

РАЗРАБОТКА И АНАЛИЗ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ПАРАМЕТРОВ КОСМИЧЕСКОЙ ПОГОДЫ И ЭЛЕКТРОННОЙ КОНЦЕНТРАЦИИ ВЕРХНЕЙ АТМОСФЕРЫ

А.С. Халматова, В.П. Лебедев

Институт солнечно-земной физики Сибирского отделения Российской академии наук, Иркутск, Россия,
alenahalmatova@iszf.irk.ru

DEVELOPMENT AND ANALYSIS OF NEURAL NETWORK MODELS FOR FORECASTING SPACE WEATHER PARAMETERS AND ELECTRON CONTENT IN THE UPPER ATMOSPHERE

A.S. Khalmatova, V.P. Lebedev

Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia, alenahalmatova@iszf.irk.ru

Аннотация. Разработан и протестирован программный комплекс краткосрочного прогноза индекса солнечной активности $F10.7$ на базе гибридной нейросетевой модели VMD-TiDE с последующим расчетом максимальной электронной концентрации N_mF2 области F2 ионосферы. Прогноз $F10.7$ интегрирован с региональной (LEMI) и глобальной (IRI2016) ионосферными моделями. Установлено, что сглаживание обучающего ряда скользящим средним с окном 81 день при отключении встроенного учета 27-дневного периода вращения Солнца снижает ошибку прогноза N_mF2 в 3–4 раза по сравнению с сырыми данными. Разработан метод оценки $RMSE$ прогноза по единственному параметру — среднегодовому значению $F10.7$.

Ключевые слова: космическая погода, прогноз индекса $F10.7$, нейронные сети, ионосфера, электронная концентрация N_mF2 , ионосферная модель IRI-2016.

Abstract. This study presents the development and testing of a software for the short-term forecasting of the $F10.7$ solar activity index, based on a hybrid VMD-TiDE neural network model, followed by the calculation of the maximum electron concentration N_mF2 in the ionospheric F2 region. The $F10.7$ forecast is integrated with both regional (LEMI) and global (IRI2016) ionospheric models. It was established that smoothing the training dataset with the use of an 81-day moving average while disabling the built-in mechanism for accounting for the Sun's 27-day rotation period reduces the N_mF2 forecast error by a factor of 3 to 4 compared to using the raw data. Furthermore, a method was developed for estimating the forecast $RMSE$ based on a single parameter: the annual mean value of $F10.7$.

Keywords: space weather, $F10.7$ index forecasting, neural networks, ionosphere, peak electron density N_mF2 , IRI-2016 ionospheric model.

ВВЕДЕНИЕ

Резкие скачки в наклонном полном электронном содержании на пути от спутника к приемнику вызывают изменения фазы несущей волны и амплитуды сигнала. Такие колебания, известные как ионосферные мерцания, могут приводить к затуханию сигнала и проблемам с сильными вариациями фазы, что значительно нарушает работу GPS-приемников. Своевременный прогноз состояния ионосферы является одной из практически значимых задач физики солнечно-земных связей.

Ключевым входным параметром для большинства эмпирических моделей верхней атмосферы и ионосферы (IRI, NeQuick2, GSM TIP, региональные модели типа LEMI) является индекс $F10.7$ — плотность потока солнечного радиоизлучения на длине волны 10.7 см (частота 2.8 ГГц), измеряемая в единицах солнечного потока (с.е.п.). В качестве выходного параметра ионосферы рассматривается максимум электронной концентрации в области F2 — N_mF2 .

В настоящее время для прогноза $F10.7$ применяются такие методы, как линейная регрессия, авторегрессионные модели AR, ARIMA и SARIMA [Du, 2020], методы машинного обучения [Chao и др., 2017; Zhang и др., 2022]. Наилучшие результаты стабильно показывают

гибридные модели, в которых исходный ряд предварительно раскладывается методом вариационной модовой декомпозиции (VMD), а каждый из узкополосных компонентов прогнозируется отдельно: VMD-LSTM [Hao и др., 2024] и VMD-TiDE [Egorov, 2025].

РАЗРАБОТАННЫЙ ПРОГРАММНЫЙ КОМПЛЕКС

Архитектура комплекса реализована по модульному принципу на основе объектно-ориентированного подхода и связывает прогноз индекса $F10.7$ с эмпирическими моделями ионосферы (LEMI (Ratovsky и Oinats, 2011; Ratovsky, Oinats и Medvedev, 2013), IRI2016) для расчета прогнозного значения N_mF2 . Цель такой архитектуры — автоматизация всех вычислений от получения временного ряда и обучения нейросети до итогового расчета прогноза N_mF2 и оценки ее ошибки, а также возможность с минимальными правками заменить любой из ключевых компонентов: модель прогноза, метод сглаживания, ионосферную модель. Модульность обеспечивается тем, что результаты работы каждого модуля сохраняются в формате CSV-файлов, и модули могут запускаться независимо друг от друга.

Исходный ряд $F10.7$ поступает в модуль обработки временного ряда, осуществляющий сглаживание скользящим средним с пользовательским размером окна (по умолчанию 27 и 81 день); сглаживание не является обязательным этапом. Подготовленный ряд распараллеливается на прогнозную и эталонную ветки. На прогнозной ветке выбранная пользователем модель (по умолчанию VMD-TiDE) обучается на обработанных данных и далее по заданной длине истории строит прогноз, на выбранной пользователем период (обычно 90 дней). Эталонная ветка сохраняет наблюдаемые значения за тот же интервал. Результаты обеих веток поступают в модуль расчета среднеквадратичного отклонения $RMSE$, который вычисляет его как функцию дня горизонта и сохраняет график и численные значения. Далее результаты попадают в модуль расчета N_mF2 : по прогнозным и наблюдаемым значениям $F10.7$ ионосферная модель рассчитывает соответствующие значения N_mF2 (по 4 значения для $LT = 0, 6, 12, 18$ ч), и ошибка N_mF2 оценивается аналогично.

Состав комплекса:

- Модуль обработки временного ряда — сглаживание скользящим средним (окна 27 и 81 день) и медианным фильтром; декомпозиция данных для гибридных моделей.

- Модуль прогноза — VMD-TiDE по умолчанию с возможностью замены на любую другую нейросетевую модель через абстрактный интерфейс.

- Модуль расчета ошибки — $RMSE$ по выборке из 365 независимых прогнозов методом скользящего старта; построение графиков и тепловых карт ошибки в координатах год $\times$ день горизонта.

- Модуль расчета N_mF2 — взаимозаменяемые ионосферные модели LEMI, IRI2016, NeQuick2.

РАСЧЕТ $RMSE$

$RMSE$ прогноза вычислялось по выборке 365 независимых прогнозов методом скользящего старта: длина прогноза фиксирована (90 дней),

начальная дата сдвигается на 1 день; результаты — матрица 365×90 , из которой по столбцам вычисляется $RMSE$ как функция дня горизонта.

Для каждого года периода 2009–2025 гг. рассчитаны коэффициенты аппроксимации зависимости $RMSE$ от дня прогноза тремя функциями: прямой (с коэффициентами наклон и начало роста), экспонентой (амплитуда и показатель) и сигмной (амплитуда и крутизна). Зависимость каждого коэффициента от среднегодового значения $F10.7$ оценивалась через коэффициент корреляции Пирсона: прямая (наклон 0.932, начало роста 0.423), экспонента (амплитуда 0.920, показатель -0.190), сигмоида (амплитуда 0.802, крутизна -0.093).

Высокая положительная корреляция трех амплитудных коэффициентов (0.93, 0.92, 0.80) со среднегодовым $F10.7$ и практически нулевая корреляция коэффициентов формы (0.42, -0.19 , -0.09) позволяют сделать вывод, что с ростом солнечной активности растет ошибка, а форма кривой роста ошибки прогноза остается относительно стабильной от года к году, т. е. кривую $RMSE$ (горизонт) для конкретного года можно аппроксимировать, используя среднегодовую $F10.7$.

На основе линейных регрессий коэффициентов аппроксимации по среднегодовому $F10.7$ получены формулы расчета ошибки прогноза для любого года для всех трех форм аппроксимации (прямой, экспоненты, сигмной), где день горизонта d и среднегодовой $F10.7$ являются единственными входами.

Прямое сравнение (рис. 1) трех форм для 2019 и 2024 гг. показало, что в год минимума (2019 г.) все три функции близки между собой и в целом воспроизводят линейный рост ошибки, хотя несколько завышают ее. В год максимума (2024 г.) линейная аппроксимация воспроизводит реальную кривую $RMSE$ наиболее точно; экспонента сильно завышает значения как в начале, так и в конце горизонта; сигмоида завышает умеренно. Таким образом, линейная функция является оптимальной формой аппроксимации кривой ошибки прогноза по среднегодовому $F10.7$.

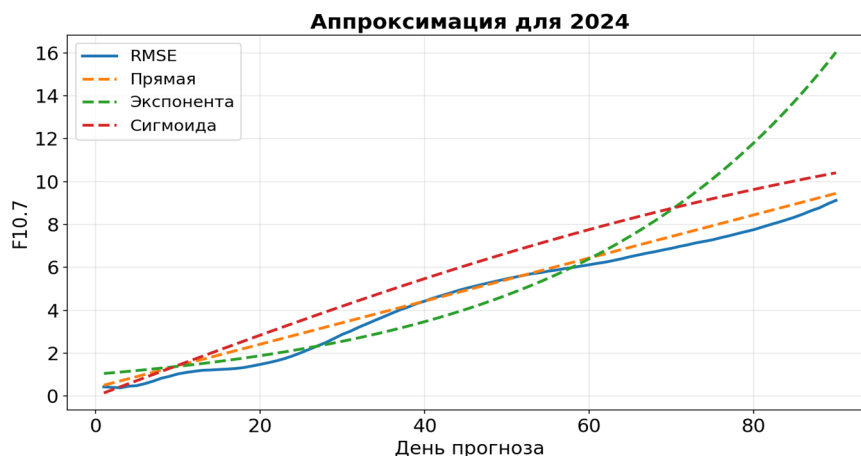


Рис. 1. Аппроксимации $RMSE$, построенные по среднегодовому $F10.7$, для 2024 г.

Прогноз F10.7, 2019

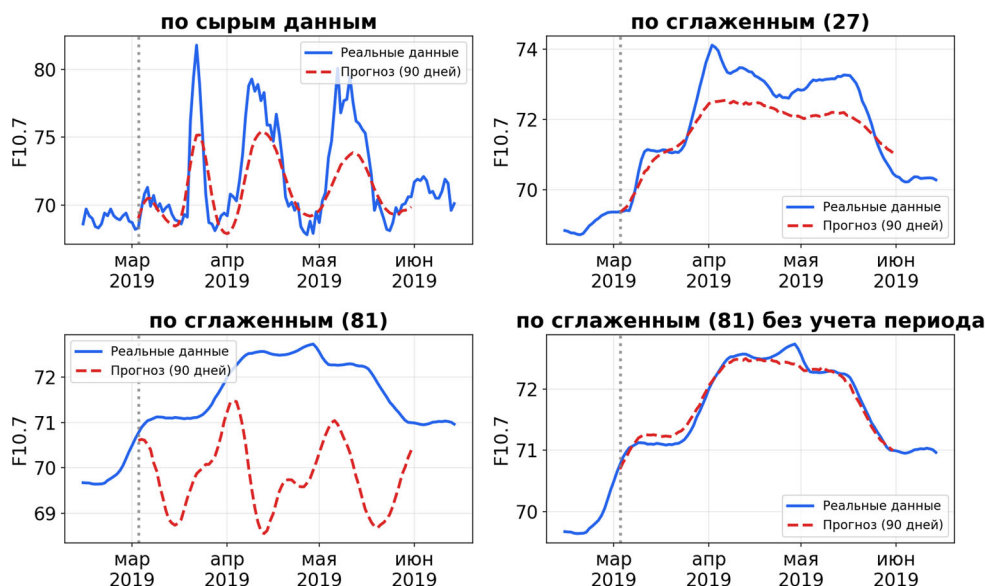


Рис.2. Сравнение прогнозов для 2019 г. по данным сырым/сглаженным 27/сглаженным 81 (с учетом периода вращения Солнца и без него)

РЕЗУЛЬТАТЫ: ПРОГНОЗ F10.7

Для прогноза, полученного по сырым данным, тренд воспроизводится корректно, но модель сглаживает пиковые значения. Краткосрочный прогноз совпадает с реальными данными, после чего наблюдается смещение прогнозной кривой относительно фактической. Зависимость $RMSE$ от длины горизонта в обоих случаях близка к линейной, однако абсолютный уровень ошибки различается принципиально: максимум $RMSE$ составляет ~2.5 с.е.п. для 2019 г. и ~30 с.е.п. для 2024 г.: рост в 12 раз объясняется большим количеством пиковых значений в максимуме солнечного цикла, которые модель не воспроизводит.

В ходе тестирования выявлено, что модель TiDE имеет встроенный учет 27-дневного периода вращения Солнца. Для сырых данных и данных, сглаженных окном 27 дней, такой учет повышает точность, поскольку период в данных явно присутствует. Однако при сглаживании окном 81 день период вращения Солнца устраняется из ряда и встроенный учет периода начинает мешать обучению: модель пытается восстановить периодичность, которой в данных нет, что нарушает даже воспроизведение тренда. Отключение встроенного учета периода при работе со сглаженными окном 81 данными уменьшает $RMSE$ более чем в три раза и сохраняет относительно высокую точность прогноза даже при высокой солнечной активности. Прогноз и ошибка по разным данным показаны на рис. 2.

РЕЗУЛЬТАТЫ: ПРОГНОЗ N_mF2

Расчет N_mF2 выполнялся по прогнозу F10.7, сглаженному окном 81 без учета периода вращения

Солнца (наилучший вариант по результатам анализа RMSE) с использованием двух ионосферных моделей: региональной LEMI, и глобальной IRI2016. Каждой дате соответствуют четыре значения N_mF2 для локальных времен LT = 0, 6, 12 и 18 ч; эталонные значения N_mF2 рассчитывались по тем же моделям, но с наблюдаемым F10.7.

Модель LEMI. Точность краткосрочного прогноза N_mF2 высока в обоих случаях (рис. 3). При низкой солнечной активности (2019 г.) прогноз сохраняет верный тренд во всем диапазоне горизонта, в солнечном максимуме (2024 г.) ошибка значительнее. Сравнение прогноза N_mF2 по сырым и сглаженным окном 81 без учета периода рядам F10.7 показывает, что сглаживание снижает максимальную ошибку в четыре раза при низкой солнечной активности и в три раза при высокой. Так, для LT = 12 ч максимальная ошибка прогноза N_mF2 составляет для 2019 г. — 0.09 (сглаженный ряд) против 0.4 (сырой); для 2024 г. — 1.4 против 4.5. Аналогичное соотношение сохраняется и для других LT.

Модель IRI2016. Графически прогнозный и наблюдаемый N_mF2 практически совпадают, что обусловлено встроенным сильным сглаживанием внутри самой модели — оно не позволяет напрямую оценить точность прогноза по графику. В численных значениях разница между прогнозным и наблюдаемым рядами составляет 0.01–0.5, что недостаточно для уверенной интерпретации на текущем этапе. Сравнение прогноза N_mF2 с реальными ионосферными наблюдениями (а не с эталоном, рассчитанным той же моделью по наблюдаемому F10.7) включено в задачи следующего этапа работ.

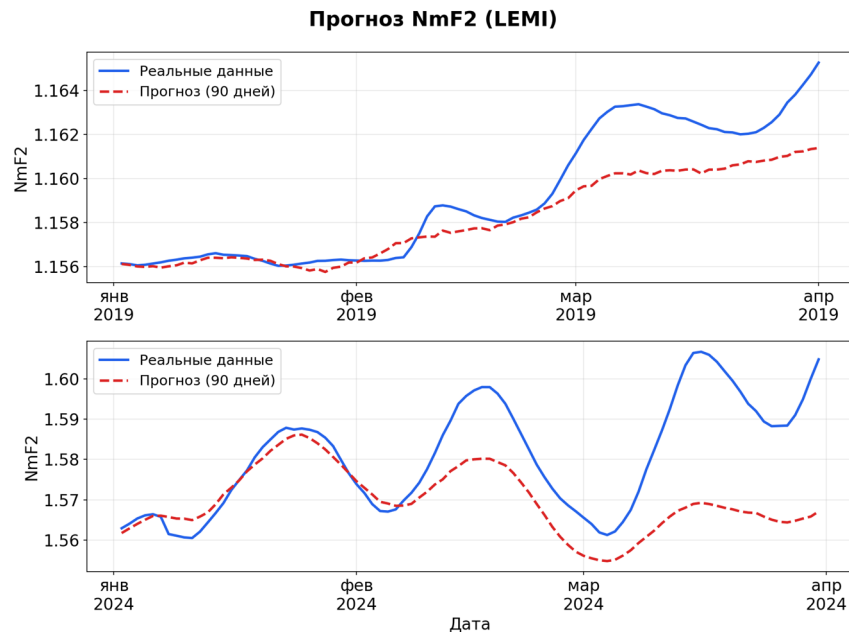


Рис. 3. Прогноз N_mF2 (LEMI) для 2019/2024 г.

ЗАКЛЮЧЕНИЕ

Проведен обзор методов прогнозирования $F10.7$; гибридные модели (декомпозиция ряда + прогноз) показывают наилучшие результаты. Разработан модульный программный комплекс прогноза $F10.7$ с последующей интеграцией результатов в ионосферные модели LEMI, IRI2016 и NeQuick. Получен прогноз N_mF2 по моделям LEMI и IRI2016; сглаживание снижает ошибку N_mF2 в 3–4 раза. Проведено обширное тестирование модели VMD-TiDE; внесены доработки, обеспечивающие автоматизацию работы с произвольным временным рядом (не только $F10.7$). Показано, что сглаживание скользящим средним с окном 81 день без учета 27-дневного периода вращения Солнца оптимально для долгосрочного прогноза $F10.7$. Создан единый метод оценки прогноза по метрике $RMSE$ с раздельным анализом по фазам солнечного цикла. Разработан метод оценки кривой $RMSE$ по единственному параметру — среднегодовому $F10.7$; наилучшая форма аппроксимации — линейная.

Работы по проведению наблюдений и первичной обработке данных выполнены при поддержке Минобрнауки России (проект FWZZ-2026-0051). Результаты получены частично с использованием оборудования Центра коллективного пользования «Ангара» (ИСЗФ СО РАН) (<http://ckp-rf.ru/ckp/3056/>).

СПИСОК ЛИТЕРАТУРЫ

- Chao X. et al. Using back propagation neural network method to forecast daily indices of solar activity $F10.7$ // Chinese J. Space Sci. 2017. V. 37, N 1. DOI: 10.11728/cjss.2017.01.001.
- Du Z. Forecasting the daily 10.7 cm solar radio flux using an autoregressive model // Solar Phys. 2020. V. 295, N 9. P. 125. <https://doi.org/10.1007/s11207-020-01689-x>.
- Egorov Y. Forecasting $F10.7$ and $F30$ indices using the variational mode decomposition algorithm and the deep-learning Time-series Dense Encoder Model // Acta Astronautica. 2025. V. 234. P. 455–461. <https://doi.org/10.1016/j.actaastro.2025.04.027>.
- Hao Y. et al. $F10.7$ daily forecast using LSTM combined with VMD method // Space Weather. 2024. V. 22, N 1. <https://doi.org/10.1029/2023SW003552>.
- Ratovsky K.G., Oinats A.V. Local empirical model of ionospheric plasma density derived from digisonde measurements at Irkutsk // Earth, Planets and Space. 2011. V. 63, N 4. P. 351–357. <https://doi.org/10.5047/eps.2011.03.002>.
- Ratovsky K.G., Oinats A.V., Medvedev A.V. Regular features of the polar ionosphere characteristics from digisonde measurements over Norilsk // Adv. Space Res. 2013. V. 51, N 4. P. 545–553. <https://doi.org/10.1016/j.asr.2011.09.026>.
- Zhang W. et al. Predicting the daily 10.7 cm solar radio flux using the long short-term memory method // Universe. 2022. V. 8, N 1. P. 30.