

## A methodology for imputing missing values in geomagnetic field variations using kNN and MICE algorithms

Sanjar A. Imashev, <https://orcid.org/0000-0003-3293-3764>, [sanzhar.imashev@gmail.com](mailto:sanzhar.imashev@gmail.com)

Research Station of the Russian Academy of Sciences in Bishkek, Bishkek city, Kyrgyzstan

Abstract [PDF ENG](#) [PDF RUS](#)

[Full text](#) [PDF RUS](#)

**Abstract.** The study presents methods for filling missing values in geomagnetic data based on the k-Nearest Neighbors (kNN) algorithm and Multiple Imputation by Chained Equations (MICE). The effectiveness of these algorithms was analyzed using data from the geomagnetic monitoring network of the Research Station of the RAS, focusing on two types of events: regular Sq-variations and geomagnetic storms. According to the results, the kNN algorithm demonstrates high accuracy in reconstructing regular variations with a Mean Absolute Error (MAE) of  $\leq 0.4$  nT. However, its accuracy significantly decreases during geomagnetic storms (MAE = 5.7 nT). In contrast, the MICE algorithm performs better in these challenging conditions, reducing the MAE to 1.1 nT by leveraging correlations between monitoring stations. A combined approach, utilizing kNN for preliminary imputation followed by MICE for refinement, proved effective both for filling missing values at the remote Karagai-Bulak station and for addressing impulsive outliers in the data. Additionally, it was shown that the proposed approach can be applied to analyze magnetic disturbances recorded at a nearby station caused by the operation of the ERGU-600 system. The results confirm the potential of the methods to automate the analysis of multidimensional data, which is particularly crucial when working with large volumes of geomagnetic data.

### Keywords:

**geomagnetic data, missing value imputation, machine learning, kNN algorithm, MICE algorithm, magnetic storms, outlier removal**

**For citation:** Imashev S.A. A methodology for imputing missing values in geomagnetic field variations using kNN and MICE algorithms. *Geosistemy perehodnykh zon = Geosystems of Transition Zones*, 2026, vol. 10, No. 2, Article 384. (In Russ.).

URL: <http://journal.imgg.ru/web/full/f2026-0-384.pdf> ; <https://doi.org/10.30730/gtr.2026.0.mfi-384> ;

<https://www.elibrary.ru/xkpbmf>

**Для цитирования:** Имашев С.А. Методика восстановления пропусков в вариациях геомагнитного поля на основе алгоритмов kNN и MICE. *Геосистемы переходных зон*, 2026, т. 10, № 2, 384. URL: <http://journal.imgg.ru/web/full/f2026-0-384.pdf> ; <https://doi.org/10.30730/gtr.2026.0.mfi-384> ; <https://www.elibrary.ru/xkpbmf>

## References

1. Schneider T. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate*. 2001,14(5):853-871. [https://doi.org/10.1175/1520-0442\(2001\)014<0853:aoicde>2.0.co;2](https://doi.org/10.1175/1520-0442(2001)014<0853:aoicde>2.0.co;2)
2. Little R.J.A., Rubin D.B. *Statistical analysis with missing data*. Third ed. Hoboken, NJ: Wiley, 2020, 449 p. <https://doi.org/10.1002/9781119482260>
3. Jadhav A., Pramod D., Ramanathan K. Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence*. 2019,33(10):913-933. <https://doi.org/10.1080/08839514.2019.1637138>
4. Love J.J. Missing data and the accuracy of magnetic observatory hour means. *Annals of Geophysics*. 2001,27:3601-3610. <https://doi.org/10.5194/angeo-27-3601-2009>
5. Vorobyeva G.R. Approach to the recovery of geomagnetic data by comparing daily fragments of a time series with equal geomagnetic activity. *Computer Optics*. 2019,43(6):1053-1063. (In Russ.). <https://doi.org/10.18287/2412-6179-2019-43-6-1053-1063>
6. Richman M.B., Trafalis T.B., Adrianto I. Missing data imputation through machine learning algorithms. In: Haupt S.E., Pasini A., Marzban C. (eds) *Artificial intelligence methods in the environmental sciences*. Dordrecht: Springer, 2009. [https://doi.org/10.1007/978-1-4020-9119-3\\_7](https://doi.org/10.1007/978-1-4020-9119-3_7)

7. Abidin N.Z., Ismail A.R., Emran N.A. Performance analysis of machine learning algorithms for missing value imputation. *International Journal of Advanced Computer Science and Applications*. 2018,9(6):442-447
8. Barkhatov N.A., Levitin A.E., Sakharov S.Yu. The method of artificial neuron networks as a procedure for reconstructing gaps in records of individual magnetic observatories from the data of other stations. *Geomagnetism and Aeronomy*. 2002,42(2):184-186.
9. Imashev S.A., Parov S.V. Modified seasonal decomposition variations of earth magnetic field induction module. *Information Technologies*. 2024,30(2):59-67. (In Russ.). <https://doi.org/10.17587/it.30.59-67>
10. Mukhamadeeva V.A., Vorontsova E.V., Lazareva E.A. Experience of geomagnetic observations at the geodynamic test ground in Bishkek. *Vestnik of KRSU = Herald of KRSU*. 2015,15(3):130-133. (In Russ.).
11. Imashev S.A., Lazareva E.A. Spatial distribution of the main geomagnetic field components based on IGRF-13 model for Kyrgyzstan territory. *Vestnik of KRSU = Herald of KRSU*. 2022,22(4):192-198. (In Russ.). <https://doi.org/10.36979/1694-500X-2022-22-4-192-198>
12. Imashev S.A., Rybin A.K. Seismic and geoacoustic responses of the earth's crust to sensing with high energy electric pulses at the territory of the Bishkek geodynamic polygon. *Nauka i tekhnologicheskiiye razrabotki*. 2023,102(2-3):63-88. (In Russ.). <https://doi.org/10.21455/std2023.2-3-3>
13. Imashev S.A. Extended isolation forest – Application to outlier detection in geomagnetic data. *IOP Conference Series: Earth and Environmental Science*. 2021,012022. <https://doi.org/10.1088/1755-1315/929/1/012022>
14. Imashev S.A., Lazareva E.A. Removal of outliers in geomagnetic field time series using the Hampel filter. *Information Technologies*. 2025,31(4):191-198. (In Russ.). <https://doi.org/10.17587/it.31.191-198>
15. Campbell W.H. *Introduction to geomagnetic fields*. Cambridge University Press, 2003, 337 p. <https://doi.org/10.1017/cbo9781139165136>
16. Imashev S.A. Method for detecting anomalies in geomagnetic field variations based on artificial neural network. *Geosystems of Transition Zones*. 2024,8(4):343-356. <https://doi.org/10.30730/gtr.2024.8.4.343-356>
17. Beretta L., Santaniello A. Nearest neighbor imputation algorithms: a critical evaluation. *BMC Medical Informatics and Decision Making*. 2016,16(S3):74. <https://doi.org/10.1186/s12911-016-0318-z>
18. Batista G.E.A.P.A., Monard M.C. An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*. 2003,17(5–6):519-533. <https://doi.org/10.1080/713827181>
19. White I.R., Royston P., Wood A.M. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*. 2011,30(4):377-399. <https://doi.org/10.1002/sim.4067>
20. Huque M.H., Carlin J.B., Simpson J.A., Lee K.J. A comparison of multiple imputation methods for missing data in longitudinal studies. *BMC Medical Research Methodology*. 2018,18:168. <https://doi.org/10.1186/s12874-018-0615-6>
21. Stekhoven D.J., Bühlmann P. MissForest – nonparametric missing value imputation for mixed-type data. *Bioinformatics*. 2012,28(1):112-118. <https://doi.org/10.1093/bioinformatics/btr597>
22. Hassanat A.B., Abbadi M.A., Altarawneh G.A., Alhasanat A.A. Solving the problem of the k parameter in the kNN classifier using an ensemble learning approach. *International Journal of Computer Science and Information Security*. 2014,12(8):33-39. <https://doi.org/10.48550/arXiv.1409.0919>
23. Lazareva E.A., Imashev S.A. [Variations of the full vector of the geomagnetic field during the launch of the electrical exploration generator unit (ERGU-600-2)]. *Sovremennyye tekhnika i tekhnologii v nauchnykh issledovaniyakh: sb. materialov XIII Mezhdunar. konf. molodykh uchenykh i studentov*. Bishkek, 2021, p. 107-112. (In Russ.).
24. Sorokin V.M., Yaschenko A.K., Novikov A.V., Imashev S.A., Lazareva E.A. Electromagnetic signal propagation into ionosphere from the radiating grounded dipole of the ERGU-600-2 electric prospecting generator facility (Northern Tien-Shan). *Dynamic Processes in Geospheres*. 2025,17(2):41-53. (In Russ.). [https://doi.org/10.26006/29490995\\_2025\\_17\\_2\\_41](https://doi.org/10.26006/29490995_2025_17_2_41)