«Оптика атмосферы и океана», **34**, № 10 (2021)
DOI: 10.15372/AOO20211008

АППАРАТУРА И МЕТОДЫ ОПТИЧЕСКОЙ ДИАГНОСТИКИ ОКРУЖАЮЩЕЙ СРЕДЫ

УДК 551.571.7

Определение содержания водяного пара в столбе атмосферы по данным КА «Электро-Л» № 3 с использованием нейронных сетей

В.Д. Блощинский, А.А. Филей, Е.И. Холодов*

Дальневосточный центр ФГБУ «НИЦ "Планета"»
680000, г. Хабаровск, ул. Ленина, 18

Поступила в редакцию 10.08.2020 г.

Рассматривается применение метода на основе искусственных нейронных сетей для оценки общего содержания водяного пара в столбе атмосферы по данным спутникового прибора МСУ-ГС геостационарного космического аппарата «Электро-Л» № 3. Результаты сравнения восстановленных оценок общего содержания водяного пара, полученных по данным МСУ-ГС, с данными спутникового прибора MODIS и результатами измерений станций AERONET показали высокую согласованность. Среднеквадратическая ошибка при сравнении с данными MODIS составила 0,311 см, с данными AERONET – 0,409 см, а корреляция – 98,2 и 84,7% соответственно. Полученные результаты говорят об эффективности предложенной методики определения общего содержания водяного пара для решения целевых задач физики атмосферы.

Ключевые слова: МСУ-ГС, Электро-Л, газ, водяной пар, искусственная нейронная сеть; MSU-GS, Electro-L, gas, water vapor, artificial neural network.

Введение

Атмосферный водяной пар – один из самых важных газов, на долю которого приходится почти две трети парникового эффекта планеты [1]. Его пространственная и временная изменчивость больше по сравнению с другими парниковыми газами, например углекислым газом или метаном. Правильная оценка водяного пара в атмосфере осложняется его вовлеченностью в формирование облаков, перенос тепла и энергетический обмен между поверхностью Земли и атмосферой [2]. Общее содержание водяного пара (ОСВП) в столбе атмосферы можно рассматривать как показатель состояния климата региона (влажный или сухой), который сильно связан с температурой поверхности и воздуха на низких высотах. Такая корреляция была показана в [3] по данным прибора GOME (Global Ozone Monitoring Experiment) космического аппарата (КА) ERS-2 (European Remote Sensing). В настоящее время региональный и глобальный мониторинг газовых составляющих атмосферы проводится по данным *in situ* измерений и измерений с КА. Наземные станции дают наиболее точные данные об ОСВП в столбе атмосферы, но их неравномерное расположение на Земле не позволяет получить полную информацию о пространственном распределении водяного пара в определенное время. Этого недостатка лишена спутниковая система оценки ОСВП [4–7].

Можно выделить четыре подхода к оценке общего содержания водяного пара по данным КА: в ближней инфракрасной (ИК) области [8, 9], в микроволновом диапазоне [10, 11], по тепловому ИК-излучению [12, 13], гиперспектральные методы [14]. Наиболее распространено использование каналов спутникового прибора в ближнем ИК-диапазоне, включающих полосу поглощения в районе 0,94 мкм и одну соседнюю полосу в атмосферном окне. Такой подход обеспечивает высокую точность оценки ОСВП, но его нельзя применять для приборов, не имеющих каналов с полосами поглощения водяного пара. Для оценки содержания водяного пара в тепловом ИК-диапазоне используются отношения разности яркостных температур каналов спутникового прибора, особенно 11 и 12 мкм. Несколько исследований [13, 15] показали, что такой подход обеспечивает достаточную точность оценки ОСВП.

В общем виде подходы к оценке параметров атмосферы, в частности содержания водяного пара в столбе атмосферы, можно разделить на две группы: физический подход, основанный на применении моделей переноса излучения и поэтапном приближении к результату; статистический подход, в котором используются методы расчета коэффициентов регрессии [6] либо искусственные нейронные сети (ИНС) [16].

ОСВП может оцениваться по данным как с полярно-орбитальных, так и с геостационарных КА. Последние представляют особый интерес по причине большого пространственного покрытия исследуемой территории с высокой частотой об-

* Владислав Дмитриевич Блощинский (v.bloshchinsky@dvrcpod.ru); Андрей Александрович Филей (andreyvm-61@mail.ru); Егор Игоревич Холодов (e.kholodov@dvrcpod.ru).

© Блощинский В.Д., Филей А.А., Холодов Е.И., 2021

новления поступающей информации. Эти показатели особенно важны для изучения пространственно-временных изменений ОСВП и получения его региональных оценок с возможностью построения многолетних трендов.

В настоящей работе рассматривается метод определения ОСВП при отсутствии облачности по данным прибора МСУ-ГС (многозональное сканирующее устройство гидрометеорологического обеспечения) с геостационарного КА «Электро-Л» № 3. Прибор позволяет измерять уходящее излучение Земли в 10 спектральных диапазонах в видимом и ИК-спектрах, которые включают полосы поглощения и пропускания водяного пара ~ 6,4 мкм. Предлагаемый метод основывается на применении искусственной нейронной сети, которая в качестве входных данных использует информацию из каналов МСУ-ГС и вертикальные профили параметров атмосферы из численной модели прогноза GFS (Global Forecast System).

Формирование набора данных для искусственной нейронной сети

Для обучения и проверки качества работы ИНС была сформирована выборка данных, которая впоследствии разделялась на обучающий, валидационный и тестовый наборы. Для наиболее адекватной оценки работы сети в тестовую выборку не включались данные, которые участвовали в обучении.

Прибор МСУ-ГС имеет точку стояния 76,6° в.д. Рассматриваемая зона покрытия лежит в пределах от –70 до 70° по широте и от 10 до 140° по долготе. В выборку включались данные измерений семи инфракрасных каналов прибора МСУ-ГС и профиль температуры (31 уровень от 1000 до 1 гПа), полученный из модели GFS. Использование измерений ИК-каналов объясняется тем, что основные полосы поглощения и пропускания водяного пара имеют длины волн больше 5 мкм, а профиль температуры был добавлен ввиду наличия корреляции ОСВП с тепловым режимом атмосферы. Для учета пространственно-временной зависимости распределения полей ОСВП [17] в выборку включалась информация о рассматриваемом месяце, зенитном угле спутника и Солнца. Значения зенитного угла спутника позволяют нейронной сети нивелировать различия измерений в каналах спутникового прибора по удаленности от надира. Информация о зенитном угле Солнца позволяет ИНС учесть зависимость регистрируемых спутником показателей яркостной температуры от освещенности подстилающей поверхности.

Поставленная задача по определению значений содержания водяного пара имеет регрессионный характер, т.е. ИНС должна вычислить ОСВП по определенным признакам. Для этого в выборку должна быть добавлена априорная информация о количестве водяного пара в столбе атмосферы для конкретной географической точки. Наиболее точные данные по ОСВП можно получить с наземных станций наблюдения, но временная изменчивость таких данных и пространственное распределение станций не позволяют собрать достаточную по объему выборку для обучения ИНС. Поэтому в качестве эталонных значений ОСВП выбирались его спутниковые оценки по данным прибора MODIS (Moderate Resolution Imaging Spectroradiometer) КА Aqua и Terra (продукты MODIS Atmospheric Profile MOD07 и MYD07). Таким образом, в набор данных для ИНС вошли 25 тыс. записей с января по апрель 2021 г., 70% из которых составили обучающую, 15% — валидационную, 15% — тестовую выборки. Распределение набранных значений ОСВП находится в диапазоне от 0 до 6,2 см.

Итоговый набор содержит в себе разные по абсолютным величинам значения признаков. Такой дисбаланс может вызвать неустойчивость работы модели ИНС, замедлить обучение и ухудшить результат. Поэтому выборка нормализовалась отдельно по каждому признаку классическим способом согласно выражению

$$X_n = \frac{\mathbf{X} - \mu}{\sigma}, \qquad (1)$$

где $\mathbf{X}$ — вектор всех значений признака из выборки; μ — среднее значение для $\mathbf{X}$; σ — стандартное отклонение значений для $\mathbf{X}$. Так удается добиться получения среднего значения каждого признака по всем данным, равного нулю, а дисперсии — единице.

Архитектура искусственной нейронной сети и ее обучение

Для поставленной регрессионной задачи определения ОСВП в настоящей работе рассматривались варианты архитектур ИНС, построенных по типу полносвязных нейронных сетей. Полносвязные ИНС состоят из входного, выходного, а также одного или нескольких скрытых слоев. Дополнительно к полносвязным слоям ИНС подключались слои активации и слои нормализации активации, которые помогают сделать обучение нейронной сети более устойчивым [18]. В настоящей работе в качестве функций активации рассматривались ReLU и Tanh, а количество используемых скрытых слоев и нейронов в них выбиралось эмпирически.

Само обучение нейронных сетей состоит в итерационном пересчете весовых коэффициентов нейронов в слоях по заданному алгоритму оптимизации. Обучение выполняется с помощью алгоритма обратного распространения ошибки, в качестве которой выступала среднеквадратическая ошибка. В качестве алгоритма оптимизации был выбран алгоритм Adam [19], который объединяет в себе идеи сохранения инерции и масштабирования градиента.

Обучение ИНС продолжается до достижения минимальных изменений в показателях метрик

машинного обучения по данным валидационной выборки на протяжении нескольких итераций. В настоящей работе в качестве метрик были выбраны среднеквадратическая ошибка и средняя абсолютная ошибка.

Оценка точности проводилась на тестовой выборке по расчетам следующих метрик: средняя абсолютная ошибка (MAE), среднеквадратическая ошибка (RMSE), среднее смещение (BIAS), коэффициент корреляции Пирсона (ρ), коэффициент детерминации (R^2).

В результате экспериментов по обучению различных архитектур ИНС и определению наиболее весомых признаков для оценки ОСВП была выявлена наилучшая, согласно указанным метрикам, конфигурация архитектуры, которая состоит из следующих элементов: входного слоя размерностью 32; трех блоков размерностями 64, 128 и 256 нейронов, каждый из которых состоит из полносвязного слоя с активацией ReLU и слоя нормализации; выходного полносвязного слоя размерностью, равной 1. Размерность входного слоя соответствует количеству входных признаков из выборки без учета тех, чье исключение почти не влияет на точность определения содержания водяного пара. Таким образом были исключены измерения 4, 7 и 8-го каналов прибора МСУ-ГС. Эта ИНС на тестовой выборке показала следующие результаты: MAE = 0,023 см, RMSE = 0,311 см, BIAS = 0,005 см, ρ = 99,206%, R^2 = 98,221%.

На рисунке (цв. вкладка) представлены примеры карт ОСВП для безоблачных пикселей, построенных с помощью разработанного алгоритма и по данным прибора MODIS. По рассчитанным метрикам точности видны хорошая корреляция и точность восстановления значений ОСВП по отношению к спутниковым оценкам по данным прибора MODIS. Из рисунка также заметна визуальная корреляция значений общего содержания водяного пара.

Валидация результатов

Тестирование на выборке не всегда дает полное представление о точности определения параметров, потому что в ней могут быть упущены некоторые сложные и спорные ситуации, возникающие при обучении ИНС. Поэтому качество предлагаемого алгоритма проверялось путем сравнения с данными из архивов AERONET (AErosol RObotic NETwork). Сравнение по данным радиозондирования не проводилось ввиду недостаточного количества измерений для доступного временного периода съемки «Электро-Л» № 3.

Сравнение проводилось с января по февраль 2021 г. Разница во времени спутниковой съемки и измерений, полученных из архивов AERONET, не превышала 15 мин, удаленность сравниваемых точек измерений не превышала 4 км. Для определения точности рассчитывались такие же метрики, как для оценки точности на тестовой выборке. Дополнительно за тот же временной период проводилось сравнение значений ОСВП, полученных по данным прибора MODIS, с данными из архивов AERONET. В первом случае для валидации было набрано 2400 точек, во втором случае – 500. В таблице представлены средние показатели метрик для всех рассматриваемых сроков.

Результаты валидации

Прибор/архив	RMSE, см	MAE, см	BIAS, см	ρ, %	R^2, %
МСУ-ГС / AERONET	0,409	0,336	0,091	93,469	84,656
MODIS / AERONET	0,401	0,323	0,077	93,841	85,128

Полученные результаты валидации говорят о хорошей корреляции значений ОСВП, восстанавливаемых по данным МСУ-ГС, с данными из архивов AERONET. Согласно спецификациям точности спутниковых оценок ОСВП [5] ошибка не должна превышать 0,3 см. Разработанный алгоритм не достигает необходимых показателей точности, но сопоставим с результатами аналогичных продуктов, получаемых по информации с других КА [5–7, 12]. Это касается и оценок ОСВП по измерениям MODIS. На результаты валидации могло отрицательно повлиять наличие пикселей с облачностью ввиду неточности используемых масок облачности.

Для получения наиболее точных оценок ОСВП необходимо применять источники данных с наименьшим показателем отклонения от реальных значений, например данные радиозондирования. Но для глобального мониторинга изменений ОСВП обучение ИНС по такой информации не представляется возможным.

Заключение

В настоящей работе рассмотрен регрессионный алгоритм на основе полносвязной ИНС, который позволяет проводить расчет ОСВП в столбе атмосферы по данным прибора МСУ-ГС геостационарного КА «Электро-Л» № 3. Алгоритм полностью автоматический, т.е. нет необходимости настраивать какие-либо параметры. Область расчета значений ОСВП ограничивается диапазонами от –70 до 70° по широте и от 10 до 140° по долготе. Для работы алгоритма требуются 5, 6, 9 и 10-й каналы прибора МСУ-ГС, информация о зенитном угле спутника и Солнца, а также температурный профиль атмосферы, полученный по данным прогностической модели GFS. По результатам сравнения со значениями ОСВП из архивов AERONET средняя абсолютная ошибка определения ОСВП составляет 0,336 см, а среднеквадратическая ошибка – 0,409 см. Такие показатели говорят о хорошей точности работы предлагаемого алгоритма по сравнению с аналогичными продуктами, использующими измерения других КА [5–7, 12].

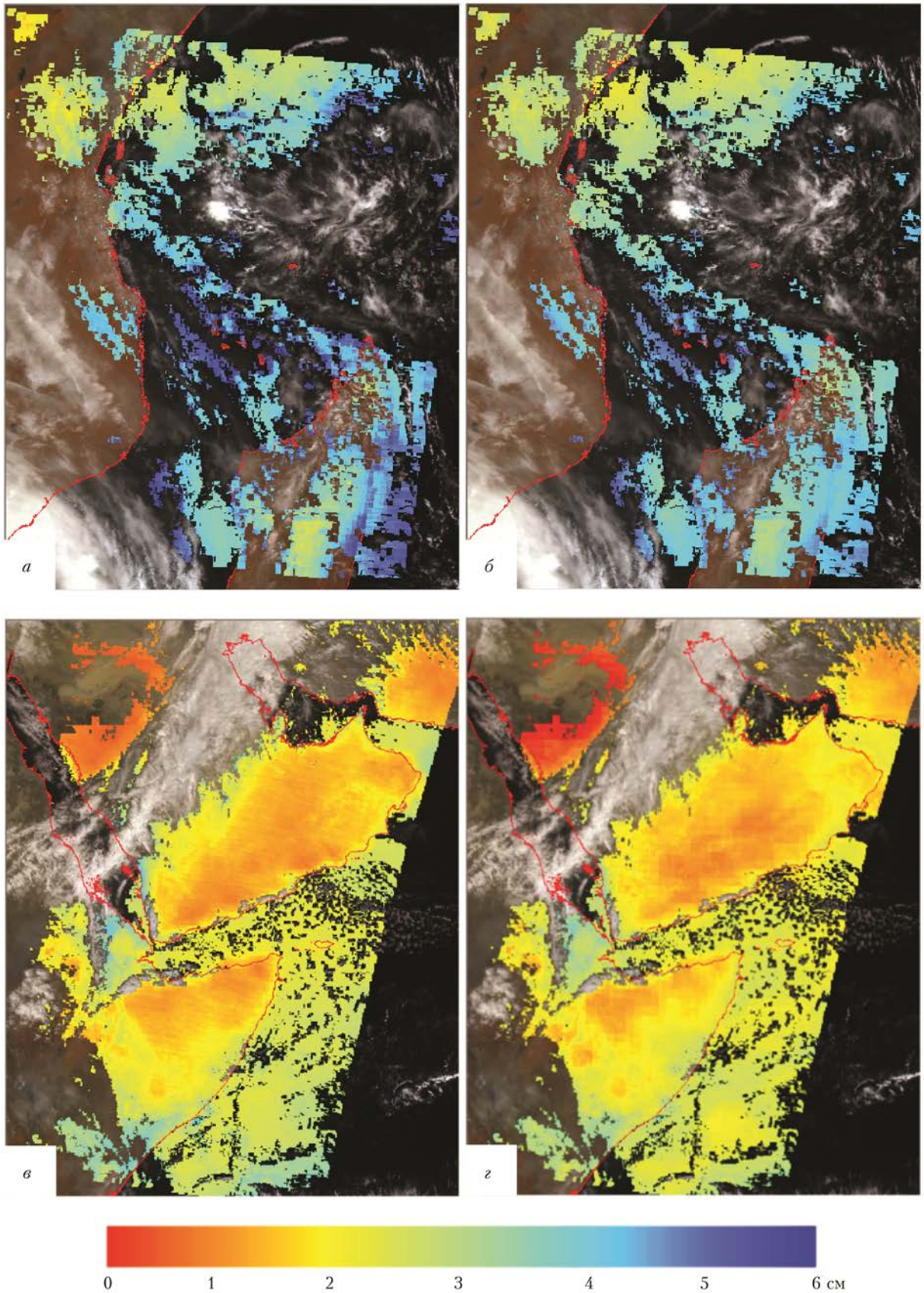
Минусом предлагаемого алгоритма можно считать необходимость в дополнительной информации

(модель GFS), помимо данных с прибора МСУ-ГС КА «Электро-Л» № 3. Также есть небольшие ограничения по диапазону восстанавливаемых значений ОСВП (0–6,5 см) ввиду особенностей работы регрессионных методов и ИНС. Представленный алгоритм имеет ограничение в работе по временному периоду, так как ИНС обучена для дневных данных с января по апрель. Но результаты валидации говорят о работоспособности алгоритма и его хорошей точности, что дает возможность снять ограничение за счет дополнительного обучения нейронной сети с привлечением соответствующих данных.

1. *Kiehl J.T., Trenberth K.E.* Earth's annual global mean energy budget // Bull. Am. Meteorol. Soc. 1997. V. 78, iss. 2. P. 197–208.
2. *Mieruch S., Noël S., Bovensmann H., Burrows J.P.* Analysis of global water vapor trends from satellite measurements in the visible spectral range // Atmos. Chem. Phys. 2008. V. 8, iss. 3. P. 491–504.
3. *Wagner T., Beirle S., Grzegorski M., Platt U.* Global trends (1996–2003) of total column precipitable water observed by Global Ozone Monitoring Experiment (GOME on ERS-2) and their relation to near-surface temperature // J. Geophys. Res. 2006. V. 111. P. D12102.
4. *Ren H., Du C., Liu R., Qin Q., Yan G., Li Z., Meng J.* Atmospheric water vapor retrieval from Landsat 8 thermal infrared images // J. Geophys. Res.: Atmos. 2015. V. 120, iss. 5. P. 1723–1738.
5. *Jun L., Timothy J.S., Xin J., Graeme M.* GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Legacy Atmospheric Moisture Profile, Legacy Atmospheric Temperature Profile, Total Precipitable Water, and Derived Atmospheric Stability Indices, document NOAA/NESDIS, version 3.0, July 2010. [Electronic resource]. URL: https://star.nesdis.noaa.gov/goesr/docs/ATBD/LAP.pdf (last access: 9.08.2021).
6. *Julien Y., Sobrino J.A., Mattar C., Jiménez-Mucoz J.C.* Near-real-time estimation of water vapor column from MSG-SEVIRI thermal infrared bands: Implications for land surface temperature retrieval // IEEE Trans. Geosci. Remote Sens. 2015. V. 53, N 8. P. 4231–4237.
7. *Gao B.-C., Kaufman Y.J.* Algorithm Technical Background Document, The MODIS Near-IR Water Vapor Algorithm, Product ID: MOD05 – Total Precipitable Water [Electronic resource]. URL: https://modis.gsfc.nasa.gov/data/atbd/atbd_mod03.pdf (last access: 9.08.2021).
8. *Albert P., Bennartz R., Preusker R., Leinweber R., Fischer J.* Remote sensing of atmospheric water vapor using the MODerate resolution imaging spectroradiometer // J. Atmos. Ocean. Technol. 2005. V. 22. P. 309–314.
9. *Gao B.-C., Kaufman Y.J.* Water vapor retrievals using MODerate resolution Imaging Spectroradiometer (MODIS) near-infrared channel // J. Geophys. Res. 2003. V. 108, iss. D13. P. 4389.
10. *Miao J., Kunzi K., Heygster G., Lachlan-Cope T.A., Turner J.* Atmospheric water vapor over Antarctica derived from Special Sensor Microwave/Temperature 2 data // J. Geophys. Res. 2001. V. 106, iss. D10. P. 10187–10203.
11. *Padmanabhan S., Reising S.C., Vivekanandan J., Iturbide-Sanchez F.* Retrieval of atmospheric water vapor density with fine spatial resolution using three-dimensional tomographic inversion of microwave brightness temperatures measured by a network of scanning compact radiometers // IEEE Trans. Geosci. Remote Sens. 2009. V. 47, iss. 11. P. 3708–3721.
12. *Kleespies T.J., McMillin L.M.* Retrieval of precipitable water from observations in the split window over varying surface temperatures // J. Appl. Meteorol. 1990. V. 29, iss. 9. P. 851–862.
13. *Li Z.-L., Jia L., Su Z., Wan Z., Zhang R.* A new approach for retrieving precipitable water from ATSR2 split-window channel data over land area // Int. J. Remote Sens. 2003. V. 24, iss. 24. P. 5095–5117.
14. *Barducci A., Guzzi D., Marcoionni P., Pippi I.* Algorithm for the retrieval of columnar water vapor from hyperspectral remotely sensed data // Appl. Opt. 2004. V. 43, iss. 29. P. 5552–5563.
15. *Sobrino J.A., Raissouni N., Simarro J., Nerry F., Petitcolin F.* Atmospheric water vapor content over land surfaces derived from the AVHRR data: Application to the Iberian Peninsula // IEEE Trans. Geosci. Remote Sens. 1999. V. 37, iss. 3. P. 1425–1434.
16. *Diouf D., Niang A., Thiria S.* Deep learning based multiple regression to predict total column water vapor (TCWV) from physical parameters in west Africa by using Keras library // International journal of Data Mining Knowledge Management Process. 2019. V. 9, N 6. P. 13–21.
17. *Palau J.L., Rovira F., Sales M.J.* Satellite observations of the seasonal evolution of total precipitable water vapour over the Mediterranean Sea // Adv. Meteorol. 2017. V. 2017, iss. 4790541.
18. *Ioffe S., Szegedy C.* Batch normalization: Accelerating deep network training by reducing internal covariate shift // Proc. of the 32nd Internat. Conf. Machine Learn. 2015. V. 37. P. 448–456.
19. *Kingma D.P., Ba J.L.* Adam: A method for Stochastic Optimization // ICLR. 2015. 15 p.

V.D. Bloshchinskiy, A.A. Filei, E.I. Kholodov. **Retrieval of water vapor content in atmospheric column from Electro-L No. 3 spacecraft data using neural networks.**

The application of a method based on artificial neural networks for assessing the total water vapor content in the atmospheric column from data of the MSU-GS satellite instrument of Electro-L No. 3 geostationary spacecraft is considered. The results of comparing the estimates of the total water vapor content from the MSU-GS data with MODIS satellite instrument data and AERONET measurements showed high agreement. The root mean square error when compared with the MODIS data was 0.311 cm, with the AERONET data, 0.409 cm, and the correlation was 98.2% and 84.7%, respectively. The results indicate the effectiveness of the method for determining the total content of water vapor for solving problems of atmospheric physics.



Примеры карт ОСВП по данным МСУ-ГС (*б*, *г*) и MODIS (*а*, *в*) за 28 января 2021 г. 07:30 UTC (*а*, *б*) и 2 марта 2021 г. 09:30 UTC (*в*, *г*)