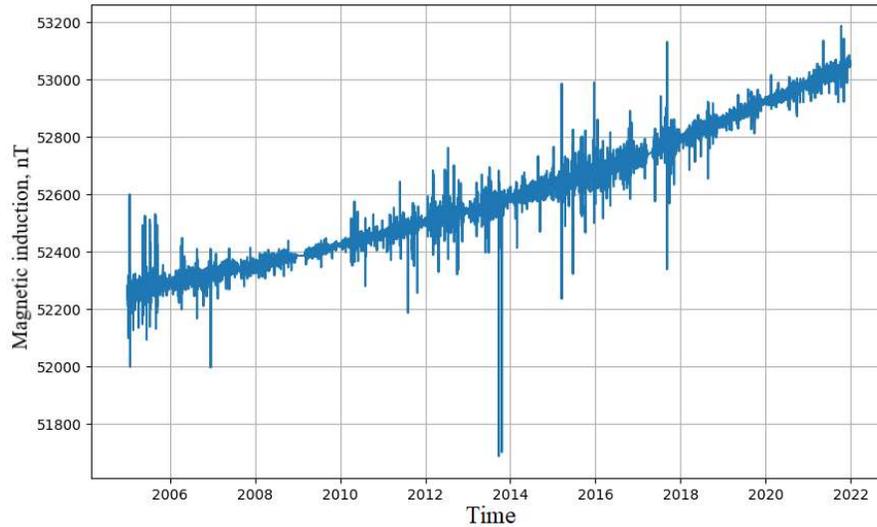
Figure 3: Graph of mean hourly values of the total magnetic field according to data from the Borok observatory

Table 1: ADF Test For The Time Series Of Mean Hourly Values Of The Full Strength Of The Total Magnetic Field According To Data From The Borok Observatory

ADF-test	0.3996
p-value	0.98

Such a trend in data can be eliminated by subtracting adjacent values, i.e., by considering a derived series of differences in neighboring indicators. To assess the possibility of applying autoregression models, we calculated autocorrelation coefficients [18]. The analysis revealed a weak correlation, albeit with a clear diurnal cyclicity. A graph of autocorrelation functions is given in Figure 4.

Furthermore, our data analysis discovered intervals with missing values, which need to be accounted for when building a model. The IAGA-2002 text files downloaded from the Intermagnet website indicate these missing values with a value of the induction component equal to "99999.00". An example of this is shown in Figure 5.
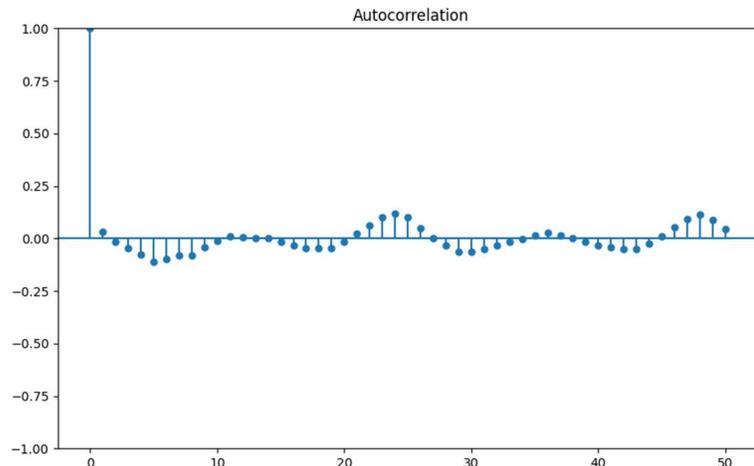

Figure 4: Autocorrelation Function Of The Time Series Of Differences In Mean Hourly Values Of The Total Magnetic Field According To Data From The Borok Observatory From 2005 To 2022 (Horizontal Axis Step – One Day)

```
DATE        TIME          DOY    HONX        HONY        HONZ        HONG     |
2022-12-20  00:00:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:01:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:02:00.000  354    99999.00    99999.00    99999.00    34389.20
2022-12-20  00:03:00.000  354    99999.00    99999.00    99999.00    34389.20
2022-12-20  00:04:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:05:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:06:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:07:00.000  354    99999.00    99999.00    99999.00    99999.00
2022-12-20  00:08:00.000  354    99999.00    99999.00    99999.00    34389.30
```

Figure 5: A Fragment Of Data From The Honolulu Observatory, USA, With Missing Values

Furthermore, In The Data Of Some Observatories, These Missing Values Coincide With Sscs, Which Reduces The Number Of Intact Samples That Could Be Used To Build Models. An Example Of Such A Coincidence Is Presented In Figure 6.
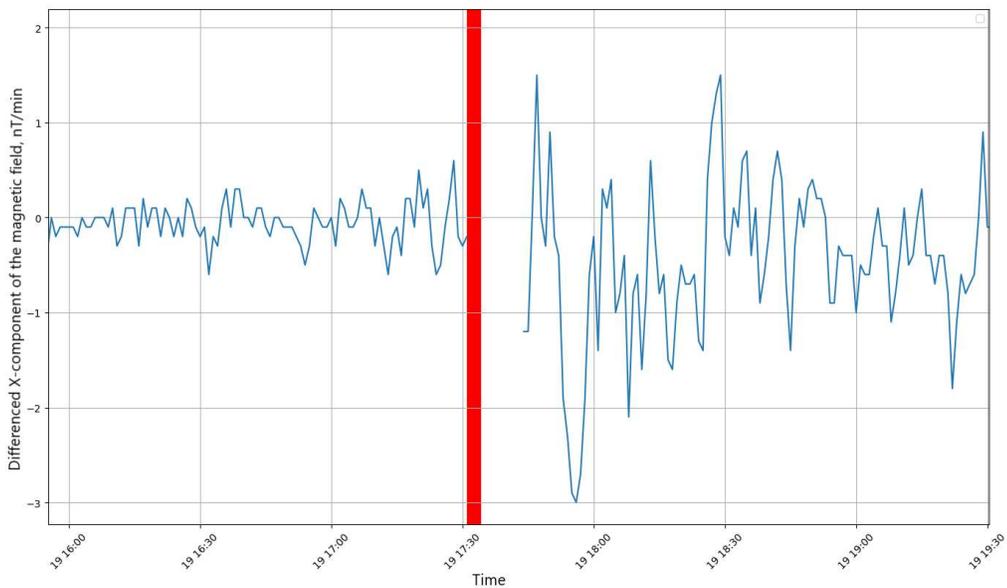

Figure 6: A Fragment Of Data (Date: 2022-08-19) From The San Juan Observatory, USA, Where A Gap In The Time Series Coincides With The SC Mark

To study the relationship between the time series of magnetic field components and SC, a general sample of data with SCs indicated was compiled and joint graphs of series with marked SCs were built.

The Intermagnet network combines a great number of observatories. However, data on SSCs published by the Ebre observatory (code EBR) typically use the observations of low-latitude laboratories. For this reason, the models were applied to data from one of these observatories – Kanoya, Japan (code KNY), collected from 2005 to 2022.

Application of machine learning methods

In this study, the automatic detection of SCs of geomagnetic storms was achieved through various neural network architectures, including recurrent neural networks (RNN). This type of neural network is designed to process organized data. All input data sequentially pass through recurrent layers, which remember the previous state and transfer it with the next set of input data. By virtue of this internal memory, RNN are efficient in analyzing time series [19]. Regular recurrent networks often experience the vanishing gradient problem, so they are rarely used in their pure form. More popular is a modification of networks with long short-term memory (LSTM) [20]. In these networks, the state of the cell passed on to the next step is controlled by three gates: the input gate, which establishes what new information should be remembered, the forget gate, which deletes part of the information, and the output gate, which defines the next value of the state.

Proceeding from LSTM network architecture, our task requires that the model is fed a sequence of values of the magnetic field

measurements time series. Based on the conducted data analysis, we decided to use a difference series of the X-component of the magnetic field vector with a one-step delay in time (one minute). From a physical point of view, this series can be interpreted as the rate of change in induction measured in nT/min. A sharp rise in this rate is what indicates possible storm commencement. A. Segarra and J.J. Curto [21] in their study established a threshold of 3 nT/min, below which samples were excluded from classification. A similar approach was employed in this study.

An additional challenge in training the binary classification model with respect to our objective is a high imbalance of data [22]. Magnetic storm SC events occur about several dozen times a year, and their detection requires examining series with one-minute resolution. Therefore, the following approach was utilized to prepare the data. For each SC mark, a 10-minute time "window" was taken, where the start of the SC was approximately in the middle of the interval with a possible random offset of 1-2 minutes. Next, the same number of "windows" which does not overlap with SC events but has a maximum rate of change in the field above the threshold were taken. The windows were checked for missing values to ensure that the selected samples did not get the time intervals with omissions. The first type of samples was assigned to the class True, the second type was classified as False, and during model training, they were randomly mixed.

Apart from time series of magnetic field changes, SC detection can be achieved using scalar features calculated based on values in the series. For example, one study [14] used the following values:

● change of level (CL) of induction components, calculated as a difference between the mean value in the series over the 10 minutes following the event and the mean value for the preceding 10 minutes;

● change of rhythm (CR) – the difference between standard deviations in 60 minutes before and after the event.

In this study, to enhance the accuracy of the model, we introduced similar features, the only difference being that the interval for the mean and standard deviation started not from the SC point itself, but from the ends of the 10-minute time window used previously as input data for the LSTM layer of the model.

To enable the combined use of vector and scalar parameters, the model was built as follows. First, the vector of values in the ten-minute window is fed into the recurrent layer, the outputs of which are joined by a fully connected network layer. The obtained value is then concatenated with the two scalar indicators, and the resulting vector of three values is forwarded to the final fully connected layer of the network. A diagram of the model is presented in Figure 7. The model was realized using the PyTorch library [23].
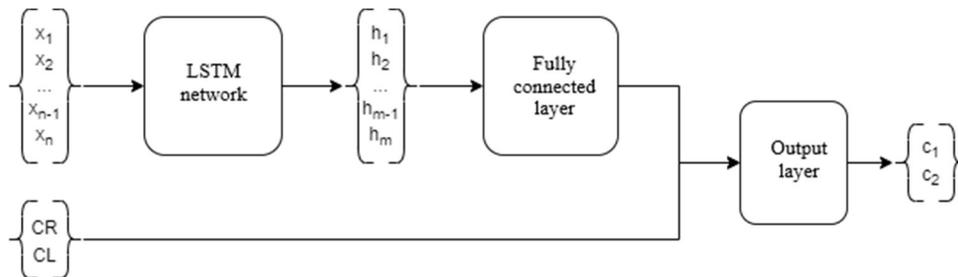

Figure 7: Scheme of the applied model

3. RESULTS

To test the models in conditions close to reality, the test sample was not balanced. A continuous time series of 2022 magnetic field data was used. The test dataset was created by applying the sliding window view to the entire series. Samples with the maximum rate of induction change below the threshold value were attributed to the False class regardless of the model's reading.

For convenient visualization and qualitative analysis of the results delivered by the model, the

graph of the differenced series of the X-component of the magnetic field, which was used as input for the model, was supplemented with red vertical marks showing the points of SSC according to the Ebre laboratory (EBR). Green vertical lines were added to mark the predictions made by the model. If the model classified some 10-minute window as True (i.e., recognized it as SC), the respective mark was placed at the start of this time window. The resulting graph is provided in Figure 8. Figure 9 shows enlarged fragments of this graph with correctly (a) and falsely (b) recognized SC

instances. The graph shows a large amount of false positive predictions, that is, marks where the model

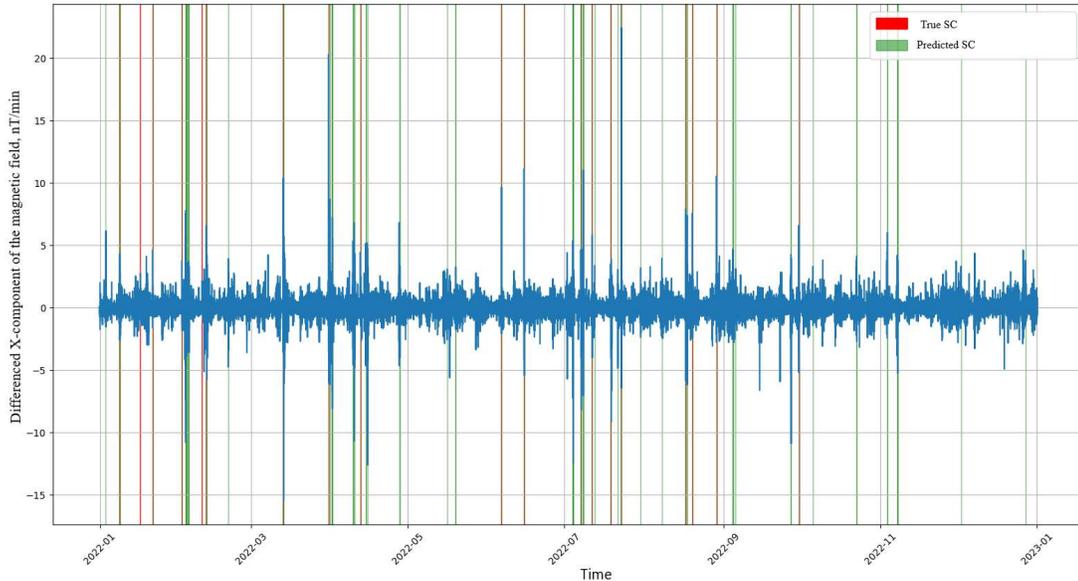predicted a storm commencement, but the EBR observatory had no such observations.

Figure 8: Graph of differenced X-component of the magnetic field for 2022 with indications of SSC events and model predictions
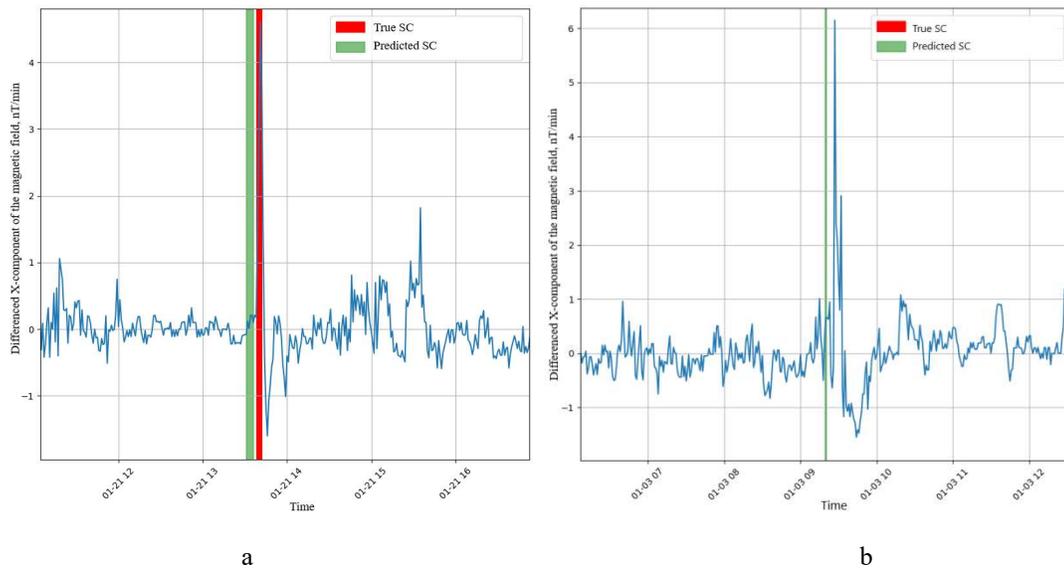

a

b

Figure 9: Enhanced fragments of the results graph: a – correctly detected SC event, b – falsely detected

Quantitative assessment of prediction results can be performed by recall and precision metrics, which are most commonly used in classification with imbalance classes, as in our case. However, for our objective, a simple comparison of real and predicted timestamps to calculate metrics will not suffice. The graphs demonstrate that with accurate detection of SC, the predicted point in time is somewhat earlier than the factual. This occurs

because the model was trained on "snippets" of values in the series (10-minute intervals) and the timestamp on the graph is placed at the start of this fragment. Thus, the metrics were calculated by a different approach: a fragment of the test sample was assigned to the True class if the time series window of this sample included at least a part of an SC event. The obtained metrics are given in Table 2. It is important to note, however, that these

metrics are not precisely indicative. For instance, recall has to show the share of accurately established True classes. The model has missed only two SC events out of sixteen, while the precision metric has a much lower value.

Table 2 – Metrics for model prediction results

Metric	accuracy	recall	precision
Value	0.999	0.337	0.413

It was hypothesized that the introduction of a new scalar indicator, the difference of standard deviations of the time series of induction rate changes averaged over a certain period after and before SC, would improve the effectiveness of SC detection. Indeed, after factual SC marks, the difference graph fluctuates more, since there is increased geomagnetic activity. Therefore, we hypothesized that a parameter similar to "change of rhythm" but applied to the derived series of differences would help reduce the number of type II errors. Nevertheless, the introduction of this new parameter into the model yielded no significant changes. The graph provided in Figure 10 still shows a significant number of false positive predictions.
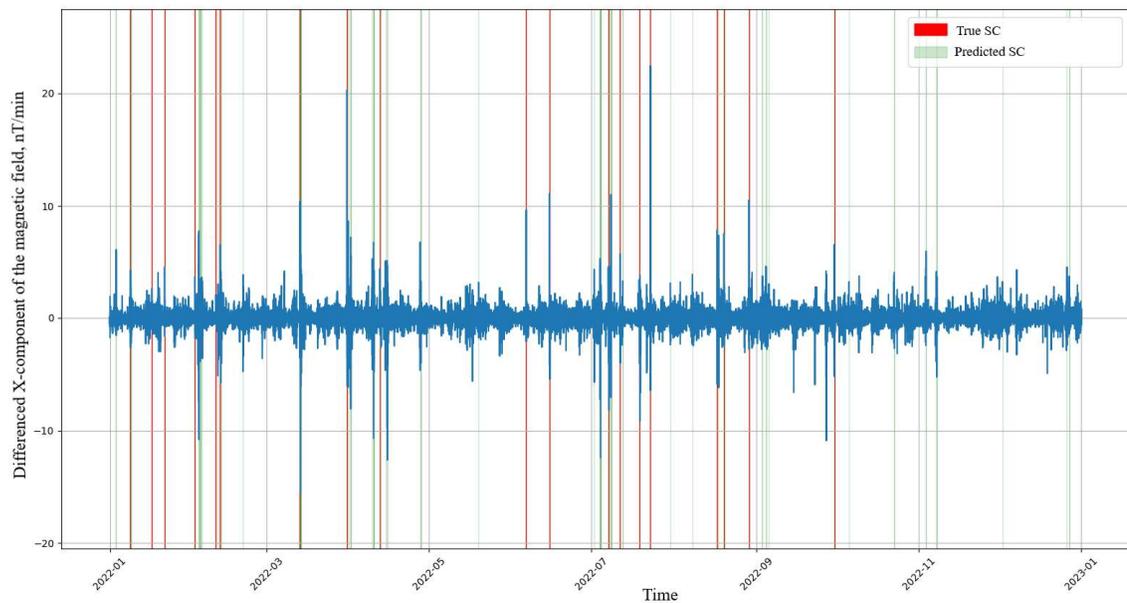

Figure 10: Results of the model with an additional scalar feature

Using data from several observatories

The next logical step was to attempt to use observation data from several low-latitude observatories together, as it is done with manual selection of SC and SI events at the Ebre Observatory (EBR).

One of the ideas was to alter the existing model so that it accepts as input a multidimensional time series compiled from data from different observatories. A similar experiment has been conducted for a two-dimensional series. Aside from the previously used data, measurements by the Alibag Observatory, India (ABG), were downloaded and processed in the same way, including the calculation of scalar indicators. Within the model, the "input size" parameter of the LSTM network was changed to make the size of one input sample (10, 2) – 10 time steps with 2 values in each. To add new scalar features, the size

of the fully connected layer was increased. A scheme of the realization of this idea is presented in Figure 11.

However, a shortcoming of this approach was discovered in the process of implementation. Omissions in data from different laboratories often did not coincide, meaning that certain SC events had to be excluded from the training sample only because the necessary fragment was missing in the observations of one of the observatories. In the same way, some SC events became impossible to identify in model testing. The situation is only worsened by the use of time series from a larger number of observatories. In the data on all SC events provided by the Ebre Observatory (EBR), this problem is addressed by the use of a variable set of five low-latitude observatories. Yet this flexibility is difficult to achieve with a single model accepting whole data from several pre-selected

observatories. For this reason, this approach was rejected at this stage of research.

The next idea was to train models of one architecture independently on data from different observatories.
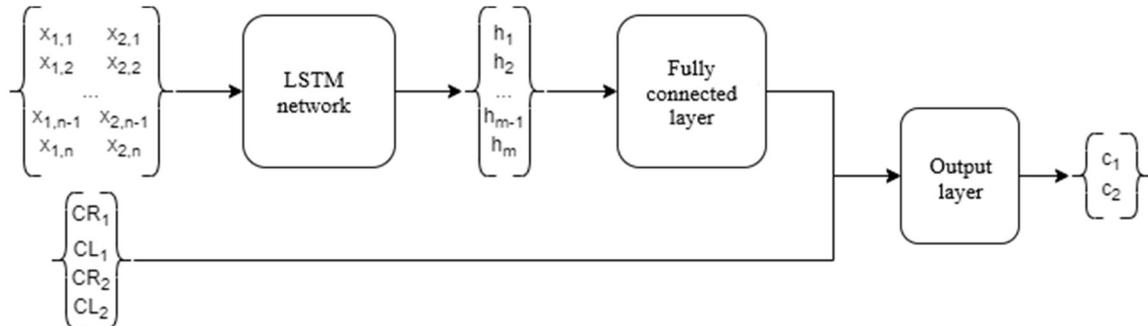

Figure 11: Diagram of the model modified for application to the data of the two observatories and subsequent "ensembling" of each observatory's predictions

Separate models were built in the same way as described above for data from the Kanoya Observatory (KNY). For this step, we took five observatories, each of which at some point was marked in the SC table of the Ebre Observatory (EBR) as deciding:

- Alibag, India (ABG),
- Guimar-Tenerife, Spain (GUI),
- Honolulu, USA (HON),
- Kanoya, Japan (KNY),
- San Juan, USA (SJG).

Thus, the results of SC/SI detection on the test sample were obtained from the data of each observatory. These results were combined by the method of *hard voting*, i.e., final marks from each model were used. The method was tested in two variants with different threshold values: the final prognosis is considered True if at least two or at least three models respond True. Initial results by observatories and final data after ensembling are provided in Table 3. The table also shows the sizes of test samples for different observatories, which differ due to omissions in raw data from some of them.

Table 3: Results of ensembling of models trained on data from different observatories

	HON	SJG	ABG	KNY	GUI	Voting (min 2)	Voting (min 3)
Number of SC/SI in the test sample	189	164	189	188	172	–	–
Accuracy	1.0	1.0	0.999	0.999	1.0	1.0	1.0
Recall	0.246	0.163	0.417	0.337	0.353	0.367	0.221
Precision	0.564	0.631	0.319	0.413	0.539	0.573	0.855

The results indicate that false positive prognoses have become much fewer. However, some real SCs that were detected previously from data from some individual observatories are not detected. The graph in Figure 12 also demonstrates the reduced number of type II errors.
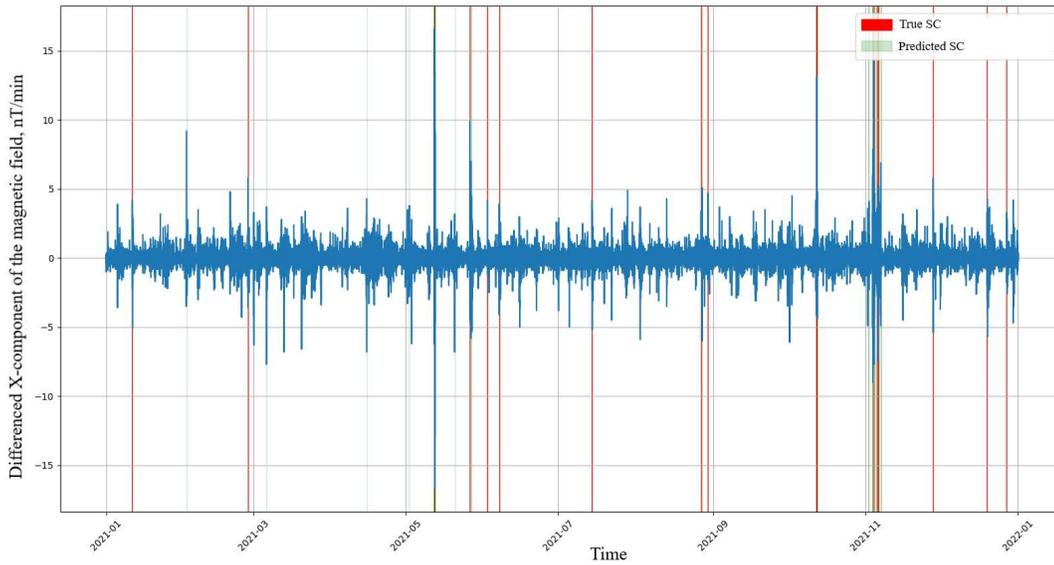
Figure 12: Graph of the predictions delivered by the model ensemble (the time series shown is from a single observatory – KNY)

Next, we tested the *soft voting* method. The *softmax* function was applied to the model's output values, providing the probabilities of the sample belonging to the True class. The probabilities estimated by five models trained on different data were averages, after which individual samples were marked as True if they surpassed the 50% threshold. Table 4 shows a comparison of metrics for the two described methods under the same model training process. The Precision metric increased with the second approach; however, the model does not recognize some real commencements of geomagnetic storms that were detected with the first approach.

Table 4 – Comparison of metrics for two approaches to model ensembling

	Hard voting	**Soft voting**
Recall	0.404	0.273
Precision	0.545	0.753

In addition to simple averaging of probabilities, we attempted weighted averaging. The weights assigned to the models were taken in a direct proportion to the metric that needed to be increased – recall. The weights were calculated using a sample for 2020-2021 and the test sample consisted of 2022 data. Table 5 shows a comparison of approaches to averaging the results of individual models. The addition of weights has indeed improved recall and the manual review of the

events shows that the number of detected events has also increased.

Table 5: Comparison of metrics for two methods of averaging probabilities

	Simple averaging	**Weighted averaging**
Recall	0.247	0.300
Precision	0.759	0.714

4. DISCUSSION

In this paper, we attempted to apply a new approach to detecting SC of geomagnetic storms based on a combination of two science-intensive methods – preliminary preparation of observation data from INTERMAGNET observatories and their deep intelligent analysis. Preliminary data processing needs to consider that these data are not strictly stationary and have gaps in the sequences of values. Intelligent analysis of the prepared data was conducted both for individual observatories and several ones participating in the establishment of SC by the EBR observatory [24]. The advantages of this approach relate to the fact that data bodies from several magnetic observatories from across the world are processed simultaneously, which prevents background errors of individual observatories in the training of detection models.

Existing studies on the detection of geomagnetic storms based on current information from various observations, which, among other methods, use neural networks to analyze the situation, either lean towards significant complication of the task or shift the focus of research from predicting geomagnetic storms and their SC to predicting other events in solar-terrestrial physics, suggesting abandoning the imbalanced distribution of observations of the Earth's magnetic field.

A complication of the task can be seen, for example, in research by the Haystack Observatory of the Massachusetts Institute of Technology and the High-Altitude Observatory of the National Center for Atmospheric Research on the creation of the TIDAS 3D system of three-dimensional regional ionospheric information by processing several types of ground- and space-based ionospheric observations [5]. This task requires continuous measurements by the GNSS, satellite radio observations, and measurements by the Millstone Hill ionospheric radar.

A shift of focus from predicting geomagnetic storms is found in studies conducted at Maryland University [25], which analyze interplanetary shock waves – disturbances normally observed in connection with the solar wind. Impacts of these waves can cause many space weather effects in the Earth's magnetopause, internal magnetosphere, ionosphere, thermosphere, and the Earth's magnetic field. The study indicates that the impact angle is a highly influential parameter that affects the geomagnetic situation and geomagnetic storms.

A change of research focus can also be observed in a paper from the National Space Science Center of the Chinese Academy of Sciences [26], where the main attention is directed toward the study of solar flares which, as the authors argue, are critical to understanding solar activity and its impact on space weather in general and magnetic storms in particular. However, the paper does not describe the method of this influence despite the application of neural network models and a decision tree.

Another interesting article by an international team of authors [27] in the sphere of solar-terrestrial physics is devoted to the dynamics of heliospheric structures that determine space weather. The emphasis is placed on ICMEs, which result from massive plasma and magnetic flux ejections from the solar corona. These ejections give rise to the largest geomagnetic storms and phenomena associated with solar energy particles,

threatening to endanger life and technology on Earth and in space. Yet once again, the paper does not touch on an algorithm of transition from research into the dynamics of heliospheric structures to magnetic storms.

Thus, drawbacks are present both in attempts to majorly complicate the task to cover many space weather phenomena and in the shift of research focus to other events and phenomena of space weather. The first approach is highly expensive in obtaining data. Furthermore, there is no evidence that these data will be sufficient to build a decent 3D model in a reasonable time and that this model will provide real-time diagnostics and produce predictions. Under the second approach to predict solar flares and heliospheric structure dynamics, which changes the focus of research but uses old methods of analysis, no qualitative changes are likely. The imbalance of data is still there, and the complexity of modeled functions is not reduced.

Several research groups utilize neural networks to diagnose various characteristics of geomagnetic storms. However, there are either individual random publications or, as in the works of University College London or the Paris Institute of Astrophysics, studies prioritizing correct calculation of geomagnetic indices rather than the detection of SC of geomagnetic storms virtually in real-time.

The approach to detecting the SC of geomagnetic storms proposed in our study is free of these disadvantages. The presented model was tested using datasets accumulated by INTERMAGNET magnetic observatories. In addition, today's neural networks are more successful with unbalanced data and allow more accurate characterization of approximations of complex multidimensional functions in solar-terrestrial physics.

5. CONCLUSIONS

The article has explored various options for solving the problem of detecting the SC of geomagnetic storms based on measurements of the Earth's magnetic field by INTERMAGNET observatories.

By integrating scalar and vector data from multiple INTERMAGNET observatory sources, our models demonstrate improved accuracy, precision, and recall over existing methods. This represents a notable advancement in the real-time monitoring and prediction of geomagnetic storms;

The application of combined data using hard and soft voting mechanisms is a novel approach that significantly enhances the robustness of SC detection models. This technique, not widely used in previous studies, offers a new pathway for handling the inherent variability and non-stationarity of geomagnetic data;

The findings from this study have immediate applications in improving early warning systems for geomagnetic storms, potentially mitigating their impact on global communication and navigation technologies;

Future research should focus on refining these models further, exploring their integration into operational platforms, and expanding their capabilities to predict other related geomagnetic phenomena.

REFERENCES

[1] A. Mott, J. Job, J.R. Vlimant, D. Lidar, M. Spiropulu, "Solving a Higgs optimization problem with quantum annealing for machine learning," Nature, vol. 550, pp. 375–379, 2017. https://doi.org/10.1038/nature24047

[2] J. J. Curto, T. Araki, and L. F. Alberca, "Evolution of the concept of Sudden Storm Commencements and their operative identification," *Earth Planet Sp*, vol. 59, pp. i–xii, 2007. https://doi.org/10.1186/BF03352059

[3] A. W. Smith, I. J. Rae, C. Forsyth, D. M. Oliveira, M. P. Freeman, D. R. Jackson, "Probabilistic forecasts of storm sudden commencements from interplanetary shocks using machine learning," *Space Weather,* vol. 18, 2020. https://doi.org/10.1029/2020SW002603

[4] M. Shinohara, T. Kikuchi, and K. Nozaki, "Automatic realtime detection of sudden commencement of geomagnetic storms," J. NICT, vol. 52, pp. 197–205, 2005. https://doi.org/10.1016/j.asr.2011.05.025

[5] J. Coster, P. J. Ericson, S. Zhang, "Imaging geomagnetic storms in the Earth's ionosphere-in 3D/MIT Haystack Observatory," Available: https://www.haystack.mit.edu/news/imaging-geomagnetic-storms-in-the-earths-ionosphere-in-3d/, March 29, 2022.

[6] G. Bernoux, A. Brunet, É. Buchlin, M. Janvier, A. Sicard, "Forecasting the Geomagnetic Activity Several Days in Advance Using Neural Networks Driven by Solar EUV Imaging," *Journal of Geophysical Research: Space Physics,* vol. 127, no. 10, 2022. https://doi.org/10.1029/2022JA030868

[7] C. Wang, Q. Ye, F. He, B. Chen, X. Zhang, "A New Method for Predicting Non-Recurrent Geomagnetic Storms," *Space Weather,* August 11, 2023. https://doi.org/10.1029/2023SW003522

[8] A. Alomari, S.A.P. Kumar, "Hybrid Classical-Quantum Neural Network for Improving Space Weather Detection and Early Warning Alerts," Conference 2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW). https://doi.org/10.1109/CCAAW57883.2023.10219316

[9] S. Shaikh, "Experts Bring An AI-Based Model – Precisely Predicts Solar Storm Worldwide," Tech latest info. Available: https://techlatest.info/listing/experts-bring-an-ai-based-model-precisely-predicts-solar-storm-worldwide/(accessed: February 18, 2024)

[10] R. Syiemlieh, E. Saikia, "Can cloud images help in predicting geomagnetic storms?," *Journal of Atmospheric and Solar Terrestrial Physics,* vol. 256, 2024. https://doi.org/10.1016/j.jastp.2024.106186

[11] A. Mourato, J. Faria, R. Ventura, "Automatic sunspot detection through semantic and instance segmentation approaches," Engineering Applications of Artificial Intelligence, vol. 129, 2024. https://doi.org/10.1016/j.engappai.2023.107636

[12] International Real-time Magnetic Observatory Network. Available: https://intermagnet.org/ (accessed: February 18, 2024)

[13] St-Louis INTERMAGNET Operations Committee, INTERMAGNET Executive Council, "INTERMAGNET Technical Reference Manual, Version 5.0.0," 2020.

[14] International Service of Geomagnetic Indices. Available: https://isgi.unistra.fr/index.php (accessed: February 18, 2024)

[15] J. Chen, W. Li, X. Zhao, J. Peng, Y. Chen, H. Deng, "Two-Stage Solar Flare Forecasting Based on Convolutional Neural Networks," Space: Science and Technology, 2022. https://doi.org/10.34133/2022/9761567

[16] P. Wintoft, M. Wik, "Exploring Three Recurrent Neural Network Architectures for Geomagnetic Predictions," Frontiers in Astronomy and Space Sciences, vol. 8, 2021.

[17] Statsmodels, "Augmented Dickey-Fuller unit root test." Available: https://www.statsmodels.org/dev/generated/sta

tsmodels.tsa.stattools.adfuller.html (accessed: February 18, 2024)

[18] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, "Time Series Analysis: Forecasting and Control," 2015.

[19] K. Sako, B. N. Mpinda, and P. C. Rodrigues, "Neural Networks for Financial Time Series Forecasting," *Entropy*, vol. 24, p. 657, 2022.

[20] R. A. Rajagukguk, R. A. A. Ramadhan, and H.-J. Lee, "A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power," *Energies*, vol. 13, p. 6623, 2020. https://doi.org/10.3390/en13246623

[21] A. Segarra, J. J. Curto, "Automatic detection of sudden commencements using neural networks," *Earth Planet Sp*, vol. 65, pp. 791–797, 2013. https://doi.org/10.5047/eps.2012.12.011

[22] M. Cristoforetti, R. Battiston, A. Gobbi, R. Iuppa, M. Piersanti, "Prominence of the training data preparation in geomagnetic storm prediction using deep neural networks," *Sci Rep*, vol. 12, p. 7631, 2022. https://doi.org/10.1038/s41598-022-11721-8

[23] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in Advances in Neural Information Processing Systems 32, pp. 8024–8035, 2019. https://doi.org/10.48550/arXiv.1912.01703

[24] J. J. Curto, J.M. Torta, J.O. Cardus, E. Sanclement, L.F. Alberca, S. Marsal, E. Blanch, L. Gaya-Pique, "Geomagnetism at Ebro: Presents and Futures", International Workshop "Challenges for Geomagnetism, Aeronomy and Seismology in the XXI Century". Retrieved from: https://dau.url.edu/handle/20.500.14342/2529

[25] M. Oliveira, "Geoeffectiveness of Interplanetary Shocks Controlled by Impact Angles: Past Research, Recent Advancements, and Future Work," Frontiers in Astronomy and Space Sciences, vol. 10, 2023. https://doi.org/10.3389/fspas.2023.1179279

[26] H. Xin-ran, Z. Qiu-zhen, C. Yan-mei, L. Si-qing, S. Yu-rong, Y. Xiao-hui, W. Zi-si-yu, "Solar Flare Short-term Forecast Model Based on Long and Short-term Memory Neural Network," Chinese Astronomy and Astrophysics, vol. 47, no. 1, pp. 108-126, 2023. https://doi.org/10.1016/j.chinastron.2023.03.003

[27] M. Akhavan-Tafti, L. Johnson, R. Sood, J. A. Slavin, T. Pulkkinen, S. Lepri, E. Kilpua, D. Fontaine, A. Szabo, L. Wolson, G. Le, T.Y. Atilaw, M. Ala-Lahti, S.L. Soni, D. Biesecker, L.K. Jian, D. Lario, "Space weather investigation Frontier (SWIFT)," *Frontiers in Astronomy and Space Sciences, Hypothesis and Theory article*, vol. 10, 2023. https://doi.org/10.3389/fspas.2023.1185603