

УДК 550.388

ИСПОЛЬЗОВАНИЕ ОБЩЕДОСТУПНЫХ АРХИВОВ ДАНЫХ ОБРАБОТКИ ИОНОГРАММ ВЕРТИКАЛЬНОГО РАДИОЗОНДИРОВАНИЯ В РОЛИ ПЕРВОНАЧАЛЬНОЙ РАЗМЕТКИ ДЛЯ МАШИННОГО ОБУЧЕНИЯ

А. О. Щирый^[0000-0001-7312-4424]

*Институт земного магнетизма, ионосферы и распространения радиоволн
им. Н. В. Пушкова РАН, г. Москва, Россия*

andreyschiriy@gmail.com

Аннотация

В работе выдвинута идея использования имеющихся больших массивов результатов обработки ионограмм вертикального радиозондирования ионосферы в качестве обучающих датасетов для построения предиктивных моделей методами машинного обучения. Рассмотрены наиболее популярные форматы сохранения результатов обработки ионограмм, а также некоторые интернет-ресурсы с архивами свободно доступных файлов этих форматов. Указанные датасеты используются для построения предиктивных моделей, в том числе временных рядов критических частот ионосферных слоев. Отмечена также возможность использования некоторых датасетов результатов обработки ионограмм для обучения моделей, предназначенных для автоматической обработки ионограмм.

Ключевые слова: *ионосфера, радиозондирование, вертикальное зондирование ионосферы, ионограмма, характеристики ионосферы, метаданные.*

ВВЕДЕНИЕ

Исследования ионосферы востребованы как для фундаментальной науки (для построения и уточнения моделей геофизических процессов), так и для прикладных целей [1–9]. По данным радиозондирования ионосферы коротковолновыми сигналами (КВ) можно получить информацию о процессах в ионосферной плазме, ее структуре и состоянии. Эти данные также крайне важны для систем КВ-радиосвязи, радиолокации и радионавигации, для компенсации

искажений, вносимых ионосферной средой распространения КВ. В частности, работа таких радиотехнических систем (РТС) КВ-диапазона, как системы дальней КВ-радиосвязи и загоризонтной радиолокации, основана на способности КВ многократно отражаться от ионосферы и земной поверхности. Ключевой задачей для обеспечения корректного функционирования таких РТС является адаптация к ионосферным условиям. При этом возможна адаптация РТС КВ-диапазона по рабочей частоте, мощности излучения и скорости передачи информации. Перспективные методы адаптации предполагают даже измерение и компенсацию передаточной функции КВ-радиолинии [9]. Информация об ионосферных процессах важна также для корректного функционирования систем глобального позиционирования (GPS, ГЛОНАСС), хотя эти системы работают не в КВ-диапазоне, но ионосфера является поглощающей средой для радиоволн их рабочих диапазонов, и учет данных об ионосферном поглощении критически важен для функционирования таких систем, в частности для обеспечения требуемой точности определения координат.

Для адаптации к ионосферным условиям необходимо производить оперативную диагностику ионосферы. Из методов такой диагностики наибольший интерес представляют вертикальное, наклонное и возвратно-наклонное радиозондирование ионосферы (соответственно ВЗИ, НЗИ, ВНЗИ) [1–6]. Далее нами будет рассмотрена работа с данными только ВЗИ.

Ионограмма ВЗИ представляет собой зависимость амплитуды радиосигнала от частоты и высоты отражения; пример ионограммы ВЗИ приведен на рис. 1, также на этом рисунке поверх ионограммы построен восстановленный по ней модельный профиль электронной концентрации. На рис. 2. представлен другой пример ионограммы ВЗИ, где кроме профиля электронной концентрации «поверх» ионограммы прорисованы автоматически выделенные и идентифицированные моды КВ-радиосигнала с отражениями от слоев E, F1, F2 («обыкновенная»), а также от F2 – «необыкновенная» мода X.

Обработка данных ВЗИ включает выделение треков мод на ионограмме, определение высот и критических частот ионосферных слоев, а также восстановление профилей электронной концентрации.

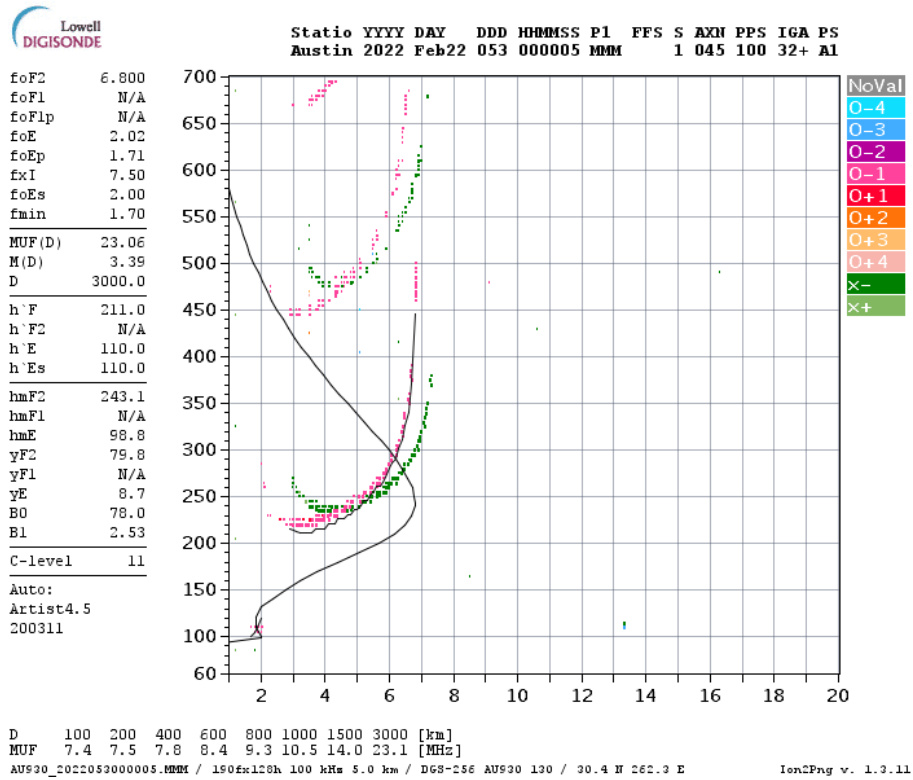


Рис. 1. Пример ВЗИ-ионограммы «дигизонда» DPS в формате картинки

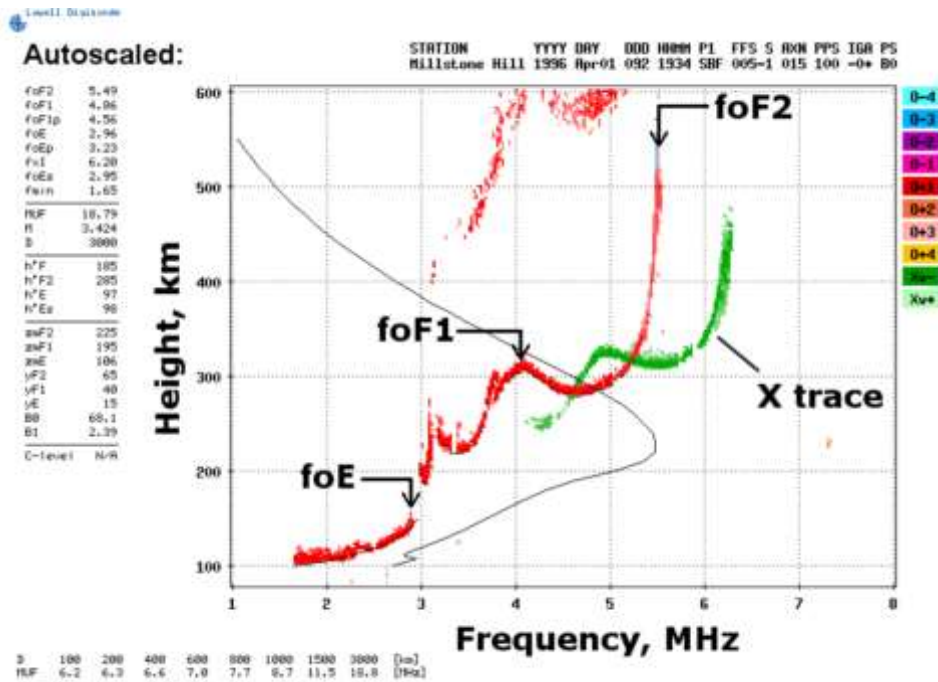


Рис. 2. ВЗИ-ионограмма «дигизонда» DPS с указанием названий мод

ОСНОВНАЯ ИДЕЯ

Развитие машинного обучения (МО) открывает новые возможности для автоматизированной обработки, анализа и интерпретации ионограмм, а также использования результатов обработки ионограмм. Интерес к большим массивам данных радиозондирования ионосферы обусловлен, кроме прочего, возможностью «тренировки» статистических моделей методами МО, тем более если наряду с «сырыми» данными радиозондирования (ионограммы) имеются и обработанные данные – выделенные треки различных мод распространения радиосигнала, критические частоты ионосферных слоев и др. (т.е. в терминах МО – имеются размеченные датасеты, или наборы данных). Напомним, что для построения моделей методами МО (моделей класса «обучение с учителем») требуются размеченные датасеты, т. е. наборы данных вида «анализируемые данные» – «результаты анализа». Таким образом, архивы данных радиозондирования, содержащие и ионограммы, и результаты их обработки, могут быть использованы для создания средствами МО статистических моделей, причем моделей предназначенных для решения двух принципиально различных типов задач. Во-первых, моделей для автоматической обработки ионограмм, прежде всего для решения задачи выделения треков – это разновидность задачи сегментации, возможно с последующей классификацией треков. Во-вторых, предиктивных моделей состояния самой ионосферы.

Конечно, уже предпринимались попытки [10–15] применения методов МО, а конкретно нейронных сетей, для решения задач обработки ионограмм ВЗИ (прежде всего для выделения треков на ионограммах). Однако в отличие от подхода, предлагаемого в настоящей работе, в указанных работах ионограммы размечаются «вручную» в ходе самого исследования, поэтому обучающие датасеты сравнительно малы (тысячи ионограмм).

ФОРМАТЫ ХРАНЕНИЯ РЕЗУЛЬТАТОВ ОБРАБОТКИ ИОНОГРАММ

Наиболее популярными форматами хранения результатов обработки ионограмм являются форматы SAO [16, 17] и SAOXML [18, 19]. Подчеркнем, что они не предназначены для хранения самих ионограмм, а нужны для хранения результатов обработки ионограмм (выделенные треки мод распространения КВ-радиосигнала, идентифицированные ионосферные слои и их высоты, вос-

становленные профили электронной концентрации, критические частоты ионосферных слоев, плазменные частоты, доплеровские сдвиги и др.). Хотя форматы SAO и SAOXML были разработаны создателями «дигизондов» – ионозондов ВЗИ типа DPS [21], тем не менее широко используются для указанных целей не только с «дигизондами», но и с ионозондами других типов. Историческим предшественником форматов SAO и SAOXML был бинарный формат ART, созданный также разработчиками «дигизондов» и встречающийся в открытых архивах до сих пор, однако совершенно не документированный. Форматы SAO и SAOXML являются текстовыми, причем SAOXML (как можно понять по названию) основан на XML.

Логическая модель SAOXML [19] показана на рис. 3; видно, что (кроме метаданных, характеризующих дату-время, место и параметры съема ионограммы) имеются и данные, характеризующие объекты на ионограмме (ветка TraceList, содержащая список выделенных треков мод распространения КВ-радиосигнала), и данные, характеризующие ионосферу, такие как высоты и критические частоты ионосферных слоев (ветка CharacteristicList), и восстановленные профили электронной концентрации (ветка ProfileList). Более архаичный формат SAO содержит аналогичную информацию и, хотя чуть менее удобен для чтения, чем SAOXML, является его функциональным аналогом.

В известных нам работах с данными форматов SAO и SAOXML, разработчики, как правило, не идут дальше создания программ просмотра (см., например, [20]).

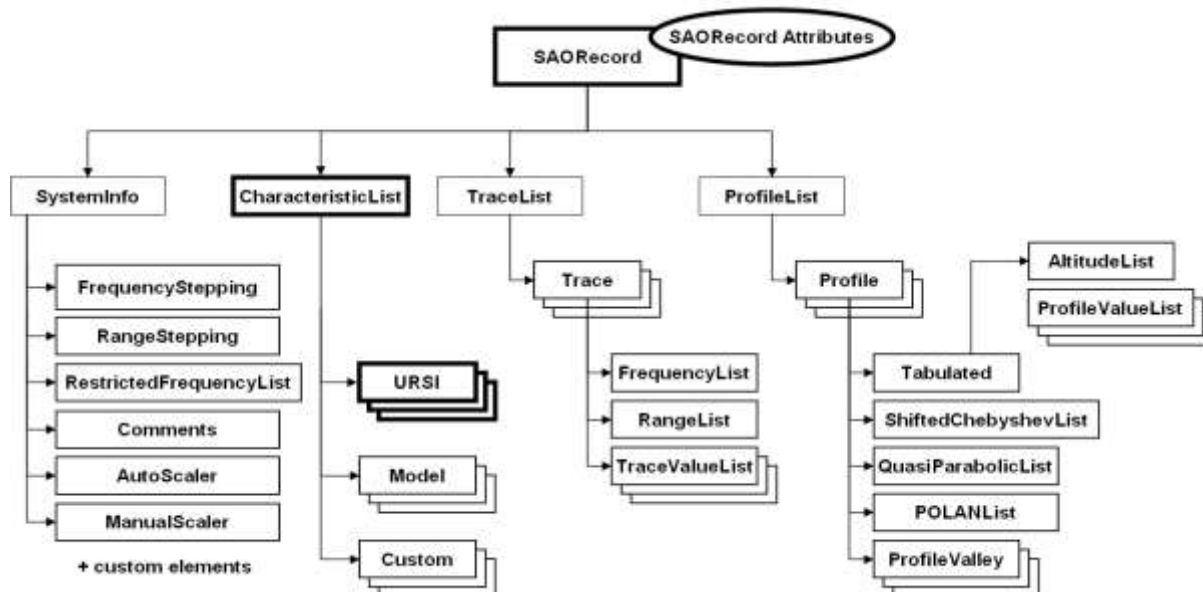


Рис. 3. Логическая модель SAOXML [19]

Извлеченные из последовательности файлов SAOXML (или SAO) разного рода характеристики ионосферы позволяют сформировать их временные ряды. На рис. 4 и 5 в качестве примера показаны временные ряды критической частоты слоя F2 (foF2) для двух различных пунктов зондирования и дат.

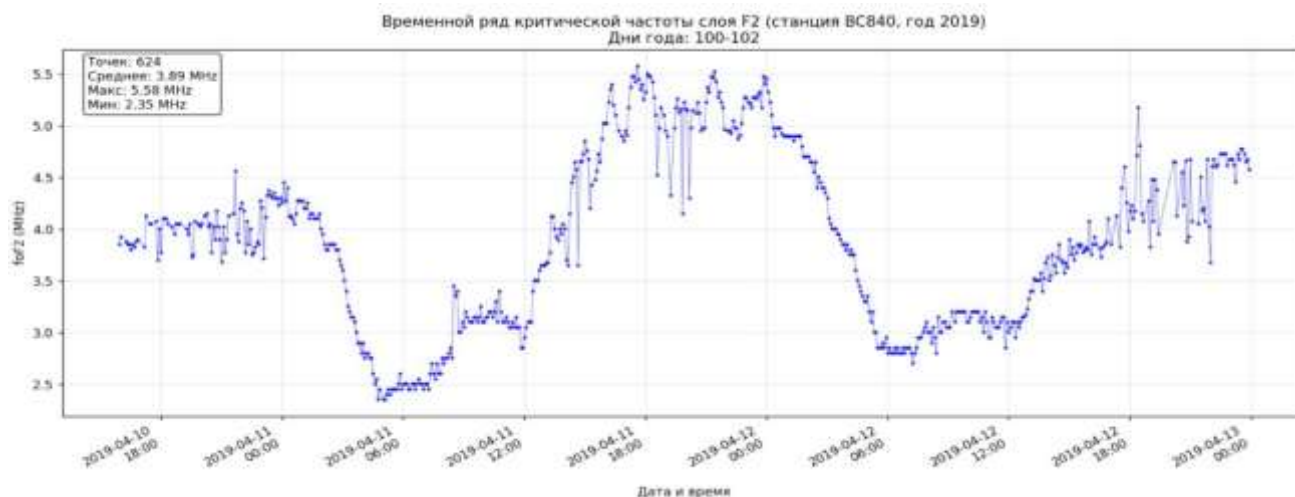


Рис. 4. Временной ряд foF2, станция BC840

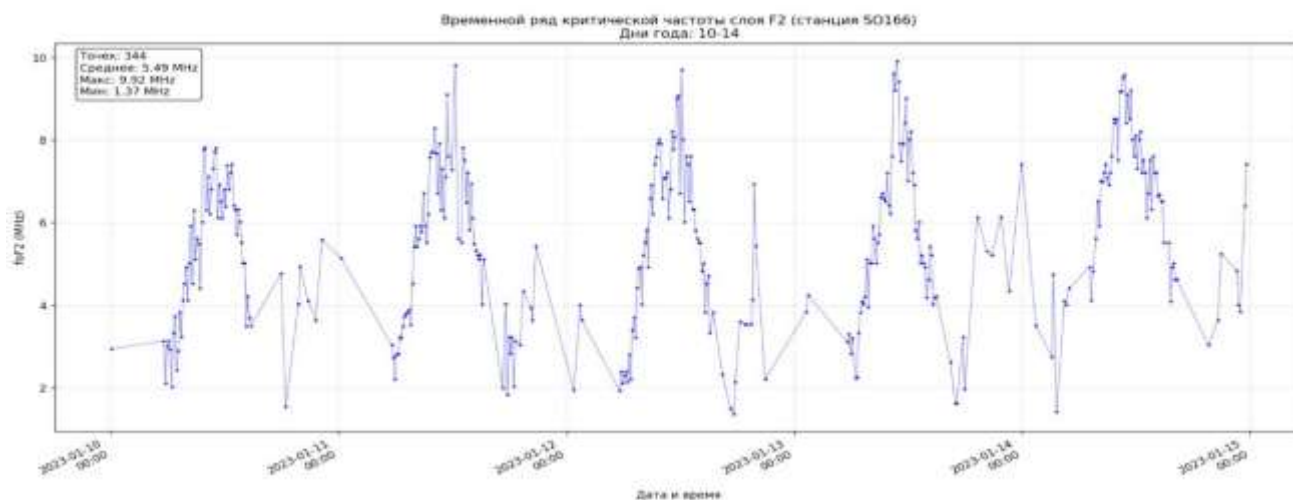


Рис. 5. Временной ряд foF2, станция SO166

Кроме указанных форматов сохранения результатов обработки ионограмм ВЗИ, разумеется, существуют и другие, например форматы австралийских ионозондов IPS [23], где результаты обработки сохраняются в файлах двух видов: CLEAN и AUTO, последний является функциональным аналогом форматов SAO и SAOXML. Что же касается формата CLEAN [24], то он содержит битовые маски выделенных сигнальных компонент «полезного» сигнала ионограммы (отнесение пикселей к разным трекам мод отсутствует); файл состоит из трех основных частей: заголовка, списка частотных каналов и массивов амплитуд для каждого из частотных каналов, при этом массив закодирован для сжатия особым алгоритмом. Просмотр данных CLEAN ионограмм показал, что кроме пикселей треков мод сигнала на ионограммах имеются остаточные шумы, поэтому для использования таких файлов в качестве размеченного датасета (очевидно, для решения задач выделения сигнала на ионограмме методами МО) необходима предобработка, заключающаяся как минимум в удалении «мелких объектов». Упомянутый же формат AUTO, по аналогии с SAO и SAOXML, содержит данные которые могут быть использованы для построения предиктивных моделей ионосферных характеристик.

РЕСУРСЫ АРХИВОВ ИОНОГРАММ И РЕЗУЛЬТАТОВ ОБРАБОТКИ

В работе [21] нами рассмотрены интернет-ресурсы со свободным доступом к данным вертикального радиозондирования ионосферы, в том числе Национального центра геофизических данных США [22], Службы космической

погоды Бюро метеорологии Австралии [23] и некоторые другие. Однако в настоящей работе особое внимание обращаем на ресурсы [22, 23], поскольку они содержат десятки миллионов файлов ионограмм и результатов обработки ионограмм. Наиболее популярные форматы данных самих ионограмм рассмотрены нами ранее в [21], наиболее актуальны (с точки зрения числа доступных ионограмм) «дигизондовские» форматы ионограмм MMM, SBF, RSF, а также форматы ионограмм австралийских ионозондов IPS.

ЗАКЛЮЧЕНИЕ

Таким образом, представленная в работе идея состоит в том, чтобы использовать имеющиеся большие массивы результатов обработки ионограмм ВЗИ в качестве обучающих датасетов для построения моделей методами МО. Были проанализированы наиболее популярные форматы сохранения результатов обработки ионограмм, а также некоторые интернет-ресурсы с архивами таких файлов. Показано, что представленные в них данные делятся на две группы: характеристики объектов ионограммы и характеристики ионосферы. С точки зрения задач МО данные первой группы могут использоваться для построения моделей, предназначенных для автоматической обработки ионограмм, а второй группы – для построения предиктивных моделей характеристик ионосферы, в том числе моделей временных рядов параметров ионосферных слоев (например, критических частот).

СПИСОК ЛИТЕРАТУРЫ

1. Радиозондирование ионосферы спутниковыми и наземными ионозондами / Под ред. С.И. Авдюшина. Труды института прикладной геофизики им. академика Е.К. Федорова. М.: ИПГ, 2008.

URL: <http://ipg.geospace.ru/publications/book-2008.pdf>

2. *Reinisch B., Sales G.* Measuring Electrodynamics of the Ionosphere by Digital Ionosondes and Other Techniques // Scientific Report No. 3. 2001.

URL: <https://apps.dtic.mil/sti/tr/pdf/ADA402979.pdf> (дата обращения 07.05.2025)

3. *Акимов В.Ф., Калинин Ю.К.* Введение в проектирование ионосферных загоризонтных радиолокаторов. М.: Техносфера, 2017. 492 с.

4. *Giuseppe Fabrizio*. High Frequency Over-the-Horizon Radar: Fundamental Principles, Signal Processing, and Practical Applications. McGraw-Hill Education, 2013.

5. *Щиры́й А.О.* Разработка и моделирование алгоритмов автоматического измерения характеристик ионосферных коротковолновых радиолиний: Автореф. дис. ... канд. техн. наук: Санкт-Петербургский гос. ун-т телекоммуникаций им. проф. М.А. Бонч-Бруевича. СПб., 2007. 19 с. – EDN: QMISDO

6. *Щиры́й А.О.* Развитие средств автоматизации наземного радиозондирования ионосферы // Фундаментальные проблемы радиоэлектронного приборостроения. 2014. №5. С. 170–173.

7. *Щиры́й А.О.* Архитектура программной части аппаратно-программного комплекса дистанционного наземного радиозондирования ионосферы // Новые информационные технологии в автоматизированных системах. 2015. №18. С. 144–152.

8. *Щиры́й А.О.* Алгоритмы и программное обеспечение автоматизации процессов измерений и обработки данных оперативной диагностики ионосферы и ионосферных радиолиний // Журнал радиоэлектроники. 2022. №10. <https://doi.org/10.30898/1684-1719.2022.10.4>

9. *Shiriy A.O.* HF channel transmit function module measurement // Proceedings of the 5th International Conference on Actual Problems of Electron Devices Engineering, APEDE 2002. Vol. 5. 2002. P. 365–369. <https://doi.org/10.1109/apede.2002.10449645>

10. *Долгачева С.А., Макарова Л.Н., Николаев А.В.* Обработка ионограмм высокоширотных станций вертикального зондирования с использованием нейронных сетей: Es и F2 слои // Physics of Auroral Phenomena. 2020. Т. 43. № 1. С. 105–108.

11. *Guo L., Xiong J.* Multi-Scale Attention-Enhanced Deep Learning Model for Ionogram Automatic Scaling // Radio Science. 2023. Vol. 58. No. 3. p.e2022RS007566. <https://doi.org/10.1029/2022RS007566>

12. *Lu Z., Hua C., Wei N., Feng J., Lou P., Liu W.* Research on classification of vertical ionogram based on deep convolution neural network // Progress in Geophysics. 2022. Vol. 37. No. 5. P. 1834–1839. <https://doi.org/10.6038/pg2022GG0073>

13. Xiao Z., Wang J., Li J., Zhao B., Hu L., Liu L. Deep-learning for ionogram automatic scaling // *Advances in Space Research*. 2020. Vol. 66. No. 4. P. 942–950.

<https://doi.org/10.1016/j.asr.2020.05.009>

14. De la Jara C., Olivares C. Ionospheric echo detection in digital ionograms using convolutional neural networks // *Radio Science*. 2021. Vol. 56. No. 8. P. 1–15.

15. Щирый А.О. Использование нейронных сетей для дальнейшего развития программной части аппаратно-программных комплексов радиозондирования ионосферы // *Электромагнитные волны и электронные системы*. 2024, Т. 29. № 5. С. 55–60. <https://doi.org/10.18127/j15604128-202405-08>

16. Standard Archiving Output (SAO) Format. 2006.

URL: <https://ulcar.uml.edu/~iag/SAO-4.3.htm>

17. Gamache R.R., Galkin I.A., Reinisch B.W. A Database Record Structure for Ionogram Data. University of Lowell Center for Atmospheric Research, UMLCAR 96-01, 1996.

18. Reinisch B.W., Galkin I.A. SAO.XML format specification v 5.0, Univ. of Mass. Lowell, Lowell. 2008. Available at <http://ulcar.uml.edu/SAOXML>

19. Galkin I.A., Khmyrov G.M., Reinisch B.W., McElroy J. The SAOXML 5: New Format for Ionogram-Derived Data. *Radio Sounding and Plasma Physics // AIP Conf. Proc.* 2008. Vol. 974. P. 160–166. <https://doi.org/10.1063/1.2885025>

20. Жбанков Г.А., Анишин М.М., Телегин В.А. Программный комплекс «Viewer_DPS4» – инструмент обработки и анализа данных ионозонда DPS-4 // *Техника радиосвязи*. 2022. Выпуск 2 (53). С. 53–65.

21. Щирый А.О., Писаренко А.А. Общедоступные архивы данных наземного радиозондирования ионосферы коротковолновыми сигналами // *Электронные библиотеки*. 2023. Т. 25. №6. С. 992–1005.

22. USA National Geophysical Data Center (NGDC) Data Services.

URL: <ftp://ftp.ngdc.noaa.gov/ionosonde/>

23. Australian Government – Bureau of Meteorology, Space Weather Services. Ionospheric data archive.

URL: https://downloads.sws.bom.gov.au/wdc/wdc_ion_archive/

24. Australian Government – Bureau of Meteorology, Space Weather Services. Ionogram Data Format. Clean ionogram data format.

URL: https://www.sws.bom.gov.au/World_Data_Centre/2/8/3

25. Australian Government – Bureau of Meteorology, Space Weather Services. Scaled Ionospheric Data Format.

URL: https://www.sws.bom.gov.au/World_Data_Centre/2/8/1

26. Australian Government – Bureau of Meteorology, Space Weather Services. Autoscaled Ionospheric Data Format.

URL: https://www.sws.bom.gov.au/World_Data_Centre/2/8/12

USING OPEN ARCHIVES OF SCALED VERTICAL RADIOSONDING IONOGRAMS AS LABELED DATA FOR TRAINING MACHINE LEARNING MODELS

A. O. Schiriy^[0000-0001-7312-4424]

Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences, Moscow, Russia

andreyschiriy@gmail.com

Abstract

The idea of using the available large arrays of ionogram processing results from vertical radiosonde of the ionosphere as training datasets for building predictive models using machine learning methods is put forward. The most common formats for saving the results of ionogram processing are considered, as well as some Internet resources with archives of freely available files of these formats. These datasets are used by us to build predictive models, including time series of critical frequencies of ionospheric layers. It is also possible to use some datasets of ionogram processing results to train models designed for automatic ionogram processing.

Keywords: *ionosphere, radiosounding, vertical sounding of the ionosphere, ionogram, ionosphere characteristics, metadata.*

REFERENCES

1. Radiozondirovanie ionosfery sputnikovymi i nazemnymi ionozondami / Pod red. S.I. Avdyushina. Trudy instituta prikladnoj geofiziki im. akademika E.K. Fyodorova. M.: IPG, 2008.
URL: <http://ipg.geospace.ru/publications/book-2008.pdf>
2. *Reinisch B., Sales G.* Measuring Electrodynamics of the Ionosphere by Digital Ionosondes and Other Techniques. Scientific Report No. 3. 2001.
URL: <https://apps.dtic.mil/sti/tr/pdf/ADA402979.pdf>
3. *Akimov V.F., Kalinin Ju.K.* Vvedenie v proektirovanie ionosfernyh zagorizontnyh radiolokatorov / pod red. S.F. Boeva. M.: Tehnosfera, 2017. 492 s.
4. *Fabrizio G.* High Frequency Over-the-Horizon Radar: Fundamental Principles, Signal Processing, and Practical Applications. McGraw-Hill Education, 2013.
5. *Schiriy A.O.* Razrabotka i modelirovanie algoritmov avtomaticheskogo izmereniya harakteristik ionosfernyh korotkovolnovyh radiolinij: Avtoref. dis. ... kand. tekhn. nauk: Spec. 05.12.04; Sankt-Peterburgskij gos. un-t telekommunikacij im. prof. M.A. Bonch-Bruevicha. SPb., 2007. 19 s.
6. *Schiriy A.O.* Razvitie sredstv avtomatizacii nazemnogo radiozondirovaniya ionosfery // Fundamental'nye problemy radioelektronnoho priborostroeniya. 2014. №5. S. 170–173.
7. *Schiriy A.O.* Arhitektura programmnoj chasti apparatno-programmnogo kompleksa distancionnogo nazemnogo radiozondirovaniya ionosfery // Novye informacionnye tekhnologii v avtomatizirovannyh sistemah. 2015. №18. S. 144–152.
8. *Schiriy A.O.* Algoritmy i programmnoe obespechenie avtomatizacii processov izmerenij i obrabotki dannyh operativnoj diagnostiki ionosfery i ionosfernyh radiolinij // Zhurnal radioelektroniki. 2022. №10.
<https://doi.org/10.30898/1684-1719.2022.10.4>
9. *Shiriy A.O.* HF channel transmit function module measurement // Proceedings of the 5th International Conference on Actual Problems of Electron Devices Engineering, APEDE-2002. 2002. Vol. 5. P. 365–369.
<https://doi.org/10.1109/apede.2002.10449645>

10. *Dolgacheva S.A., Makarova L.N., Nikolaev A.V.* Obrabotka ionogramm vysokoshirotnyh stancij vertikal'no-go zondirovanija s ispol'zovaniem nejronnyh setej: Es i F2 sloi // *Physics of Auroral Phenomena*. 2020. T. 43. № 1. S. 105–108.
11. *Guo L., Xiong J.* Multi-Scale Attention-Enhanced Deep Learning Model for Ionogram Automatic Scaling // *Radio Science*. 2023. Vol. 58. No. 3. p.e2022RS007566. <https://doi.org/10.1029/2022RS007566>
12. *Lu Z., Hua C., Wei N., Feng J., Lou P., Liu W.* Research on classification of vertical ionogram based on deep convolution neural network // *Progress in Geophysics*. 2022. Vol. 37. No. 5. P. 1834–1839. <https://doi.org/10.6038/pg2022GG0073>
13. *Xiao Z., Wang J., Li J., Zhao B., Hu L., Liu L.* Deep-learning for ionogram automatic scaling // *Advances in Space Research*. 2020. Vol. 66. No. 4. P. 942–950. <https://doi.org/10.1016/j.asr.2020.05.009>
14. *De la Jara C., Olivares C.* Ionospheric echo detection in digital ionograms using convolutional neural networks // *Radio Science*. 2021. Vol. 56. No. 8. P. 1–15.
15. *Schiriy A.O.* Ispol'zovanie nejronnyh setej dlja dal'nejshego razvitija programmnoj chasti apparatno-programmnyh kompleksov radiozondirovanija ionosfery // *Jelektromagnitnye volny i jelektronnye sistemy*. 2024, Tom 29, №5. S. 55–60. <https://doi.org/10.18127/j15604128-202405-08>
16. Standard Archiving Output (SAO) Format. 2006. URL: <https://ulcar.uml.edu/~iag/SAO-4.3.htm>
17. *Gamache R.R., Galkin I.A., Reinisch B.W.* A Database Record Structure for Ionogram Data. University of Lowell Center for Atmospheric Research, UMLCAR 96-01, 1996.
18. *Reinisch B.W., Galkin I.A.* SAO.XML format specification v 5.0, Univ. of Mass. Lowell, Lowell. 2008. Available at <http://ulcar.uml.edu/SAOXML>.
19. *Galkin I.A., Khmyrov G.M., Reinisch B.W., McElroy J.* The SAOXML 5: New Format for Ionogram-Derived Data. *Radio Sounding and Plasma Physics* // *AIP Conf. Proc.* 2008. Vol. 974. P. 160–166. <https://doi.org/10.1063/1.2885025>
20. *Zhbankov G.A., Anishin M.M., Telegin V.A.* Programmnyj kompleks «Viewer_DPS4» – instrument obrabotki i analiza dannyh ionozonda DPS-4 // *Tehnika radiosvjazi*. 2022. Vypusk 2 (53). S.53–65.

21. Schiriy A.O., Pisarenko A.A. Obshhedostupnyye arhivy dannyh nazemnogo radiozondirovaniya ionosfery korotkovolnovymi signalami // Jelektronnye biblioteki. 2023. T. 25. №6. S. 992–1005.

22. USA National Geophysical Data Center (NGDC) Data Services.
URL: <ftp://ftp.ngdc.noaa.gov/ionosonde/>

23. Australian Government – Bureau of Meteorology, Space Weather Services. Ionospheric data archive.
URL: https://downloads.sws.bom.gov.au/wdc/wdc_ion_archive/

24. Australian Government – Bureau of Meteorology, Space Weather Services. Ionogram Data Format. Clean ionogram data format.
URL: https://www.sws.bom.gov.au/World_Data_Centre/2/8/3

СВЕДЕНИЯ ОБ АВТОРЕ



ЩИРЫЙ Андрей Олегович – кандидат технических наук, старший научный сотрудник Института земного магнетизма, ионосферы и распространения радиоволн РАН им. Н. В. Пушкова (ИЗМИРАН), Москва.

Andrey Olegovich SCHIRIY – Candidate of Technical Sciences (PhD equivalent), senior researcher at the Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation of the Russian Academy of Sciences (IZMIRAN), Moscow.

email: andreyschiriy@gmail.com

ORCID: 0000-0001-7312-4424

Материал поступил в редакцию 21 января 2026 года