

Методика восстановления пропусков в вариациях геомагнитного поля на основе алгоритмов kNN и MICE

Имашев Санжар Абылбекович, <https://orcid.org/0000-0003-3293-3764>, sanzhar.imashev@gmail.com

Научная станция РАН в г. Бишкеке, Бишкек, Киргизия

Резюме [PDF RUS](#) [PDF ENG](#) **Полный текст** [PDF RUS](#)

Резюме. В статье представлены методы восстановления пропусков в геомагнитных данных, которые основаны на алгоритмах k-ближайших соседей (k-Nearest Neighbors, kNN) и многократного заполнения пропусков методом цепных уравнений (Multiple Imputation by Chained Equations, MICE). Анализ эффективности алгоритмов проведен на данных сети геомагнитного мониторинга Научной станции РАН по двум типам событий: регулярным Sq-вариациям и геомагнитным бурям. Согласно полученным результатам, алгоритм kNN демонстрирует хорошую точность при восстановлении регулярных вариаций с показателем MAE ≤ 0.4 нТл (Mean Absolute Error – средняя абсолютная ошибка), но его точность значительно снижается в условиях магнитных бурь (MAE = 5.7 нТл). Алгоритм MICE лучше справляется с обработкой пропусков в таких сложных условиях, снижая MAE до 1.1 нТл за счет учета межстанционных корреляций. Комбинированный подход, в котором используется алгоритм kNN для предварительного восстановления пропусков, а MICE – для последующего уточнения, показал эффективность как при заполнении пропусков в данных самой удаленной станции, так и при устранении импульсных выбросов в данных. Кроме того, показано, что предложенный подход может применяться для анализа магнитных возмущений, регистрируемых на близлежащей станции и вызванных работой установки ЭРГУ-600. Полученные результаты подтверждают возможности предложенных методов машинного обучения для автоматизации анализа многомерной информации, что особенно актуально при обработке больших массивов геомагнитных данных.

Ключевые слова:

геомагнитные данные, восстановление пропусков, машинное обучение, алгоритм kNN, алгоритм MICE, магнитные бури, устранение выбросов

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>

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>

Список литературы

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. Воробьева Г.П. Подход к восстановлению геомагнитных данных путем сопоставления суточных фрагментов временного ряда с равной геомагнитной активностью. *Компьютерная оптика*. 2019,43(6):1053-1063. <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
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. Бархатов Н.А., Левитин А.Е., Сахаров С.Ю. Метод искусственных нейронных сетей как способ восстановления пробелов в записях отдельных магнитных обсерваторий по данным других станций. *Геомагнетизм и аэрономия*. 2002,42(2):195-198.
9. Имашев С.А., Паров С.В. Модифицированное сезонное разложение вариаций модуля индукции магнитного поля Земли. *Информационные технологии*. 2024,30(2):59-67. <https://doi.org/10.17587/it.30.59-67>
10. Мухамадеева В.А., Воронцова Е.В., Лазарева Е.А. Опыт проведения геомагнитных наблюдений на Бишкекском геодинамическом полигоне. *Вестник Кыргызско-Российского Славянского университета*. 2015,15(3):130-133.
11. Имашев С.А., Лазарева Е.А. Пространственное распределение составляющих главного геомагнитного поля на основе модели IGRF-13 для территории Кыргызстана. *Вестник Кыргызско-Российского Славянского университета*. 2022,22(4):192-198. <https://doi.org/10.36979/1694-500X-2022-22-4-192-198>
12. Имашев С.А., Рыбин А.К. Сейсмические и геоакустические отклики земной коры на зондирования мощными электрическими импульсами на территории Бишкекского геодинамического полигона. *Наука и технологические разработки*. 2023,102(2-3):63-88. <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. Имашев С.А., Лазарева Е.А. Удаление выбросов во временных рядах геомагнитного поля на основе фильтра Хампеля. *Информационные технологии*. 2025,31(4):191-198. <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/gtrz.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. Лазарева Е.А., Имашев С.А. Вариации полного вектора геомагнитного поля во время пусков электроразведочной генераторной установки (ЭРГУ-600-2). *Современные техника и технологии в научных исследованиях: сб. материалов XIII Междунар. конф. молодых ученых и студентов*. Бишкек, 2021, с. 107-112.
24. Сорокин В.М., Яценко А.К., Новиков В.А., Имашев С.А., Лазарева Е.А. Распространение электромагнитного сигнала в ионосферу от излучающего заземленного диполя электроразведочной генераторной установки ЭРГУ-600-2 (Северный Тянь-Шань). *Динамические процессы в геосферах*. 2025,17(2):41-53. https://doi.org/10.26006/29490995_2025_17_2_41