2. НОВЫЕ ТЕХНОЛОГИИ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ И РАБОТЫ С ДАННЫМИ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ (ДДЗ)

УДК 528.873

DOI: 10.23885/2500-123X-2020-2-5-56-62

ОБНАРУЖЕНИЕ БЕРЕГОВОЙ ЛИНИИ ПО ДАННЫМ ДЗЗ СРЕДНЕГО ПРОСТРАНСТВЕННОГО РАЗРЕШЕНИЯ

Е.И. Волынец, А.В. Волынец, Е.А. Паниди

Санкт-Петербургский государственный университет
evgenia59-007@mail.ru, alvolynets1992@mail.ru, panidi@ya.ru

Аннотация. Исследование посвящено разработке метода обнаружения береговой линии и ее характеристик по материалам спутниковых съемок среднего пространственного разрешения. Предлагается использовать полуавтоматический метод на основе неконтролируемой классификации мультиспектрального снимка среднего разрешения по водным индексам с последующим экспертным уточнением классов и применением методов машинного обучения. Приводится описание методики и анализ результатов обработки изображений на примере участка прибрежной зоны восточной части Финского залива.

Ключевые слова: Дистанционное зондирование, ГИС, машинное обучение, береговая линия, спутниковое картографирование

Введение. Задача обнаружения береговой линии и определения ее динамики имеет решающее значение в принятии мер по предотвращению береговой эрозии и управлению прибрежными зонами.

Методы и данные дистанционного зондирования (ДДЗ) широко применяются в мониторинге прибрежных зон, поскольку обеспечивают исследователей радиометрической информацией с возможностью полу- или полной автоматизации процесса обработки данных. Популярность использования ДДЗ со средним пространственным разрешением обусловлена отсутствием данных съемки сенсоров с высоким пространственным разрешением в свободном доступе.

Проблемой для картографирования береговой линии с использованием спутниковых данных среднего разрешения является наличие “смешанных” пикселей – смеси различных пространственных классов из-за относительно грубого пространственного разрешения. Часто решения сводятся к конкретной задаче (положению) или съемочной системе, при этом ограничивается применимость методики.

Цель исследования заключается в разработке метода обнаружения береговой линии по материалам спутниковых съемок среднего пространственного разрешения.

Район исследований. Исследование проводилось на участке прибрежной зоны восточной части Финского залива, характеризующейся специфическим режимом морфодинамики, вызванным колебаниями уровня моря и интенсивной антропогенной деятельностью. В качестве тестовых областей

по результатам экспертной оценки были выбраны 2 участка (рис.1) – сев. побережье о. Котлин и пляж от пос. Репино до г. Сестрорецк. Выбор объясняется активной аккумуляцией песчаного материала в первой области интереса и преобладанием процессов отступления береговой линии вследствие влияния интенсивных абразионных процессов, развитии интенсивной рекреационной инфраструктуры во второй области. Наиболее существенное воздействие на береговую зону восточной части Финского залива оказывает Комплекс защитных сооружений Санкт-Петербурга от наводнений (КЗС), а также осенние и зимние шторма [1].

Исходные данные. В работе был использован снимок Landsat 8 от 25 июля 2019 года, соответствующий уровню обработки Level-2 в проекции UTM 35N зоны. В качестве опорной была использована береговая линия, оцифрованная по мультиспектральному снимку высокого пространственного разрешения (2 м) аппарата Ресурс-П1 уровня обработки 2А [2] от 26 июля 2019 года, предоставленного Научным центром оперативного мониторинга Земли ОАО “Российские космические системы”. Интервал между съемками был выбран максимально коротким для минимизации смещения границ водных объектов.

Методы. В исследовании для определения положения береговой линии рекомендуется полуавтоматический метод на основе неконтролируемой классификации мультиспектрального снимка среднего пространственного разрешения по водным индексам с последующим экспертным уточнением классов

и применением методов машинного обучения. Схема основных этапов исследования приведена на рис.2:

Обзор использованных водных индексов. Согласно схеме (рис.2), для определения положения береговой линии были использованы 4 водных индекса с двумя вариантами пороговых значений

с целью построения бинарных растров. Для выбора и оценки применимости водных индексов проводилась оценка точности между береговыми линиями, полученными с разных индексных изображений. Аналогичная оценка точности проводилась для выбора оптимального порогового значения.

ROI – 1 (сев.побережье о.Котлин), ROI – 2 (пляж от пос.Репино до г.Сестрорецк)

Рис. 1. Тестовые области, снимок Landsat

Рис. 2. Схема основных этапов исследования

Методы, основанные на определениях водных индексов, решают задачу различимости водных и неводных пикселей на основе разделения спек-

тральных каналов и операций с ними

Для определения береговой линии в исследовании были применены водные индексы (табл.1):

Таблица 1. Используемые водные индексы

Индекс	Формула	Комментарий
NDWI	$(\text{Green}-\text{NIR})/(\text{Green}+\text{NIR})$	Застроенные участки земли могут получить положительные значения [3]
MNDWI	$(\text{Green}-\text{SWIR})/(\text{Green}+\text{SWIR})$	Уточнение NDWI. Не всегда возможно отличить смешанные территории растительного покрова и водоемов [4]
AWEI	$4*(\text{Green}-\text{MIR}) - (0,25*\text{NIR} + 2,75*\text{SWIR})$	Для районов исследования объектов с малой величиной альбедо [5]
EWI	if (NDVI<0 and NDWI >0), 'water' else 'land'	Повышение точности извлечения водных поверхностей при наличии шумов, характерных для речных бассейнов. Результат применения – бинарная дискретная маска [6]

Построение всех индексных изображений было рассчитано по формулам, приведенным в таблице, с помощью QGIS 3.10.

Определение порога для построения бинарных растров. В исследовании было рассмотрено два метода для определения порогового значения при построении бинарных растров – классический метод фиксированного значения (0) и метод глобального порога (метод Отсу) [7].

Определение глобального порога Отсу было выполнено с помощью пакета skimage языка Python на основе индексного изображения MNDWI. Новое расчетное значение порога получилось равным 0.235. Для оценки полученных результатов была проведена оценка точности на основе вычисления ближайших расстояний от сгенерированных береговых линий до наземных точек, оцифрованных с опорного снимка Ресурс-П1. Все вычисления были выполнены

в QGIS 3.10. Было обнаружено, что при использовании глобального порога в болотистой местности (ROI-1) береговая линия переместилась в сторону открытой воды и количество водных пикселей сократилось (рис. 3).

Оценка точности показала следующие результаты: средняя, максимальная ошибки и стандартное отклонение для всех точек 119.53, 569.83, 130.89 или 159.05, 586.70, 147.25 м при использовании фиксированного и глобального порога соответственно. Преимущество фиксированного порога сохраняется и при расчетах оценки точности для каждого тестового участка отдельно. Таким образом, береговая линия, сгенерированная с индексного изображения MNDWI с фиксированным порогом, имеет большую точность. В связи с этим, был выбран фиксированный порог при построении бинарных растров.

Рис. 3. Береговые линии для ROI-1. Красным с прозрачностью показаны области, которые переклассифицированы как суша после применения глобального значения, голубым – новая береговая линия, желтым – береговая линия с индексного изображения MNDWI с порогом 0

Разработка модели для генерации береговых линий. Для автоматической генерации береговых линий была разработана модель в QGIS 3.10 Model builder, реализующая этапы:

- построения бинарных изображений с помощью калькулятора растров;
- морфологической фильтрации “closing” для заполнения небольших по размерам пустот, сохраняя при этом форму и размер объектов на изображении;
- отсеивания (удаление незначительных растровых полигонов (в пикселах), заменяя значением соседних, более значительных растровых полигонов);
- конвертации “растр-в-вектор”;
- фильтрации векторного слоя с выбором только водных объектов;
- записи результата обработки в файл.shp.

Машинное обучение. Определение береговых линий лишь по данным водных индексов в болотистых районах показывает существенные ошибки. В настоящем исследовании для уточнения классификации после использования водных индексов были применены методы машинного обучения, имплементированные на языке Python в библиотеке «scikit-learn».

В исследовании данные представляют собой числовые значения, по которым необходимо определить

принадлежность к классу по выборке с помеченными данными. Выбор модели проводился по схеме [8], согласно которой был использован алгоритм классификации RandomForestClassifier.

В основе реализации алгоритма Random Forest для уточнения типа береговой линии была использована комбинация полученных водных индексных изображений. При этом за основу классификации был взят результат, полученный со снимка MNDWI, как наиболее точного, который был визуально конкретизирован с выделением класса болота.

Обучающая выборка включала в себя как переходные от класса суши к классу воды полигоны, так и однородные, содержащие пиксели только одного класса (вода, суша, болото), а тестовые области состояли только из переходных полигонов, т.к. основная задача состояла в корректном определении класса. В общей сложности было извлечено около 7000 обучающих точек со всех индексных изображений. Количество тестовых точек составляло примерно 35 % от числа обучающих.

Результаты и обсуждение. Индексные изображения. Для оценки применимости водных индексов при нахождении береговой линии была проведена оценка точности положения береговых линий, сгенерированных по индексным изображениям NDWI, MNDWI, AWEI, EWI (табл. 2, рис. 4):

Таблица 2. Точность сгенерированных береговых линий

водный индекс	минимальная ошибка, м	максимальная ошибка, м	Средняя ошибка, м	стандартное отклонение
ROI-1 (2334 точек)				
AWEI	0.498542	572.056521	152.301619	138.824979
MNDWI	0.498542	569.833353	129.460873	132.906867
NDWI	1.128436	608.568476	197.338150	151.756939
EWI	1.909613	628.033812	206.793344	153.896891
ROI-2 (226 точек)				
AWEI	0.890899	75.068103	22.373226	12.663186
MNDWI	0.732103	62.248399	16.928537	10.127193
NDWI	0.890899	91.968581	18.347650	13.981650
EWI	0.890899	91.968581	22.706786	15.805697

Очевидно, что для обоих тестовых участков MNDWI показал лучшую точность в сравнении с другими индексами. Однако все сгенерированные береговые линии имели существенные ошибки в болотистых районах, где суша и вода были плохо классифицированы.

Машинное обучение. Если проследить за ходом кривой изменения водных индексов (NDWI/MNDWI) и сопоставить точки с их расположением на местности, то можно заметить, что болотистая местность (точки с 561 по 610, рис. 5 а) определяется

достаточно четко: явно отрицательная стабильная область значений, соответствующих классу суши, сменяется скачкообразной кривой со значениями, ближе к пороговому, для класса заболоченностей и завершается вновь стабильно положительным ходом кривой для класса воды. В общем случае, при отсутствии переходных зон и наличии достаточно четкой границы между классами суши и воды, наблюдается значительный интервал в значениях индекса при переходе порогового значения от воды к суше, что хорошо видно на тестовом участке с пля-

жем (рис. 5 б). При этом следует отметить, что разница значений между водными и неводными пикселями для индекса MNDWI больше чем для индекса NDWI на участке с пляжем, а на заболоченном участ-

ке разница значений пикселей примерно одинакова для обоих индексов. Кроме того, разница значений пикселей между незаболоченными и заболоченными территориями также выше для индекса MNDWI.

Рис. 4. Опорная (а) и сгенерированные (б) береговые линии за 2019 год, ROI-1,2. Желтым обозначена береговая линия MNDWI, оранжевым – AWEI, зеленым – EWI, фиолетовым – NDWI, красным – референсная береговая линия

Рис. 5. Ход кривых изменения величины водных индексов на обучающих выборках

Результаты обучения модели показали, что самым значимым параметром при разделении набора данных на классы является значения пиксела по MNDWI.

Оценка точности модели машинного обучения проводилась методом `accuracy_score` из библиотеки `sklearn.metrics`: точность предсказаний модели полу-

чилась равной 96.0 %, точность прогноза классификатора – 95.3 %.

Результаты классификации с использованием модели показали значительное повышение точности определения как положения береговой линии, так и определения заболоченных территорий (рис. 6).

Общая картина распределения ошибочно классифицированных пикселей позволяет сделать вывод, что предложенный метод показывает достаточную эффективность в определении

заболоченностей на данных тестовых участках – средняя ошибка определения болот составила 4.5 % (9 % и 1 % для ROI-1 и ROI-2 соответственно).

Рис. 6. Общая картина результатов классификации после применения алгоритма Random Forest на тестовой выборке ROI-1. Желтым отмечены ошибочно классифицированные значения

Заключение. В настоящем исследовании для определения положения береговой линии предлагается модель, реализующая полуавтоматический метод неконтролируемой классификации мультиспектрального снимка среднего пространственного разрешения по водным индексам с последующим экспертным уточнением классов и применением алгоритмов машинного обучения совместно с методами ГИС-анализа.

Сравнение применимости водных индексов показало, что в тестовых областях наилучшие результаты определение береговой линии дает Модифицированный нормализованный разностный водный индекс MNDWI с фиксированным пороговым значением равным нулю.

Для уточнения типа береговой линии, в особенности на заболоченных участках, где водные индексы показали плохую точность, был применен алгоритм машинного обучения Random Forest. Исследование показало, что применение выбранного алгоритма ведет к повышению точности. Однако, следует отметить, что значительное влияние на точность результата оказывает точность определения заболоченных территорий для обучения модели, которое в данной работе производилось вручную.

Полученные результаты свидетельствуют о необходимости продолжения исследований, нацеленных на поиск наиболее эффективных параметров алгоритма и комбинации водных индексов.

Список литературы

1. Российская Академия Наук, Институт океанологии им. П.П. Ширшова. Научное обеспечение сбалансированного планирования хозяйственной деятельности на уникальных морских береговых ландшафтах и предложения по его использованию на примере Азово-Черноморского побережья. Под редакцией Р.Д. Косьяна, Геленджик 2013
2. Перечень стандартных продуктов, изготавливаемых по материалам ДЗЗ с КА типа «Ресурс-П». URL-адрес: http://www.ntsomz.ru/for_clients/standart_ip/standart_ip_resurs_p
3. McFeeters S.K. 1996, The use of normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17, pp.1425–1432
4. Xu H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery // *International Journal of Remote Sensing*. 2006
5. Feyisa, Gudina Legese & Meilby, Henrik & Fensholt, Rasmus & Proud, Simon. (2014). Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sensing of Environment*. 140. 23–35.
6. Ayana E.K., Philpot W.D., Melesse A.M., Steenhuis T.S. Assessing the potential of MODIS/Terra version 5 images to improve near shore lake bathymetric surveys. *Int. J. Applied Earth Observ. Geoinform.* 2015, 36.
7. Otsu N. 1979. “A Threshold Selection Method from Gray-Level Histograms.” *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1).
8. Библиотека scikit-learn, схема для выбора подходящей классификации. URL-адрес: https://scikit-learn.org/stable/tutorial/machine_learning_map/#choosing-the-right-estimator

USING MEDIUM RESOLUTION SATELLITE REMOTE SENSING DATA FOR COASTLINE DETECTION AND INTERPRETATION

E.I. Volynets, A.V. Volynets, E.A. Panidi

St Petersburg State University

evgenia59-007@mail.ru, alvolynets1992@mail.ru, panidi@ya.ru

Abstract. Remote sensing data are widely used in coastal zones monitoring, since they provide radiometric information with the possibility to automate the data processing. Due to the lack of high resolution satellite images in free access using the medium resolution satellite images is widespread. The study is dedicated to the development of a coastline detection method based on medium resolution satellite images. It is proposed to use a semi-automatic method based on uncontrolled classification of a mid-resolution image by water indices, followed by expert refinement of classes and the use of machine learning methods. The shorelines of the eastern part of the Gulf of Finland have been extracted from Landsat 8 Level-2 image, using proposed method. The position accuracy of the generated shorelines has been analyzed using manually digitized shoreline from high-resolution image Resurs-P1 processing level 2A. The results showed that in the test areas the best output for extracting the coastline are given by the MNDWI with a fixed threshold value equal to zero. The Random Forest machine learning algorithm was used to refine the type of coastline, especially in wetlands where water indices showed poor accuracy. The study showed a significant increase in the position accuracy with the use of the algorithm. However, the accuracy of manually classified wetlands for training the model has a significant impact on the result.

Keywords: remote sensing, GIS, machine learning, coastline, satellite mapping.

References:

1. Russian Academy of Sciences, P.P. Shirshov Institute of Oceanology. Scientific support for balanced business planning on unique marine coastal landscapes and suggestions for its use on the example of the Azov-Black Sea coast. Edited by R.D. Kosyan, Gelendzhik 2013 (In Russian)
2. Resurs-P Standard Data Products. http://www.ntsomz.ru/for_clients/standart_ip (In Russian)
3. *McFeeters S.K.* 1996, The use of normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17, pp.1425–1432 (In English)
4. *Xu H.* Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery // *International Journal of Remote Sensing*. 2006 (In English)
5. Feyisa, Gudina Legese & Meilby, Henrik & Fensholt, Rasmus & Proud, Simon. (2014). Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sensing of Environment*. 140. 23–35. (In English)
6. *Ayana E.K., Philpot W.D., Melesse A.M., Steenhuis T.S.* Assessing the potential of MODIS/Terra version 5 images to improve near shore lake bathymetric surveys. *Int. J. Applied Earth Observ. Geoinform.* 2015, 36. (In English)
7. *Otsu N.* 1979. "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1). (In English)
8. Scikit-learn library, flowchart for choosing the right estimator. https://scikitlearn.org/stable/tutorial/machine_learning_map/#choosing-the-right-estimator (In English)