ISSN 0027-1349, Moscow University Physics Bulletin, 2024, Vol. 79, Suppl. 2, pp. S854–S865. © Allerton Press, Inc., 2024.

Comparative Analysis of the Procedures to Forecast the Kp Geomagnetic Index by Machine Learning

I. M. Gadzhiev[1,2*], O. G. Barinov[1], S. A. Dolenko[1], and I. N. Myagkova[1]

[1]*Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia*

[2]*Faculty of Physics, Lomonosov Moscow State University, Moscow, Russia*

Received October 1, 2024; revised October 10, 2024; accepted October 20, 2024

Abstract—Geomagnetic disturbances are one of the most important factors in space weather, the role of which will increase with the development of the space industry and the global digital industry, both on Earth and in near-Earth space. Geomagnetic activity is usually characterized by special indices. One of the most common geomagnetic indices is the Kp index, first introduced by Julius Bartels in 1939. In this study, we explore the possibility of predicting the following Kp index values during the next day using machine learning models based on the hourly values of the parameters of solar wind and interplanetary magnetic field, and of the hourly Dst index. We use such ML models as linear regression, gradient boosting and multilayer perceptrons. We test to what extent the use of history of time series improves the performance of ML models. We draw conclusions about the optimal procedure of creating and applying of a machine learning model to solve the Kp index forecasting problem. The best results by most of the quality metrics were demonstrated by CatBoost and perceptron with two hidden layers. The most significant input features detected were preceding values of the Kp index itself, solar wind velocity and density, modulus and z-component of the interplanetary magnetic field.

Keywords: Kp index, regression, machine learning, CatBoost, LightGBM, neural networks

DOI: 10.3103/S002713492470231X

1. INTRODUCTION

Geomagnetic disturbances (or as they are commonly called, magnetic storms) are well known to be one of the most important factors of space weather. On Earth, magnetic storms can lead to the appearance of geomagnetically induced currents in conductive structures [1, 2], disruptions in power grids [3], radio communication systems, and satellite navigation systems, as well as increased corrosion in pipelines [4]. Geomagnetic storms can also cause dangerous situations in space, for example [5–7]. As advanced high-tech systems become more widespread, such disruptions become more significant for the economy and daily life, raising the question of reliable geomagnetic disturbance forecasting [8].

The intensity of geomagnetic disturbances is usually described using special geomagnetic indices. One of such indexes is the Kp index. It was officially introduced by Julius Bartels in 1949 [9, 10] to measure solar particle radiation by its magnetic effects. The Kp index dates back to 1932, and it is an important parameter for studying long-term changes in climatic phenomena in the upper atmosphere, in geospace, in the solar wind, as well as in the Sun.

The Kp index is a planetary geomagnetic index that characterizes the global disturbance of the Earth's magnetic field every three hours. The Kp index is determined as the average value of the disturbance levels of two horizontal components of the geomagnetic field observed at 13 magnetic observatories located in the subauroral zone between 48° and 63° of northern and southern geomagnetic latitudes, and the standardized local K indices from these 13 observatories are used to determine the Kp index. It is also known that the main cause of disturbances in the Earth's magnetosphere are processes occurring in the Sun and in the heliosphere [11].

The Kp index has a unique feature—the longest observational history after the AA index, which is calculated based on data from two magnetic stations. The Kp index has been available since 1932, while

*E-mail: ismailgadzhievff@gmail.com

the AA index has been available since 1868. Such a long series allows for the statistical investigation of the relationship between geomagnetic activity and processes in the solar wind, interplanetary space, and the Earth's magnetosphere, and further allows for the prediction of the Kp index based on the identified relationships [12, 13].

With a fundamentally different approach, the existence of a long-term Kp index time series suggests that machine learning methods, including artificial neural networks (NNs), could be used for its prediction [14–20]. One should note the wide scope of machine learning methods used: multilayer feed-forward NN [14, 17–19], support vector machines [15], NARX models [20], LSTM recurrent NN [16], gradient boosting over decision trees [21, 22]. Several national Kp forecasting operational services also use NN, e.g., those in Sweden [23–25] and in China [26, 27].

One should also note that the specificity of space weather forecasting tasks, among other things, lies in the fact that statistical indicators characterizing the results of forecasting strongly depend not only on the algorithms and forecasting techniques used, but also on the time intervals at which training and testing (evaluation) are performed. This is due to the significant dependence of the dynamics of the Earth's magnetosphere on the current phase and other characteristics of the solar activity cycle. For this reason, an adequate comparison of the results obtained by different authors is difficult.

In addition, this study aimed not at obtaining the highest possible forecast quality, but at developing an optimal forecasting procedure (methodology). For this reason, we do not try to compare the results obtained with the results of other authors, but compare the results of using various forecasting procedures.

2. DATA

The Kp index ranges from 0 (no disturbances) to 9 (extremely strong disturbances) and includes 28 discrete values: 0, 0+, 1−, 1, 1+, …, 9−, 9. The Kp index is calculated with 3-hour period, specifically at 03 : 00, 06 : 00, and so forth.

For modeling and forecasting purposes, the Kp index is often multiplied by 10. In this case, a plus corresponds to plus 3 units, and a minus to minus 3 units, converting the values to 0, 3, 7, 10, 13, …, 87, 90. In this study, we utilize Kp index data provided by the German Research Center for Geosciences in Potsdam (GFZ) [28].

In this study we make an attempt of building a model for forecasting the following Kp index values during the next day (24 h) using machine learning (ML) models based on the hourly values of the parameters of solar wind (SW) and interplanetary magnetic field (IMF), and of the hourly Dst geomagnetic index [29]. IMF and SW parameters are measured in the experiment on spacecraft *ACE* (advanced composition explorer) [30], located at the Lagrange point L1 of the Sun–Earth system. In addition to predicting future values of the Kp index, analyzing such a model could help detect features and patterns important for Kp index prediction, which might have some value from a physical point of view. For this purpose, we test different machine learning (ML) models.

Due to the nature of geomagnetic processes, where events can cause changes with some delay, including the history of measurements prior to the prediction moment in the input data for the model may improve the quality of the forecast. To account for the history, we use the time series delay embedding technique. The vector of variables at the prediction moment $x(t_i)$ is concatenated with the vectors for the N preceding moments $\{x(t_i), x(t_{i-1}), \ldots, x(t_{i-N})\}$ and then fed into a model. Here, N is referred to as the depth of the delay embedding. We test delay embeddings ranging from $N = 0$ (no delay) to 24 h, with step size of 3 h.

In our previous research [21, 22], we attempted to address the forecasting task using a classification approach. We categorized the Kp index into three classes corresponding to weak, medium, and strong disturbances. However, this approach did not generally outperform the regression approach and, in many cases, yielded even worse results (though still better than random). This could be attributed to the possibility that retaining more continuous information about the target variable is crucial for the performance of a model. Additionally, analysis of the resulting classification models revealed that they were not significantly influenced by the history of the variables, which contradicts common theoretical understanding of the process. In this study, we adopt a more straightforward approach by building regression models to directly predict Kp index values, rather than classifying the categories. It should be noted that this approach is hampered by the difference between the data period of the Kp index (3 h) and the data period for all the other input features used in the modeling.

In this study, we use variable history from 1998 to 2021 to train the models. The range of values of all variables (including the Kp index) for 2022–2023 was used as the test set to assess the quality of the models.

As outlined in the Section 1, we utilize IMF data, SW data, Dst index data, and data on the Kp index itself to forecast the Kp index value from 3 to 24 h ahead with 3-h step.

To account for variations in the index due to the Earth's rotation around its axis and around the Sun,

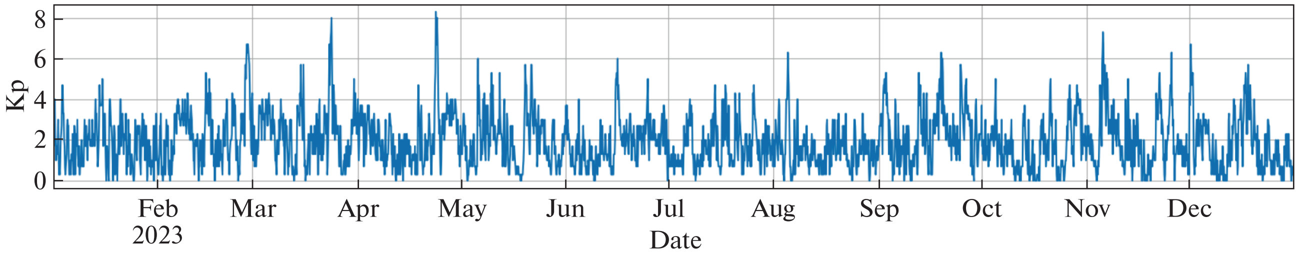

Fig. 1. Dynamics of the Kp index for year 2023. The displayed value is Kp*10.

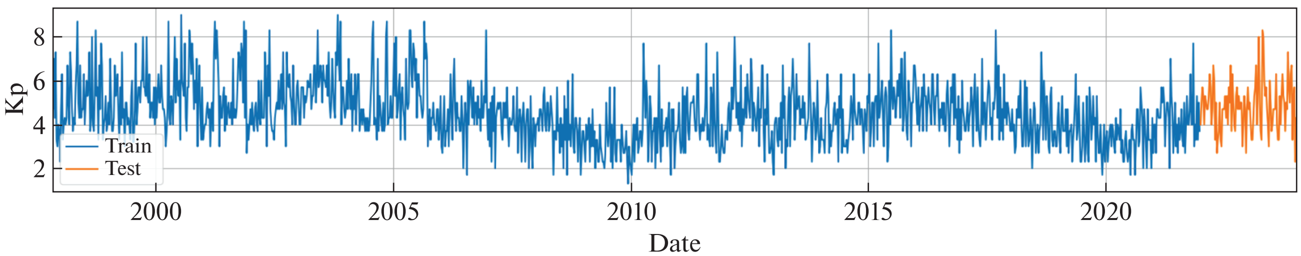

Fig. 2. Smoothed dynamics of the Kp index (Kp*10) split into the training and test sets.

we also include information about the hour within a day and the day within a year. This information is represented as sine and cosine values with daily and annual periods and it is also fed into the machine learning algorithms.

Table 1 details all the input data used to solve the regression problem. From a physics perspective, it is of interest to determine which of the features are most important for predicting the Kp index values.

It should be noted that, compared to our previous study [21], we have additionally taken into consideration the variable corresponding to the solar wind temperature at the Lagrange point (Trr_SWP).

We have also extended the data array used for this study with the data for the year 2023, characterized by a relatively large number of strong geomagnetic disturbances due to the current phase of the solar cycle. Figure 1 illustrates the dynamics of the Kp index for the year 2023.

Given some preceding data, the goal is to predict the next 8 values of the Kp index. We refer to the Kp index at horizon i as the ith next Kp index value. Because the Kp index has a three-hour period (while all other variables have an hourly period), this statement expands to the following:

1. At $00:00$, $03:00$, $06:00$, ... (hereinafter T_0 hours): predict the Kp index from 3 to 24 h ahead inclusively, in 3-h steps.

2. At $02:00$, $05:00$, $08:00$, ... (hereinafter T_1 hours): predict the Kp index from 1 to 22 h ahead inclusively, in 3-h steps.

3. At $01:00$, $04:00$, $07:00$, ... (hereinafter T_2 hours): predict the Kp index from 2 to 23 h ahead inclusively, in 3-h steps.

So, predicting Kp index at the first horizon means predicting Kp index from 1 to 3 h ahead (1 h ahead for T_1 hour, 2 h for T_2 hour, 3 h for T_0 hours), for ith horizon from $3i-2$ to $3i$ hours ahead ($3i-2$ hours ahead for T_1 hour, $3i-1$ for T_2 hour, $3i$ for T_3 hour ahead).

3. QUALITY ASSESSMENT

We use historical data from 1998 to 2021 to train the models. The range of values for all variables (including the Kp index) for 2022–2023 was used as the test set to assess the quality of the models.

Figure 2 shows weekly maximums of the Kp index split into the training and test sets.

We use root mean squared error (RMSE) to measure quality of the models on the test set and to compare them to each other.

Besides RMSE, we calculate R^2 and classification accuracy. To calculate multiclass classification accuracy, we round predictions of the models to the nearest value in the Kp index discrete grid (28 possible categories).

We compare these metrics with the results of the trivial model for which the predicted Kp value is equal to the preceding value. We consider the model's quality unsatisfactory if its RMSE is higher than that of the trivial model.

Table 1. Description of the variables used for the Kp index forecasting

No.	Variable	Physical meaning	Unit
1	Kp*10	Kp index	Dimensionless
2	Dst	Dst geomagnetic index	nT
3	B_x	x-component of the magnetic field at L1 point	nT
4	B_gsm_y	y-component of the magnetic field (in the GSM system) at L1 point	nT
5	B_gsm_z	z-component of the magnetic field (in the GSM system) at L1 point	nT
6	B_magn	Magnetic field module	nT
7	SW_spd	Solar wind speed at the Lagrange point	km/s
8	H_den_SWP	Solar wind density at the Lagrange point	cm^{-3}
8	Trr_SWP	Solar wind temperature at the Lagrange point	K
10	daySin	$\sin\left(2\pi \times \frac{[\text{day of the year}]}{365}\right)$	Dimensionless
11	hourSin	$\sin\left(2\pi \times \frac{[\text{hour of the day}]}{24}\right)$	Dimensionless
12	dayCos	$\cos\left(2\pi \times \frac{[\text{day of the year}]}{365}\right)$	Dimensionless
13	hourCos	$\cos\left(2\pi \times \frac{[\text{hour of the day}]}{24}\right)$	Dimensionless

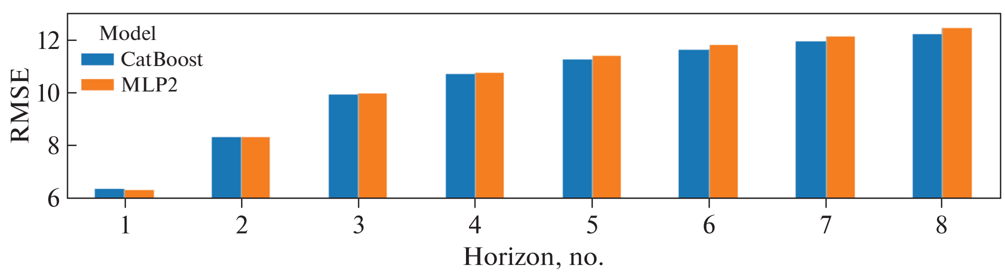

Fig. 3. Comparison plot for CatBoost and perceptron with 2 hidden layers based on RMSE (Kp*10).

Another way to judge the model's quality from an application perspective is to relate RMSE to Kp index change units (the Kp index changes in steps of 3).

We calculate all the metrics and draw conclusions for each horizon separately (8 horizons in total). But for each horizon we average results from all the hour types (T_0, T_1, T_2).

4. MODELS

The following is the list of machine learning models and their hyperparameters we test:

1. Linear regression with L2 regularization (ridge).
2. Gradient boosting regression over decision trees [31] (CatBoost [32] and LightGBM [33] implementations) with maximum tree depth 3, learning rate 0.1, maximum number of iterations: 10 000, early stopping after 50 iterations if validation loss does not decrease.
3. Multilayer perceptron [34] (MLP) with similar training setup—learning rate: 0.03, with a reducing learning rate on validation loss plateau after 30 iterations, Adam optimizer and early stopping after 30 iterations if validation loss does not decrease.
   (a) Perceptron with one hidden layer (MLP1): 128 neurons in the hidden layer.

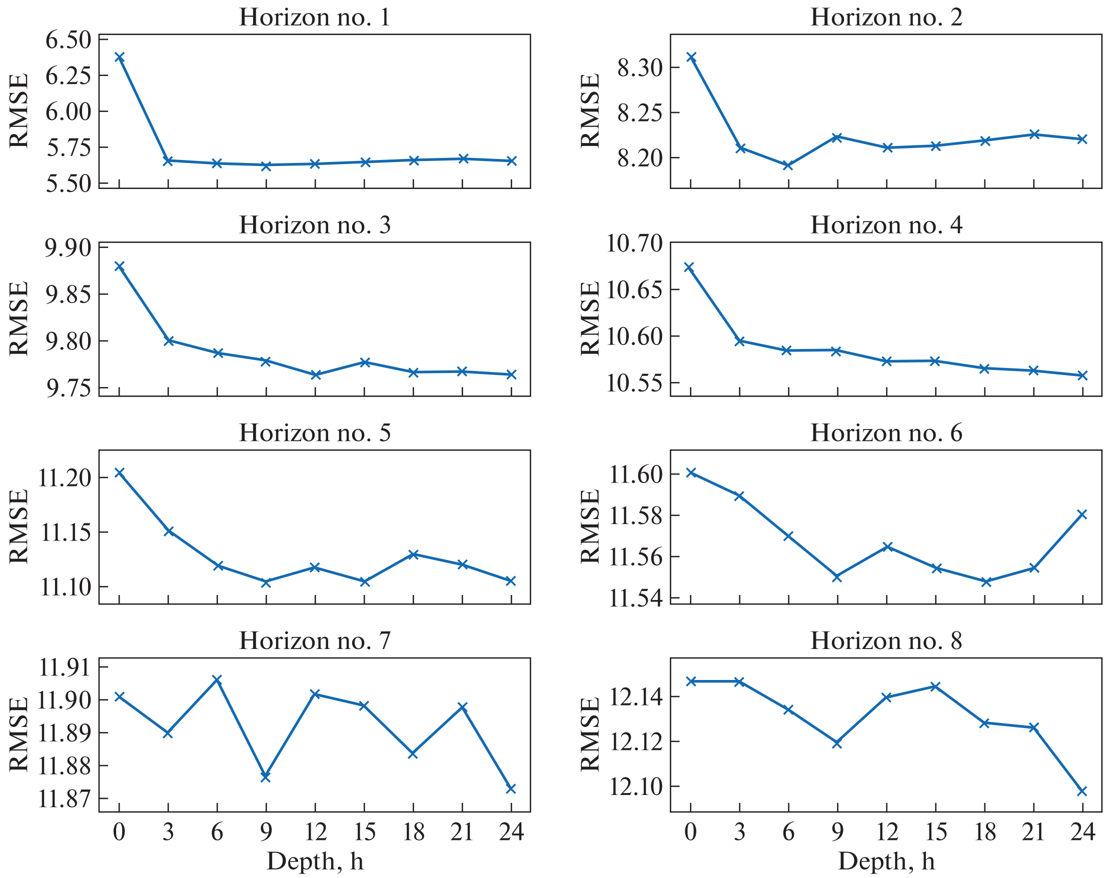

Fig. 4. Dependence of RMSE (Kp*10) of the CatBoost model on delay embedding depth for each of the forecasting horizons.

   (b) Perceptron with two hidden layers (MLP2): 256 and 128 neurons in the hidden layers.

4. Trivial Model (TM)—inertial forecast, where the predicted value of Kp is equal to the last known value.

For the gradient boosting and the multilayer perceptron models, a validation set is needed to perform early stopping when validation loss stops decreasing. We use data from 2020 to 2021 as the validation set.

One can argue this choice of validation set is not optimal because models may overfit to this period, which may be anomalous. A potential solution could be to randomly sample the validation set from the entire training data from 1998 to 2021. On the other hand, the Kp index has a distribution drift, and choosing nonintersecting periods for the training and validation sets could help simulate that and ensure the model is robust. We conducted additional experiments to determine which choice of validation set (random or fixed period) is better for the task, and it turned out that the first option (random sampling) results in worse test set metrics.

We train a separate model for each horizon (using a single model for all hour types within each horizon). Additional experiments show that this setup is better than training a separate model for each hour type and horizon, as it yields approximately the same results while requiring three times fewer resources.

5. RESULTS

5.1. The Role of Delay Embedding

We first train the models without the delay embedding (i.e., using just the values of all the variables at the prediction moment) to set a baseline for all the horizons and to compare performance of the models. Table 2 shows the results for each model and each horizon number (Tables 2a, 2b, and 2c show results for horizons numbers 1–3, 4–6, and 7–8, respectively). All the results can be compared to the inertial forecast at the bottom of each table.

For the first horizon, the best RMSE is 6.312 (approximately 2 change units of Kp), and accuracy is 22.7% (compared to RMSE 9.031 and accuracy 18% for the inertial forecast). For the last horizon, the best RMSE is 12.147, and accuracy is 11.8% (compared to 15.775 and 10.5% for the inertial forecast).

Fig. 5. Feature importance for CatBoost with delay embedding 24 h for horizon 1 (3 h ahead). x-Axis represents feature, y-axis—its time lag.

It can be seen that the best model for horizons number 1–3 is the multilayer perceptron with two hidden layers (MLP2), while for the rest of the horizons, the best model is CatBoost. Figure 3 shows the comparison plot for CatBoost and perceptron.

The next step is adding the delay embedding. As mentioned earlier, we test the delay embedding depths from 0 (no delay embedding, as in the results above) to 24 h (corresponding to 8 preceding values of the Kp index and 24 preceding values of other variables) with the step of 3 h. Given that the full results listing becomes large in this case (9 different depths, 8 different horizons, 5 different models), we report only the best model with the best delay embedding depth for each of the horizons. Table 3 shows the results with comparison to the best models without using delay embedding.

We see that adding variable history decreases RMSE for all the horizons. This result is very expected due to the nature of geomagnetic processes. For the first horizon, we obtain RMSE of 5.615 and accuracy of 26.7% (compared to RMSE of 6.312 and accuracy of 22.7% for the baseline). For the eighth horizon, we achieve RMSE of 12.098 and accuracy of 12.1% (compared to RMSE of 12.147 and accuracy of 11.8% for the baseline).

5.2. Interpreting the Best Model

Another conclusion (see Table 3) is that CatBoost (with various delay embedding depths) shows the best RMSE at almost all the horizons, except for the first and sixth horizons. In this chapter we study this model in more detail.

Figure 4 shows the dependence of RMSE on the delay embedding depth for each of the horizons. We see that for the nearest horizons, a lower delay embedding depth is needed, while for the farthest horizons, the optimal depth is higher. For some of the last horizons increasing the delay embedding depth even more may promise better results.

CatBoost with the delay embedding of 24 h produces the best results, or results comparable to the best, for almost all the horizons. This is likely due to the algorithm's inherent feature selection ability. Figure 5 shows features importance for this model for the first horizon (with feature names in columns, time lags in rows and importance in values). Feature importance is calculated based on how many times a feature is used in splits. The most important features seem to be Kp itself (preceding value), B_gsm_z and its two time lags, B_magn, SW_spd and H_den_SWP.

Fig. 6. Feature importance for CatBoost with delay embedding 24 h for horizon 2 (6 h ahead). x-Axis represents feature, y-axis—its time lag.

Figure 6 shows features importance for the model for the second horizon (6 h ahead). We notice that the same variables are important as for the first horizon, but the preceding value of the Kp index becomes less important and the importance of the other variables (B_gsm_z and its two time lags, B_magn, SW_spd and H_den_SWP and the Kp index time lags) increases.

Figure 7 shows features importance for the model for the fourth horizon (12 h ahead). We see that the importance of the preceding Kp index value has become even less (contrary to its lags) and Dst, B_gsm_y become more important. B_magn becomes more important than the Kp index value itself.

From the feature importance analysis we find that the longer the horizon, the less important is the preceding Kp value and the more important is the absolute value of the magnetic field (B_magn).

Figure 8 shows how feature importance of Kp, B_magn, SW_spd, H_den_SWP changes with the horizon and illustrates the mentioned fact in a more verbose way.

Figure 9 demonstrates application of the model (CatBoost with delay embedding depth 24 h) during the period from February to April 2023, for predictions made at T_0 type hours. This period contains strong magnetic storms with Kp index up to 8 and higher.

We see that the model is capable of predicting Kp index peaks only for the three first horizons (9 h ahead maximum). For the other horizons the model does not predict the peaks or predicts the disturbances with a delay.

6. CONCLUSIONS

In this paper we investigated the potential of predicting Kp index values 24 h ahead using machine learning models. These models are based on hourly data from solar wind parameters, interplanetary magnetic field parameters, and the Dst index. We utilized linear regression, gradient boosting, and multilayer perceptrons in our analysis to predict next 8 values of Kp index. We measured performance of the models on the independent test set using R^2, accuracy of rounded predictions and RMSE.

We showed, that a trivial model, for which the predicted value of the Kp index value is equal to the last known value of the Kp index could be outperformed by ML models using just values of the variables at the prediction moment.

Fig. 7. Feature importance for CatBoost with delay embedding 24 h for horizon 4 (12 h ahead). x-Axis represents feature, y-axis—its time lag.

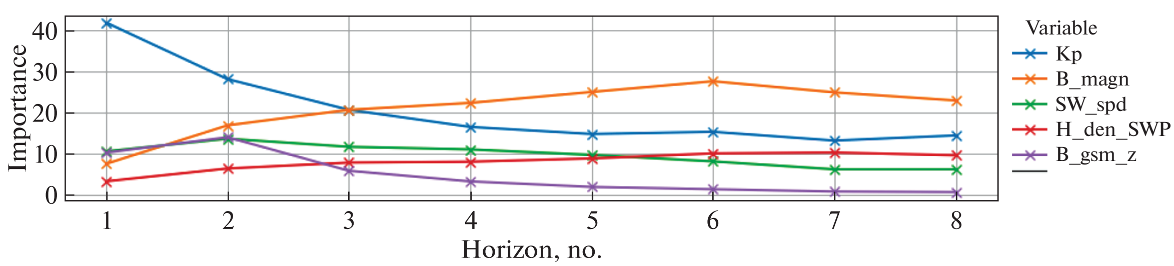

Fig. 8. Feature importance for CatBoost with delay embedding 24 h vs forecasting horizon.

We improved these baseline results by adding history of variables before the prediction moment to the training data using delay embedding technique (the current vector of variables is concatenated with the N preceding vectors). We tested different delay embedding depths—from 0 to 24 h.

The best model (almost for all horizons) appeared to be CatBoost implementation of gradient boosting over decision trees. We studied the dependence of RMSE on the delay embedding depth and found that the farther the horizon, the greater depth is needed. We found that using the depth of 24 h produces the best or comparable to the best results for all the horizons.

We explored the features that are most important for CatBoost with the delay embedding depth 24 h to make a prediction. The most important features appeared to be Kp itself (preceding value), B_gsm_z and its two time lags, B_magn, SW_spd, and H_den_SWP.

Compared to our previous research, we built regression models instead of classifying the Kp index

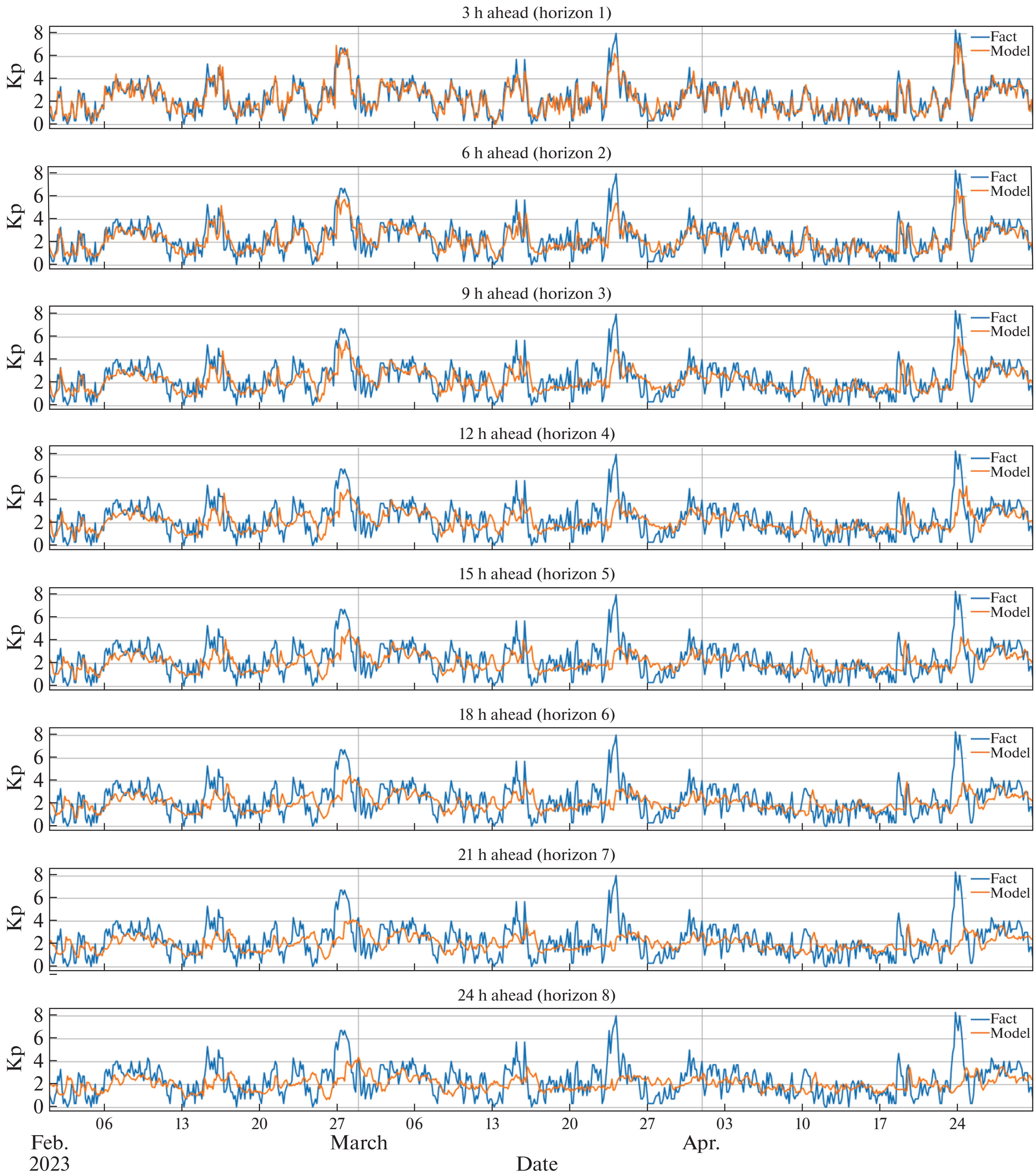

Fig. 9. Ground truth Kp index values (blue line) and CatBoost predictions at T_0 hours with delay embedding depth 24 h (orange line) during the period from February to April 2023.

into three categories (weak, medium, and strong disturbances). These regression models can be adapted for classification, but the reverse is not true, which makes the regression models more general-purpose.

Based on model interpretation, the new models appear to be more physically reasonable than the classification models, as they require the historical data of the variables to make good predictions, which aligns with theoretical expectations. One more difference from feature importance perspective is that the

Table 2. Accuracy, R^2, RMSE (Kp*10) for each model and horizon without delay embedding

(a) Horizons number 1–3

Horizon number	1			2			3		
	Accuracy	R^2	RMSE	Accuracy	R^2	RMSE	Accuracy	R^2	RMSE
CatBoost	22.9%	0.748	6.372	18.4%	0.572	8.312	15.1%	0.395	**9.88**
LGBM	22.7%	0.747	6.394	18.7%	0.569	8.335	15%	0.391	9.908
MLP1	22.5%	0.75	6.347	18.3%	0.57	8.324	15.2%	0.392	9.901
MLP2	22.7%	0.753	**6.312**	18.4%	0.573	**8.297**	14.9%	0.386	9.954
Ridge	20.2%	0.701	6.949	15.9%	0.518	8.818	13.6%	0.356	10.195
Inertial	18%	0.495	9.031	13%	0.163	11.62	12.3%	−0.023	12.848

(b) Horizons number 4–6

Horizon number	4			5			6		
	Accuracy	R^2	RMSE	Accuracy	R^2	RMSE	Accuracy	R^2	RMSE
CatBoost	14%	0.293	**10.677**	12.3%	0.222	**11.204**	11.8%	0.166	**11.6**
LGBM	13.7%	0.289	10.712	12.5%	0.215	11.254	11.8%	0.161	11.637
MLP1	13.8%	0.285	10.739	12.9%	0.21	11.287	12.1%	0.147	11.733
MLP2	14.6%	0.288	10.715	13%	0.202	11.347	12.3%	0.141	11.77
Ridge	12.8%	0.256	10.959	12.1%	0.18	11.504	11.7%	0.124	11.888
Inertial	12%	−0.162	13.693	10.8%	−0.281	14.379	10.4%	−0.384	14.943

(c) Horizons number 7–8

Horizon number	7			8		
	Accuracy	R^2	RMSE	Accuracy	R^2	RMSE
Catboost	11.6%	0.122	**11.901**	11.8%	0.085	**12.147**
LGBM	11.4%	0.114	11.958	11.6%	0.078	12.194
MLP1	11.7%	0.109	11.987	11.6%	0.075	12.215
MLP2	11.7%	0.095	12.082	12%	0.048	12.393
Ridge	11.3%	0.078	12.192	11.2%	0.046	12.404
Inertial	10.6%	−0.473	15.415	10.5%	−0.543	15.775

Table 3. The best models and delay embedding depths in hours for each horizon. The rightmost column shows RMSE for the baseline without the delay embedding (RMSE D.0)

Horizon number	Horizon, h	Best model	Accuracy	R^2	RMSE	RMSE D.0
1	1–3	MLP2 3H	0.275	0.809	5.557	6.312
2	4–6	CatBoost 6H	0.187	0.584	8.191	8.297
3	7–9	CatBoost 12H	0.151	0.409	9.764	9.88
4	10–12	CatBoost 24H	0.136	0.31	10.559	10.677
5	13–15	CatBoost 9H	0.128	0.236	11.105	11.204
6	16–18	MLP1 3H	0.121	0.177	11.519	11.6
7	19–21	CatBoost 24H	0.117	0.127	11.873	11.901
8	22–24	CatBoost 24H	0.121	0.093	12.098	12.147

classification models depend on the Dst index, but regressions models do not.

The next directions of our research include:

1. Use of synthetic Kp index with hourly frequency.

2. Exploiting convolutional and recurrent neural networks and increasing the number of layers and neurons for multilayer perceptrons.

3. Collecting more data and adding new variables, which could be potential Kp forecasting factors.

FUNDING

This study was supported by the Russian Science Foundation, project no. 23-21-00237, https://rscf.ru/en/project/23-21-00237/.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. V. B. Belakhovsky, V. A. Pilipenko, Ya. A. Sakharov, and V. N. Selivanov, Bull. Russ. Acad. Sci.: Phys. **87**, 236 (2023). https://doi.org/10.3103/S1062873822700988
2. V. A. Pilipenko, A. D. Gvishiani, A. A. Solovyev, and I. N. Rosenberg, Zemlya Vselennaya **6**, 20 (2023).
3. Q. Qiu, J. A. Fleeman, and D. R. Ball, IEEE Electrif. Mag. **3** (4), 22 (2015). https://doi.org/10.1109/mele.2015.2480615
4. D. V. Kostarev, V. A. Pilipenko, and O. V. Kozyreva, Nauka i Tekhnologii Trubo-Provodnogo Transporta Nefti i Nefteproduktov **13** (1), 352 (2023).
5. *Space Storms and Space Weather Hazards*, Ed. by I. A. Daglis, NATO Science Series II: Mathematics, Physics and Chemistry, Vol. 38 (Kluwer, Dordrecht, 2001). https://doi.org/10.1007/978-94-010-0983-6
6. D. Vassiliadis, in *Space Weather—Physics and Effects*, Springer Praxis Books (Springer, Berlin, 2007), p. 403. https://doi.org/10.1007/978-3-540-34578-7_14
7. R. M. McGranaghan, E. Camporeale, M. Georgoulis, and A. Anastasiadis, J. Space Weather Space Clim. **11**, 50 (2021). https://doi.org/10.1051/swsc/2021037
8. D. G. Cole, Space Sci. Rev. **107**, 295 (2003). https://doi.org/10.1023/A:1025500513499
9. J. Bartels, N. H. Heck, and H. F. Johnston, Terrestrial Magnetism and Atmospheric Electricity **44**, 411 (1939). https://doi.org/10.1029/TE044i004p00411
10. J. R. Bartels, IATME Bull. **12**, 97 (1949).
11. S.-I. Akasofu and S. Chapman, *Solar-Terrestrial Physics* (Clarendon Press, Oxford, 1972).
12. H. A. Elliott, J. Jahn, and D. J. Mccomas, Space Weather **11**, 339 (2013). https://doi.org/10.1002/swe.20053
13. J. Wang, Q. Zhong, S. Liu, J. Miao, F. Liu, Zh. Li, and W. Tang, Space Weather **13**, 831 (2015). https://doi.org/10.1002/2015sw001251
14. F. Boberg, P. Wintoft, and H. Lundstedt, Physics and Chemistry of the Earth, Part C: Solar, Terrestrial amp; Planetary Science **25**, 275 (2000). https://doi.org/10.1016/s1464-1917(00)00016-7
15. E.-Y. Ji, Y.-J. Moon, J. Park, J.-Y. Lee, and D.-H. Lee, J. Geophys. Res.: Space Phys. **118**, 5109 (2013). https://doi.org/10.1002/jgra.50500 https://agupubs.onlinelibrary.wiley.com/doi/pdf/-10.1002/jgra.50500.
16. Ya. Tan, Q. Hu, Zh. Wang, and Q. Zhong, Space Weather **16**, 406 (2018). https://doi.org/10.1002/2017sw001764
17. S. Wing, J. R. Johnson, J. Jen, C.-I. Meng, D. G. Sibeck, K. Bechtold, J. Freeman, K. Costello, M. Balikhin, and K. Takahashi, J. Geophys. Res.: Space Phys. **110**, A04203 (2005). https://doi.org/10.1029/2004ja010500
18. R. Bala and P. Reiff, Space Weather **10**, S06001 (2012). https://doi.org/10.1029/2012sw000779
19. R. Bala and P. Reiff, Space Weather **12**, 417 (2014). https://doi.org/10.1002/2014sw001075
20. J. R. Ayala Solares, H.-L. Wei, R. J. Boynton, S. N. Walker, and S. A. Billings, Space Weather **14**, 899 (2016). https://doi.org/10.1002/2016sw001463
21. I. M. Gadzhiev, I. V. Isaev, O. G. Barinov, S. A. Dolenko, and I. N. Myagkova, Moscow Univ. Phys. Bull. **78**, S96 (2023). https://doi.org/10.3103/s002713492307007x
22. I. M. Gadzhiev, O. G. Barinov, I. N. Myagkova, and S. A. Dolenko, Geomagn. Aeron. **64**, 415 (2024). https://doi.org/10.1134/S0016793224600140
23. Kp forecast, Swedish Space Weather Center. https://www.spaceweather.se/forecast/kp.
24. P. Wintoft, M. Wik, J. Matzka, and Yu. Shprits, J. Space Weather Space Clim. **7**, A29 (2017). https://doi.org/10.1051/swsc/2017027
25. P. Wintoft and M. Wik, Space Weather **16**, 1972 (2018). https://doi.org/10.1029/2018sw001994
26. Y. Liu, B. X. Luo, and S. Q. Liu, Manned Spaceflight **19** (2), 70 (2013).
27. S. Liu and J. Gong, Space Weather **13**, 599 (2015). https://doi.org/10.1002/2015sw001298

28. German Research Center for Geosciences in Potsdam (GFZ). https://www.gfz-potsdam.de/en/kp-index.
29. Kyoto Dst index service. https://wdc.kugi.kyoto-u.ac.jp/dstdir/.
30. The Ace Science Center. https://izw1.caltech.edu/ACE/ASC/.
31. J. H. Friedman, Ann. Stat. **29**, 1189 (2001). https://doi.org/10.1214/aos/1013203451
32. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, in *Advances in Neural Information Processing Systems, Montreal, 2018*, Ed. by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, 2018), Vol. 31, p. 6638. https://proceedings.neurips.cc/paper_files/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf.
33. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, in *Advances in Neural Information Processing Systems*, Ed. by I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, 2017), Vol. 30, p. 3146. https://proceedings.neurips.cc/paper_files/paper/2017/ file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.
34. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, in *Paralleled Distributed Processing* (Defense Technical Information Center, Cambridge, 1986), Vol. 1, p. 318. https://doi.org/10.21236/ada164453

Publisher's Note. Allerton Press, Inc. remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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