

УДК 550.385.1+550.385.4

ПРОГНОЗИРОВАНИЕ РЯДОВ ИНДЕКСОВ ГЕОМАГНИТНОЙ АКТИВНОСТИ DST И AP
НЕЙРОСЕТЬЮ С LSTM СЛОЯМИ

Б. В. Козелов

доктор физико-математических наук, доцент, главный научный сотрудник

Полярный геофизический институт (ПГИ)

184209 г.Апатиты, Мурманская обл., Россия

e-mail: bob-koz@yandex.ru

<https://orcid.org/0000-0003-2738-2443>

Поступила в редакцию

После доработки

Принята к публикации

В работе описывается реализация нейросетевого подхода с использованием слоев LSTM (Long Shot-Term Memory) к прогнозированию рядов индексов геомагнитной активности Dst и ap на основе данных о предыдущей активности и параметрах солнечного ветра. Модель тренирована на данных базы OMNI за 2000-2023 гг. По предыстории значений z -компоненты межпланетного магнитного поля (Bz ММП), скорости солнечного ветра (V_{tot}), плотности плазмы солнечного ветра (Np) и индекса Dst за 80 часов прогнозируются следующие за ними 12 часовых значений Dst и ap . Программы работают на сайте <http://aurora.pgia.ru/AI/>.

FORECASTING SERIES OF DST AND AP GEOMAGNETIC ACTIVITY INDICES BY
NEURAL NETWORK WITH LSTM LAYERS

Boris Kozelov

Doctor of Physical and Mathematical Sciences, Associate Professor, Chief Researcher

Polar Geophysical Institute (PGI)

184209 Apatity, Murmansk region, Russia

e-mail: bob-koz@yandex.ru

<https://orcid.org/0000-0003-2738-2443>

The paper describes the implementation of a neural network approach using LSTM (Long Shot-Term Memory) layers to forecasting series of geomagnetic activity indices *Dst* and *ap* based on data on previous activity and solar wind parameters. The model is trained on the OMNI database data for 2000-2023. Based on the history of the values of the z-component of the interplanetary magnetic field (*Bz* IMF), solar wind speed (*Vtot*), solar wind plasma density (*Np*) and the *Dst* index for 80 hours, the following 12-hour values of *Dst* and *ap* are predicted. The programs work on the website <http://aurora.pgia.ru/AI/>.

1. ВВЕДЕНИЕ

В моделях ионосферы и верхней атмосферы, важных для многих прикладных вопросов, в качестве параметров используются индексы геомагнитной активности, которые получают на основе наблюдений наземными приборами [Vorobjev et al., 2013; Шубин и Деминов, 2019]. Для прогноза состояния моделируемых сред необходимо прогнозировать индексы геомагнитной активности. Задача прогнозирования рядов геомагнитной активности вперед на экваториальных и средних широтах с хорошей точностью решается при наличии данных спутников о солнечном ветре, который в значительной степени определяет геомагнитную активность. Было показано, что одним из самых геоэффективных параметров солнечного ветра является его южная компонента $B_z < 0$ и связанное с ней азимутальное электрическое поле $E_y = V \cdot B_z$, где V – радиальная скорость солнечного ветра [Burton et al., 1975]. Также скачки плотности плазмы в отдельных случаях влияет на геомагнитную активность. Многие работы посвящены поиску формальной функциональной связи между параметрами солнечного ветра и индексами геомагнитной активности на основе физических представлений [Дремухина и др., 2018; Newel et al., 2007; Borovsky, 2014; Borovsky and Birn, 2014; Kan, 1979]. В последние годы показано, что трудно формализуемые связи могут учитываться с помощью нейросетевого подхода [Бархатов и Ревунов, 2010; Barkhatov et al., 2020]. При этом особый интерес представляют рекуррентные сети, сети с памятью.

LSTM (long short-term memory, дословно - долгая краткосрочная память) — тип рекуррентной нейронной сети, способный обучаться долгосрочным зависимостям. LSTM были представлены в работе [Hochreiter and Schmidhuber, 1997], впоследствии усовершенствованы другими исследователями. В работе [Козелов, 2023] описана реализация сети со слоями LSTM для предсказания солнечной активности (числа солнечных пятен и поток радиоизлучения на волне 10.7 см) на следующий оборот Солнца. В данной работе так организованные слои использованы в нейронной сети для предсказания значений индексов геомагнитной активности Dst и ap на основе данных о предыдущей геомагнитной активности и параметрах солнечного ветра. Эти индексы характеризуют степень возмущенности геомагнитного поля в зоне экваториальных и средних широт, соответственно. Индекс ap является производным от Kp индекса, где он преобразован в численную линейную шкалу, в отличие от квазилогарифмического Kp . Ранее сети со слоями LSTM успешно использовались в работах [Gruet et al., 2018; Tan et al., 2018] для предсказания Dst и Kp индексов. В данной работе использовались несколько другие параметры сети, в частности использовались 80 часов предыстории не только самих индексов, но и параметров солнечного ветра, что

охватывает все обсуждаемые в литературе существенные масштабы связей. Другие параметры моделей обсуждаются далее.

2. ДАННЫЕ

За основу использованы данные из базы OMNI за 2000-2023 годы с часовым разрешением: z-компонента межпланетного магнитного поля (Bz ММП), скорость солнечного ветра (V_{tot}), плотность плазмы солнечного ветра (Np), Dst индекс, ap индекс.

Из последовательности этих данных выбраны все отрезки длиной $m = 80$ часов для (Bz , V_{tot} , Np , Dst) и следующие за ними 12 часовых значений Dst и ap . Отрезки, в которые попадают пробелы данных в базе OMNI исключены из набора. Полученный набор разбит на тренировочный и тестовый, состоящие из 151412 и 26505 отрезков. Из этих наборов формировались также данные для тренировки на меньшей предыстории ($m = 60, 40$ и 20 часов), см. обсуждение дальше.

Необходимо отметить, что экстремальных пиковых значений индексов ($Dst < -30$ нТ и $ap > 30$) относительно мало — они лежат на «хвостах» распределений. По такому ряду нейронная сеть обучается предсказывать преимущественно значения из «ядра» распределения. Однако повышенные значения геомагнитной активности представляют особый интерес для предсказания. Поэтому для тренировки сети использовались для лучшего предсказания маловероятных значений модификации набора выходных значений: значения Dst и ap нормировались в диапазон от 0.01 до 1.0, после чего возводились в некоторую степень p (см. далее табл.1) для смещения максимума распределения. Тем самым нормированные распределения целевых величин, на которых проводилось обучение, становились ближе к симметричному виду. На предсказанных сетью значениях делалось обратное преобразование.

3. НЕЙРОННАЯ СЕТЬ

Нейронная сеть реализована на языке Python с использованием библиотеки keras пакета TensorFlow. В сети использованы слои LSTM (Long Shot-Term Memory), позволяющие кроме общей связи между входными и выходными данными, запоминать ближайшую предысторию в рядах данных. Общая схема блок-сети представлена на рис.1, показаны составляющие сеть слои из библиотеки.

Подаваемые на вход значения последовательно кодируются во внутреннем состоянии первой LSTM сети, которое подается на вторую LSTM сеть, которая предсказывает следующие значения. То есть сеть по m векторам значений ($Bz, Vtot, Np, Dst$) обучалась предсказывать вектор из 12 последующих значений Dst или ap . Эти значения сравнивались с истинными, в процессе обучения минимализировалась метрика «средняя абсолютная ошибка» MAE – mean absolute error. Также для обучения имеет значение, насколько полно используется информация в исходном ряду, что в данном случае характеризует параметр «потерь» (loss='mean squared error').

Из тренировочного набора 30% выделялось для проверки (валидации) в процессе обучения. Обучение прерывалось по условию, что на протяжении 5 эпох не убывает функция потерь loss_val='mean squared error' на проверочном наборе. Модель быстро обучается и повышение числа эпох обучения приводит признакам переобучения — росту MAE. Для проверки качества обучения проводилось тестовое дообучение модели с уменьшением скорости обучения, оно не приводило к заметному уменьшения MAE.

Зависимость результатов обучения от числа нейронов скрытого слоя в LSTM-слоях, а также показателя p , приведены в табл. 1.

На рис.2 и 3 приведены значения MAE для истинных и предсказанных Dst и ap индексов для тренировочного и тестового наборов для лучших параметров p для каждого индекса (использованы параметры модели, соответствующие подчеркнутым значениям в табл.1). В целом, средние ошибки находятся на уровне менее 8 нТ, несмотря на отдельные пики для больших значений. Интересно, что на использованных наборах данных метрика MAE всегда меньше для тестовых данных, чем для тренировочных. Это объясняется, видимо, тем, что в тестовом интервале оказались пики заметно меньше амплитуды, чем в тренировочном.

Результаты работы модели на тестовом ряду значений индексов для предсказаний на 1 час показаны на рис.4. На рис.5 более детально показан короткий участок со сравнением с истинными значениями предсказаний на 1 и на 12 часов. Видна тенденция уточнения деталей предсказаний на меньшем интервале предсказания.

4. ОБСУЖДЕНИЯ

Внешним параметром, которым необходимо задать для слоев LSTM, является n - число нейронов скрытого слоя. Это число зависит от зависимостей в данных и подбирается экспериментально. Экспериментально для использованного набора данных установлено, что

при $n < 12$ качество работы сети заметно ухудшается - не удается добиться малого значения потерь *loss* при обучении модели, а при $n > 48$ чаще возникают артефакты (ложные пики) в выходной последовательности. Оказалось, что одновременно одной сетью два ряда предсказываются хуже. Для нахождения оптимальных значений проводились численные эксперименты со значениями $n = 12, 18, 24, 30, 36, 42, 48$, приведенными в табл.1. В приведенных в табл.1 выбраны по две модели с наименьшей средней ошибкой предсказания для *Dst* и *ap* по MAE метрике (выделены шрифтом).

Надо отметить, что характеристика MAE не всегда совпадает с субъективной оценкой совпадения предсказанного и истинного ряда, т. к. для субъективной оценки более важно совпадение пиков, которые в среднее значение отклонения дают малый вклад даже в модифицированном ряду. Субъективно лучшие предсказания дают модели, значения MAE для которых в табл.1 подчеркнуты, они использованы для построения рис.2-5.

В данной работе для предсказания использовались ряды данных 80 часов предыстории, что обусловлено данными с сайта https://lasp.colorado.edu/space_weather/dsttemerin/, которые предполагается использовать для прогноза. Надо отметить, что в работах группы [Бархатов и Ревунов, 2010; Barkhatov et al., 2020] использовались более короткие ряды с предысторией 8 часов. Но в той работе заранее были отобраны только события изолированных бурь. Здесь мы проводили тестирование моделей на более коротких рядах, результаты показаны в табл. 2. Видно, что для *Dst* использование длительной предыстории позволяет несколько уменьшить MAE предсказания. Для *ap* такой зависимости нет, ошибка MAE даже несколько уменьшается с уменьшением длины предыстории, но незначительно, чтобы делать выводы. Субъективно же предсказание с длительной предысторией лучше описывает флуктуации *ap* индекса порядка нескольких десятков часов.

Известно, что обычно включение в прогностическую модель дополнительной эмпирической информации улучшает прогноз. Так в работе [Gruet et al., 2018] для предсказания значений *Dst* индекса от 1 до 6 часов вперед по параметрам солнечного ветра используется предыстория на 6 часов, однако для лучшего описания пиковых значений добавлена дополнительная ветвь в нейросети для значений $Dst > 6$. Скрытый слой состоит из 20 нейронов. Также улучшения прогноза добиваются включением дополнительной информации (оценки поглощения из данных GPS).

Комбинированная модель с использованием слоя LSTM строится в работе [Tan et al., 2018] для предсказания *Kp* индекса на 3 часа вперед по параметрам солнечного ветра и межпланетного магнитного поля по предыстории за 6 часов. В модели строятся отдельные

ветви при отсутствии магнитной бури и для условий бури. Использование дополнительных ветвей приводит к неплохому прогнозу, даже при небольшой предыстории.

5. ВЫВОДЫ

Описана реализация нейросетевого подхода с использованием слоев LSTM (Long Shot-Term Memory) к прогнозированию рядов индексов геомагнитной активности Dst и ap . Используются данные из базы OMNI за 2000-2023 с часовым разрешением. По предыстории значений параметров солнечного ветра (Bz ММП, скорость V_{tot} , плотность Np) и Dst за 80 часов прогнозируются следующие за ними 12 часовых значений Dst и ap . Исследованы зависимости «средней абсолютной ошибки» предсказания модели от числа нейронов в скрытых слоях LSTM слоев и длины предыстории в данных. Для моделирования оптимально использовать разные модели при числе нейронов скрытого слоя 18 - 24 (для Dst) и 30 — 42 (для ap). Длина использованной предыстории в данных обуславливается рядом данных, которые предполагается использовать для прогноза. Сокращение предыстории ухудшает прогноз.

Программы работают на сайте http://aurora.pgia.ru/AI/?id=dst_pred. Для оперативного прогноза данные по Bz ММП, V_{tot} и Np в виде картинок загружаются с сайта https://lasp.colorado.edu/space_weather/dsttemerin/ и оцифровываются. Данные индекса Dst геомагнитной активности были взяты с сайта World Data Center for Geomagnetism, Kyoto (<http://wdc.kugi.kyoto-u.ac.jp/>). Текущие значение ap берутся с сайта https://www-app3.gfz-potsdam.de/kp_index/Kp_ap_nowcast.txt. Прогноз обновляется через 3 часа.

Для лучшего предсказания маловероятных экстремальных значений предложено модифицировать набор выходных значений показательной функцией к более симметричному виду. Для дальнейшего улучшения прогноза в сеть в последующих работах можно ввести дополнительные ветви, использующие дополнительную эмпирическую информацию.

6. БЛАГОДАРНОСТИ

Автор благодарит GSFC/SPDF OMNIWeb за подготовку использованных в статье данных.

7. ФИНАНСИРОВАНИЕ РАБОТЫ

Работа поддержана грантом РНФ и Министерства образования и науки Мурманской области №22-12-20017 «Пространственно-временные структуры в околоземном космическом

пространстве Арктики: от полярных сияний через особенности самоорганизации плазмы к прохождению радиоволн».

СПИСОК ЛИТЕРАТУРЫ

- Бархатов Н.А., Ревунов С.Е. Искусственные нейронные сети в задачах солнечно-земной физики// Н.Новгород: «Поволжье», 2010.
- Дремухина Л. А., Лодкина И. Г., Ермолаев Ю. Связь параметров солнечного ветра разных типов с индексами геомагнитной активности // Космические исследования, 2018, Т. 56, № 6 С. 410-419.
- Козелов Б.В. Предсказание временных рядов солнечной активности с помощью искусственной нейронной сети LSTM // Труды Кольского научного центра РАН. Серия: Естественные и гуманитарные науки. 2023. Т.2 №2. С.19-24.
- Шубин В. Н., Деминов М. Г. Глобальная динамическая модель критической частоты F2-слоя ионосферы // Геомагнетизм и аэрономия, 2019, том 59, № 4, с. 461–473.
- Barkhatov N.A., Vorobjev V.G., Revunov S.E. et al. Neural network classification of substorm geomagnetic activity caused by solar wind magnetic clouds // Journal of Atmospheric and Solar-Terrestrial Physics, 2020, V. 205, № 105301.
- Borovsky J.E. Canonical correlation analysis of the combined solar-wind and geomagnetic-index data sets // J. Geophys. Res., V.119, doi 10.1002/2013JA019607, 2014.
- Borovsky J.E., Birn J. The solar wind electric field does not control the dayside reconnection rate // J. Geophys. Res. Space Physics., V.119, doi:10.1002/2013JA019193, 2014.
- Burton R.K., McPherron R.L., Russel C. J. An empirical relationship between interplanetary conditions and Dst // Geophys. Res., V.80, P. 4204- 4214, 1975.
- Hochreiter S., Schmidhuber J. Long short-term memory // Neural computation. 1997. V. 9(8). P. 1735–1780.
- Kan J.R., Lee L.C. Energy coupling function and solar wind-magnetosphere dynamo // Geophys. Res. Lett., V.6, P.577- 580, 1979.
- Newel P.T., Sotirelis T., Liou K., et al. A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables // J. Geophys. Res., V.112, A01206, doi:10.1029/2006JA012015, 2007.
- Vorobjev V.G., Yagodkina, O.I., Katkalov, Y. Auroral Precipitation Model and its applications to ionospheric and magnetospheric studies // J. Atm. S-Terr. Phys. 2013. V. 102(9). P. 157-171.
- Gruet M.A., M. Chandorkar, A. Sicard, E. Camporeale, "Multiple-hour ahead forecast of the Dst index using a combination of long short-term memory neural network and Gaussian process," Space Weather, 16, 1882–1896 (2018). <https://doi.org/10.1029/2018SW001898>
- Tan Y., Q. Hu, Z. Wang, Q. Zhong, "Geomagnetic index Kp forecasting with LSTM," Space Weather, 16, 406–416 (2018). <https://doi.org/10.1002/2017SW001764>.

Таблица 1. Зависимость средней ошибки MAE на обучающей и тестовой выборке от параметров для моделей предсказания *Dst* и *ap* индексов

Индекс	p	n	MAE обучение	MAE тест
<i>Dst</i>	4	12	11,1	9,70
	4	18	7,75	6.81
	4	24	7,55	6,74
	4	30	8,28	7,18
	4	36	8,87	7,68
	3	24	7,75	6,82
	5	24	8,61	7,46
<i>ap</i>	1/5	18	6,73	5,34
	1/5	24	6,38	5,08
	1/5	30	6,27	4,79
	1/5	36	6,23	4,79
	1/5	42	6,13	4,77
	1/5	48	6,26	4,89
	1/3	36	6,32	4,83
	1/3	42	6,13	4,88
	1/4	30	6,10	4,85
	1/4	36	6.22	4,89
	1/4	42	6,25	4,99

Примечания. Параметры: n - число нейронов скрытого слоя; p - степень смещения максимума распределения прогнозируемого индекса. Длина предыстории $m=80$. Две модели с наименьшей ошибкой для каждого индекса выделены шрифтом, подчеркиванием выделены субъективно лучшие модели.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»

Таблица 2. Зависимость средней ошибки MAE на обучающей и тестовой выборке от длины предыстории для моделей предсказания *Dst* и *ap* индексов.

Индекс	m	p	n	MAE обучение	MAE тест
<i>Dst</i>	80	4	24	7,55	6,74
	60	4	24	8,99	7,81
	40	4	24	8,45	7,40
	20	4	24	9,32	8,13
<i>ap</i>	80	1/5	36	6,31	4,88
	60	1/5	36	6,32	4,97
	40	1/5	36	6,16	4,88
	20	1/5	36	6,19	4,74

Примечания. Параметры: n - число нейронов скрытого слоя; p - степень смещения максимума распределения прогнозируемого индекса; m - длина предыстории.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»

Подписи к рисункам к работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности Dst и ap нейросетью с LSTM слоями»

Рис.1. Блок-схема использованной нейронной сети.

Рис.2. Абсолютная ошибка предсказанных нейронной сетью и истинных значений Dst индекса в ряду: a – обучающий набор; b – тестовый набор. Средняя ошибка на наборе приведена выше графика.

Рис.3. Абсолютная ошибка предсказанных нейронной сетью и истинных значений ap индекса в ряду: a – обучающий набор; b – тестовый набор. Средняя ошибка на наборе приведена выше графика.

Рис.4. Сравнение истинных и прогнозируемых рядов значений индексов на тестовом наборе данных: a – Dst ; b – ap . Серая сплошная линия — наблюдения; черная сплошная линия — прогноз на 1 час.

Рис.5. Сравнение истинных и прогнозируемых значений индексов на фрагменте тестового набора данных: a – Dst ; b – ap . Серая сплошная линия— наблюдения; черная сплошная линия — прогноз на 1 час, черный пунктир — прогноз на 12 часов.

РИСУНКИ

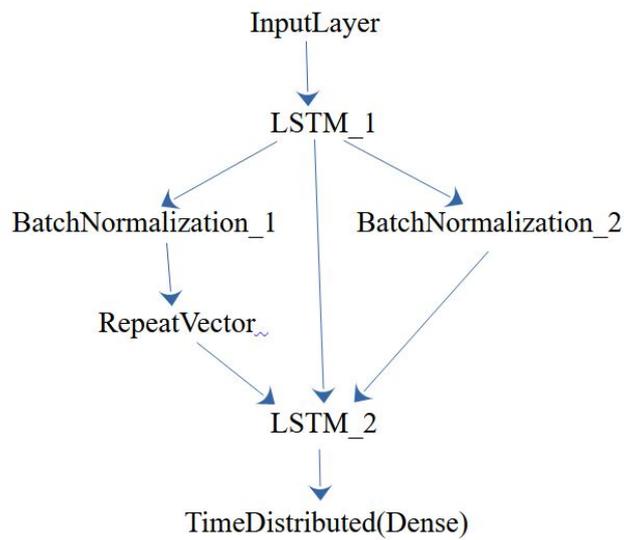
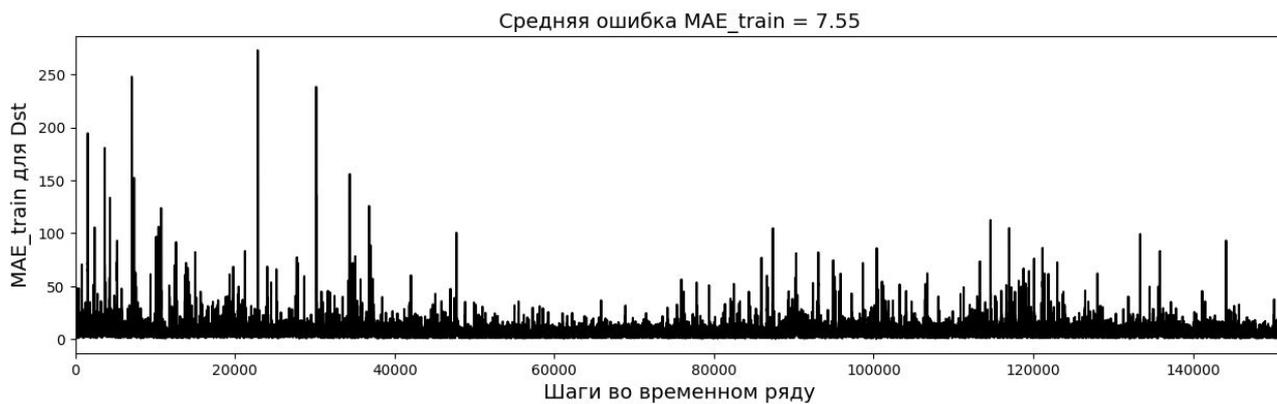


Рис.1.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»

а



б

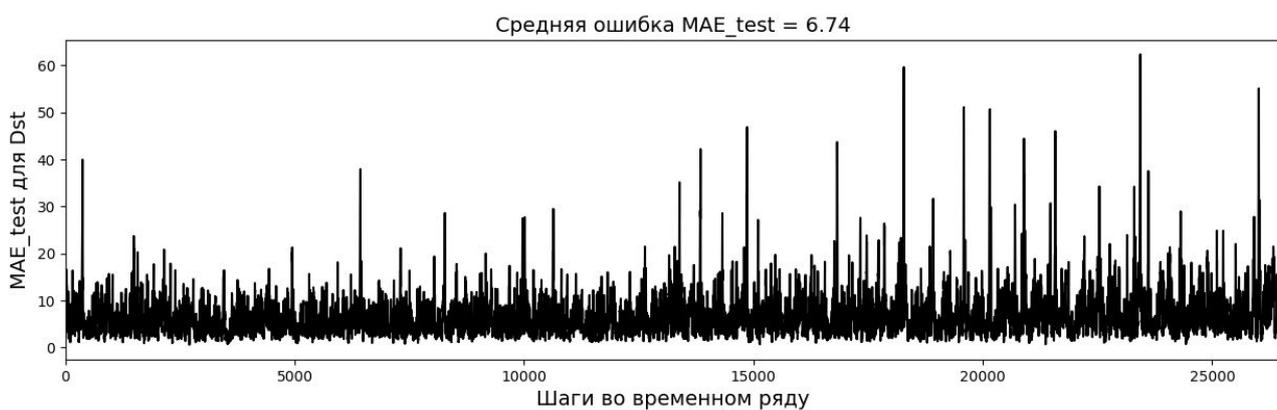
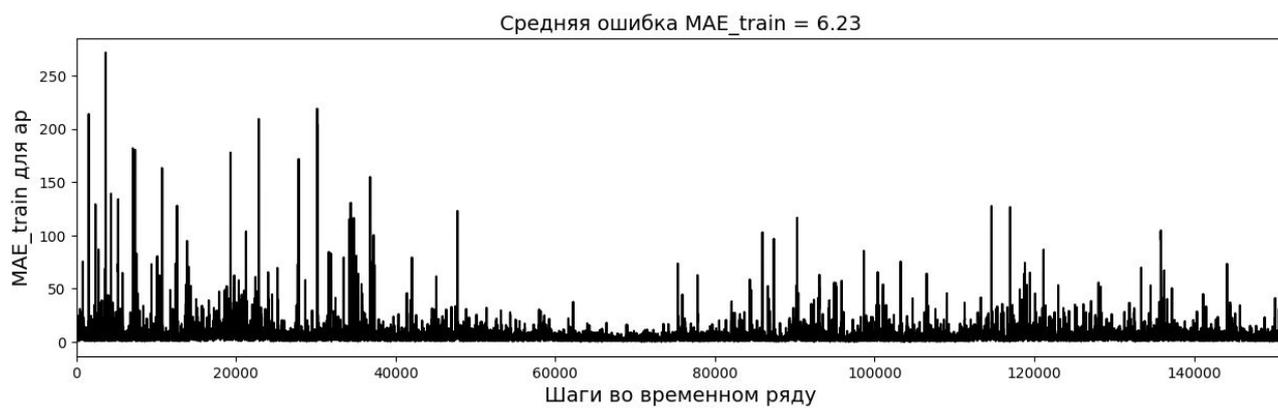


Рис.2.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»

а



б

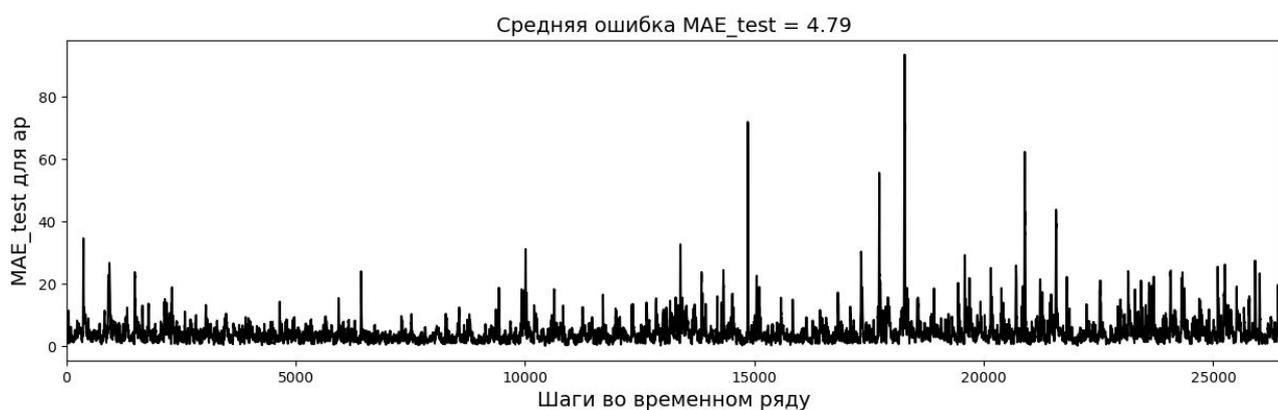
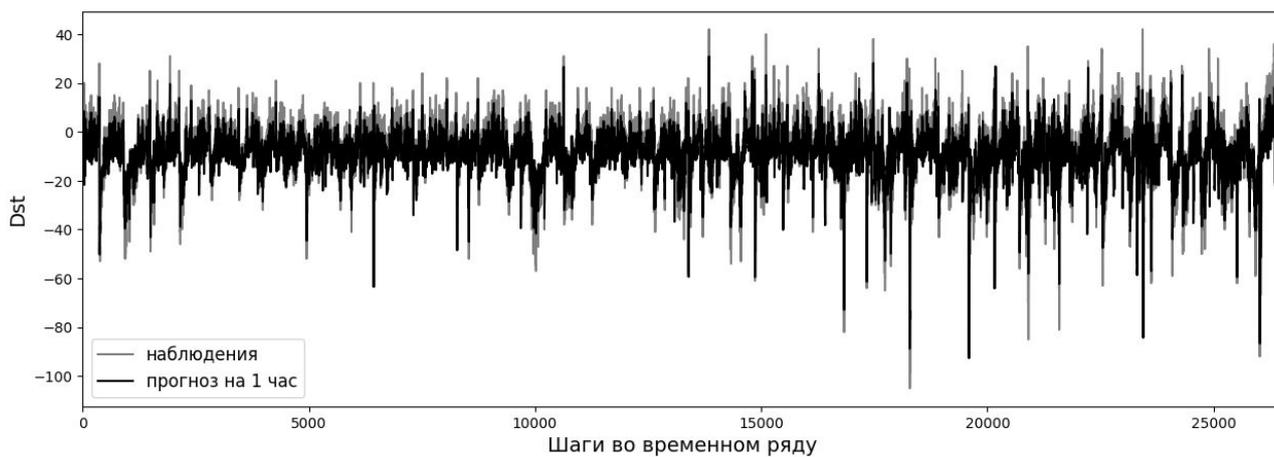


Рис.3.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»

а



б

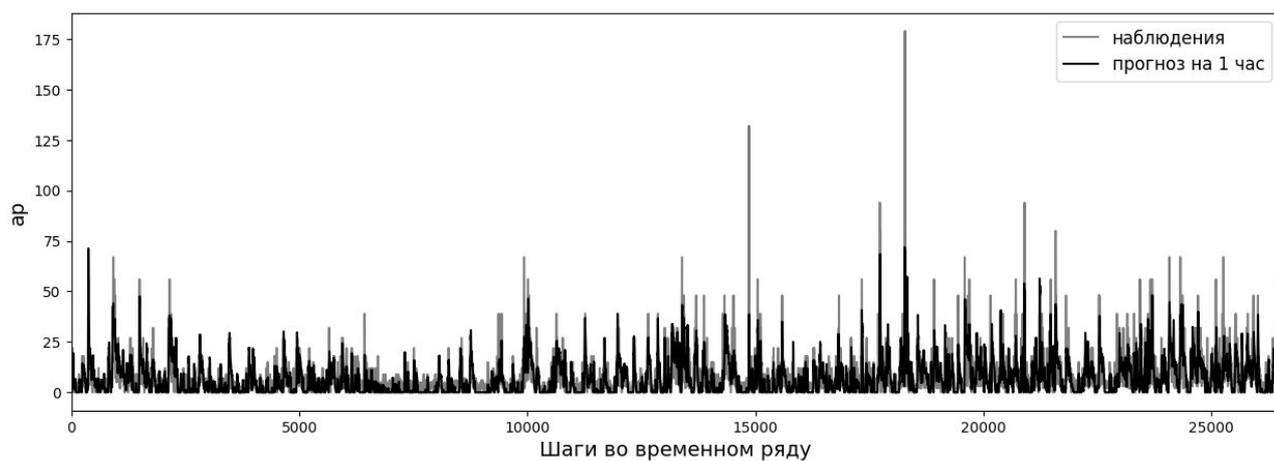
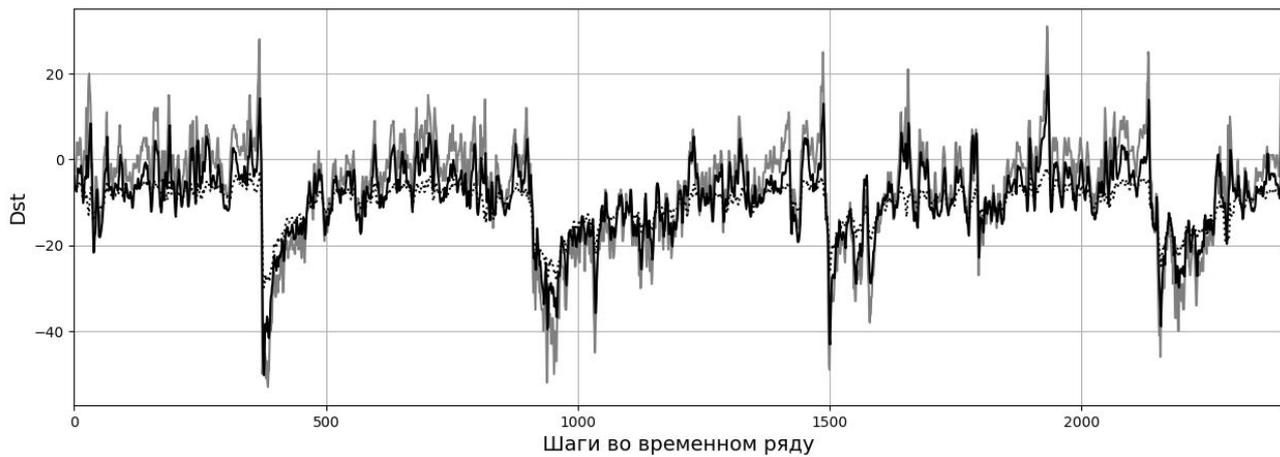


Рис.4.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности Dst и ap нейросетью с LSTM слоями»

а



б

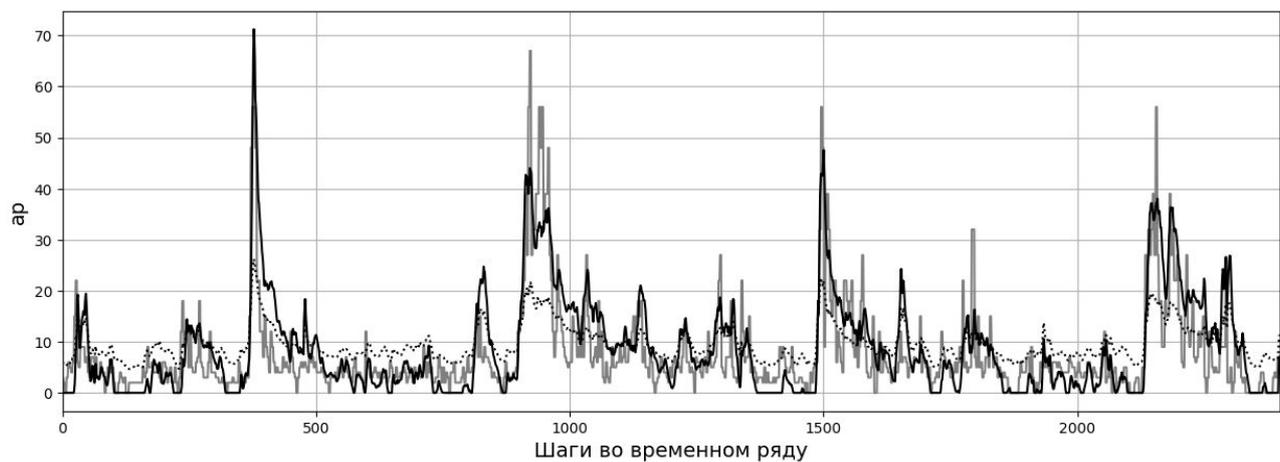


Рис.5.

К работе Б. В. Козелов «Прогнозирование рядов индексов геомагнитной активности *Dst* и *ap* нейросетью с LSTM слоями»