

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ КРИТИЧЕСКОЙ ЧАСТОТЫ f_oF2

APPLICATION OF CONVOLUTION NEURAL NETWORKS FOR CRITICAL FREQUENCY f_oF2 PREDICTION

Б.Г. Салимов*Институт солнечно-земной физики СО РАН,
Иркутск, Россия, salimov@iszf.irk.ru***О.И. Бернгардт** *Институт солнечно-земной физики СО РАН,
Иркутск, Россия, bergn@iszf.irk.ru***А.Е. Хмельнов** *Институт динамики систем и теории управления
им. В.М. Матросова СО РАН,
Иркутск, Россия, hmelnov@icc.ru***К.Г. Ратовский***Институт солнечно-земной физики СО РАН,
Иркутск, Россия, ratovsky@iszf.irk.ru***О.А. Кусонский***Институт геофизики УрО РАН,
Екатеринбург, Россия, zavlab@arti.igfuoran.ru***B.G. Salimov***Institute of Solar-Terrestrial Physics SB RAS,
Irkutsk, Russia, salimov@iszf.irk.ru***O.I. Bergardt***Institute of Solar-Terrestrial Physics SB RAS,
Irkutsk, Russia, bergn@iszf.irk.ru***A.E. Khmelnov***Matrosov Institute for System Dynamics and Control Theory
SB RAS,
Irkutsk, Russia, hmelnov@icc.ru***K.G. Ratovsky***Institute of Solar-Terrestrial Physics SB RAS,
Irkutsk, Russia, ratovsky@iszf.irk.ru***O.A. Kusonsky***Institute of Geophysics, Ural Branch RAS,
Ekaterinburg, Russia, zavlab@arti.igfuoran.ru*

Аннотация. Ионосфера оказывает важное влияние на качество работы средств радиосвязи, радиолокации и глобального позиционирования. Одной из важных характеристик, описывающих состояние ионосферы, является критическая частота f_oF2 . Ее прогноз позволяет обеспечить режимы эффективной работы технических радиосредств, а также рассчитать поправки, необходимые для повышения точности их функционирования. Традиционно для прогноза f_oF2 применяются различные физические и эмпирические модели. В данной работе предлагается эмпирическая методика прогноза, использующая методы машинного обучения и историю наблюдений. В ее основе лежит регрессионный подход к прогнозу по известной суточной квазипериодичности ионосферных параметров, связанной с солнечной освещенностью. Алгоритмически этот подход реализуется в виде сверточных нейронных сетей с двумерной сверткой. Исходные данные для анализа представляются в виде двумерных матриц солнечное время — дата. Модель обучена на данных среднеширотного ионозонда в Иркутске (РФ) и протестирована на данных нескольких среднеширотных ионозондов: Арти (РФ), Варшава (Польша), Мохе (Китай). Показано, что основной вклад в прогнозное значение f_oF2 вносят данные ближайших нескольких дней перед прогнозом, вклад остальных дней сильно убывает. Построенная модель обладает метриками качества прогноза f_oF2 (коэффициент корреляции Пирсона 0.928, корень среднеквадратичной ошибки 0.598 МГц, средняя абсолютная ошибка в процентах 10.45 %, коэффициент детерминации 0.861) и может быть применена для ее прогноза в средних широтах.

Ключевые слова: ионосфера, машинное обучение, нейронные сети, f_oF2 .

Abstract. Ionosphere has an important impact on the quality of radio communication, radar, and global positioning. One of the essential characteristics describing the state of the ionosphere is its critical frequency f_oF2 . Its prediction provides effective modes of operation of technical radio equipment as well as enables calculation of the corrections needed to improve the accuracy of its functioning. Different physical and empirical models are generally used for f_oF2 prediction. This paper proposes an empirical prediction technique based on machine learning methods and observational history. It relies on a regression approach to the prediction based on the known daily quasi-periodicity of ionospheric parameters related to solar illumination. Algorithmically, this approach is implemented in the form of convolutional neural networks with two-dimensional convolution. The input data for the analysis is presented as two-dimensional solar time — date matrices. The model was trained on data from the mid-latitude ionosonde in Irkutsk (RF) and tested using data from several mid-latitude ionosondes: Arti (RF), Warsaw (Poland), Mohe (China). It is shown that the main contribution to the prediction value of f_oF2 is made by the data on the nearest few days before the prediction; the contribution of the remaining days strongly decreases. This model has the following forecast quality metrics (Pearson correlation coefficient 0.928, root mean square error 0.598 MHz, mean absolute error in percent 10.45 %, coefficient of determination 0.861) and can be applied to f_oF2 forecast in middle latitudes.

Keywords: ionosphere, machine learning, neural networks, f_oF2 .

ВВЕДЕНИЕ

Ионосфера — динамическая среда, с одной стороны, находящаяся под воздействием солнечного излучения, процессов в нижележащей атмосфере и вышележащей магнитосфере и плазмосфере, а с другой стороны, обладающая определенной инерционностью, связанной с процессами ионизации и рекомбинации составляющих частиц и с процессами переноса вещества. Таким образом, при прогнозе ионосферных характеристик важно учитывать зависимости как локальные, связанные с инерционностью ионосферы, так и квазипериодические, связанные с суточным ходом солнечного излучения. Поэтому для прогноза ионосферных характеристик эффективно использовать модели, учитывающие состояние ионосферы в прошлом, а также историю параметров, характеризующих основное воздействие на ионосферу сверху, — солнечную и магнитную активность.

В последнее время при решении сложных задач с большим количеством неизвестных часто используют методы машинного обучения, являющиеся синтезом различных математических методов — от теории аппроксимации до теории оптимального управления. Термин «машинное обучение» на сегодня обобщает широкоизвестные простые методы типа разнообразных регрессионных методик и более сложные методы классического обучения (с учителем и без него) и глубокого, с использованием сложных end-to-end нейронных сетей, обучения с подкреплением и т. д. [Goodfellow et al., 2016]. В последнее время машинное обучение в тех или иных формах часто используется при решении геофизических задач [Yu, Ma, 2021] и задач прогноза состояния ионосферы. В работе [Sivavaraprasad et al., 2022] представлена нелинейная авторегрессионная нейронная сеть с внешним входом, прогнозирующая значение полного электронного содержания (ПЭС) в ионосфере. В качестве входных данных для этой модели используются значения ПЭС, геомагнитного индекса A_p , данные солнечной активности, время суток, географические координаты и др.

Критическая частота f_oF2 является одной из основных характеристик ионосферы, используемых при решении прикладных задач. В задачах вертикального зондирования — это максимальная частота отраженной от ионосферы радиоволны обыкновенной поляризации [Харгривс, 1982]. При наклонном распространении радиоволны с помощью f_oF2 можно оценить величину максимально применимой для радиосвязи частоты. При глобальном позиционировании с помощью глобальных навигационных спутниковых систем f_oF2 влияет на величину ионосферной поправки, необходимой для повышения точности позиционирования. Пример суточного хода f_oF2 по данным среднеширотного ионозонда ИСЗФ СО РАН (Иркутск, $52^\circ 16' N$, $104^\circ 17' E$) показан на рис. 1. Видна существенная суточная квазипериодическая составляющая, связанная с солнечной освещенностью: максимальные значения наблюдаются в дневное время, минимальные — в ночное.

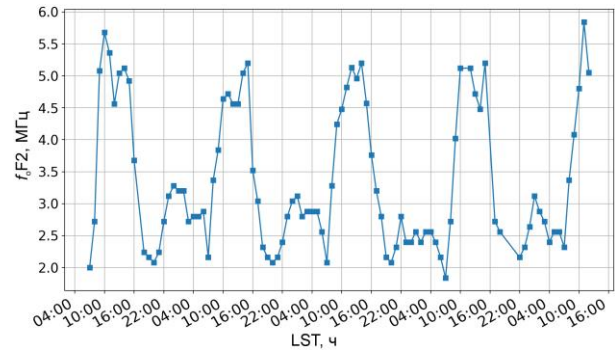


Рис. 1. Суточный ход критической частоты f_oF2 по данным среднеширотного ионозонда (Иркутск, $52^\circ 16' N$, $104^\circ 17' E$), как функция локального солнечного времени (LST)

В последнее время для прогноза f_oF2 используются модели машинного обучения: рекуррентные нейронные сети [Бархатов и др., 2005] и сверточные нейронные сети [Boulch et al., 2018]. Широко известна также эмпирическая модель IRI (International Reference Ionosphere), предназначенная для оценки и прогнозирования ионосферных параметров, в том числе f_oF2 [Bilitza et al., 2011]. Модель IRI имеет две разновидности для прогнозирования f_oF2 : одна разработана организацией International Radio Consultative Committee (модель CCIR) [CCIR, 1967], другая — организацией International Union of Radio Science (URSI) [Rush et al., 1989]. Модель IRI доступна в виде исходного кода на языке FORTRAN [http://irimodel.org/IRI-2016]. Она регулярно обновляется, последней версией является IRI-2016. Для уточнения модели IRI может быть взята методика компенсации ошибок NECTAR, которая использует рекуррентную нейронную сеть, в том числе для интерполяции временных и пространственных данных [Galkin et al. 2015].

В настоящее время интерпретируемость нейронных сетей — важное направление исследований, ставящее своей целью не только получение результата, но и объяснение того, почему и как он следует из исходных данных [Lundberg, Lee, 2017]. Нейронные сети обычно сложны для интерпретации и с точки зрения конечного пользователя представляют собой черный ящик. Качественно выяснить, какое влияние оказывают те или иные изменения входных параметров, по такой обученной нейронной сети обычно затруднительно. В этой работе нами проведено построение сверточной нейронной сети, способной прогнозировать значения f_oF2 на сутки вперед по истории данных, и качественная интерпретация ее коэффициентов в терминах регрессионного анализа.

ИСХОДНЫЕ ДАННЫЕ ДЛЯ ОБУЧЕНИЯ СЕТИ

В качестве исходных данных для анализа использовались временные ряды f_oF2 ручной обработки ионограмм, полученных на среднеширотном Иркутском дигизонде DPS-4 [https://ckp-rf.ru/catalog/ckp/3056], а также индексы геомагнитной возмущенности Dst и интенсивности солнечного излучения $F10.7$

на длине волны 10.7 см за период с 2009 по 2016 г. полученные из базы данных OMNI [<https://omniweb.gsfc.nasa.gov/form/dx1.html>]. Дигизонд DPS-4, размещенный в Иркутске, состоит из аналого-цифровой приемно-передающей части, четырех приемных антенн и передающей антенной системы «скрещенные вертикальные ромбы». Дигизонд обладает возможностью одновременной регистрации в автоматическом режиме таких параметров радиосигнала, как амплитуда, частота, высота (дальность), углы прихода, фаза, поляризация и доплеровский сдвиг частоты радиоволн. Такой набор параметров, цифровая согласованная обработка принятых сигналов, когерентное доплеровское интегрирование, а также эффективные алгоритмы и программы автоматической и полуавтоматической обработки данных до получения геофизических параметров существенно отличают дигизонд DPS-4 от ионозондов с использованием непрерывного сигнала с линейной частотной модуляцией [Смирнов, Степанов, 2004; Ратовский и др., 2004]. Автоматическое и полуавтоматическое выделение и обработка треков на ионограммах, необходимые для получения геофизических параметров, таких как f_oF2 , реализованы при помощи программного комплекса SAO-Explorer [Ратовский и др., 2004].

Все использованные в работе данные приведены к временному разрешению 1 ч. Данные наблюдений f_oF2 имеют характерные пропуски, где 1-часовые составляют ~67 %, 2-часовые — ~15 % всех пропусков. Имеются также более длинные пропуски, самый большой 4120 ч находится в обучающей выборке. Всего доля пропущенных значений составляет ~12 % всей выборки данных. Для заполнения пропущенных данных была использована линейная интерполяция. Приведение к временному разрешению 1 ч проводилось следующим образом:

- для f_oF2 и Dst , имеющих временное разрешение выше 1 ч (15 мин и 1 мин соответственно) проводилось прореживание данных до необходимого временного разрешения;
- для $F10.7$, имеющих временное разрешение 1 сут, использовалась интерполяция суточных данных гладкой квадратичной кривой.

Одним из основных методов валидации моделей машинного обучения является разделение исходных данных на три непересекающихся выборки — обучающую, валидационную и тестовую. Первая используется для обучения модели и определения ее параметров, вторая — для контроля качества модели в процессе обучения, третья — для окончательной оценки качества и адекватности обученной модели на независимом наборе данных. Способы разделения исходных данных на обучающую, валидационную и тестовую выборки различны и зависят от особенностей изучаемых данных. В данной работе использовались прогнозные модели двух типов: исходная и итоговая, различающиеся архитектурой.

Для обучения исходных моделей данные были разделены на две выборки — обучающую и валидационную — в пропорциях 80 и 20 % соответственно.

Для обучения итоговой модели, использующей результаты обработки исходных моделей, разделе-

ние происходило на три датасета — обучающий, валидационный и тестовый — в пропорциях 60, 20 и 20 % соответственно.

Основным механизмом формирования ионосферы является ионизация атмосферы солнечным излучением. Поэтому параметры ионосферы характеризуются квазипериодическими изменениями, связанными с длительностью основного 11-летнего солнечного цикла и суточным ходом освещенности. Для подтверждения адекватности построенной модели необходимо, чтобы она была работоспособна на различных участках солнечного цикла. Поэтому разделение на обучающую и валидационную выборки для исходной модели проводилось по фазе 11-летнего солнечного цикла: 1 января 2009 г. – 3 апреля 2015 г. — обучающая выборка, 3 апреля 2015 г. – 31 декабря 2016 г. — валидационная.

Для обработки обучающая и валидационная выборки были стандартизованы и преобразованы от векторнозначной 3-мерной функции времени (значения f_oF2 , $F10.7$, Dst с временным разрешением 1 ч) к последовательности матриц 90×24 для каждого из параметров (90 дней до текущего момента, 24 часа относительно текущего часа), характеризующих значения соответствующих параметров за предыдущие три месяца. Последовательность матриц строилась таким образом, чтобы каждому часу каждого дня измерений соответствовала своя тройка матриц. Двадцать четыре значения f_oF2 , следующих за текущим моментом, являлись целевыми, подлежащими прогнозу по значениям трех матриц (f_oF2 , $F10.7$, Dst), соответствующих текущему моменту.

Методика формирования этого набора данных заключается в следующем.

Этап 1. Весь обучающий датасет по типам данных f_oF2 , $F10.7$, Dst стандартизуется (масштабируется). Масштабированные данные $z(t)$ каждого типа были получены следующим образом из исходных $x(t)$:

$$z(t) = (x(t) - \mu) / s, \quad (1)$$

где μ — среднее значение обучающей выборки $x(t)$; s — стандартное отклонение обучающей выборки $x(t)$.

Этап 2. Для трех временных рядов f_oF2 , $F10.7$, Dst для каждого часа выбран отрезок данных, размером 2160 часовых значений, предшествующих этому часу, включая данный час. В качестве целевых значений для обучения модели были выбраны 24 значения f_oF2 , следующие за данным часом. Каждая выборка размером 2160 значений была преобразована в двумерную матрицу размером 90×24 , где каждая строка представляет день от 0 до 89, каждый столбец — час от 0 до 23.

Этап 3. Для обучения была сгенерирована четырехмерная матрица признаков данных размером $65736 \times 90 \times 24 \times 3$ и целевая матрица 65736×24 , которая содержит суточные наборы частот f_oF2 , иными словами 65736 (количество часов в полном наборе данных) трехмерных блоков размером $90 \times 24 \times 3$ и 65736 векторов длиной 24 (количество прогнозных часов после текущего часа). Сформированные данные разделены на две выборки объемом 52608 (обучающая) и 13128 (валидационная). Четвертое измерение

матрицы признаков данных имеет размерность 3 и отвечает за параметры f_0F2 , $F10.7$, Dst .

На рис. 2 показан пример стандартизованных данных f_0F2 , $F10.7$, Dst для временного интервала 2 января 2009 г., 00:00 LST — 1 апреля 2009 г., 23:00 LST, соответствующий моменту времени 1 апреля 2009 г., 23:00 LST в массиве многомерных данных (после выполнения Этапа 3). По вертикали отложены дни, по горизонтали часы. В f_0F2 отчетливо наблюдается суточный ход, связанный с уровнем освещенности — времени локального дня соответствуют высокие значения f_0F2 , локальной ночи — низкие.

АРХИТЕКТУРА НЕЙРОННОЙ СЕТИ, ЕЕ ОБУЧЕНИЕ И СВОЙСТВА

Базовыми предположениями при построении нейронной сети были следующие:

- значение f_0F2 определяется из линейной комбинации предшествующих значений f_0F2 , Dst и $F10.7$ с неизвестными весовыми коэффициентами;
- весовые коэффициенты могут различаться для разных текущих моментов наблюдений, но зависят только от сдвига от локального времени текущего момента.

Эти предположения выполняются в сверточной нейронной сети, показанной на рис. 3.

Исходная сеть представляет собой три независимых параллельных сверточных слоя (L.I) (для каждого из входных параметров f_0F2 , Dst , $F10.7$), результаты прохождения через которые объединяются (L.II) далее в матрицу размерностью $90 \times 24 \times 3$. Полученные матрицы выпрямляются (преобразуются в вектор размерностью 6480) (L.III) и подаются на вход решающего слоя (L.IV) — однослойной нейронной сети с линейной функцией активации, наиболее часто используемой в задачах прогноза и выполняющей функцию линейной комбинации результатов работы сверточных слоев (см. рис. 3). Сверточные слои L.I добавлены для улучшения качества фильтрации случайных выбросов и для дополнительного адаптивного сглаживания входных данных, получаемых

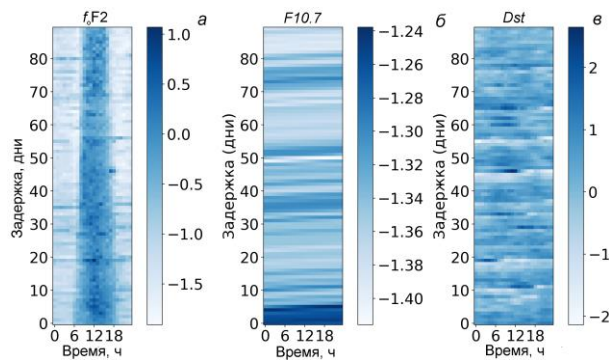


Рис. 2. Пример исходных стандартизованных данных по f_0F2 , $F10.7$, Dst со 2 января 2009 г., 00:00 LST по 1 апреля 2009 г., 23:00 LST (используемых для обучения нейронной сети), соответствующий моменту времени измерений 1 апреля 2009 г., 23:00 LST. Для данного момента измерений ось времени соответствует локальному солнечному времени

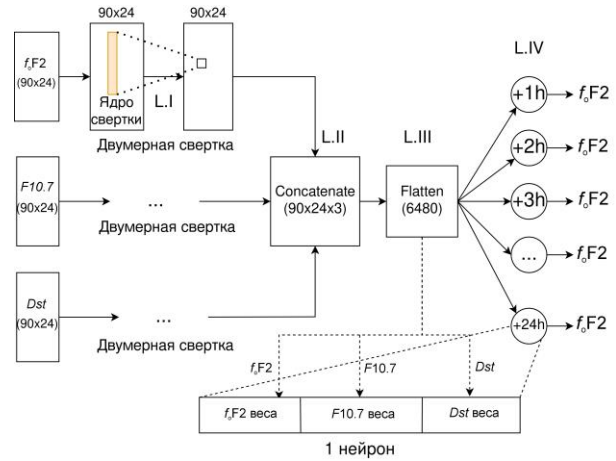


Рис. 3. Архитектура нейронной сети

с разным временным разрешением. В них используется ядро свертки размером 30 дней \times 1 ч.

Такая архитектура сети эквивалентна задаче линейной регрессии с очень большим количеством коэффициентов, которые находятся методами машинного обучения.

При обучении сети использовался аналог метода градиентного спуска — метод адаптивного момента Adam, ускоряющий скорость обучения [Kingma, Ba, 2014]. В качестве функции потерь использовалось MSE — среднеквадратичное отклонение прогнозного значения от реального. Эта функция потерь традиционна для решения задач прогноза числовых значений. Код нейронной сети реализован на языке Python с использованием библиотеки TensorFlow [Abadi et al., 2016].

После обучения нейронной сети был проведен анализ ее коэффициентов в целях упрощения архитектуры этой, достаточно общей, нейронной сети. В целях упрощения архитектуры мы исследовали два основных вопроса: 1) зависят ли коэффициенты сети от локального времени текущего момента или их можно считать в первом приближении стационарными и тем самым упростить нейронную сеть? 2) каковы структура коэффициентов нейронной сети и вклад наблюдаемых данных за предыдущие дни/часы в прогнозное значение f_0F2 ?

Первым шагом при обучении было повышение точности прогноза сетью не только за счет подбора коэффициентов нейронной сети, но и за счет использования нескольких сетей идентичной архитектуры, обученных независимо на различных отрезках данных (ансамблевых методов). Простейшим ансамблевым методом, позволяющим в некоторых случаях увеличить качество прогноза, является бэггинг [Breiman, 1994; Oritz, Maclin, 1999], заключающийся в независимом обучении ансамбля сетей и суперпозиции результатов, прогнозируемых каждой сетью неким фиксированным алгоритмом, обычно усреднением.

Модели для последующего бэггинга обучались следующим образом.

В качестве валидационной выборки использовались 20 % датасета в его конце. Каждая из 200 моделей обучалась на 80 % данных, случайным образом выбранных из оставшейся части. В результате было получено 200 моделей, обученных на различных датасетах, с различными весовыми коэффициентами.

Поскольку в нашем случае каждая сеть линейна, проведение бэггинга (усреднение результатов работы 200 нейронных сетей) эквивалентно усреднению коэффициентов нейронных сетей с поправкой на возможную инверсию результата выходным слоем и в рамках бэггинга позволяет ожидать увеличения качества прогноза. С другой стороны, физический смысл усредняемых коэффициентов нейронной сети — это величина вклада тех или иных предшествующих данных f_0F2 , Dst , $F10.7$ в прогнозное значение f_0F2 . С качественной точки зрения, усреднение этих коэффициентов по ансамблю сетей позволит уменьшить их вариации, связанные с неточностями обучения или шумом в исходных данных, и таким образом более точно определить величину вклада предшествующих данных. Для исследования влияния количества усреднений на качество прогноза в работе анализировались ансамбли объемом 1 (без бэггинга), 10, 30, 100 и 200 моделей.

Для реализации бэггинга проводится усреднение коэффициентов нейронов выходного слоя с учетом возможной инверсии выходным слоем по следующему алгоритму:

- в цикле для каждой сети из ансамбля на основе ближайшего к моменту наблюдения дня был выявлен знак решающего коэффициента для каждого часа: -1 , если значение веса отрицательно; 1 — если положительно;
- при значении знака -1 значения коэффициентов сети, соответствующие данному часу, инвертировались (рис. 4, а);
- итоговый коэффициент итоговой матрицы коэффициентов получался усреднением этого коэффициента по ансамблю обученных сетей.

Описанный выше алгоритм реализует следующие преобразования (на примере обработки коэффициентов, отвечающих за f_0F2) итогового слоя:

$$W_{i,j}^l = \frac{1}{N} \sum_{k=1}^N \left(INV \left(w_{i,j,k}^l, w_{i,0,k}^l \right) \right), \quad (2)$$

где k — номер сети в ансамбле; i, j — часы и дни задержки от -23 до 0 и от 0 до 89 соответственно; l — номер выходного нейрона от $+1$ до $+24$, каждый из которых отвечает за прогноз частоты f_0F2 на l -й час из 24 часов прогноза; N — количество обученных моделей в обрабатываемом ансамбле; $INV \left(w_{i,j,k}^l, w_{i,0,k}^l \right)$ — функция условного инвертирования в зависимости от знака мультипликативного коэффициента; $W_{i,j}^l$ — усредненная в рамках ансамбля моделей матрица коэффициентов выходных нейронов, отвечающих за измерения f_0F2 ; $w_{i,j,k}^l$ — исходная матрица коэффициентов выходных нейронов, отвечающих за f_0F2 для одной (k -й) модели. Функция инвертирования веса выглядит следующим образом:

$$INV \left(w, w_0 \right) = \left\{ \begin{array}{l} w, \text{ если } w_0 \geq 0 \\ -w, \text{ если } w_0 < 0 \end{array} \right\}, \quad (3)$$

где w — значение весового коэффициента; w_0 — значение весового коэффициента в ближайший к моменту наблюдения день, при этом номер нейрона,

номер сети из ансамбля и номер часа у этого коэффициента те же, что у w . Знак «+» перед номером выходного нейрона используется для акцентирования внимания на то, что нейрон выдает прогнозное значение критической частоты f_0F2 .

В исходной нейронной сети каждый выходной нейрон отвечает за соответствующий час прогноза после момента наблюдения. Нумерацию выходных нейронов примем от $+1$ до $+24$. Моментом наблюдения мы назовем момент времени, начиная с которого мы хотим получить прогноз на последующие 24 часа. Взаимосвязь между нумерацией нейронов выходного слоя, прогнозируемым часом и моментом наблюдения показана на рис. 4, б.

Перед получением итогового усредненного прогноза по ансамблю моделей, описываемого формулами (2), (3), было обучено необходимое количество моделей (см. архитектуру на рис. 3) и проведен анализ зависимости результатов прогноза от их числа (1, 10, 30, 100 и 200 моделей).

Рассмотрим дальнейшие вычисления на примере ансамбля из 200 моделей. Нами были независимо обучены 200 моделей. Для каждого из трех входных параметров исходной нейронной сети (f_0F2 , $F10.7$, Dst) были получены 24 матрицы коэффициентов размером 90×24 , соответствующих часам прогноза после момента наблюдений. Эти матрицы явились результатом усреднения соответствующих матриц в каждой из 200 моделей по формулам (2), (3). Назовем эти матрицы усредненными нейронами. На рис. 5 приведены коэффициенты нейрона, отвечающего за $+12$ -й час прогноза относительно момента наблюдения по данным f_0F2 , измеренным в -12 -й (а) и -18 -й (б) часы относительно момента наблюдения. По горизонтальной оси отложена задержка перед моментом наблюдения (в сутках), по вертикальной — значение весового коэффициента. Схема получения данных коэффициентов показана на рис. 6, а. Из рис. 5 можно сделать вывод, что на значение f_0F2 прогноза на $+12$ -й час относительно момента наблюдения сильнее влияют данные в предшествующие дни, отвечающие за -12 -й час относительно момента наблюдения, чем данные в предшествующие дни, отвечающие за -18 -й час относительно момента наблюдения, так как значения соответствующих весовых коэффициентов в среднем больше. Это означает, что основной вклад в прогнозное значение f_0F2 для часа прогноза вносят наблюдения, сдвинутые от него на 24 часа назад в дни предшествующих наблюдений.

При анализе коэффициентов отдельных нейронов, отвечающих за часы прогноза, отмечено, что основной вклад в прогнозное значение f_0F2 вносят наблюдения в предшествующий час, сдвинутый на 24 часа назад от часа прогноза, а также в соседние с ним часы в дни предшествующих наблюдений.

Анализ всех 24 часов прогноза показал аналогичную картину.

Таким образом, значения в матрице коэффициентов смещаются в зависимости от номера нейрона (часа прогноза) и будут иметь максимум на соответствующем часе. Таким образом, при переходе в систему отсчета времени, связанную с часом прогноза,

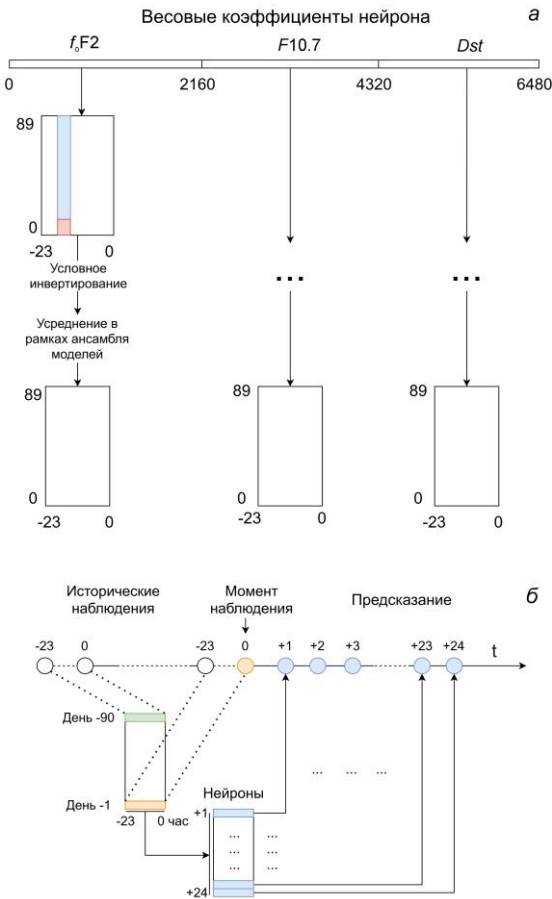


Рис. 4. Схема усреднения весовых коэффициентов в рамках ансамбля обученных моделей для одного из 24 выходных нейронов (а) и схема прогноза нейросетевой модели (б)

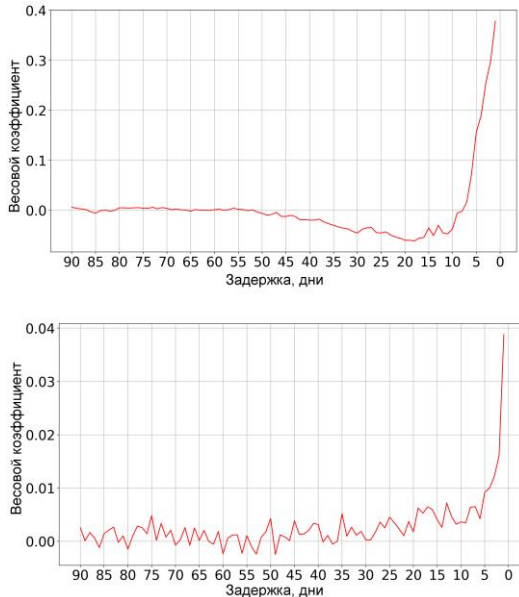


Рис. 5. Коэффициенты нейрона, отвечающего за +12-й час прогноза после момента наблюдения. Коэффициенты соответствуют -12-му (а) и -18-му (б) часам наблюдений перед моментом наблюдений в предшествующие дни

эти функции не будут зависеть от часа прогноза. Поэтому для анализа можно переводить все коэффициенты матриц в систему отсчета, связанную не с локальным временем момента наблюдения, а с часом

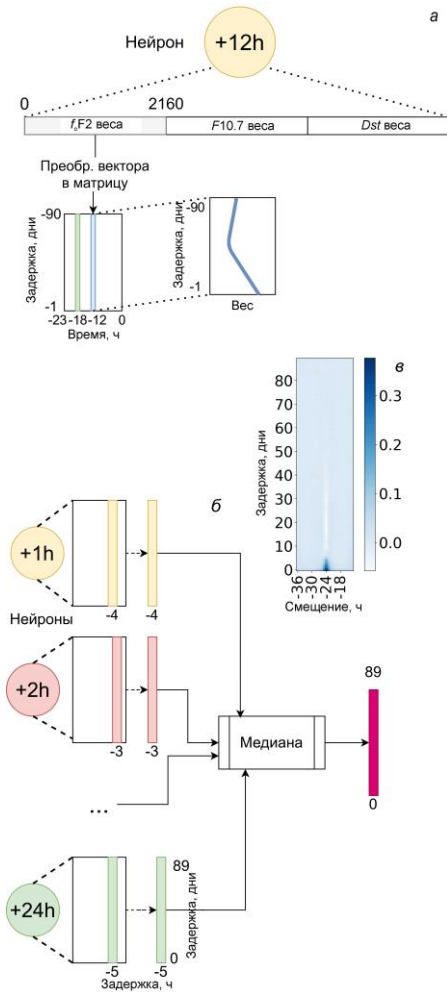


Рис. 6. Выделение коэффициентов нейрона, соответствующих обработке наблюдений f_0F2 в -18-й и -12-й часы перед временем момента наблюдений, в предшествующие дни для +12-го часа прогноза (а); схема получения коэффициентов отклика, отвечающих за влияние предшествующего часа, отстоящего на 29 часов назад от часа прогноза (б); медианная матрица для предшествующих часов, отстоящих от -36 до -13 часов от часа прогноза (матрица значений отклика θ для f_0F2) (в)

прогноза (сдвинутые матрицы). За нулевую отметку смещения примем смещение -24 часа от часа прогноза. Соответственно интервал соседних часов обозначим от -36 до -13 от часа прогноза.

Продемонстрированная нами независимость формы этих сдвинутых матриц от прогнозного часа позволяет усреднить матрицы по ансамблю, составленному из 24-х сдвинутых матриц, каждая из которых соответствует своему часу прогноза.

На рис. 6, б показано формирование медианных коэффициентов по столбцам для смещения -29 часов от часа прогноза, каждая ячейка новой матрицы является медианным значением соответствующих ячеек 24-х исходных сдвинутых матриц.

Для подтверждения слабой зависимости коэффициентов сдвинутой матрицы от часа прогноза, продемонстрируем, как выглядят коэффициенты, по которым мы считаем медиану. На рис. 7 показаны весовые коэффициенты сдвинутых матриц (для разных часов прогноза) для f_0F2 , соответствующие времени, отстоящему на 29 ч назад (а) и 24 часа назад (б)

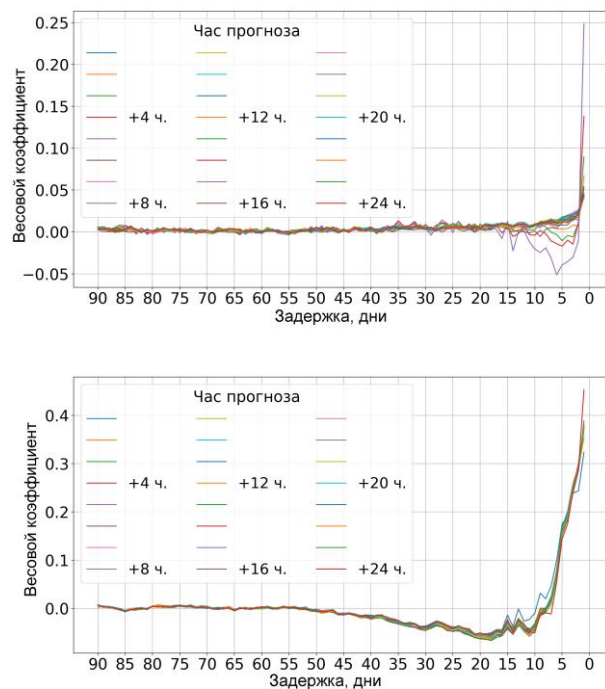


Рис. 7. Весовые коэффициенты сдвинутых матриц коэффициентов, соответствующие времени отстоящему: на 29 часов назад от часа прогноза, в зависимости от номера предыдущего дня (ось X) и часа прогноза (цвет) (а); на 24 часа назад от часа прогноза, в зависимости от номера предыдущего дня (ось X) и часа прогноза (цвет) (б)

от часа прогноза в предыдущие дни. Цвета кривых соответствуют различным часам прогноза.

Видно, что коэффициенты сдвинутых матриц в первом приближении не зависят от часа прогноза, а зависят только от относительной задержки между часом прогноза и временем наблюдения предшествующих данных, участвующих в расчете прогнозного значения. Поэтому мы имеем право статистически обработать сдвинутые матрицы по ансамблю, составленному из сдвинутых матриц для каждого из 24-х часов прогноза.

Стационарность сдвинутой матрицы позволяет уточнить ее коэффициенты, проведя их дополнительную статистическую обработку, упростить тем самым архитектуру сети и уменьшить количество свободных коэффициентов. Поэтому в дальнейшем мы будем анализировать лишь полученную нами медианную сдвинутую матрицу коэффициентов.

Таким образом, прогнозное значение f_oF2 в первом приближении является суперпозицией значений f_oF2 в предыдущие моменты времени и величина их вклада в прогнозное значение не зависит от локального времени, а лишь от относительной задержки между используемыми для прогноза данными и моментом, на который мы этот прогноз осуществляем.

Значение этих коэффициентов описываются трехмерной сдвинутой матрицей размерностью $90 \times 24 \times 3$ в координатах (задержка по дням, смещение по времени, параметр), где третья координата отвечает за одно из измерений f_oF2 , $F10.7$, Dst . Эту сдвинутую матрицу будем далее называть матрицей откликов.

Рассмотрим подробнее вычисление этой матрицы на примере обработки измерений f_oF2 :

$$\theta_{k,j} = \text{median}_l (W_{SHIFT(l,k),j}^l), \quad (4)$$

где θ — матрица отклика для f_oF2 размером (90×24) ; W — матрица усредненных по ансамблю нейронных сетей коэффициентов, полученных при вычислении по формуле (2); k — относительные смещения предшествующего часа от часа прогноза от -36 до -13 ; l — час прогноза от $+1$ до $+24$; j — номер дня задержки от 0 до 89 ; $SHIFT(l, k)$ — функция вычисления сдвинутого номера часа в зависимости от смещения и часа прогноза, преобразующая относительные смещения в часовые задержки относительно нуля, которые необходимо получить от матрицы W ; median — функция, возвращающая медианный вектор-столбец размером 90 в рамках обработанных нейронов. К примеру, на рис. 7, а показаны 24 кривые, соответствующие разным прогнозным часам, по которым вычисляется медианная кривая для времени, отстоящего на 29 ч назад от часа прогноза.

Поскольку коэффициенты смещений могут выходить за пределы матрицы W , при вычислениях она преобразуется функцией $SHIFT(l, k)$, исправляющей эту проблему:

$$SHIFT(l, k) = \begin{cases} l+k, & \text{если } l+k \in [-23, 0] \\ (l+k)+24, & \text{если } l+k < -23. \\ (l+k)-24, & \text{если } l+k > 0 \end{cases} \quad (5)$$

ВЫЧИСЛЕНИЕ КОЭФФИЦИЕНТОВ МОДЕЛИ ДЛЯ ВХОДНЫХ ИНДЕКСОВ $F10.7$ И Dst

Ранее, рассматривалось только влияние измеренных значений f_oF2 на ее прогнозные значения. В рамках этой модели мы рассчитали вклад индексов $F10.7$ и Dst в прогнозное значение f_oF2 . Матрицы-откликов для $F10.7$ и Dst показаны на рис. 8.

Поскольку входные данные измерений стандартизованы по формуле (1), мы можем оценить относительный вклад значений f_oF2 , $F10.7$ и Dst в прогнозное значение f_oF2 , сравнивая абсолютные значения коэффициентов матриц отклика между собой [Bring, 1994]. Сравнение показывает, что влияние индексов $F10.7$ и Dst на прогноз f_oF2 в среднем невелико по сравнению с данными f_oF2 , что может быть объяснено особенностью архитектуры нашей достаточно простой сети.

На панелях а–е видно, что на результат прогноза f_oF2 наиболее сильно должны влиять значения $F10.7$ и Dst , наблюдаемые в день, предшествующий прогнозируемому суткам. На панелях д, е показаны весовые коэффициенты, полученные аналогично рис. 7, которые демонстрируют стационарность матриц откликов.

На панелях ж, з показаны матрицы отклика на $F10.7$ и Dst при прогнозе f_oF2 , вычисленные по тому же алгоритму, что и отклик на f_oF2 . Из анализа рисунка сделать вывод, что на прогноз f_oF2 сильнее влияют значения $F10.7$ и Dst , измеренные в ближайший к моменту наблюдения день. Учет влияния $F10.7$ и Dst на прогноз f_oF2 в день, ближайший к дню

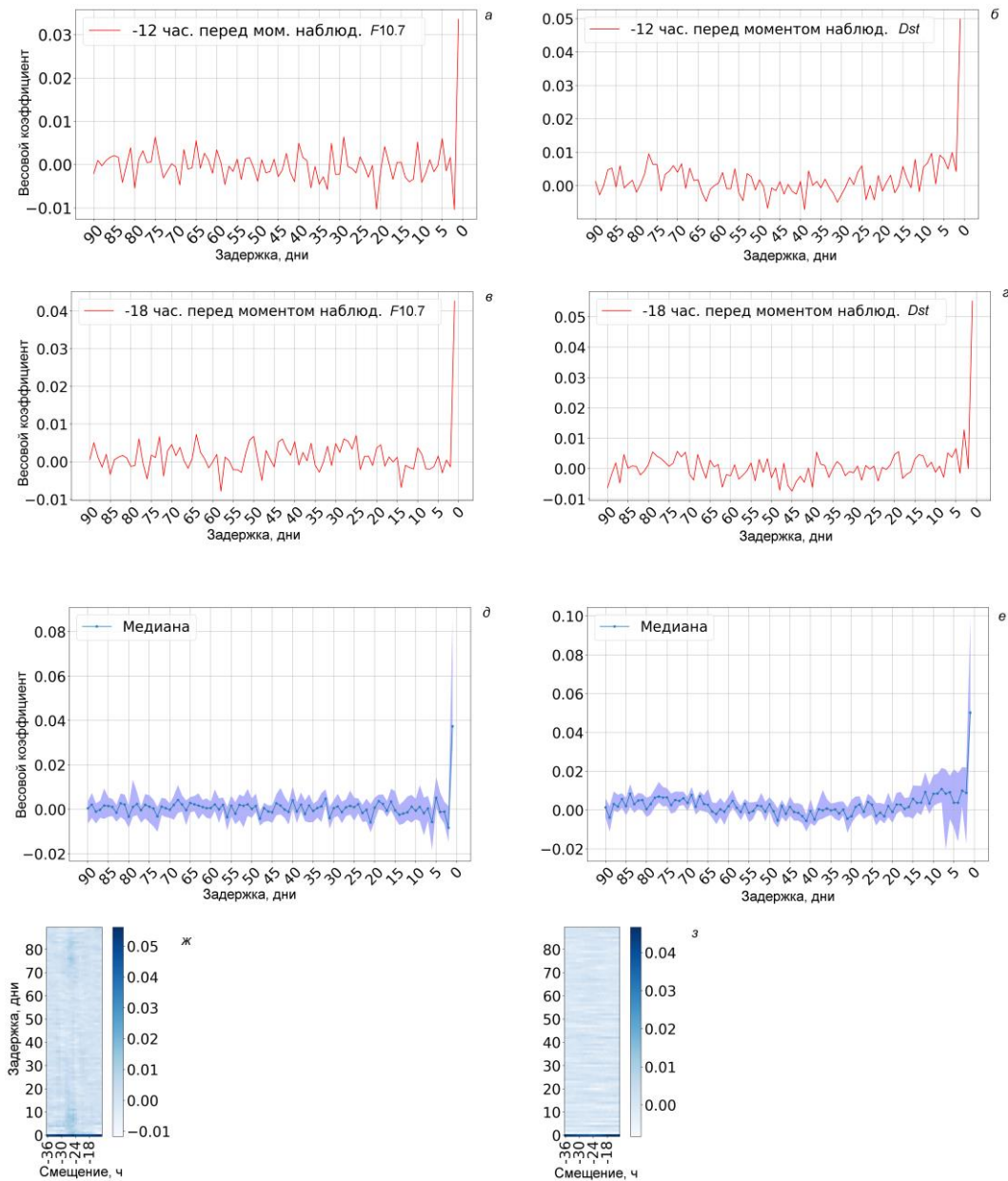


Рис. 8. Весовые коэффициенты нейрона, отвечающего за +12-й час прогноза, соответствующие –12-му (а, б) и –18-му (в, г) часу перед моментом наблюдения для F10.7 (а, в) и Dst (б, г). Медианные весовые коэффициенты сдвинутых матриц коэффициентов, соответствующие времени отстоящему на 24 часа назад от часа прогноза, в предыдущие дни для F10.7 (д) и Dst (е). Матрицы отклика на F10.7 (ж) и Dst (з) при прогнозе f_0F2

прогноза, должен вестись более аккуратно, в этой работе он не исследуется.

ИТОГОВАЯ МОДЕЛЬ ПРОГНОЗА f_0F2

Для использования модели нейронной сети проводится заново ее обучение с учетом особенностей характеристик матриц отклика — их практической независимости от момента наблюдения, продемонстрированной выше. Поэтому мы можем зафиксировать коэффициенты выходного слоя и матрицы откликов, полученные на соответствующих ансамблях (1, 10, 30, 100, 200 независимо обученных моделей), а также заново обучить сеть для уточнения формы ядра свертки.

Для итоговой архитектуры модели (см. рис. 3) коэффициенты выходного слоя были получены на предыдущем этапе, описанном выше, они не меняются в процессе обучения. Необходимо отметить, что в выходном слое итоговой модели матрицы коэффициентов отклика сдвинуты для каждого нейрона таким образом, чтобы максимум отклика соответствовал часу прогноза. Сдвиг осуществлялся по формуле

$$W_{i,j}^{l,k} = \theta_{SHIFTOUT(l,i,j)}^k, \quad (6)$$

где θ^k — матрица отклика k -го признака, размером (90×24) ; k — номер признака, признак один из f_0F2 , F10.7, Dst; W — матрица коэффициентов выходного слоя; i — индекс столбца матрицы весов от 0 до 23; l — час прогноза от +1 до +24; j — индекс дня задер-

Таблица 1

Показатели качества модели IRI и нейросетевых моделей при различных способах получения коэффициентов модели

Количество независимых обучений для усреднения коэффициентов	Pearson	RMSE, МГц	MAPE, %	R ²
1	0.93	0.595	10.502	0.863
10	0.929	0.596	10.486	0.862
30	0.929	0.596	10.468	0.862
100	0.93	0.596	10.532	0.862
200	0.928	0.598	10.445	0.861
Модель IRI	0.876	0.823	13.803	0.738

жки от 0 до 89; $SHIFTOUT(l, i)$ — функция вычисления сдвинутого номера часа в зависимости от индекса матрицы отклика, отвечающего за относительные смещения предшествующего часа от часа прогноза, и часа прогноза (7). Диапазон выходных значений функции $SHIFTOUT(l, i)$ от -36 до -13:

$$SHIFTOUT(l, i) = \begin{cases} i-l-23, & \text{если } i-l-23 \in [-36, -13] \\ (i-l-47), & \text{если } i-l-23 > -13 \\ i-l+1, & \text{если } i-l-23 < -36 \end{cases} \quad (7)$$

Таким образом, в последнем слое итоговой нейронной сети выходные нейроны отличаются коэффициентами, так как матрица отклика для каждого нейрона сдвинута особым образом по формуле (6).

Для каждого времени прогноза исходные данные преобразуются по методике, изложенной в разделе «Исходные данные для обучения сети». Это матрицы $f_0F2, F10.7, Dst$.

При обучении итоговой модели матрицы отклика собирались из соответствующего ансамбля матриц коэффициентов исходных моделей (1, 10, 30, 100, 200 моделей) и сдвигались в соответствии с формулой (6) для генерации матриц коэффициентов выходного слоя итоговой нейронной сети.

При обучении сети (нахождении коэффициентов свертки) исходная выборка данных была разделена в следующих пропорциях: 60 % — обучающая выборка, 20 % — валидационная выборка, 20 % — тестовая выборка.

При оценке качества работы модели использовались четыре метрики: коэффициент корреляции Пирсона, квадратный корень среднеквадратичной ошибки (RMSE), средняя абсолютная ошибка в процентах (MAPE) и коэффициент детерминации (R^2). Сравнивались реальные данные тестовой выборки, данные, прогнозируемые результирующей моделью, и данные прогнозируемые моделью IRI. На рис. 9, а показаны результаты прогноза суточного хода f_0F2 . В табл. 1 даны показатели качества итоговой модели при генерации матриц отклика на основе ансамблей 1, 10, 30, 100 и 200 независимых моделей, а также сравнение с показателями качества модели IRI. В работе использовалась модель IRI-2016 с опциями URSI.

Приведенные данные позволяют сделать вывод, что улучшения качества модели при увеличении количества независимых обучений коэффициентов матриц отклика практически не происходит. Таким образом,

стационарность матрицы отклика (ее независимость от момента наблюдения) является основным фактором, влияющим на метрики качества данной модели.

Для сравнения в [Салимов, Хмельнов, 2020] приведены показатели качества прогнозирования f_0F2 на 24 часа вперед для рекуррентной нейронной LSTM сети на данных Иркутска: коэффициент корреляции Пирсона — 0.923, RMSE — 0.605 МГц, MAPE — 10.07 %. Данные показатели качества схожи с показателями качества нашей модели, что позволяет сделать вывод о важности учета предшествующих ионосферных наблюдений при построении прогнозных моделей.

Для анализа поведения модели на данных с пропусками был проведен эксперимент: в тестовую выборку были специально добавлены пропуски таким образом, чтобы их распределение было неотличимо от распределения пропусков в обучающей выборке (достигнутое значение p -value — 0.11 по критерию Манна-Уитни при сравнении распределений пропусков измененной тестовой и обучающей выборок). Метрики качества модели, основанной на усредненных коэффициентах из 200 обучений, для такой тестовой выборки следующие: коэффициент корреляции Пирсона — 0.923; RMSE — 0.598 МГц; MAPE — 10.45 %; R^2 — 0.86. Данные показатели практически совпадают с метриками качества модели на неизменной тестовой выборке (см. предпоследнюю строку табл. 1), что позволяет сделать вывод об устойчивости модели к пропускам данных, наиболее характерным для текущих измерений на Иркутском дигизонде.

Для анализа качества работы модели в зависимости от часа прогноза были вычислены соответствующие метрики отдельно для каждого часа прогноза. Они показаны на рис. 10. Можно сделать вывод, что наиболее качественный прогноз дается для 1-го часа, далее к 24-му часу метрики качества немного ухудшаются.

На рис. 9 в, г показаны графики прогноза f_0F2 во время геомагнитных бурь (четверо суток после и сутки до снижения геомагнитного индекса K_p ниже 6). Видно, что модель хуже реагирует на пик геомагнитной бури и с некоторым опозданием адаптируется к успокаивающимся условиям. Таким образом, предложенная модель описывает скорее невозмущенную ионосферу и в сильно возмущенных условиях не рекомендуется к применению из-за роста ошибки.

На рис. 11 показаны результаты работы модели в различные сезоны. Видно, что в первом приближении суточный ход отслеживается моделью удовлетворительно вне зависимости от сезона.

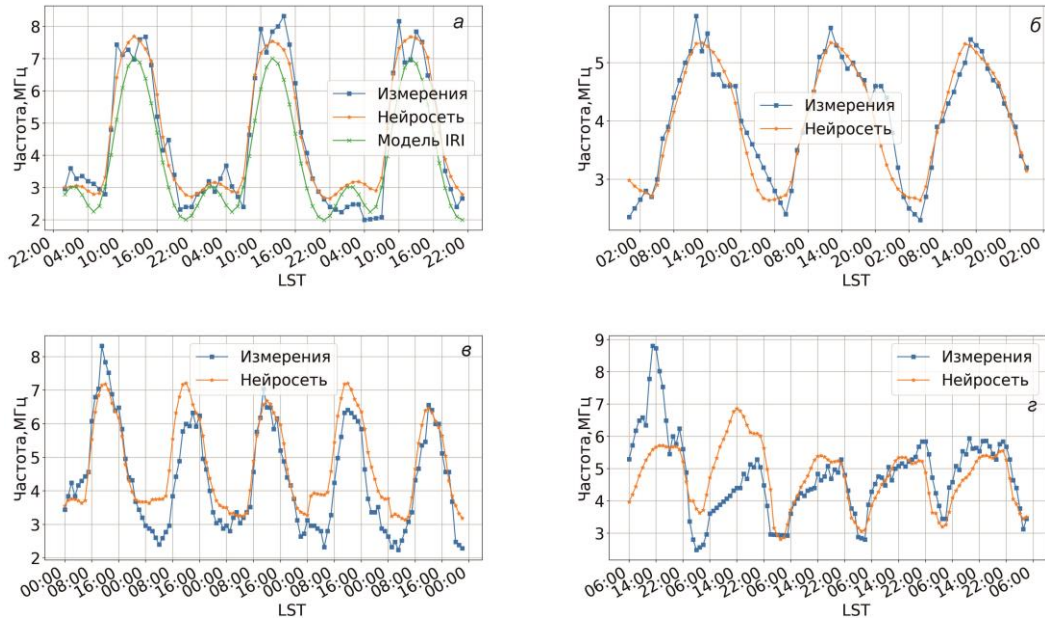


Рис. 9. Прогноз f_oF_2 с помощью нейронной сети и модели IRI и суточный ход f_oF_2 с 14 декабря 2015 г. 00:00 LST по 16 декабря 2015 г. 22:00 LST (а), с 25 октября 2016 г. 00:00 LST по 29 октября 2016 г. 23:00 LST (б), с 8 мая 2016 г. 06:00 LST по 13 мая 2016 г. 05:00 LST (в) для Иркутска для итоговой модели с коэффициентами, рассчитанными на ансамбле из 200 независимых обучений; графики прогноза частоты f_oF_2 , ее суточного хода для Арты ($56^{\circ}25' N$, $58^{\circ}32' E$) с 13 апреля 2017 г. 00:00 LST по 16 апреля 2017 г. 00:00 LST (г)

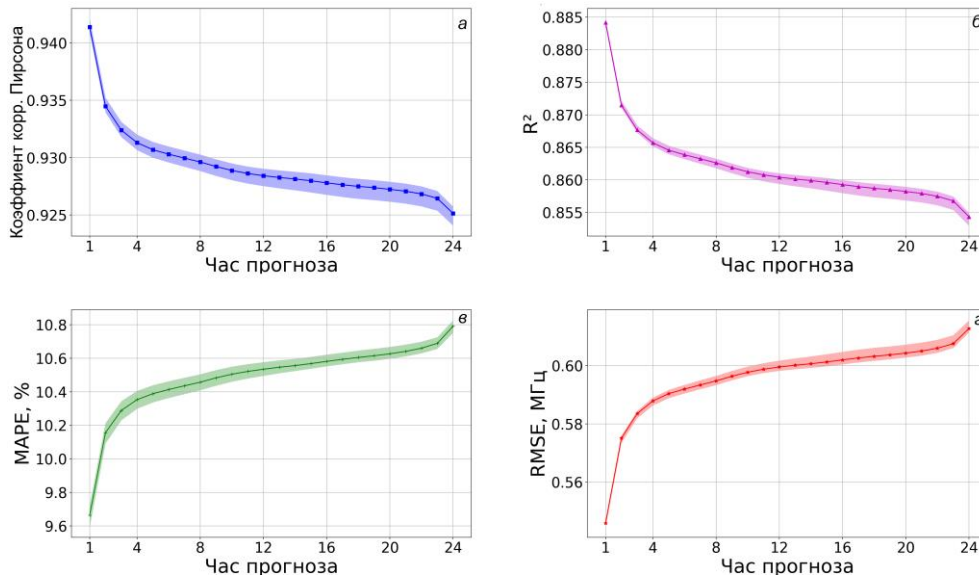


Рис. 10. Метрики качества моделей в зависимости от часа прогноза. Линии — среднее арифметическое значение по моделям, цвет — разброс для разных итоговых моделей, основанных на откликах, вычисленных с использованием 1, 10, 30, 100, 200 исходных моделей

ПРОВЕРКА МОДЕЛИ НА ДАННЫХ ДРУГИХ СРЕДНЕШИРОТНЫХ ИОНОЗОНДОВ

При создании моделей прогноза важна их обобщающая способность, которая может быть проверена показателями качества на различных тестовых выборках, не пересекающихся с обучающей выборкой. В данной работе в качестве таких тестовых выборок дополнительно к тестовой выборке Иркутска были выбраны данные трех среднелиричных ионозондов

близких по широте к Иркутску, но существенно разнесенных по долготе: Арты (Россия, $56^{\circ}25' N$, $58^{\circ}32' E$), Варшава (Польша, $52^{\circ}13' N$, $21^{\circ}02' E$), Мохе (Китай, $52^{\circ}58' N$, $122^{\circ}31' E$) за 2017 год.

Проверка осуществлялась следующим образом.

- Преобразование данных соответствующего ионозонда к нужному формату (набор матриц 90×24 и векторов размером 24).
- Масштабирование по формуле (1) полученных данных с помощью коэффициентов стандартизации μ и s , полученных при стандартизации обучающих выборок Иркутска. Эти коэффициенты одинаково приме-

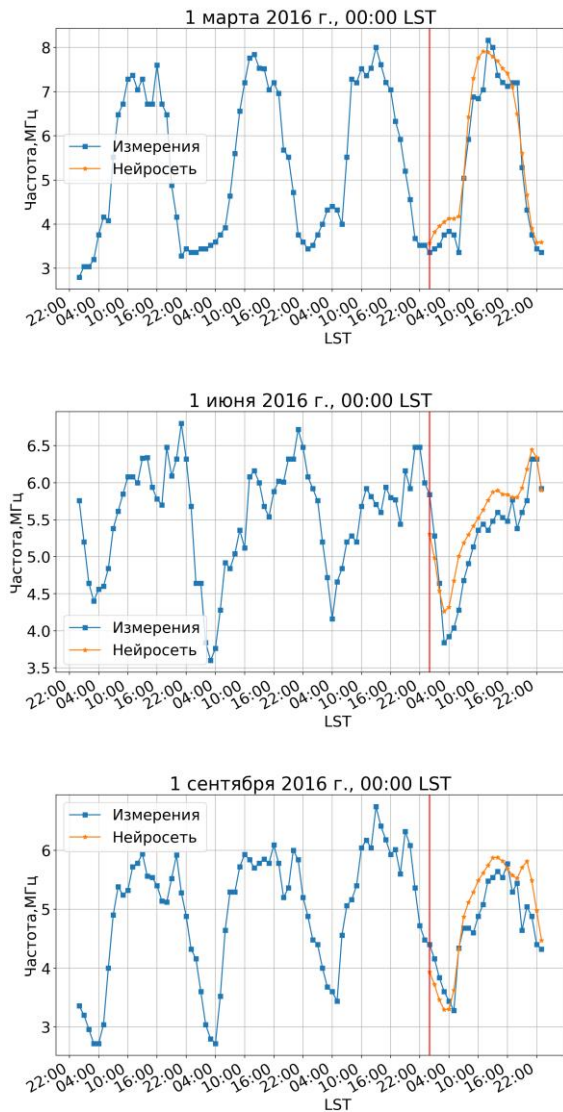


Рис. 11. Прогноз критической частоты f_oF2 над Иркутском после точек измерения 1 марта 2016 г., 00:00 LST (а); 1 июня 2016 г., 00:00 LST (б); 1 сентября 2016 г., 00:00 LST (в). Точки измерения совпадают с началом вертикальной линии

Таблица 2

Коэффициенты стандартизации обучающих выборок Иркутска

Признак		f_oF2 , МГц	$F10.7$	Dst
Исходная модель («Модель А»)	μ	5.403	110.795	-9.814
	s	1.989	31.779	15.051
Итоговая модель	μ	5.068	100.864	-9.458
	s	1.672	26.995	14.628

няются для данных всех радаров и приведены в табл. 2. Для исходной и итоговой моделей эти данные будут отличаться из-за различающегося способа разделения выборок.

- Использование этих данных для прогноза с помощью предложенной нами нейросетевой модели.
- Для сравнения проведен прогноз f_oF2 моделью, построенной и обученной без предположения о стационарности матрицы отклика (Модель А).
- Сравнение прогнозируемой частоты f_oF2 со значением, измеренным в данном пункте.

Показатели качества прогноза, полученные в результате тестирования данных среднеширотных ионозондов, даны в табл. 3. Графики прогноза f_oF2 и ее суточного хода для Арты показаны на рис. 9, б.

Из таблицы следует, что показатели качества результирующей модели слабо зависят от количества моделей в ансамбле, однако эти показатели лучше результатов исходной Модели А. Соответственно можно сделать вывод о том, что учет стационарности матриц отклика дает прирост в качестве прогноза. Результат говорит также о хорошей общности построенных моделей, абсолютная ошибка которых составляет ~ 0.6 МГц, и слабо зависит от долготы станции. Это позволяет сделать вывод о пригодности обученной нами модели для прогноза f_oF2 по данным среднеширотных ионозондов.

ВЫВОДЫ

Исследовалось влияние предшествующих моменту наблюдения 90 сут на прогноз критической частоты f_oF2 . Получена прогнозная модель f_oF2 на сутки вперед на основе сверточных сетей по данным f_oF2 , $F10.7$ и Dst . Коэффициенты модели получены усреднением коэффициентов, полученных по 200 независимым обучением исходной модели, последующим усреднением коэффициентов сети в предположении о стационарности матриц отклика, а также дополнительным обучением итоговых моделей для уточнения коэффициентов входного сверточного слоя.

Приведен метод усреднения коэффициентов модели по ансамблю из независимо обученных моделей и ансамблю из 24 сдвинутых матриц с учетом их стационарности, являющийся для данной архитектуры сети неким аналогом ансамблевого метода bagging.

Показано, что вклад параметров в прогнозное значение f_oF2 в первом приближении не зависит от локального времени (при прогнозе на 24 часа и менее), а только от временного интервала между соответствующим измеренным значением и моментом, на который считается прогноз. Это говорит о стационарности матриц отклика. Показано, что основной вклад в прогнозное значение f_oF2 вносят данные ближайших нескольких дней перед прогнозом, вклад остальных дней существенно убывает, что понятно с физической точки зрения.

Показано, что учет стационарности матрицы отклика позволяет улучшить метрики качества прогноза модели по сравнению с моделью, не использующей усреднение весовых коэффициентов (см. табл. 3).

Показано, что при увеличении количества независимых моделей, используемых для усреднения коэффициентов от 1 до 200, качество прогноза практически не улучшается.

Итоговая модель обеспечивает следующие метрики качества прогноза: коэффициент корреляции Пирсона — 0.928; корень среднеквадратичной ошибки (RMSE) — 0.598 МГц; средняя абсолютная ошибка в процентах (MAPE) — 10.45 %; коэффициент детерминации — 0.861. В сравнении с моделью IRI предложенная нами модель выдает лучшие показатели качества (средняя абсолютная ошибка в процентах 13.803 % для модели IRI и 10.45 % для предлагаемой нами модели). Что касается приемлемой точности

Таблица 3

Показатели (метрики) качества моделей по данным различных среднеширотных ионозондов. Модель А построена и обучена без предположения о стационарности матрицы отклика (исходная модель). Показатели качества Модели А усреднены по набору из 200 независимых обучений модели

	Модель	Pearson	RMSE, МГц	MAPE, %	R^2
Арти	Модель А	0.848	0.669	14.242	0.683
	1 (итоговая модель)	0.874	0.613	13.506	0.735
	10 (итоговая модель)	0.872	0.596	12.818	0.750
	30 (итоговая модель)	0.870	0.609	13.194	0.739
	100 (итоговая модель)	0.867	0.629	13.966	0.721
	200 (итоговая модель)	0.875	0.591	12.584	0.754
Мохе	Модель А	0.821	0.632	12.622	0.624
	1 (итоговая модель)	0.847	0.556	11.080	0.710
	10 (итоговая модель)	0.844	0.554	10.763	0.712
	30 (итоговая модель)	0.842	0.560	11.042	0.705
	100 (итоговая модель)	0.842	0.567	11.353	0.699
	200 (итоговая модель)	0.845	0.554	10.687	0.713
Варшава	Модель А	0.809	0.708	14.279	0.613
	1 (итоговая модель)	0.841	0.626	12.398	0.698
	10 (итоговая модель)	0.838	0.622	12.072	0.702
	30 (итоговая модель)	0.841	0.621	12.192	0.703
	100 (итоговая модель)	0.837	0.634	12.599	0.691
	200 (итоговая модель)	0.839	0.622	12.015	0.703

модели, требование к величине RMSE зависит от конкретной решаемой задачи. Очевидно, чем она меньше, тем лучше. Для качественных задач оценки распространения радиоволн достаточной точностью, видимо, можно считать $RMSE \sim 0.7$ МГц, что позволяет разделять магнитоионные компоненты в среднеширотной ионосфере.

На основе данных значительно разнесенных по долготе ионозондов (Арти, Варшава, Мохе) показано, что построенная модель может быть использована для прогноза ионосферы по данным других среднеширотных ионозондов без изменения коэффициентов этой модели, что говорит о хорошей общности модели.

Исследование выполнено при финансовой поддержке Минобрнауки России. Результаты получены с использованием оборудования Центра коллективного пользования «Ангара» [<https://ckp-rf.ru/catalog/ckp/3056>].

СПИСОК ЛИТЕРАТУРЫ

- Бархатов Н.А., Ревунов С.Е., Урядов В.П. Технология искусственных нейронных сетей для прогнозирования критической частоты ионосферного слоя F2. *Изв. вузов. Радиофизика*. 2005. Т. 48, № 1. С. 1–15.
- Ратовский К.Г., Потехин А.П., Медведев А.В., Куркин В.И. Современный цифровой ионозонд DPS-4 и его возможности. *Солнечно-земная физика*. 2004. Вып. 5. С. 102–104.

Салимов Б.Г., Хмельнов А.Е. Предсказание критической частоты ионосферы f_oF_2 с помощью нейронной рекуррентной LSTM сети. *Материалы конференции «Ляпуновские чтения»*. Иркутск, 7–11 декабря 2020 г. Иркутск: ИДСТУ СО РАН, 2020. С. 60–61.

Смирнов В.Ф., Степанов А.Е. Новые возможности в исследованиях высокоширотной ионосферы: дигизонд DPS-4 — первые результаты по измерениям локализации и динамики крупномасштабных структур ионосферы в Якутске. *Солнечно-земная физика*. 2004. Вып. 5 (118). С. 105–106.

Харгривс Дж. К. Верхняя атмосфера и солнечно-земные связи. *Введение в физику околоземной космической среды*. Ленинград, Гидрометеоздат, 1982. 351 с.

Abadi M., Barham P., Chen J., et al., TensorFlow: A system for large-scale machine learning. *Proc. OSDI*. 2016. P. 265–283. DOI: [10.5281/zenodo.4724125](https://doi.org/10.5281/zenodo.4724125).

Bilitza D., Mckinnell L.-A., Reinisch B., Fuller-Rowell T. The International Reference Ionosphere (IRI) today and in the future. *J. Geodesy*. 2011. Vol. 85. DOI: [10.1007/s00190-010-0427-x](https://doi.org/10.1007/s00190-010-0427-x).

Boulch A., Cherrier N., Castaings T. *Ionospheric activity prediction using convolutional recurrent neural networks*. 2018. DOI: [10.48550/arXiv.1810.13273](https://doi.org/10.48550/arXiv.1810.13273).

Breiman L. Bagging Predictors. *Technical Report*. 1994. No. 421.

Bring J. How to standardize regression coefficients. *The American Statistician*. 1994. Vol. 48, no. 3. P. 209–213. DOI: [10.2307/2684719](https://doi.org/10.2307/2684719).

Consultative Committee on International Radio (CCIR) *Atlas of Ionospheric Characteristics Report 340*. International Telecommunication Union, Geneva, Switzerland, 1967.

Galkin I.A., Reinisch B., Vesnin A.M., Huang X. Assimilation of sparse continuous groundbased ionosonde data into IRI using NECTAR model morphing. *The 1st URSI Atlantic Radio Science Conference (URSI AT-RASC)*. Las Palmas, 2015. P. 1–8, DOI: [10.1109/URSI-AT-RASC.2015.7303112](https://doi.org/10.1109/URSI-AT-RASC.2015.7303112).

Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016. 800 p.

Kingma D.P., Ba J.A. A Method for Stochastic Optimization. *International Conference on Learning Representations*. 2014. DOI: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).

Lundberg S., Lee S.-I. A Unified Approach to Interpreting Model Predictions. arXiv. 2017. DOI: [10.48550/ARXIV.1705.07874](https://doi.org/10.48550/ARXIV.1705.07874).

Opitz D., Maclin R. Popular ensemble methods: an empirical study. *J. Artificial Intelligence Res.* 1999. Vol. 11. P. 169–198. DOI: [10.1613/jair.614](https://doi.org/10.1613/jair.614).

Rush C., Fox M., Bilitza D., et al. Ionospheric mapping – an update of f_oF2 coefficients. *Telecommun. J.* 1989. Vol. 56. P. 179–182.

Sivavaraprasad G., Lakshmi Mallika I., Sivakrishna K., Venkata Ratnam D. A novel hybrid Machine learning model to forecast ionospheric TEC over low-latitude GNSS stations. *Adv. Space Res.* 2022. Vol. 69, iss. 3. P. 1366–1379. DOI: [10.1016/j.asr.2021.11.033](https://doi.org/10.1016/j.asr.2021.11.033).

Yu S., Ma J. Deep learning for geophysics: Current and future trends. *Rev. Geophys.* 2021. Vol. 59, iss. 3. e2021RG000742. DOI: [10.1029/2021RG000742](https://doi.org/10.1029/2021RG000742).

URL: <http://irimodel.org/IRI-2016> (дата обращения 27 июня 2019 г.).

URL: <https://ckp-rf.ru/catalog/ckp/3056> (дата обращения 3 октября 2018 г.).

URL: <https://omniweb.gsfc.nasa.gov/form/dx1.html> (дата обращения 31 августа 2020 г.).

Как цитировать эту статью:

Салимов Б.Г., Бернгардт О.И., Хмельнов А.Е., Ратовский К.Г., Кусонский О.А. Применение сверточных нейронных сетей для прогнозирования критической частоты f_oF2 . *Солнечно-земная физика*. 2023. Т. 9, № 1. С. 60–72. DOI: [10.12737/szf-91202307](https://doi.org/10.12737/szf-91202307).