

## НОВЫЙ ПОДХОД К СОЗДАНИЮ ГЛОБАЛЬНОЙ МОДЕЛИ ПОЛНОГО ЭЛЕКТРОННОГО СОДЕРЖАНИЯ ИОНОСФЕРЫ

<sup>1</sup>А.В. Жуков, <sup>1,2</sup>Ю.В. Ясюкевич, <sup>1</sup>С.А. Серебrenникова,  
<sup>1</sup>А.М. Веснин, <sup>1</sup>А.В. Киселёв

<sup>1</sup>Институт солнечно земной физики СО РАН, Иркутск, Россия  
zhukov@iszf.irk.ru

<sup>2</sup>Иркутский государственный университет, Иркутск, Россия

## NEW APPROACH TO CREATING GLOBAL MODEL OF THE IONOSPHERIC TOTAL ELECTRON CONTENT

<sup>1</sup>A.V. Zhukov, <sup>1,2</sup>Yu.V. Yasyukevich, <sup>1</sup>S.A. Serebrennikova, <sup>1</sup>A.M. Vesnin, <sup>1</sup>A.V. Kiselev

<sup>1</sup>Institute of Solar-Terrestrial Physics SB RAS, Irkutsk, Russia  
zhukovaleksei@iszf.irk.ru

<sup>2</sup>Irkutsk State University, Irkutsk, Russia

**Аннотация.** Разработан новый подход моделирования в глобальном масштабе полного электронного содержания ионосферы на основе методов машинного обучения. В качестве входных данных использовались глобальные ионосферные карты GIM за 1998–2017 гг., получаемые на основе GPS/ГЛОНАСС-измерений. Оценена релевантность различных управляющих параметров. В данной работе выполнено построение нескольких параметров, включая индексы геомагнитной активности и солнечного потока, в качестве наиболее значимых были выделены  $F10.7$  и день года.

**Ключевые слова:** полное электронное содержание, модель, GPS, ГЛОНАСС, машинное обучение.

**Abstract.** A new global modeling approach has been developed for the total electron content of the ionosphere based on machine learning methods. The GIM global ionospheric maps for 1998–2017 obtained by GPS / GLONASS measurements were used as input data. The relevance of various control parameters is evaluated. In this paper, several parameters were constructed including geomagnetic and solar flux indices;  $F10.7$  and the day of the year were selected as the most significant.

**Keywords:** total electron content, model, GPS, GLONASS, machine learning.

### ВВЕДЕНИЕ

В последние годы все активнее развиваются средства мониторинга околоземного космического пространства. Большой объем экспериментальных данных предоставили глобальные навигационные спутниковые системы (ГНСС) [Afraimovich et al., 2013]. На основе данных ГНСС строятся глобальные ионосферные карты [Hernández-Pajares et al., 2009]. Подобные данные привели к существенному повышению качества ионосферных моделей. Если модель Клобучара, используемая в GPS [Klobuchar, 1987] позволяет устранять до ~50 % ошибки, то модель BDGIM [Yuan et al., 2019] до ~70 %. Кроме того, появились модели долгосрочного прогноза, требующие малое число входных параметров, такие как, например, GEM TEC [Ivanov et al., 2011], требующая на вход данные о индексе солнечной активности  $F10.7$ .

При этом, подходы, используемые при решении ионосферных задач, как правило, не включают те значительные достижения, которые были получены в последние годы в области компьютерного интеллекта. В настоящее время существуют достаточно эффективные методы машинного обучения, которые могут использоваться для решения задач изучения и прогнозирования ионосферы [Zhukov et al., 2018]. В настоящей работе осуществляется построение глобальной модели полного электронного содержания на основе данных карт GIM [Hernández-Pajares et al., 2009] с применением методов машинного обучения.

### МЕТОДОЛОГИЯ

Одной из сложностей, возникающих при анализе и построении изображений, таких как карты TEC, является проблема сокращения размерности. Реше-

нием данной проблемы может служить разложение карты на набор компонент. Таким методом является, например, метод сингулярного разложения [Golub, van Loan, 1996], который позволяет получить ортогональные компоненты карт с учетом необходимой точности приближения. Преимущество данного подхода по отношению к некоторым нейросетевым методам сокращения размерности (например, автоэнкодер) является наибольшая интерпретируемость компонент и существенно меньшие требования к вычислительным ресурсам.

Как показали первичные эксперименты и опыт предыдущих исследований, наиболее простым способом является получение разложений для каждого всемирного времени. Это объясняется наличием известных периодических эффектов, таких как экваториальная аномалия.

Следующим этапом валидационная и тестовая выборки раскладываются по полученным компонентам. Для каждого момента времени строится модель машинного обучения. Гиперпараметры этой модели подбираются с учетом точности ее работы на валидационной выборке.

Так как в данном случае необходимо по одним и тем же данным получить множество ответов (коэффициент для каждой компоненты), то предлагается использовать метод поддерживающий данный режим работы, такой как Random Forest [Segal, Xiao, 2011]. Другим плюсом данного метода является возможность оценки значимости каждого входного параметра.

В качестве потенциальных входных параметров предлагается использовать следующие индексы геомагнитной и солнечной активности:  $K_p$ ,  $R$ ,  $A_p$ ,  $F10.7$ ,  $AE$ ,  $AL$ ,  $AU$ . Также для учета сезонных особенностей

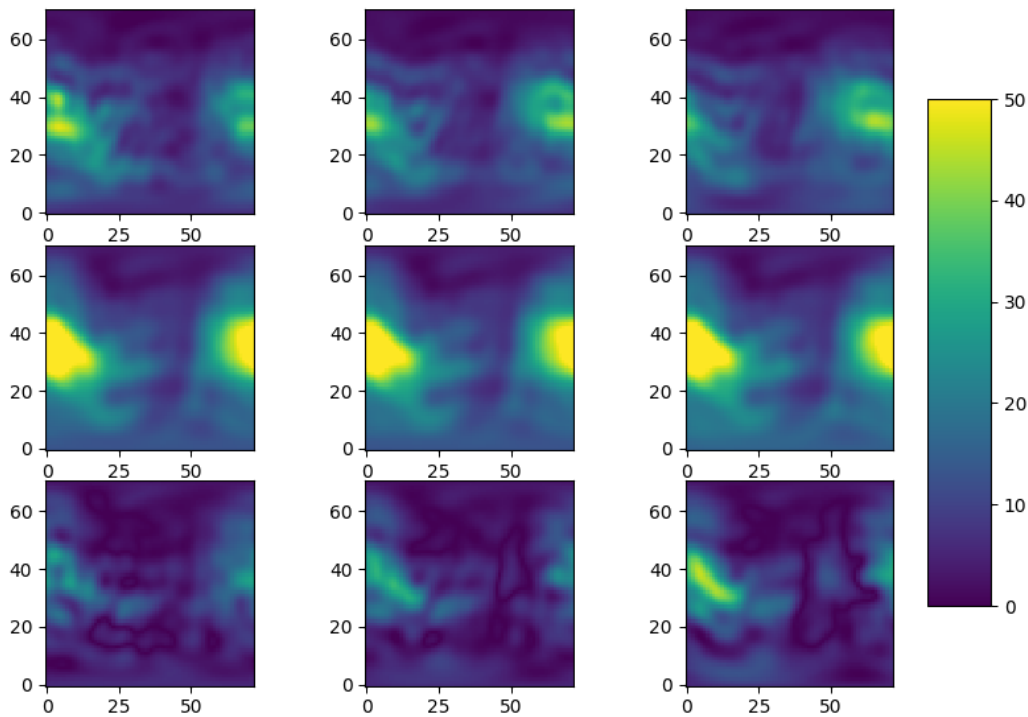


Рис. 1. Изображения оригинальных карты (верхний ряд), изображения соответствующих карт полученных с помощью предлагаемой модели (средний ряд), и разница между ними (нижний ряд)

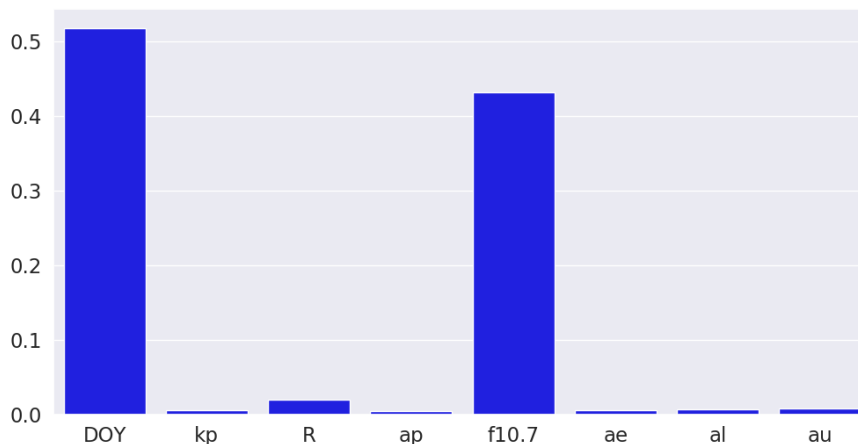


Рис. 2. Относительная значимость входных параметров модели

в качестве входного параметра используется день года (далее DOY). Важность каждого из параметров оценивается встроенными средствами Random Forest. Итоговое значение получено путем усреднения по всем моделям (для каждого времени).

Финальная проверка точности полученной модели осуществляется оценкой ошибки на тестовой выборке.

### ЭКСПЕРИМЕНТ

Для проведения расчетов были использованы карты GIM [цитирование GIM] имеющие пространственное разрешение  $71 \times 72$  и временное два часа. Все карты были разбиты по времени на три выборки 1998-2015 гг. для обучения, 2016 г. для валидации модели и подбора ее параметров и 2017 для итогового тестирования. В качестве потенциальных параметров модели были использованы различные индексы, полученные из базы данных SPDF OMNIWeb [King, Papitashvili, 2005].

Построив разложения для каждого момента времени для описания 3 сигма вариаций значений карт на тренировочной выборке, было получено максимальное количество компонент равное 138. В дальнейшем оно использовалось для всех разложений.

На основе разложений для каждого времени построены модели Random Forest со следующими основными гиперпараметрами: количество решающих деревьев 100, максимальная глубина дерева 7. С помощью этих моделей получены изображения для трех карт из тестовой выборки показанные на рис. 1.

Оценка релевантности параметров полученной модели представлена на рис. 2. Как можно заметить, наибольшую информативность имеет день года и индекс солнечной активности  $F10.7$ .

### ЗАКЛЮЧЕНИЕ

Таким образом, был разработан новый подход моделирования в глобальном масштабе полного

электронного содержания ионосферы на основе методов машинного обучения. Как было показано, данный подход позволяет оценивать значимость параметров, что является значимым преимуществом. Требуется дальнейшее исследование предложенной модели. Модель может быть усовершенствована добавлением новых входных параметров, в том числе учитывающих изменение индексов во времени.

Работа выполнена при поддержке РФФИ (грант № 18-35-20038).

#### **СПИСОК ЛИТЕРАТУРЫ**

Afraimovich E.L., Astafyeva E.I., Demyanov V.V., et al. Review of GPS/GLONASS studies of the ionospheric response to natural and anthropogenic processes and phenomena // *Journal of Space Weather and Space Climate*. 2013. V. 3. A27. DOI: [10.1051/swsc/2013049.2013](https://doi.org/10.1051/swsc/2013049.2013).

Hernández-Pajares M., Juan J.M., Sanz J., et al. The IGS VTEC maps: a reliable source of ionospheric information since 1998 // *J Geod.* 2009. V. 83. P. 263–275. DOI: [10.1007/s00190-008-0266-1](https://doi.org/10.1007/s00190-008-0266-1).

Ivanov V.B., Gefan G.D., Gorbachev O.A. Global Empirical Modeling of the Total Electron Content of the Ionosphere for Satellite Radio Navigation Systems // *J. Atmos. Solar–Terr. Phys.* 2011. V. 73. P. 1703–1707. DOI: [10.1016/j.jastp.2011.03.010](https://doi.org/10.1016/j.jastp.2011.03.010).

Klobuchar J. Ionospheric Time-Delay Algorithms for Single-Frequency GPS Users // *IEEE Transactions on Aerospace and Electronic Systems*. 1987. N 3. P. 325–331. DOI: [10.1109/TAES.1987.310829](https://doi.org/10.1109/TAES.1987.310829).

Yuan Y., Wang N., Li Z., Huo X. The BeiDou global broadcast ionospheric delay correction model (BDGIM) and its preliminary performance evaluation results // *NAVIGATION*. 2019. V. 66. P. 55–69. DOI: [10.1002/navi.292](https://doi.org/10.1002/navi.292).

Zhukov A., Sidorov D., Mylnikova A., Yasyukevich Y. Machine learning methodology for ionosphere total electron content nowcasting // *International Journal of Artificial Intelligence*. 2018. V. 16, N 1. P. 144–157. DOI: [10.13140/RG.2.2.19349.83685](https://doi.org/10.13140/RG.2.2.19349.83685).

Hernández-Pajares M., et al. The IGS VTEC maps: a reliable source of ionospheric information since 1998 // *J. Geodesy*. 2009. V. 83. N. 3-4. P. 263–275. DOI: [10.1007/s00190-008-0266-1](https://doi.org/10.1007/s00190-008-0266-1).

Golub G.H., van Loan C.F. The singular value decomposition and unitary matrices // *Matrix Computations*. 1996. P. 70–71.

Segal M., Xiao Y. Multivariate random forests // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2011. V. 1, N 1. P. 80–87.

King J.H., Papitashvili N.E. Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data // *J. Geophys. Res.: Space Phys.* 2005. V. 110. N A2. DOI: [10.1029/2004JA010649](https://doi.org/10.1029/2004JA010649).