

*М. Е. Сиденко<sup>1,2</sup>, С. А. Рылов<sup>2</sup>✉*

## **Влияние спектральных характеристик и методов предобработки спутниковых снимков при выделении водной поверхности**

<sup>1</sup>Новосибирский государственный университет, г. Новосибирск, Российская Федерация

<sup>2</sup>Федеральный исследовательский центр информационных и вычислительных технологий,  
г. Новосибирск, Российская Федерация  
e-mail: RylovS@mail.ru

**Аннотация.** В работе рассматривается задача выделения водной поверхности на оптических спутниковых снимках в условиях паводков. Был использован подход с предварительной кластеризацией изображения и последующей пороговой классификацией по спектральным индексам. Проведено сравнение качества результатов, полученных на данных со спутников Sentinel-2 и Landsat 8/9 при различных вариантах предварительной обработки данных. Получены результаты для двух вариантов расчета спектральных индексов, на оригинальных данных и на данных, полученных после линейного процентного растяжения. Также были рассмотрены два варианта расчета значений индексов для кластеров, при попиксельном усреднении и при вычислении индекса на усредненных значениях яркостей спектральных каналов. Экспериментальные результаты показали, что наилучшими индексами для данной задачи оказались MSAVI, ATSAVI, NDWI, NDVI, NIR.

**Ключевые слова:** аэрокосмические изображения, атмосферная коррекция, спектральные индексы, линейное растяжение, выделение водных объектов

*М. Е. Sidenko<sup>1,2</sup>, S. A. Rylov<sup>2</sup>✉*

## **The influence of spectral characteristics and preprocessing methods on water segmentation on satellite images**

<sup>1</sup>Novosibirsk State University, Novosibirsk, Russian Federation

<sup>2</sup>Federal Research Center for Information and Computational Technologies,  
Novosibirsk, Russian Federation  
e-mail: RylovS@mail.ru

**Abstract.** The paper considers the problem of water surface segmentation on optical satellite images under flood conditions. An approach with preliminary image clustering and subsequent threshold classification by spectral indices was used. A comparison of the quality of the results obtained on the data from Sentinel-2 and Landsat 8/9 satellites with different preliminary data processing options was performed. The results were obtained for two options for calculating the spectral indices, on the original data and on the data obtained after linear percentage stretching. Two options for calculating the index values for clusters were also considered, with pixel-by averaging and with index calculation on the averaged values of spectral bands. The experimental results showed that the best indices for this problem are MSAVI, ATSAVI, NDWI, NDVI, NIR.

**Keywords:** aerospace imagery, atmospheric correction, spectral indexes, linear stretching, allocation of water bodies

## ***Введение***

Полноценное решение задачи оперативного мониторинга паводковой ситуации не представляется возможным без использования данных дистанционного зондирования Земли. В настоящее время существует огромное количество космических аппаратов, осуществляющих оптическую съемку поверхности Земли, однако при решении задачи мониторинга паводковой ситуации наиболее подходящими являются спутники серии Sentinel-2 и Landsat-8/9. Это обусловлено их достаточно высоким пространственным разрешением, глобальным покрытием, регулярностью съемки, высоким качеством данных и доступностью.

Задача сегментации разливов в связи с паводками сложна тем, что вода в затопляемых районах имеет крайне широкий спектр возможных спектральных характеристик, из-за примесей растительности и почвы [1]. Проблему также усугубляет возможная высокая облачность в районах затопления или горный рельеф местности [2]. При выделении воды в таких условиях часто возникают ошибки сегментации водной поверхности на тенях от облаков или рельефа местности, увлажненных почвах и краях облаков, на которых часто наблюдается смещение каналов из-за их высоты и относительного движения [3]. Поэтому разработка алгоритма выделения воды, способного работать в сложных условиях является актуальной задачей.

Одним из важнейших этапов предобработки спутниковых снимков является атмосферная коррекция, особенно для использования пороговых методов сегментации с фиксированным порогом. Однако имеющиеся данные не всегда имеют атмосферную коррекцию и даже качественную радиометрическую коррекцию. В связи с этим интересна разработка алгоритмов выделения водной поверхности, не требующих обязательной атмосферной коррекции.

Цель данного исследования заключается в определении влияния уровня предобработки спутниковых снимков серии Sentinel-2 и Landsat-8/9 на работу методов выделения водной поверхности.

## ***Материалы и методы***

Сибирским Центром «НИЦ «Планета» были предоставлены снимки разливов рек, сделанные аппаратами Sentinel-2 в количестве 10 штук, из которых 4 снимка имеют уровень коррекции L1C, 3 снимка - уровень L2A, и еще 3 снимка с обоими уровнями коррекции. Также были получены снимки Landsat-8/9 в количестве 5 штук, из которых два имели уровень коррекции L1TP, один с уровнем L2SP и еще два с обоими уровнями коррекции.

Для каждого снимка было проведено линейное процентное растяжение данных исходных каналов Red, Green, Blue, NIR в диапазон [0, 255], с обрезкой 1% значений с краев. Данная операция призвана снизить влияние спектральных искажений и сдвигов на качество работы алгоритмов выделения водной поверхности.

В данной работе рассматривается подход с предварительной кластеризацией изображения и последующей классификацией полученных кластеров на

водные и неводные. Такой подхода позволяет повысить устойчивость результатов и снизить влияние выбросов [3, 4]. После линейного растяжения была проведена кластеризация изображений методом  $k$ -средних, с параметром  $k=50$ . Каждый из кластеров затем был вручную размечен на классы «вода», «почва», «непроницаемые облака». Кластеры, на которых алгоритмы не смогли разделить водные части от неводных, а также кластеры облаков были исключены из дальнейшего рассмотрения. Итоговый объем выборки составил 49 водных и 436 почвенных кластера для Sentinel-2, 26 водных и 60 почвенных кластера для Landsat.

Для каждого полученного кластера затем были вычислены следующие спектральные индексы: NDWI [5], NDVI, Datt, Datt2 [6], EVI [7], MSAVI [8], ATSAVI [9]. Выбор этих индексов обусловлен их известностью при решении задачи выделения водных объектов [10]. Индексы были посчитаны как на оригинальных данных, так и на данных после линейного растяжения. Спектральные индексы рассчитывались в двух вариантах – как усредненные значения индекса, вычисленного для всех пикселей кластера, и как отношение средних значений спектральных каналов по пикселям кластера. Второй вариант мог бы быть удобен для дальнейшего расчета новых индексов по имеющимся средним значениям яркостей спектральных каналов. При невозможности вычисления индекса в связи с делением на ноль, значение индекса в таких пикселях считалось равным нулю.

Для определения влияния линейного растяжения и методики подсчета индекса на разделимость классов были использованы пороговый метод для одномерного случая и алгоритм классификации случайного леса для многомерного случая. Для порогового метода метрикой качества была выбрана F1 с макроусреднением для минимизации влияния несбалансированности выборки. Порог выбирался исходя из максимизации F1 метрики.

### ***Результаты***

На рис. 1 представлены гистограммы распределения значений индекса NDWI, вычисленного на оригинальных данных для разных уровней атмосферной коррекции. Явно виден качественный характер различий этих распределений. На рис. 2 представлены гистограммы распределения значений индекса NDWI, посчитанного после применения линейного растяжения. Можно видеть, что качественное различие между гистограммами было устранено.

В ходе применения порогового алгоритма классификации к снимкам были выявлены признаки, показавшие лучшие значения по метрике F1. Вне зависимости от методики подсчета, пятерка лучших индексов получилась следующей: MSAVI, ATSAVI, NDWI, NDVI, NIR. В табл. 1 представлены результаты для оригинальных данных и при применении линейного растяжения в случае расчета индекса как среднего значения по пикселям кластера. В табл. 2 представлены результаты для случая расчета индексов по усредненным значениям яркостей спектральных каналов.

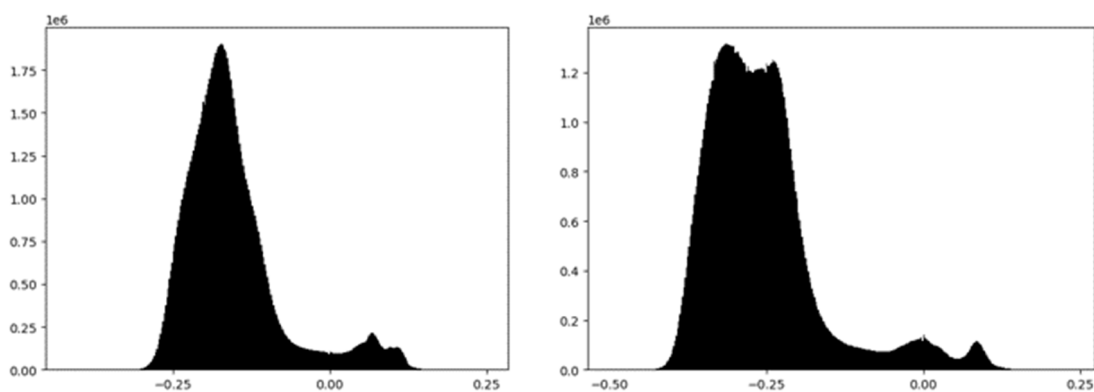


Рис. 1. Гистограммы индекса NDWI для изображения Sentinel-2: уровень предобработки L1C (слева) и уровень предобработки L2A (справа)

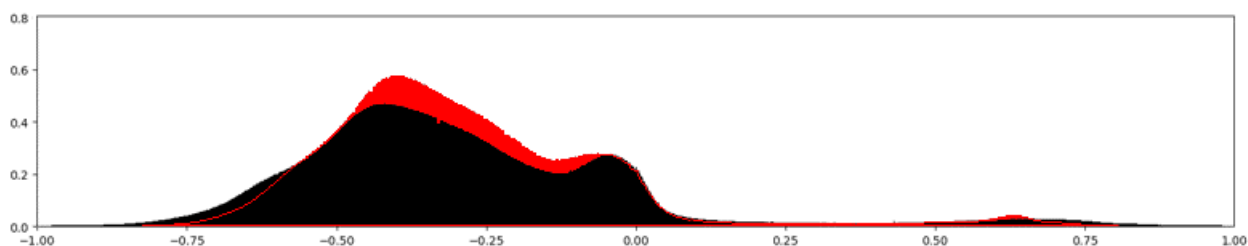


Рис. 2. Гистограммы индекса NDWI после линейного растяжения: для уровня предобработки L1C и уровня предобработки L2A (красным цветом)

Сравнив приведенные таблицы, можно увидеть, что результаты разделения с предварительным линейным процентным растяжением оказываются в среднем немного выше, чем на оригинальных данных. Также можно видеть, что при расчете индекса как среднего значения по пикселям кластера результаты оказались выше для всех индексов. В свою очередь, для многомерного случая влияние предварительного растяжения каналов на точность классификации требует более тщательного изучения, так как на исследуемом объеме данных алгоритм классификации случайного леса успешно справился с классификацией всех кластеров независимо от применения линейного растяжения и метода подсчета значений индексов. В этом случае по критерию Джини (Gini) пятерку наиболее значимых признаков составили: MSAVI, ATSAVI, NDWI, NDVI, NIR.

### *Заключение*

В работе рассматривалась задача выделения водной поверхности на оптических спутниковых снимках в условиях паводков. Был использован подход с предварительной кластеризацией изображения и последующей пороговой классификацией по спектральным индексам. Исследовалось влияние атмосферной коррекции и методов предобработки данных на качество выделения воды.

Проведено сравнение результатов, полученных при двух вариантах расчета спектральных индексов – на оригинальных данных и на данных после линейного процентного растяжения. Также использовались два варианта расчета значений

индексов для кластеров – при попиксельном усреднении и при вычислении индекса на усредненных значениях яркостей спектральных каналов.

Экспериментальные результаты, полученные на данных со спутников Sentinel-2 и Landsat 8/9, показали, что линейное растяжение позволяет снизить качественное различие между гистограммами водного индекса NDWI в случае наличия и отсутствия атмосферной коррекции. При этом применение линейного растяжения слабо влияет на получаемые результаты сегментации пороговым алгоритмом по выбранному спектральному индексу. При расчете индекса как среднего значения по пикселям кластера результаты оказались выше для всех индексов, соответственно такой способ является предпочтительным. Также были получены результаты сравнения качества результатов для различных спектральных индексов, самые высокие результаты показали MSAVI, ATSAVI, NDWI, NDVI, NIR (в порядке убывания). Применение алгоритма классификации случайного леса с использованием множества признаков показало высокую эффективность.

В дальнейшем планируется увеличить объем данных для подсчета статистики с целью выбора наилучших признаков и порогов для создания максимально универсального алгоритма выделения водной поверхности и добавление специализированных алгоритмов для обработки особых случаев [11].

*Таблица 1*

Результаты применения порогового алгоритма с методикой расчета индекса как усредненного значения по всем пикселям кластера

| Линейное растяжение | F1-macro | Оригинальные данные | F1-macro |
|---------------------|----------|---------------------|----------|
| MSAVI               | 0,898    | MSAVI               | 0,888    |
| NDWI                | 0,832    | NDWI                | 0,836    |
| ATSAVI              | 0,875    | ATSAVI              | 0,798    |
| NDVI                | 0,862    | NDVI                | 0,786    |
| NIR                 | 0,729    | NIR                 | 0,693    |

*Таблица 2*

Результаты применения порогового алгоритма с методикой расчета индекса по усредненным значениям яркостей спектральных каналов

| Линейное растяжение | F1-macro | Оригинальные данные | F1-macro |
|---------------------|----------|---------------------|----------|
| MSAVI               | 0,828    | MSAVI               | 0,861    |
| NDWI                | 0,825    | NDWI                | 0,829    |
| ATSAVI              | 0,841    | ATSAVI              | 0,786    |
| NDVI                | 0,837    | NDVI                | 0,780    |
| NIR                 | 0,729    | NIR                 | 0,693    |

## ***Благодарности***

Авторы выражают благодарность Сибирскому центру ФГБУ «НИЦ «Планета» за предоставленные спутниковые снимки и комментарии по их обработке.

## **БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

1. Portalés-Julià E., Mateo-García G., Gómez-Chova L. Understanding Flood Detection Models Across Sentinel-1 and Sentinel-2 Modalities and Benchmark Datasets. 2025. [Электронный ресурс]. URL: <https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=5118486> (дата обращения: 20.05.2025).
2. Amitrano D., Martino G.Di, Simone A.Di, Imperatore P. Flood detection with SAR: A review of techniques and datasets // *Remote Sensing*. – 2024. – Vol. 16. – No. 4. – P. 656.
3. Рылов С.А., Новгородцева О.Г., Пестунов И.А., Дубровская О.А. Технология обработки данных с космических аппаратов "Канопус-В", "Ресурс-П" и "Метеор-М" для мониторинга и картографирования паводковой ситуации // Региональные проблемы дистанционного зондирования Земли : Материалы II Международной научной конференции, Красноярск, 22–25 сентября 2015 года. – Красноярск: СУФ, 2015. – С. 207-212.
4. Rylov S.A., Pestunov I.A. Fast hierarchical clustering of multispectral images and its implementation on NVIDIA GPU // *Journal of Physics: Conference Series (JPCS)*. – 2018. – Vol. 1096. – P. 012039.
5. McFeeters, S.K. Using the Normalized Difference Water Index (NDWI) within a geographic information system to detect swimming pools for mosquito abatement: a practical approach // *Remote Sensing*. – 2013. – Vol. 5. – No. 7. – P. 3544-3561.
6. Main R., Cho M.A., Mathieu R., O'Kennedy M.M., Ramoelo A., Koch S. An investigation into robust spectral indices for leaf chlorophyll estimation // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2011. – Vol. 66. – No. 6. – P. 751-761.
7. Huete A.R., Liu H.Q., Batchily K., Leeuwen W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS // *Remote sensing of environment*. – 1997. – Vol. 59. – No. 3. – P. 440-451.
8. Qi J., Chehbouni A., Huete A.R., Kerr Y.H., Sorooshian S. A modified soil adjusted vegetation index // *Remote sensing of environment*. – 1994. – Vol. 48. – No. 2. – P. 119-126.
9. Baret F., Guyot G. Potentials and limits of vegetation indices for LAI and APAR assessment // *Remote sensing of environment*. – 1991. – Vol. 35. – Iss. 2-3. – P. 161-173.
10. Кучма М.О. Применение физических и нейросетевых методов в задаче оперативного детектирования водной поверхности // *Метеорология и гидрология*. – 2024. – № 4. – С. 67-77.
11. Рылов С.А. Использование активного обучения для выделения водных объектов на спутниковых изображениях // Региональные проблемы дистанционного зондирования Земли: Материалы VII Международной научной конференции, Красноярск, 29 сентября – 02 октября 2020 года / СФУ, ИКИТ. – Красноярск: СФУ, 2020. – С. 122-126.

© М. Е. Сиденко, С. А. Рылов, 2025