

ISSN 2541-8912 (Print)
ISSN 2713-2161 (Online)

ГЕОСИСТЕМЫ ПЕРЕХОДНЫХ ЗОН

GEOSYSTEMS
of Transition Zones

2024 Том 8 № 4

Решением Научно-экспертного совета Морской коллегии
при Правительстве Российской Федерации
доктор геолого-минералогических наук, доцент, член редколлегии
нашего журнала, Ренат Белалович Шакиров награжден медалью
«За достижения в морской науке».

Редколлегия и редакция журнала «Геосистемы переходных зон» поздравляют
Рената Белаловича и его коллектив с этой неординарной наградой
и желают ему дальнейших успехов в морских научных исследованиях
и экспертной работе на международных площадках.



© Автор, 2024 г.
Контент доступен по лицензии Creative Commons Attribution License 4.0 International (CC BY 4.0)



© The Author, 2024.
Content is available under Creative Commons Attribution License 4.0 International (CC BY 4.0)

УДК 004.89:004.855.5:550.385

<https://doi.org/10.30730/gtr.2024.8.4.343-356> (In English)
<https://www.elibrary.ru/fhzskv> (In English)
<http://journal.imgg.ru/web/full/f2024-4-6.pdf>

Методика обнаружения аномалий в вариациях величины геомагнитного поля на основе искусственной нейронной сети*

С. А. Имашев

E-mail: sanzhar.imashev@gmail.com

Научная станция РАН в г. Бишкеке, Бишкек, Киргизия

Резюме. Предложена методика обнаружения аномалий в геомагнитных данных, основанная на архитектуре классического автоэнкодера. В качестве обучающих данных выбраны суточные изменения величины геомагнитного поля в спокойные дни за 2020, 2021 и 2022 гг. по базовой станции Ак-Суу сети геомагнитного мониторинга Научной станции РАН в г. Бишкеке. Нейронная сеть имеет 5 скрытых слоев с общим количеством обучаемых параметров, равным $\sim 3.5 \cdot 10^6$. Обученная модель хорошо воспроизводит типичные признаки нормальных данных, тогда как в случае данных, содержащих различные аномалии, демонстрирует ухудшение качества восстановления. Это свойство автоэнкодера использовалось для разделения данных на два класса: норма и аномалия. Ошибка восстановления в виде средней абсолютной погрешности (MAE) служила мерой аномальности. В частности, значение MAE, равное 0.109, использовалось в качестве границы раздела классов. Проверка модели на тестовых данных по станции Ак-Суу за 2017, 2018 и 2019 гг. показала хорошие результаты. В частности, такие метрики бинарной классификации, как полнота (recall) и F1-мера, имели высокие значения: 0.965, 0.918 для данных 2017 г., 0.982, 0.933 для 2018 г. и 0.970, 0.935 для 2019 г. соответственно.

Ключевые слова: аномалия, геомагнитное поле, вариационные ряды, нейросеть, автоэнкодер, матрица ошибок

Method for detecting anomalies in geomagnetic field variations based on artificial neural network

Sanjar A. Imashev

E-mail: sanzhar.imashev@gmail.com

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

Abstract. The paper proposes a method for anomaly detection in geomagnetic data based on the classical autoencoder architecture. The training data consisted of daily variations in the geomagnetic field on quiet days for 2020, 2021, and 2022, collected from the Ak-Suu base station of the geomagnetic monitoring network of the Research Station of the Russian Academy of Sciences in Bishkek. The neural network has five hidden layers with a total of $\sim 3.5 \cdot 10^6$ trainable parameters. The trained model accurately reproduces typical features of normal data, whereas in the presence of anomalies it shows a decline in reconstruction quality. This property of the autoencoder was used to classify the data into two categories: normal and anomalous. The reconstruction error, measured as the Mean Absolute Error (MAE), was used as the anomaly metric. In particular, the MAE value of 0.109 was used as the threshold for class separation. Testing the model on the data from the Ak-Suu station for 2017, 2018, and 2019 demonstrated good results. Binary classification metrics such as recall and F1-score were notably high: 0.965 and 0.918 for the 2017 data, 0.982 and 0.933 for the 2018 data, and 0.970 and 0.935 for the 2019 data, respectively.

Keywords: anomaly, geomagnetic field, variational series, neural network, autoencoder, confusion matrix

* Перевод данной статьи на английский язык напечатан в журнале «Геосистемы переходных зон», 2024, т. 8, № 4, с. 343–356. <https://doi.org/10.30730/gtr.2024.8.4.343-356>; <https://www.elibrary.ru/fhzskv>

Full text of this article in English published in the “Geosystems of Transition Zones” journal, 2024, vol. 8, No. 4. pp. 343–356. <https://doi.org/10.30730/gtr.2023.8.4.343-356>; <https://www.elibrary.ru/fhzskv>

Для цитирования: Имашев С.А. Методика обнаружения аномалий в вариациях величины геомагнитного поля на основе искусственной нейронной сети [Электронный ресурс]. *Геосистемы переходных зон*, 2024, т. 8, № 4. <http://journal.imgg.ru/web/full/f2024-4-6.pdf>; <https://doi.org/10.30730/gtrz.024.8.4.343-356>

Финансирование

Работа выполнена в рамках государственного задания ФГБУН Научная станция Российской академии наук в г. Бишкеке (тема № 1021052806454-2-1.5.1).

Введение

Основной задачей режимных геомагнитных наблюдений, проводимых на Научной станции РАН в г. Бишкеке (НС РАН), является исследование взаимосвязи вариаций модуля индукции магнитного поля Земли и деформационных процессов в земной коре [1]. Изучение геомагнитного поля важно также при исследовании таких явлений космической погоды, как вспышки на Солнце и корональные выбросы массы, и проблем обеспечения точности навигационных систем, включая авиацию и морскую навигацию [2].

Как правило, аномальными считаются те наблюдения, которые существенно отличаются от остальных, при этом закономерно предположение о том, что они являются следствием влияния сторонних механизмов [3]. Обнаружение и анализ аномалий позволяют получить полезную информацию о характеристиках процесса генерации данных. Методы обнаружения аномалий применяются при выявлении сетевых атак, случаев мошенничества с кредитными картами и неисправностей оборудования, диагностике медицинских патологий и во многих других областях [4]. В последнее время нейронные сети глубокого обучения все шире применяются для обнаружения аномалий в данных [5]. Среди множества архитектур и подходов в решении подобных задач выделяются архитектуры, основанные на применении так называемых автоэнкодеров (автокодировщиков) [6]. Автоэнкодеры применяются в таких областях, как сжатие данных [7], выявление новизны в данных [8], оценка петрофизических свойств по каротажу скважин [9], идентификация состояния сна у новорожденных на основе анализа данных электроэнцефалографии [10], мо-

For citation: Imashev S.A. Method for detecting anomalies in geomagnetic field variations based on artificial neural network. *Geosistemy perehodnykh zon = Geosystems of Transition Zones*, 2024, vol. 8, No. 4, pp. 343–356. <https://doi.org/10.30730/gtrz.2024.8.4.343-356>; <https://www.elibrary.ru/fhzskv>

Funding

The work was carried out within the framework of the state task of Research Station of the Russian Academy of Sciences in Bishkek city (No. 1021052806454-2-1.5.1).

нитинг состояния лопастей ветряных турбин [11], распознавание речи и лиц [12, 13] и даже выявление необоснованных финансовых расходов депутатов [14].

Аномалии во временных рядах геомагнитного поля могут быть вызваны как природными (магнитные бури, электромагнитные помехи в виде грозových разрядов), так и техногенными (перемещение транспорта вблизи датчиков магнитометров, всплески сетевого напряжения, работа электроразведочной генераторной установки (ЭРГУ) [15] при проведении электромагнитных наблюдений и т.д.) факторами. Применение традиционных методов обнаружения аномалий, таких как статистические методы или правила на основе пороговых значений, в случае геомагнитных данных может сопровождаться трудностями, связанными с вариативностью аномалий, обусловленной многообразием источников возмущающего воздействия и степенью его интенсивности. В данной работе представлена методика выявления аномалий в вариациях величины геомагнитного поля на основе архитектуры классического автоэнкодера.

Автоэнкодер – эффективный инструмент, способный автоматически выявлять сложные паттерны и структуры в данных [8]. Например, автоэнкодеры применялись для выделения аномальных участков в электрокардиограммах, содержащих искажения в виде дрейфа базовой линии и высокочастотного шума [16], что в конечном итоге позволяет улучшить диагностику патологий сердца. В статье [17] показаны преимущества автоэнкодеров по сравнению с алгоритмами на основе метода анализа главных компонент для задач обнаружения аномалий в телеметрических данных космических аппаратов.

Данные

В настоящее время основу сети геомагнитного мониторинга НС РАН составляют 7 стационарных пунктов (рис. 1). Дискретизация измерений на этих станциях составляет 20 с.

Временные вариации геомагнитного поля включают в себя: а) долгопериодные вариации (тренды) вследствие изменений, связанных с такими процессами, как дрейф магнитных полюсов или изменения в ядре Земли, б) суточные вариации в величине и направлении геомагнитного поля, вызванные взаимодействием солнечного ветра с магнитосферой Земли [2]. Вариации, связанные с географическим ме-

стоположением, а также локальные аномалии, обусловленные местным изменением геомагнитного поля вследствие наличия магнитных минералов в земной коре, мы рассматривать не будем, так как, во-первых, исследуемые аномалии будут анализироваться в рамках одной станции наблюдения, а во-вторых, они будут аддитивной постоянной в значении полного вектора индукции магнитного поля Земли.

В качестве примера временного ряда вариаций на рис. 2 представлен годовой ход величины геомагнитного поля для базовой станции Ак-Суу за 2021 г.

Из рис. 2 видно, что данные содержат как долгосрочный тренд вековых вариаций, так и



Рис. 1. Карта расположения стационарных пунктов сети геомагнитных наблюдений НС РАН: 1 – Ак-Суу, 2 – Шавай, 3 – Чункурчак, 4 – Таш-Башат, 5 – Иссyk-Ата, 6 – Кеgetы, 7 – Карагаy-Булак.

Fig. 1. Location map of the permanent stations of the geomagnetic observation network of the RS RAS: 1, Ak-Suu; 2, Shavay; 3, Chunkurchak; 4, Tash-Bashat; 5, Issyk-Ata; 6, Kegety; 7, Karagay-Bulak.

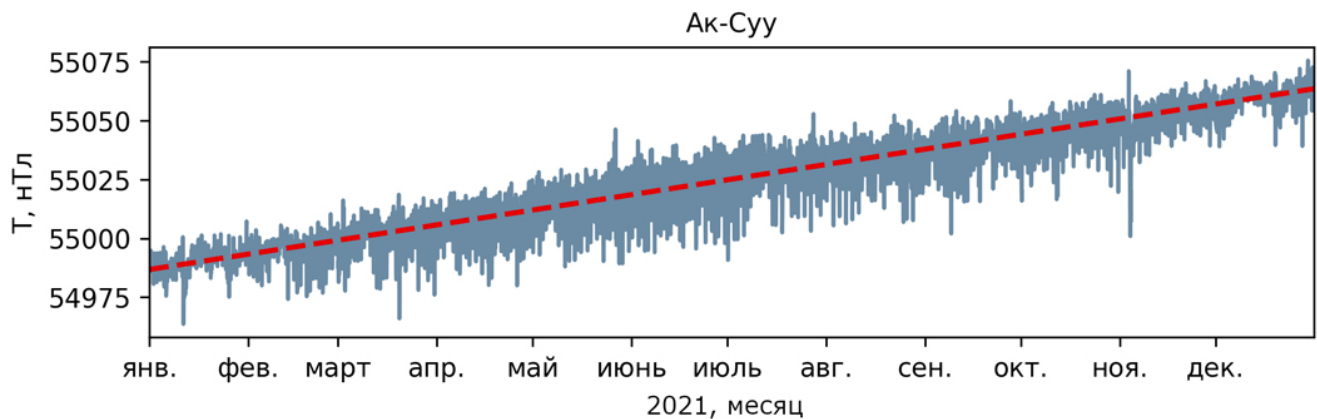


Рис. 2. Временной ряд вариаций величины геомагнитного поля на станции Ак-Суу за 2021 г. (штриховая линия – вековой тренд).

Fig. 2. Time series of the geomagnetic field variations for the Ak-Suu station for 2021 (dashed line indicates secular trend).

сезонные вариации суточного размаха (разность между максимальным и минимальным значением за сутки), возникающие из-за изменений в интенсивности солнечного освещения. Различия в ионосферных условиях в летнее и зимнее время отражаются в геомагнитном поле в размахе суточных вариаций. Для примера на рис. 3 приведены типичные суточные вариации (в отсутствие геомагнитных аномалий) на станции Ак-Суу в различные периоды года (весна, лето, осень) после вычитания векового тренда.

Классическим подходом в выявлении аномалий может быть сравнение анализируемой величины с неким эталоном, при этом степень (величина) их различия может служить мерой аномальности данных. Однако широкий диапазон суточных вариаций геомагнитных данных в магнитоспокойные дни для различных сезонов не дает возможности сконструировать единый эталонный (модельный) суточный профиль для сравнения с наблюдениями, сделанными в течение года. Это накладывает на методику выявления аномалий определенные ограничения, связанные с формализацией признаков, позволяющих отличить нормальные данные от аномальных. Также возникают трудности в разработке строгого алгоритма, реализующего это сравнение.

Методы искусственного интеллекта предоставляют нам мощные инструменты, которые превосходят классические алгоритмы благодаря своей способности обрабатывать

большие объемы данных, выявлять сложные паттерны, адаптироваться к новым данным [6, 18]. Это делает их особенно эффективными для таких сложных задач, как выявление аномалий геомагнитного поля. В частности, с этой целью можно использовать архитектуру автоэнкодеров и провести обучение на данных, которые считаются нормальными и тщательно подобраны экспертом. В таком случае при использовании нейронной сети сильная вариативность нормальных данных уже не будет представлять собой ограничение, так как нам не надо будет алгоритмизировать метод выявления аномалий в явном виде.

Наиболее значимыми по интенсивности видами геомагнитных аномалий являются магнитные бури различной интенсивности, которые могут продолжаться от нескольких часов до нескольких дней [2]. В связи с этим в данной работе за единицу входных данных принимаются вариации величины геомагнитного поля за сутки (наподобие приведенных на рис. 3) и, соответственно, оценка аномальности будет применяться интегрально ко временному ряду длительностью в 24 ч.

Также необходимо отметить, что к обучающим данным применялась процедура z-score нормализации (стандартизация), которая заключается в том, что из суточных вариаций вычитается их среднее и делится на стандартное отклонение, т.е. данные приводятся к стандартному нормальному распределению со средним значением, равным 0, и стандарт-

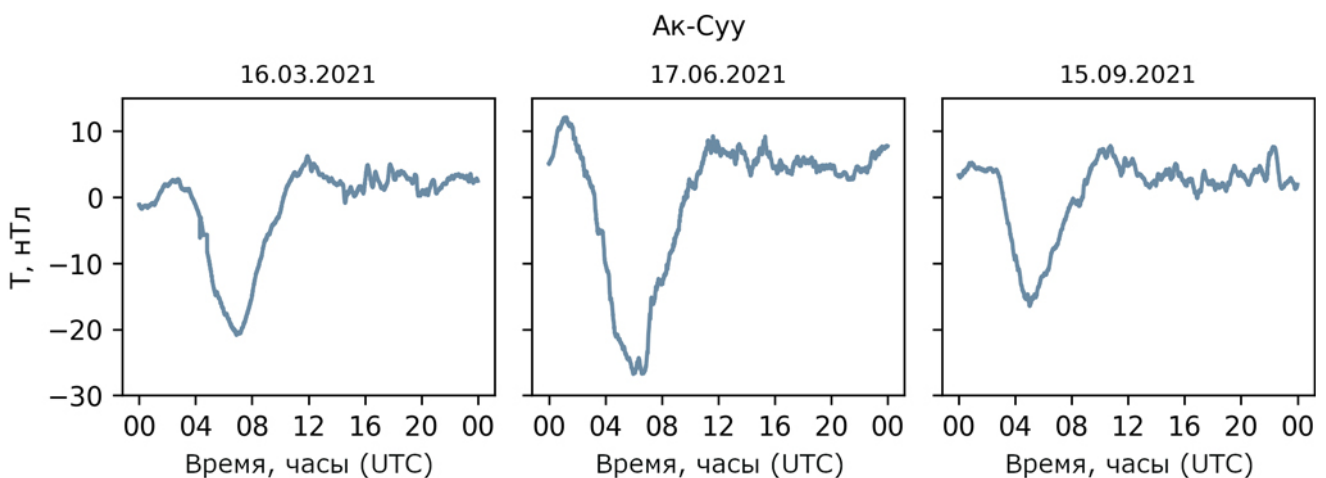


Рис. 3. Типичные суточные вариации величины геомагнитного поля для различных периодов года на станции Ак-Суу.

Fig. 3. Typical daily variations of the geomagnetic field for different periods of the year at the Ak-Suu station.

ным отклонением, равным 1. Нормализация данных часто используется при обучении нейронных сетей и является одним из ключевых этапов в предварительной подготовке входных данных. В частности, когда входные данные имеют одинаковый масштаб, тогда градиенты, используемые для обновления весов нейронной сети, становятся более стабильными. Нормализованные данные также позволяют нейросети быстрее сходиться к оптимальному решению, так как все параметры будут обновляться более согласованно. В целом можно сказать, что нормализация способствует более стабильному и быстрому обучению, улучшает сходимость алгоритмов оптимизации и снижает риск переобучения [18].

Для обеспечения производительности алгоритма было решено также уменьшить размерность сигнала за счет уменьшения частоты дискретизации (*downsampling*) до одного измерения в минуту. При этом общее количество отсчетов в суточном профиле составляет $N = 24 \cdot 60 = 1440$. При этом в качестве агрегирующей функции была использована медиана, что дает робастную оценку суточного профиля с минимальным влиянием различных выбросов, которые могут возникать в геомагнитных данных, например, во время грозовой деятельности [19].

На рис. 4 приведены примеры нормированных суточных профилей $T_{\text{норм}}$ на станции Ак-Суу: данные нормальные и содержащие аномалии. Все данные, подаваемые на вход модели, имеют безразмерную величину.

Далее был сформирован набор данных по станции Ак-Суу за 3 года (2020, 2021, 2022), в который вошли 316 примеров нормальных данных для обучающей выборки и 780 примеров аномальных данных, которые были впоследствии использованы для оценки границы разделения классов норма/аномалия.

Для проверки качества работы нейросети использовались данные по станции Ак-Суу за 3 года (2017, 2018 и 2019), которые были предварительно разделены экспертом на два класса: нормальные и аномальные.

Выбор станции Ак-Суу в качестве источника данных обусловлен тем, что эта станция является базовой и меньше всего подвержена влиянию различных техногенных помех, что немаловажно для отбора примеров нормальных дней для задачи обучения.

Методика

В данной работе использована архитектура классического автоэнкодера, который состоит из двух основных блоков: энкодера и декодера [6] (рис. 5).

Энкодер представляет собой нейронную сеть, которая сжимает входные данные в скрытое представление меньшей размерности, называемое латентным пространством. Декодер предназначен для восстановления данных из скрытого представления обратно в исходное пространство.

Отличительная особенность автоэнкодеров по сравнению с другими архитектурами искусственных нейронных сетей заключа-

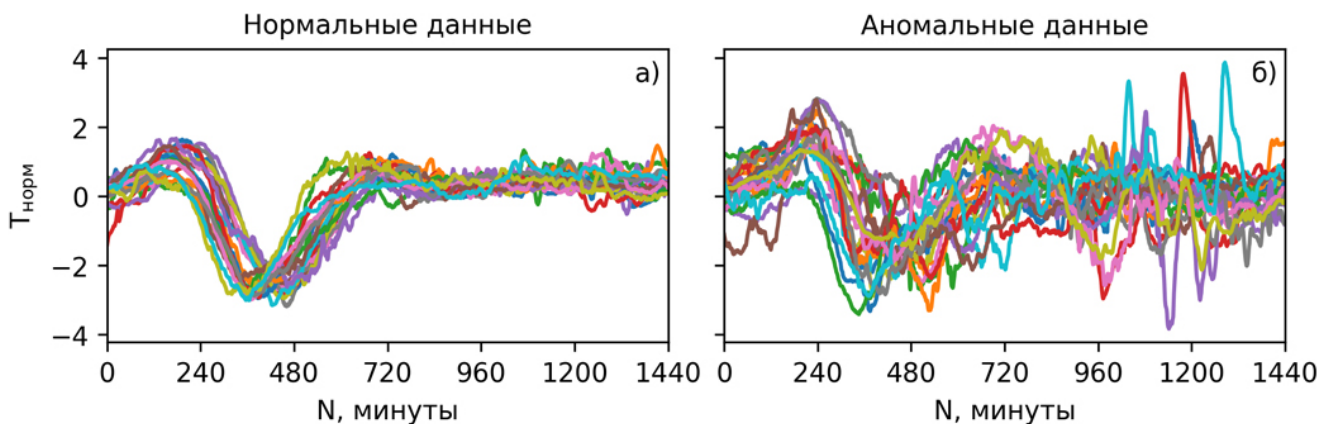


Рис. 4. Примеры нормированных суточных вариаций величины геомагнитного поля: нормальные (а) и аномальные (б) данные
Fig. 4. Examples of normalized daily variations of the geomagnetic field: normal (a) and anomalous (b) data.

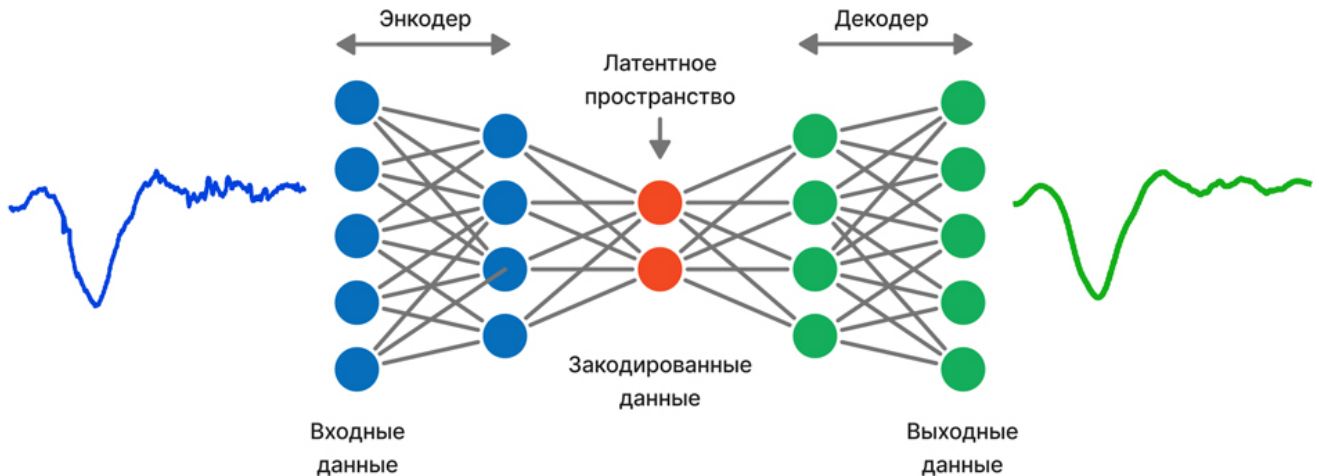


Рис. 5. Схема архитектуры классического автоэнкодера и его основные составляющие. Входные и выходные данные – суточные вариации величины геомагнитного поля.

Fig. 5. Schematic diagram of the classical autoencoder architecture and its main components. Input and output data are daily variations of the geomagnetic field.

ется в том, что количество нейронов на входе и на выходе, как правило, совпадает. При этом основной задачей автоэнкодера является получение на выходном слое результата, наиболее близкого к входному. Однако простое копирование входных данных на выход было бы бесполезным, если бы не было скрытого представления с определенными полезными свойствами. Это достигается созданием ограничений для задачи копирования, когда латентное пространство имеет размеры меньше, чем исходное входное. Такое автокодирование, называемое неполным, вынуждает автоэнкодер сжимать данные и улавливать основные структуры и паттерны нормальных данных, игнорируя при этом незначительные шумы. В ходе обучения модель будет стараться воспроизводить характерные признаки обучающих данных и далее, сталкиваясь с неизвестными аномалиями, будет демонстрировать ухудшение в качестве восстановления.

Алгоритм

Общий принцип методики выявления аномалий на основе архитектуры автоэнкодера построен на двух процедурах: обучение и расчет ошибки восстановления. Сначала автоэнкодер обучается на данных, которые считаются нормальными, минимизируя разницу между входными и восстановленными данными. Далее, для каждого входного наблюдения из полной

выборки вычисляется ошибка восстановления (разница между исходными данными и восстановленными), при этом высокие значения ошибки будут указывать на наличие возможных аномалий.

Исходя из вышесказанного, алгоритм работы автоэнкодера можно представить в следующем виде:

Этап 1. Сбор и подготовка данных. Из общей выборки, для обеспечения необходимого количества данных, экспертом отбираются нормальные (не содержащие каких-либо существенных аномалий) измерения за определенный продолжительный период. Эти данные проходят процедуру нормализации, так как имеют в своем значении вклад векового тренда и сезонных вариаций. В целом, нормализация обучающих данных является стандартной процедурой в задачах машинного обучения. Как упоминалось ранее, это позволяет стабилизировать и ускорить процесс обучения, способствует борьбе с проблемой затухания градиентов (vanishing gradient problem), а также улучшает обобщающую способность модели [18].

Этап 2. Обучение модели. Модель обучается на примерах из множества нормальных данных, при этом минимизируется функция потерь (loss function), которая в нашем случае была выбрана в виде среднеквадратичной ошибки MSE (mean squared error).

Этап 3. Вычисление ошибок восстановления. Полная выборка данных за выбранный период будет содержать два набора: данные из обучающей выборки (нормальные данные) и данные, которые являются аномальными. Для каждого наблюдения из этих двух выборок вычисляется ошибка восстановления в виде среднего квадрата разности между элементами входных и восстановленных данных. При этом распределение значений ошибки восстановления для указанных двух наборов данных будет различаться: для нормальных ошибка в среднем будет меньше, чем для аномальных.

Этап 4. Установка порога. Исходя из гистограмм распределения ошибок восстановления, для нормальных и аномальных данных устанавливается порог ошибки восстановления, выше которого данные будут считаться аномальными.

Этап 5. Выявление аномалий. На вход обученной модели подаются новые данные, и образцы с ошибкой восстановления выше порогового значения будут помечены как аномалии.

Гиперпараметры

Необходимо добавить, что на этапе обучения также происходит настройка гиперпараметров (параметров, которые настраиваются вручную перед началом обучения и определяют структуру, или способ обучения модели). К таким гиперпараметрам относятся: скорость обучения (learning rate), размер батча (batch size), количество эпох (epochs), функции активации. При этом корректируется изначальная архитектура сети в целом, т.е. размеры входного и выходного слоев, а также скрытых слоев, включая слой латентного состояния. Батч представляет собой небольшую подвыборку из полного обучающего набора, используемую для одного шага обновления параметров модели в процессе обучения. Малые размеры батча улучшают обобщающую способность модели за счет регуляризации и требуют меньших вычислительных ресурсов [18]. Эпоха представляет собой один полный проход через весь обучающий набор данных, т.е. в процессе одной эпохи модель обучается на всех примерах из обучающей выборки.

В качестве функции активации использовалась ректификационная функция ReLU (Rectified Linear Unit) $R(x)$ (рис. 6), которая преобразует отрицательные значения входа x в нули, оставляя положительные значения без изменений [20]. Исследования показали, что сверточные нейронные сети обучаются значительно быстрее при использовании ReLU в качестве функции активации [21]. Это объясняется несколькими основными преимуществами: 1) простота реализации ReLU ускоряет вычисления, что напрямую повышает скорость обучения нейронной сети; 2) форма ReLU помогает избежать проблемы затухающих градиентов, которая может возникать при использовании функций активации, таких как сигмоида, или гиперболический тангенс; 3) использование ReLU приводит к разреженным активациям, при которых часть нейронов на выходе дает нулевые значения, что улучшает обобщающую способность модели и снижает риск переобучения.

Финальные значения основных гиперпараметров были следующими: скорость обучения – 10^{-5} , количество батчей 8 и количество эпох обучения 2000.

Также необходимо упомянуть о таком методе регуляризации, как дропаут (dropout). Эта техника широко используется в нейронных сетях для предотвращения переобучения (overfitting) и представляет собой случайное «выключение» некоторых нейронов во время

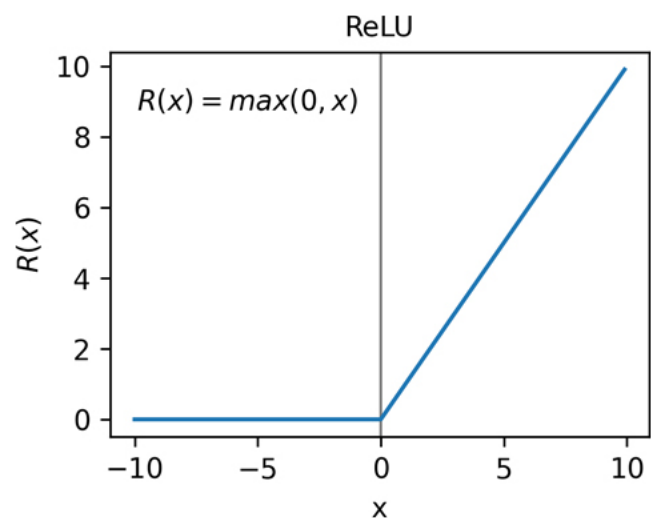


Рис. 6. Функция активации ReLU.

Fig. 6. ReLU activation function.

обучения. Это позволяет нейронной сети быть более устойчивой и лучше проявлять обобщающие свойства, за счет предотвращения чрезмерного приспособления (запоминания) к обучающим данным. Коэффициент дропаут обычно находится в диапазоне от 0.1 до 0.4, это означает, что 10–40 % нейронов выключаются на каждой итерации. Следует уточнить, что во время тестирования или предсказания дропаут не используется [18].

Архитектура модели

В таблице приведены основные параметры архитектуры классического автоэнкодера (рис. 5), используемого для задачи выявления геомагнитных аномалий. Входной слой состоит из 1440 нейронов. Энкодер включает 2 скрытых слоя с функциями активации ReLU и состоит из нейронов, количество которых последовательно уменьшается (1024 и 256). Латентное представление (bottleneck) состоит из 64 нейронов с активацией ReLU и выключением 10 % нейронов (dropout 0.1). Декодер состоит из 2 скрытых слоев с функциями активации ReLU. Количество нейронов в декодере симметрично относительно латентного представления и равно, соответственно, 256 и 1024. Выходной слой имеет 1440 нейронов, чтобы соответствовать размеру входного сигнала.

Общее количество обучаемых параметров составило 3.5 млн.

Обучение и валидация

Обучение нейронной сети происходит в результате итерационного изменения значений ее параметров. При этом основной целью такого подбора параметров является достижение минимального отклонения между сигналом на выходе нейронной сети и входным сигналом из обучающего набора. Это отклонение, как упоминалось ранее, было рассчитано с помощью функции потерь в виде среднеквадратичной ошибки MSE.

Особое внимание необходимо уделить важности валидационной выборки в машинном обучении, особенно в контексте нейронных сетей глубокого обучения. Основная цель включения валидационной выборки в процесс обучения состоит в оценке производительности модели в ходе обучения и оптимальной настройке гиперпараметров [18]. Оценка работы модели на валидационной выборке позволяет также отслеживать переобучение. Например, если качество работы нейросети на валидационной выборке начинает ухудшаться после достижения наилучших результатов на обучающей выборке, это может свидетельствовать о начале переобучения. Мы использовали эту особенность для получения оптимальной модели, сохраняя ее состояние на итерации с минимальным значением функции потерь на валидационной выборке. Размер валидационной выборки составлял 20 % от объема обучающего набора.

Таблица. Архитектура автоэнкодера
Table. Autoencoder architecture

Тип слоя	Название	Кол-во нейронов	Функция активации / регуляризация
<i>Энкодер</i>			
Входной	Input	1440	ReLU
Скрытый 1	Encoder 1	1024	ReLU
Скрытый 2	Encoder 2	256	ReLU
<i>Латентное пространство</i>			
Скрытый 3	Bottleneck	64	ReLU / Dropout (0.1)
<i>Декодер</i>			
Скрытый 4	Decoder 1	256	ReLU
Скрытый 5	Decoder 1	1024	ReLU
Выходной	Output	1440	

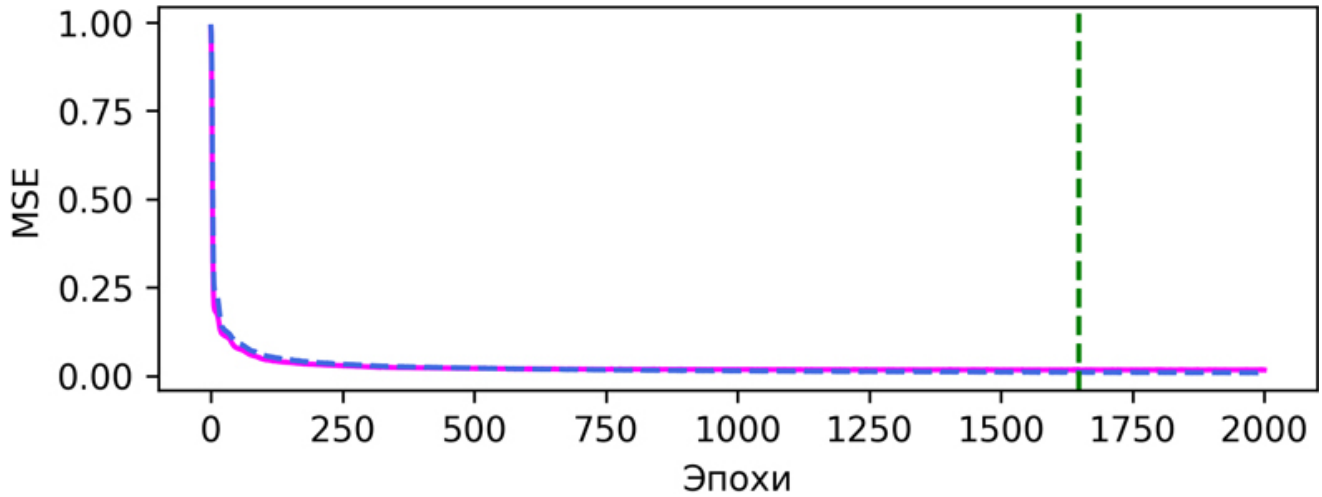


Рис. 7. Среднеквадратичная ошибка MSE для обучающей и валидационной выборок. Штриховая линия – обучение, сплошная – валидация.

Fig. 7. Mean squared error (MSE) for training and validation sets. Dashed line indicates training, solid line indicates validation.

В качестве алгоритма оптимизации, ответственного за обновление параметров модели в процессе обучения с целью минимизации функции потерь, был выбран Adam (adaptive moment estimation) [22]. Применение этого оптимизатора позволяет в большинстве случаев избежать тщательной настройки скорости обучения, так как алгоритм автоматически адаптирует ее для каждого параметра на основе истории градиентов. Как упоминалось ранее, в процессе обучения сохранялось минимальное значение функции потерь для валидационной выборки и состояние модели на этой эпо-

хе, чтобы затем использовать ее как рабочую версию (рис. 7). Глобальный минимум функции потерь для валидационной выборки был достигнут на 1648-й эпохе.

Критерии выявления аномалий

На рис. 8 представлены примеры исходного и восстановленного моделью суточных профилей $T_{\text{норм}}$ для нормальных (рис. 8 а) и аномальных (рис. 8 б) данных, а также их ошибки восстановления (рис. 8 в, г), рассчитанные в виде абсолютной погрешности AE (absolute error).

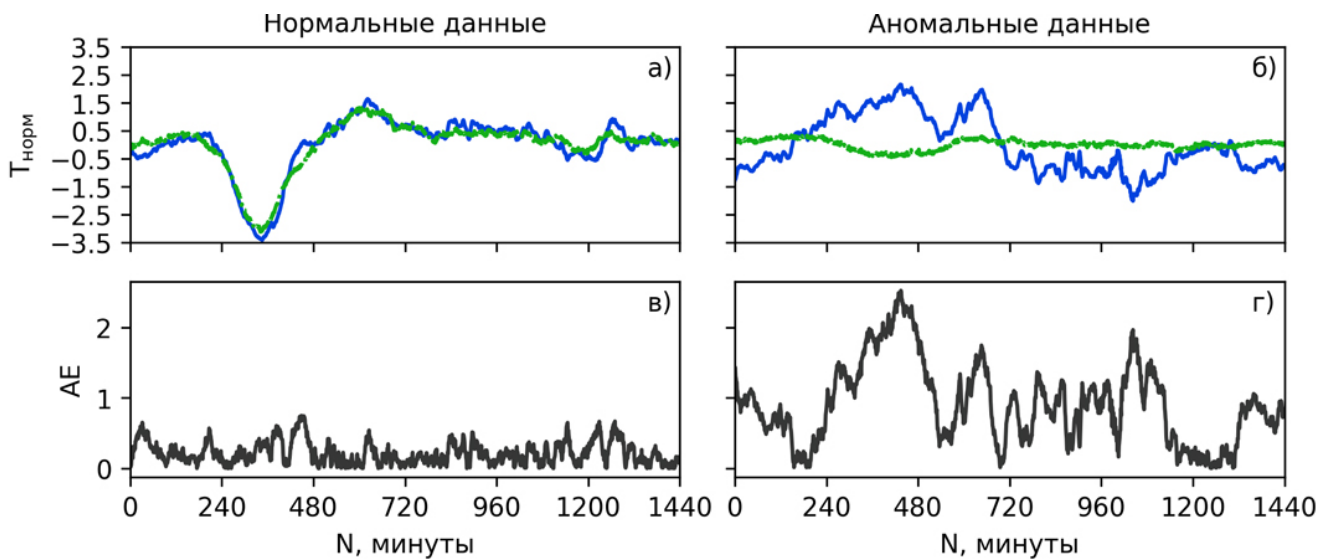


Рис. 8. Исходный (синяя линия) и восстановленный моделью (зеленая линия) суточные профили, и их ошибки восстановления в виде абсолютной погрешности AE для нормальных и аномальных данных.

Fig. 8. Initial (blue line) and reconstructed (green line) daily profiles and their reconstruction errors calculated as the absolute error (AE) for normal and anomalous data.

Ранее упоминалось, что обучение проходило на нормальных данных, следовательно, когда на вход модели подается пример нормальных значений (рис. 8 а), ошибка восстановления будет минимальной (рис. 8 в). Когда же на вход модели подаются данные, содержащие значительные аномалии (рис. 8 б), модель не сможет восстановить их (вернее, попытается сделать из аномальных данных нормальные), что приведет к большой ошибке восстановления (рис. 8 г).

Следующим шагом является определение порога меры аномальности на основе ошибки восстановления. Для этого через обученную модель были пропущены все данные за период 2020–2022 гг. (316 примеров нормальных данных и 780 аномальных). Анализ распределения ошибок восстановления на основе среднеквадратичного отклонения показал, что в этом случае распределения для нормальных и аномальных данных сильно перекрывают друг друга, что затрудняет выделение границы раздела классов. Поэтому для этого этапа было решено ошибку восстановления представить в виде средней абсолютной погрешности MAE (mean absolute error) (рис. 9).

Из рис. 9 видно, что обучающая выборка (данные, не содержащие аномалий) имеет распределение ошибки восстановления, близкое по форме к гауссову распределению, что дает основание задать порог разделения в единицах стандартных отклонений. Данные, содержащие аномалии, как и следовало ожидать, име-

ют ошибки восстановления в несколько раз больше по сравнению с нормальными значениями. Исходя из анализа распределений, можно в качестве границы выбрать $\mu + \sigma$ или $\mu + 2\sigma$ (μ – среднее значение, σ – среднеквадратичное отклонение). Так как на практике существуют примеры, когда эксперт не может однозначно отнести анализируемое наблюдение к одному из двух классов, окончательный вариант порога необходимо выбирать на основе того, что мы хотим выделить в первую очередь. Если основной интерес представляют нормальные данные, необходимо увеличить порог, и, наоборот, если основной задачей является выделение как можно большего количества аномальных наблюдений, то порог необходимо снижать. Также необходимо отметить, что граница пересечения этих двух классов занимает довольно узкую область (0.05–0.1), тогда как большинство аномалий имеет ошибку восстановления выше 0.1.

Результаты и обсуждение

Для проверки качества работы обученной нейронной сети было решено проанализировать данные за 2017, 2018 и 2019 гг., которые не участвовали в обучении модели. Когда данные нужно отнести к одному из двух классов, это является задачей бинарной классификации. В таком случае для оценки эффективности классификатора чаще всего используется так называемая матрица несоответствий или ошибок (confusion matrix).

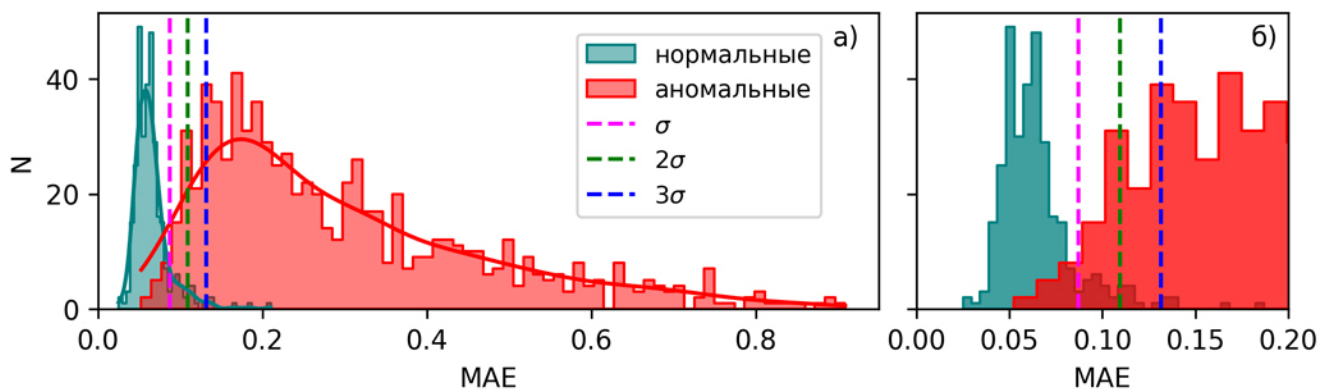


Рис. 9. Распределение ошибок восстановления (MAE) для нормальных и аномальных данных: а) общий масштаб, б) увеличенный. Штриховыми линиями показаны $\mu + \sigma$, $\mu + 2\sigma$ и $\mu + 3\sigma$ (μ – среднее значение, σ – среднеквадратичное отклонение), рассчитанные по нормальным данным.

Fig. 9. Distribution of reconstruction errors (MAE) for normal and anomalous data: a) general scale, b) enlarged scale. Dashed lines indicate $\mu + \sigma$, $\mu + 2\sigma$, and $\mu + 3\sigma$ (μ , mean value; σ , standard deviation) calculated from normal data.

Для описания комбинаций, которые могут получаться при сопоставлении ответов модели и истинных меток наблюдений, установленных экспертом, используются следующие понятия (элементы матрицы ошибок).

TP – истинно положительный объект (true positive). Данные содержат аномалии, и модель правильно идентифицирует их как аномальные.

FP – ложноположительный объект (false positive). Данные являются нормальными, а модель ошибочно идентифицирует их как аномальные.

TN – истинно отрицательный объект (true negative). Данные являются нормальными, и модель правильно идентифицирует их как нормальные.

FN – ложноотрицательный объект (false negative). Данные содержат аномалии, а модель ошибочно идентифицирует их как нормальные.

Для оценки качества работы классификатора применяются различные метрики, основные из которых: accuracy (точность – доля правильных случаев классификации), precision (точность – доля верно классифицированных объектов среди всех объектов, которые к этому классу отнес классификатор, т.е. доля настоящих аномалий среди всех наблюдений, которые модель отнесла к аномали-

ям), recall (полнота – отношение верно классифицированных объектов класса к общему числу элементов этого класса, т.е. доля верно идентифицированных аномальных данных среди всех примеров аномалий) и F1-score (F1-мера – среднее гармоническое между precision и recall). Эти метрики вычисляются по следующим формулам:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN};$$

$$Precision = \frac{TP}{TP + FP};$$

$$Recall = \frac{TP}{TP + FN};$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

На рис. 10 приведены примеры такой матрицы для данных 2019 г. Если в качестве порога выбрать значение 0.087 ($\mu + \sigma$) (рис. 10, слева), то для данных 2019 г. метрики бинарной классификации имеют следующие значения: accuracy – 0.841, precision – 0.827, recall – 0.996 и F1 – 0.904, что можно считать хорошим результатом. При пороговом значении 0.109 ($\mu + 2\sigma$) метрики немного улучшаются (исходя из интегральной метрики F1): accuracy – 0.899, precision – 0.901, recall – 0.970 и F1 – 0.935. Это является следствием того, что меньшее ко-



Рис. 10. Матрица ошибок бинарного классификатора для данных по станции Ак-Суу за 2019 г. Слева – порог 0.087 ($\mu + \sigma$), справа – порог 0.109 ($\mu + 2\sigma$).

Fig. 10. The binary classifier error matrix for the Ak-Suu station data for 2019. The threshold of 0.087 ($\mu + \sigma$) is on the left, and the threshold of 0.109 ($\mu + 2\sigma$) is on the right.

личество наблюдений попало в группу FP, поскольку перешло в группу TN (рис. 10, справа).

Как видим, существует определенный компромисс между значениями в группах FP и FN, что напрямую связано с выбранным пороговым значением, используемым для классификации. Если пороговое значение понижается, модель будет более склонна классифицировать примеры как аномалии, что может увеличить количество FP и уменьшить FN. И наоборот, при повышении порогового значения модель более склонна классифицировать примеры как нормальные, что может уменьшить количество FP и увеличить FN. Поэтому зачастую в качестве решающей метрики используется F1-мера, которая дает оценку баланса между метриками FN и FP. Выбор порога, который максимизирует F1-меру, дает компромисс между точностью (precision) и полнотой (recall), обеспечивая оптимальные результаты работы модели для конкретной задачи [23]. Поэтому для дальнейшего анализа выберем скорректированный порог со значением 0.109.

Для полноты анализа построим примеры и их восстановленные аналоги для каждой группы из элементов матрицы ошибок: TN, FN, FP и TP (рис. 11). Для группы TN (истинные нормальные данные) восстановленный аналог практически повторяет исходное наблюдение (рис. 11 а), что дает малую ошибку восстановления. Такая же картина наблюдается для группы FN (аномалия, ошибочно помеченная как норма): восстановленный аналог

точно воспроизводит исходное наблюдение (рис. 11 б). Этот пример был отмечен экспертом как содержащий аномалии вследствие наличия небольшой неоднородности в районе 480-й минуты. Для группы FP (нормальные данные, ошибочно помеченные как аномальные) восстановленный аналог сгладил высокочастотные флуктуации поля на конце суточного профиля (рис. 11 в) (960–1440-й отсчеты), что увеличило значение ошибки восстановления до 0.139, так что оно стало немного выше порогового значения 0.109. Это дало основание алгоритму пометить данное наблюдение как аномальное, несмотря на то что эксперт изначально посчитал его нормой. Для группы TP (истинные аномальные данные) восстановленный аналог сильно отличается от исходного наблюдения (рис. 11 г), он содержит сильные флуктуации начиная с середины сигнала, что отразилось в увеличении ошибки восстановления. Из графиков также видно, что с ростом амплитуд флуктуаций от нормальных (TN) до аномальных (TP), которые в основном связаны с магнитными бурями, растет и ошибка восстановления MAE. Так, максимальные суточные значения планетарных индексов, характеризующих глобальную возмущенность магнитного поля Земли – K_p и a_p , существенно различаются в случае нормальных (рис. 11 а) и аномальных (рис. 11 г) данных: $K_{p(TN)} = 1.667$, $a_{p(TN)} = 6$ нТл и $K_{p(TP)} = 4$, $a_{p(TP)} = 27$ нТл соответственно.

Средняя ошибка восстановления для рассмотренных групп имеет следующее значение:

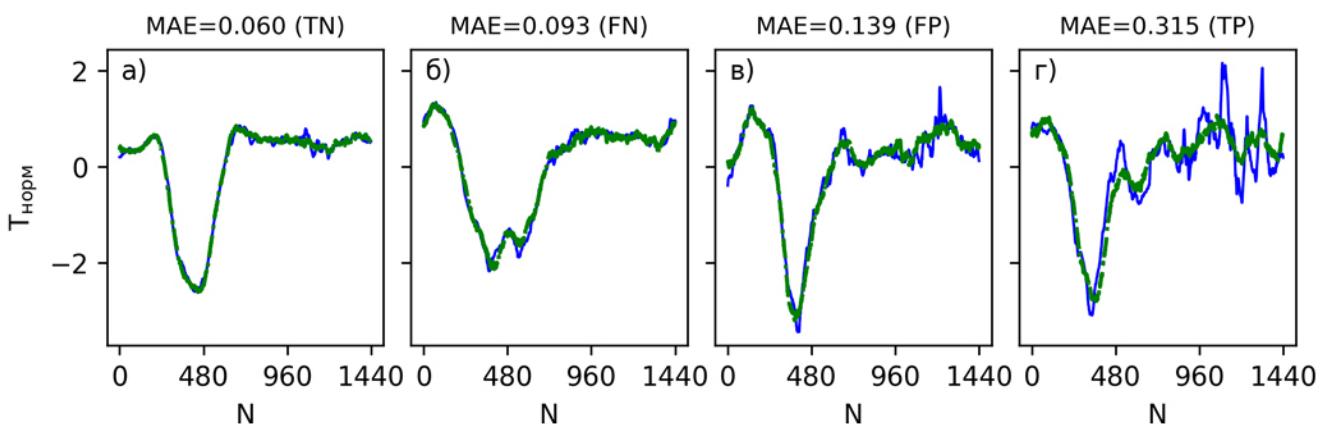


Рис. 11. Примеры из групп TN, FN, FP, TP и их восстановленные аналоги. Приведены значения ошибок восстановления (MAE). Сплошная линия – исходный пример, штрихпунктирная – восстановленный.

Fig. 11. Examples from the TN, FN, FP, and TP groups and their reconstructed analogs. The values of the reconstruction errors (MAE) are provided. The solid line indicates the original example, the dashed line indicates the reconstructed example.

$MAE_{TN} = 0.084$, $MAE_{FN} = 0.097$, $MAE_{FP} = 0.133$ и $MAE_{TP} = 0.364$.

Исходя из этого, можно сделать вывод, что модель хорошо выделяет наблюдения со значительными аномалиями, тогда как в случае небольших аномальных участков она может ошибаться. Однако общая оценка выявления аномальных данных довольно высокая, так как и полнота и F1-мера показывают высокие значения ($recall = 0.970$ и $F1 = 0.935$).

Аналогичный анализ для 2018 и 2017 гг. дал следующие показатели, соответственно:

метрики, рассчитанные по матрице ошибок: $accuracy - 0.896$ и 0.866 , $precision - 0.890$ и 0.876 , $recall - 0.982$ и 0.965 , $F1 - 0.933$ и 0.918 ;

средние ошибки восстановления для групп: $MAE_{TN} - 0.086$ и 0.083 , $MAE_{FN} - 0.091$ и 0.097 , $MAE_{FP} - 0.128$ и 0.151 , $MAE_{TP} - 0.354$ и 0.404 .

Таким образом, проверка работы модели на тестовых данных за 2017, 2018 и 2019 гг., значения которых она не видела в процессе обучения, показала хорошие результаты. В частности, высокие значения F1-меры дают основание считать, что модель достаточно хорошо выделяет аномалии на уровне качества, сопоставимого с работой эксперта.

Также необходимо отметить, что модель, обученная по данным станции Ак-Суу, подходит и для выявления аномалий в данных близлежащей станции Шавай. В силу относительно малого (~17 км) расстояния между станциями наклон векового тренда, размах и

форма суточных вариаций, а также проявление магнитных бурь на этих станциях практически идентичны, а все различия обусловлены локальными динамическими изменениями [1, 24]. На рис. 12 приведены графики вариаций величины геомагнитного поля на этих двух станциях за 1–15 июня 2019 г., а также временной ряд ошибок восстановления.

Как видно из рис. 12, ошибки восстановления для временных рядов двух соседних станций практически идентичны (рис. 12, справа) и накладываются друг на друга. При этом умеренное значение ошибки восстановления ~0.33 соответствует магнитной буре 08.06.2019, а относительно небольшие значения ошибки ~0.18 (04.06.2019 и 13.06.2019) соответствуют небольшим аномалиям, проявляющимся в нарушении типичного суточного хода поля. Идентичный ход геомагнитного поля на двух станциях дает на выходе нейронной сети практически одинаковую ошибку восстановления, что позволяет использовать модель, обученную на данных станции Ак-Суу, и для ближайшей станции Шавай.

Заключение

Представлена методика выявления аномалий в вариациях величины геомагнитного поля на основе классического автоэнкодера. В качестве обучающих данных были выбраны суточные вариации величины геомагнитного поля в магнитоспокойные дни по базовой

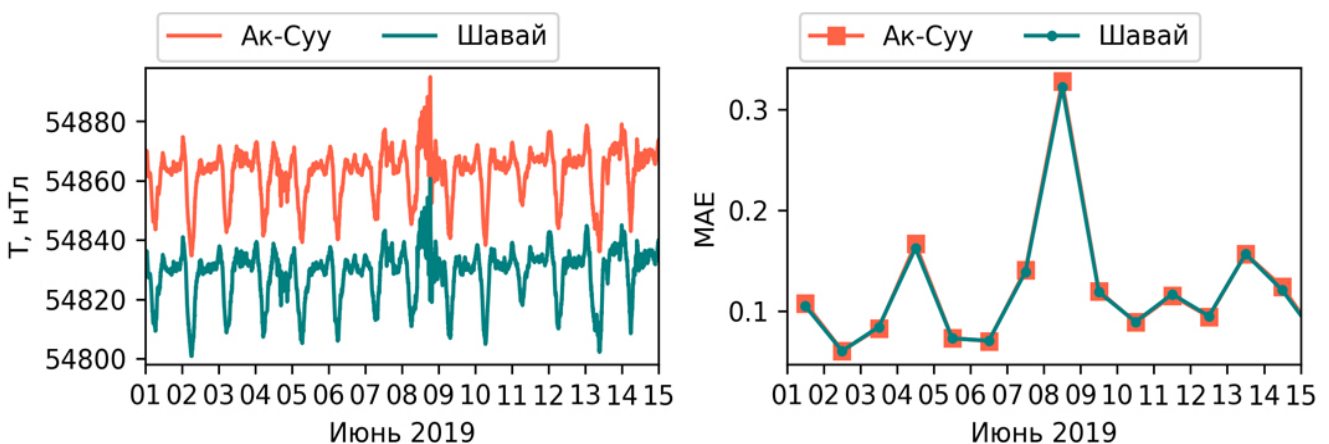


Рис. 12. Вариации величины геомагнитного поля (слева) и ошибок восстановления (справа) с 1 по 15 июня 2019 г. для станций Ак-Суу и Шавай.

Fig. 12. Variations of the geomagnetic field (on the left) and reconstruction errors (on the right) for June 1–15, 2019 for the Ak-Suu and Shavay stations.

станции Ак-Суу за три года (2020, 2021, 2022). Для увеличения производительности и уменьшения сложности модели частота дискретизации данных была уменьшена с 20 с до 1 мин. Нормализация обучающих данных проведена путем вычитания среднего значения с последующим делением на среднеквадратичное отклонение. В обучающую выборку вошли 316 примеров нормальных данных, а 780 примеров аномальных данных были впоследствии использованы для оценки границы разделения классов норма/аномалия. Обученная модель воспроизводила характерные признаки нормальных данных, а далее, сталкиваясь с неизвестными аномальными данными, демонстрировала ухудшение в качестве восстановления. При этом ошибка восстановления в виде средней абсолютной погрешности MAE является оценкой меры аномальности. Пороговым значением уровня MAE было выбрано значение 0.109, которое делит данные на 2 класса (нормальные и аномальные). Метрики бинарной классификации, рассчитанные по матрице ошибок, свидетельствуют о положительном качестве модели, в частности, для данных по станции Ак-Суу за 2017, 2018 и 2019 гг. Значения полноты (метрика recall) для этих лет (0.965, 0.982 и 0.970 соответственно) дают основание считать, что обученная нейронная сеть выделяет аномалии с качеством, сопоставимым с работой эксперта. Также необходимо отметить, что модель, обученная по данным станции Ак-Суу, хорошо подходит и для выявления аномалий в данных близлежащих станций, например станции Шавай. Это позволяет использовать одну модель для группы близкорасположенных станций и тем самым сократить количество обучаемых моделей, используемых для анализа аномальных данных станции сети геомагнитного мониторинга. Дальнейшее развитие этой работы предполагает апробацию модели для выделения аномалий, связанных с геомагнитными бурями, и сравнение их со значениями различных планетарных индексов, характеризующих глобальную возмущенность магнитного поля Земли, а также чувствительность метода к уровню техногенных аномалий в виде импульсных выбросов и ступеней.

Список литературы

1. Мухамадеева В.А., Воронцова Е.В., Лазарева Е.А. **2015**. Опыт проведения геомагнитных наблюдений на Бишкекском геодинамическом полигоне. *Вестник Кыргызско-Российского Славянского университета*, 15(3): 130–133.
2. Campbell W.H. **2003**. *Introduction to Geomagnetic Fields*. Cambridge Univ. Press, 337 p. <https://doi.org/10.1017/cbo9781139165136>
3. Hawkins D.M. **1980**. *Identification of outliers*. Springer, 188 p. <https://doi.org/10.1007/978-94-015-3994-4>
4. Chandola V., Banerjee A., Kumar V. **2009**. Anomaly detection: A survey. *ACM Computing Surveys*, 41(3): 1–58. <https://doi.org/10.1145/1541880.1541882>
5. Pang G., Shen C., Cao L., Hengel A. **2020**. Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 54(2): 1–38. <https://doi.org/10.1145/3439950>
6. Goodfellow I., Bengio Y., Courville A. **2016**. *Deep learning*. MIT Press, 800 p.
7. Yildirim O., Tan R.S., Rajendra Acharya U. **2018**. An efficient compression of ECG signals using deep convolutional autoencoders. *Cognitive Systems Research*, 52: 198–211. <https://doi.org/10.1016/j.cogsys.2018.07.004>
8. Marchi E., Vesperini F., Squartini S., Schuller B. **2017**. Deep recurrent neural network-based autoencoders for acoustic novelty detection. *Computational Intelligence and Neuroscience*, 1(4694860). 14 p. <https://doi.org/10.1155/2017/4694860>
9. Valentin M.B., Bom C.R., Martins Compan A.L., Correia M.D., Menezes de Jesus C., de Lima Souza, de Albuquerque Marcio P., de Albuquerque Marcelo P., Faria E.L. **2018**. Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate. *Journal of Petroleum Science and Engineering*, 170: 315–330. <https://doi.org/10.1016/j.petrol.2018.06.038>
10. Fraiwan L., Lweesy K. **2017**. Neonatal sleep state identification using deep learning autoencoders. *2017 IEEE 13th International Colloquium on Signal Processing & Its Applications (CSPA)*, Penang, Malaysia. <https://doi.org/10.1109/cspa.2017.8064956>
11. Wang L., Zhang Z., Xu J., Liu R. **2018**. Wind turbine blade breakage monitoring with deep autoencoders. *IEEE Transactions on Smart Grid*, 9(4): 2824–2833. <https://doi.org/10.1109/tsg.2016.2621135>
12. Pal A., Baskar S. **2015**. Speech emotion recognition using deep dropout autoencoders. *2015 IEEE International Conference on Engineering and Technology*, Penang, Malaysia. <https://doi.org/10.1109/icetech.2015.7275003>
13. Gao S., Zhang Y., Jia K., Lu J., Zhang Y. **2015**. Single sample face recognition via learning deep supervised

- autoencoders. *IEEE Transactions on Information Forensics and Security*, 10(10): 2108–2118. <https://doi.org/10.1109/tifs.2015.2446438>
14. Gomes T.A., Carvalho R.N., Carvalho R.S. **2017**. Identifying anomalies in parliamentary expenditures of Brazilian chamber of deputies with deep autoencoders. *2017 16th IEEE International Conference on Machine Learning and Applications, Cancun, Mexico*. <https://doi.org/10.1109/icmla.2017.00-33>
 15. Имашев С.А., Рыбин А.К. **2023**. Сейсмические и геоакустические отклики земной коры на зондирования мощными электрическими импульсами на территории Бишкекского геодинамического полигона. *Наука и технологические разработки*, 102(2-3): 63–88. <https://doi.org/10.21455/std2023.2-3-3>
 16. Chen S., Meng Z., Zhao Q. **2018**. Electrocardiogram recognition based on variational autoencoder. *Machine Learning and Biometrics*. <https://doi.org/10.5772/intechopen.76434>
 17. Sakurada M., Yairi T. **2014**. Anomaly detection using autoencoders with nonlinear dimensionality reduction. *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*. <https://doi.org/10.1145/2689746.2689747>
 18. Bishop C.M., Bishop H. **2024**. *Deep learning: Foundations and concepts*. Cham: Springer Intern. Publ., 649 p. <https://doi.org/10.1007/978-3-031-45468-4>
 19. Imashev S.A. **2021**. Extended isolation forest – Application to outlier detection in geomagnetic data. *IOP Conference Series: Earth and Environmental Science*, 929, 012022. <https://doi.org/10.1088/1755-1315/929/1/012022>
 20. Nair V., Hinton G.E. **2010**. Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning*. Madison, USA: Omnipress, p. 807–814.
 21. Glorot X., Bordes A., Bengio Y. **2011**. Deep sparse rectifier neural networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Fort Lauderdale, FL, USA, 15: 315–323.
 22. Kingma D., Ba J. **2015**. Adam: A method for stochastic optimization. In: *International Conference on Learning Representations*. San Diego, USA. p. 13.
 23. Bishop C.M. **2006**. *Pattern recognition and machine learning*. New York: Springer, 778 p. <https://doi.org/10.1007/978-0-387-45528-0>
 24. Имашев С.А., Лазарева Е.А. **2022**. Пространственное распределение составляющих главного геомагнитного поля на основе модели IGRF-13 для территории Кыргызстана. *Вестник Кыргызско-Российского Славянского университета*, 22(4): 192–198. <https://doi.org/10.36979/1694-500X-2022-22-4-192-198>
- ## References
1. Mukhamadeeva V.A., Vorontsova E.V., Lazareva E.A. **2015**. Experience of geomagnetic observations at the geodynamic test ground in Bishkek. *Vestnik of KRSU = Herald of KRSU*, 15(3): 130–133.
 2. Campbell W.H. **2003**. *Introduction to Geomagnetic Fields*. Cambridge Univ. Press, 337 p. <https://doi.org/10.1017/cbo9781139165136>
 3. Hawkins D.M. **1980**. *Identification of outliers*. Springer, 188 p. <https://doi.org/10.1007/978-94-015-3994-4>
 4. Chandola V., Banerjee A., Kumar V. **2009**. Anomaly detection: A survey. *ACM Computing Surveys*, 41(3): 1–58. <https://doi.org/10.1145/1541880.1541882>
 5. Pang G., Shen C., Cao L., Hengel A. **2020**. Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 54(2): 1–38. <https://doi.org/10.1145/3439950>
 6. Goodfellow I., Bengio Y., Courville A. **2016**. *Deep learning*. MIT Press, 800 p.
 7. Yildirim O., Tan R.S., Rajendra Acharya U. **2018**. An efficient compression of ECG signals using deep convolutional autoencoders. *Cognitive Systems Research*, 52: 198–211. <https://doi.org/10.1016/j.cogsys.2018.07.004>
 8. Marchi E., Vesperini F., Squartini S., Schuller B. **2017**. Deep recurrent neural network-based autoencoders for acoustic novelty detection. *Computational Intelligence and Neuroscience*, 1(4694860). 14 p. <https://doi.org/10.1155/2017/4694860>
 9. Valentín M.B., Bom C.R., Martins Compan A.L., Correia M.D., Menezes de Jesus C., de Lima Souza, de Albuquerque Marcio P., de Albuquerque Marcelo P., Faria E.L. **2018**. Estimation of permeability and effective porosity logs using deep autoencoders in borehole image logs from the brazilian pre-salt carbonate. *Journal of Petroleum Science and Engineering*, 170: 315–330. <https://doi.org/10.1016/j.petrol.2018.06.038>
 10. Fraiwan L., Lweesy K. **2017**. Neonatal sleep state identification using deep learning autoencoders. *2017 IEEE 13th International Colloquium on Signal Processing & Its Applications (CSPA)*, Penang, Malaysia. <https://doi.org/10.1109/cspa.2017.8064956>
 11. Wang L., Zhang Z., Xu J., Liu R. **2018**. Wind turbine blade breakage monitoring with deep autoencoders. *IEEE Transactions on Smart Grid*, 9(4): 2824–2833. <https://doi.org/10.1109/tsg.2016.2621135>
 12. Pal A., Baskar S. **2015**. Speech emotion recognition using deep dropout autoencoders. *2015 IEEE International Conference on Engineering and Technology*, Penang, Malaysia. <https://doi.org/10.1109/icetech.2015.7275003>
 13. Gao S., Zhang Y., Jia K., Lu J., Zhang Y. **2015**. Single sample face recognition via learning deep supervised autoencoders. *IEEE Transactions on Information Fo-*

- rensis and Security*, 10(10): 2108–2118. <https://doi.org/10.1109/tifs.2015.2446438>
14. Gomes T.A., Carvalho R.N., Carvalho R.S. **2017**. Identifying anomalies in parliamentary expenditures of Brazilian chamber of deputies with deep autoencoders. *2017 16th IEEE International Conference on Machine Learning and Applications, Cancun, Mexico*. <https://doi.org/10.1109/icmla.2017.00-33>
 15. Imashev S.A., Rybin A.K. **2023**. Seismic and geoaoustic responses of the Earth's crust to sensing with high energy electric pulses at the territory of the Bishkek geodynamic polygon. *Science and technological developments*, 102(2-3): 63–88. <https://doi.org/10.21455/std2023.2-3-3>
 16. Chen S., Meng Z., Zhao Q. **2018**. Electrocardiogram recognition based on variational autoencoder. *Machine Learning and Biometrics*. <https://doi.org/10.5772/intechopen.76434>
 17. Sakurada M., Yairi T. **2014**. Anomaly detection using autoencoders with nonlinear dimensionality reduction. *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*. <https://doi.org/10.1145/2689746.2689747>
 18. Bishop C.M., Bishop H. **2024**. *Deep learning: Foundations and concepts*. Cham: Springer Intern. Publ., 649 p. <https://doi.org/10.1007/978-3-031-45468-4>
 19. Imashev S.A. **2021**. Extended isolation forest – Application to outlier detection in geomagnetic data. *IOP Conference Series: Earth and Environmental Science*, 929, 012022. <https://doi.org/10.1088/1755-1315/929/1/012022>
 20. Nair V., Hinton G.E. **2010**. Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning*. Madison, USA: Omnipress, p. 807–814.
 21. Glorot X., Bordes A., Bengio Y. **2011**. Deep sparse rectifier neural networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. Fort Lauderdale, FL, USA, 15: 315–323.
 22. Kingma D., Ba J. **2015**. Adam: A method for stochastic optimization. In: *International Conference on Learning Representations*. San Diego, USA. p. 13.
 23. Bishop C.M. **2006**. *Pattern recognition and machine learning*. New York: Springer, 778 p. <https://doi.org/10.1007/978-0-387-45528-0>
 24. Imashev S.A., Lazareva E.A. **2022**. Spatial distribution of the main geomagnetic field components based on IGRF-13 model for Kyrgyzstan territory. *Vestnik of KRSU = Herald of KRSU*, 22(4): 192–198. <https://doi.org/10.36979/1694-500X-2022-22-4-192-198>

Об авторе

Имашев Санжар Абылбекович (<https://orcid.org/0000-0003-3293-3764>), кандидат физико-математических наук, ведущий научный сотрудник, Научная станция РАН в г. Бишкеке, Бишкек, Киргизия, sanzhar.imashev@gmail.com

Поступила 15.08.2024

Принята к публикации 20.09.2024

About the Author

Imashev, Sanjar A. (<https://orcid.org/0000-0003-3293-3764>), Cand. Sci. (Phys. and Math.), Leading Researcher, Research Station of the Russian Academy of Sciences in Bishkek, Bishkek city, Kyrgyzstan, sanzhar.imashev@gmail.com

Received 15 August 2024

Accepted 20 September 2024