



УДК 556

ИСПОЛЬЗОВАНИЕ ГЕОГРАФИЧЕСКИХ ИНФОРМАЦИОННЫХ СИСТЕМ И ВЫСОКОУРОВНЕВОГО ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON ДЛЯ ОПРЕДЕЛЕНИЯ СОСТОЯНИЯ ГОРЬКОВСКОГО ВОДОХРАНИЛИЩА

Лыкова Елизавета Алексеевна, студент бакалавр
Московский Государственный Технический Университет им. Н. Э. Баумана
105005, г. Москва, 2-я Бауманская ул., 5

Лыкова Елена Сергеевна, доцент кафедры экономики
ФГБОУ ВО «Волжский государственный университет водного транспорта».
603950, г. Нижний Новгород, ул. Нестерова, 5

Аннотация. В данной статье был изучен вопрос использование геопространственных данных и высокоуровневого языка программирования Python для подсчета площади Горьковского водохранилища. Недостатком использования языка программирования Python при в обработке спутниковых изображений является большая погрешность из-за некорректности снимков, предлагаемых свободными источниками. Преимуществами использования – быстрый анализ результатов за счет готовых библиотек.

Ключевые слова: геопространственные данные, векторный формат изображений, язык программирования Python, площадь Горьковского водохранилища, предсказание результатов, статистика площади водохранилища, график предсказания результата.

На сегодняшний день геоинформационные системы применяют в различных областях жизнедеятельности человека: при организации сельского хозяйства, планировании тур маршрута, анализе экологических тенденций и в других сферах. Геоинформационные системы включают в себя системы сбора, хранения и анализа информации в виде графической интерпретации. Таким образом, геоинформационные системы позволяют значительно облегчить работу с цифровыми картами.

В данной статье будет рассмотрена работа с геопространственными данными, которые позволяют вычислить площадь Горьковского водохранилища. Форматы геопространственных данных различное множество, однако выделим самые распространенные:

- векторный формат – набор файлов, состоящие из геометрических данных, их специальных атрибутов и прочее;
- georaskage – данные хранятся в файле, который представляет собой контейнер базы данных SQLite;
- geoJSON – данные хранятся в стандартном текстовом формате JSON.

Для анализа площади водоема будем использовать векторный формат данных. В таком случае геопространственные данные представляют собой векторные объекты[2]. Так,

например, точкой может быть представлены дома, ломанной – реки, полигоном – водохранилища, озера. Также в векторном виде с помощью дуг, узлов, вершин задается топологическая информация, которая позволяет обнаружить связи между объектами.

Специальная библиотека GeoPandas, доступная при написании кода на высокоуровневом языке программирования Python, упрощает задачу загрузки геоданных. С помощью этой библиотеки удастся одной командой загружать данные различных форматов в объект GeoDataFrame, особенностью которого является наличие дополнительного столбца с данными о местоположении в каждой строке.

При написании кода на языке Python подсчет площади Горьковского водохранилища был осуществлен за счет подсчета пикселей загруженного спутникового изображения за разные года. После загрузки растрового изображения, были считаны его первые два слоя, которые отвечали за цветовые каналы. После этого была создана маска, которая отделяла пиксели, отвечающие за воду на изображении, от пикселей местности. Расчет конечной площади водной поверхности осуществлялся подсчетом итоговой суммы пикселей, прошедших через созданную маску. Последним шагом для получения искомой площади было перемножения параметра src.res, который определяет размер одного пикселя в пространственных единицах, на рассчитанное суммарное количество пикселей-воды.

Таким образом, расчет площади Горьковского водохранилища на карте может быть произведен с помощью написанной функции

```
def calculate_water_area(image_path):
    # Открываем растровое изображение
    with rasterio.open(image_path) as src:
        # Читаем первые два канала
        band1 = src.read(1) # читаем первый слой
        band2 = src.read(2) # читаем второй слой
        # Здесь мы создаем маску для синего цвета
        # Простой способ: выбрать порог значений
        water_mask = (band1 < 100) & (band2 < 100) # условие для определения воды
        # Подсчитываем площадь
        water_area_pixels = np.sum(water_mask)
        # Пиксели в физические единицы площади
        pixel_area = (src.res[0] * src.res[1]) # площадь одного пикселя в квадратных
        метрах
        water_area_sqm = water_area_pixels * pixel_area
    return water_area_sqm
```

Горьковское водохранилище было образовано 1957 году на территории нескольких областей: Ивановской, Ярославской, Нижегородской и Костромской. Известно, что на данный момент площадь водохранилища равна 1590 км² средней глубиной в 3,65 м, а максимальной – 22 м [1]. Найдем площадь водохранилища за 01.10.2024, которое приведено на рисунке 1.



Рис. 1. Изображение со спутника Горьковского водохранилища 01.11.2024.

После обработки, программа выдала:

Площадь воды за 10.2024: 1058134000.00 м²

Таким образом, абсолютная погрешность составила 33%, что значительно влияет на доверие полученным с помощью реализованной программой результатам. Такую большую погрешность можно объяснить тем, что изображения со спутника были использованы с помощью открытых для общего пользования сервисов. Если использовать платные версии, тогда растровое изображение будет более точным, а также может включать дополнительную уточняющую информацию.

Аналогичным образом был произведен расчет площади воды по фотографиям за несколько лет: 2013, 2018, 2019, 2020, 2023, 2024, где удалось найти наиболее четкие изображения Горьковского водохранилища со спутника. Программы выдала:

Площадь воды за 07.2013: 493197000.00 м²
Площадь воды за 07.2024: 441985000.00 м²
Площадь воды за 08.2018: 446638000.00 м²
Площадь воды за 10.2024: 1058134000.00 м²
Площадь воды за 11.2023: 1057790000.00 м²
Площадь воды за 04.2018: 918452000.00 м²
Площадь воды за 05.2018: 458658000.00 м²
Площадь воды за 08.2019: 468539000.00 м²
Площадь воды за 10.2020: 1053696000.00 м²
Площадь воды за 04.2020: 1017106000.00 м²

Проанализировать полученные результаты можно с помощью построения графика, представленного на рисунке 2.



Рис.2.График подсчета площади Горьковского водохранилища с 2013-2024 год.

С помощью графика легко можно проанализировать полученные данные. Таким образом, площадь Горьковского водохранилища за 07.2013, 08.2018, 08.2019 и 07.2024 находилась примерно на одном уровне. Наибольший показатель площади был выявлен в 2013 году за сравниваемые летние месяцы. Площадь Горьковского водохранилища за 10.2020, 11.2023, 10.2024 также находится примерно на одном уровне и достигает максимального значения в 2020 и 2023 годах за сравниваемые осенние месяцы. Площадь за летние месяцы значительно меньше площади, полученная в результате обработки спутниковых снимков за осенние периоды.

Попробуем предсказать площадь Горьковского водохранилища на летний месяц 2025 года, для этого за основу возьмем данные, полученные за 07.2013, 08.2018, 08.2019 и 07.2024. Для решения поставленной задачи используем готовую библиотеку `sklearn.linear_model`, которая позволит быстро реализовать линейную регрессию. Линейная регрессия – один из методов статического анализа, используемый для предсказания зависимой переменной на основе независимых переменных. В нашем случае функция будет выглядеть подобным образом:

$$y = ax + b, \quad (1)$$

где y – предсказание зависимой переменной (площадь водоема в 2025 году),

x – независимая переменная (год), a – угловой коэффициент, b – свободный член.

Реализованная библиотекой модель старается так подобрать угловой коэффициент и свободный член, чтобы ошибка между предсказанным значением и фактическим было минимальным (метод наименьших квадратов)[3]. Полученный график нахождения предсказанного значения приведен на рисунке 3. Таким образом, полученное значение равно:

Предсказанная площадь воды в 07-08 2025 года: 430 км²



Рис.3.График нахождения площади Горьковского водохранилища на 2025 год.

Подводя итог, можно сказать, что работа с геопространственными данными с использованием высокоуровневого языка Python возможно. Однако стоит обратить внимание на внешние сервисы, предлагаемые растровые спутниковые изображения. Доступные в общем пользовании предоставляют изображения с большой погрешностью, а потому могут быть неприменимы для нахождения точных результатов. Тем не менее, наличие огромного количества встроенных библиотек позволяют быстро проанализировать данные и даже сделать предсказание на будущее, что значительно облегчает анализ геоданных.

Список литературы:

1. Трубе Л.Л., География Горьковской области. – Горький: Волго-Вятское книжное издательство, 1978. – 176 с.
2. Кариев Ч.А., Масштабируемая векторная графика, 2007. – 123 с.
3. Дж. Себер., Линейный регрессионный анализ. – Мир, 1980. – 456 с.
4. Экономико-математические методы и модели в управлении морским транспортом: Учеб. для студентов морских вузов. / Е.Н. Воевудский, Н.А. Коневцева, Г.С. Махуренко, И.П. Тарасова; под ред. Е.Н. Воевудского. - М.: Транспорт, 1988. - 384 с.

USING GEOGRAPHIC INFORMATION SYSTEMS AND THE HIGH-LEVEL PROGRAMMING LANGUAGE PYTHON TO DETERMINE THE STATE OF THE GORKY RESERVOIR

Likova A. Elisaveta, Likova S. Elena,

Abstract. This article explored the use of geospatial data and the high-level Python programming language to calculate the area of the Gorky Reservoir. The disadvantage of using the Python programming language when processing satellite images is the large error due to the incorrectness of the images offered by free sources. The advantages of using it are fast analysis of results due to ready-made libraries.

Keywords: geospatial data, vector image format, Python programming language, Gorky Reservoir area, prediction of results, reservoir area statistics, prediction graph of results.