KazRivDyn: Toolkit for Measuring the Dynamics of Kazakhstan Rivers with a Graphics Based on Google Earth Engine

Assel Ospan[1], Madina Mansurova[2*], Erkin Kakimzhanov[3], Baurzhan Aldakulov[4]

[1]Al-Farabi Kazakh National University, Kazakhstan, asselyaospan@gmail.com
[2*]Al-Farabi Kazakh National University, Kazakhstan, mansurova.madina@gmail.com
[3] Al-Farabi Kazakh National University, Kazakhstan, erkinkakimzhanov@gmail.com,
[4]Al-Farabi Kazakh National University, Kazakhstan, aldakulovb@mail.ru

Abstract – Now it is possible to control the change in the width of the rivers of Kazakhstan using remote sensing. This article presents a platform called KazRivDyn, developed on the Google Earth Engine cloud computing platform, to monitor changes in the width of Kazakhstan's rivers over the past 20 years, with a graph for more accurate data. Due to the fact that in Kazakhstan there is a problem of reducing the volume of water in rivers, identify the general trend of changing the volume of water, as well as turn prevention to prevent such phenomena as drought and pollution. This platform has been applied to the pool. This platform flows through two countries, the darkest as the width of the river has changed since 1984. KazRivDyn is a publicly available tool and can be used to solve scientific problems related to rivers, as well as to create applications for operational water resources management. The results obtained are close to measurements taken using manual methods, and the application works for all rivers in Kazakhstan.

Keywords – *KazRivDyn, Google Earth Engine, river dynamics, water resources, river monitoring.*

I. INTRODUCTION

After the launch of the Google Earth Engine (hereinafter GEE) platform, it became possible to use satellite images for scientific purposes, as well as apply different types of algorithms to add additional tools to the program.

Google Earth Engine is a cloud-based, planet-wide geospatial analysis platform that leverages Google's immense computing power to tackle critical social challenges such as deforestation, drought, natural disasters, disease, food security, water management, climate monitoring and protection. Environment. It is unique in the field as an integrated platform designed to empower not only traditional remote sensing scientists, ut also a much wider audience that lacks the technical capabilities required to leverage traditional supercomputers or large-scale cloud computing resources [1].

When developing the KazRivDyn tool, the following were used: the GEE cloud computing platform; RivWidthCloud algorithm - method for automatically extracting river width from remote sensing images; Global River Widths from Landsat (GRWL) database to distinguish river pixels from water pixels outside the river; Timelapse for time-lapse interactive time-lapse animation of a river. The purpose of this article is to present a ready-made tool as an application where you can observe river changes, with the ability for users to set the timeline themselves. Based on the results obtained, it is possible to build a general trend in the dynamics of rivers, and, if necessary, take measures. In this program, changes in centerline and width along river channels were used as the basis for river dynamics. [2].

Next, we describe the algorithm, verify its output, and discuss possible next steps for KazRivDyn.

Prior to the development of this application, a literature review was conducted to compare different flow monitoring methods and determine the most appropriate methods for different flow sizes in different areas. Web resources and technical databases such as ScienceDirect, JSTOR, Wiley Online Library and Google Scholar were searched and related literature downloaded.

There are various methods for quantifying and monitoring surface waters. These methods are grouped into four categories [3]: (1) direct measurement methods, (2) velocity area measurement methods, (3) flow restriction or limitation methods, and (4) non-contact measurement methods.

The previous 3 measurement methods involve manual field measurements and are detailed in [4-6]. To date, the RivWidth [2] algorithm based on the Global River Widths Landsat (GRWL) [7] has been successfully used. According to the results of work [2], it was proved that RivWidth corresponds to accurate manual measurement methods. Therefore, our application is being developed using the RivWidth algorithm.

II. METHODOLOGY

The development of the KazRivDyn toolkit includes the following stages:

- Using the automated GEE algorithm to extract river width from remote sensing images [2];
- Global River Widths from Landsat (GRWL) database to distinguish river pixels from water pixels outside the river;
- Insertion of a Timelapse function for time-lapse interactive time-lapse animation of a river.
- Establishing a timeline from 1984 to 2020.
- Building a graph of changes in the width of the river over time.

KazRivDyn toolkit is based on the automatic RivWidthCloud algorithm and Landsat's Global River Widths dataset (GRWL) to distinguish river pixels from non-constant water pixels. RivWidthCloud is a software suite developed on the GEE cloud computing platform. RivWidthCloud automatically extracts the centerline and width of a river from optical satellite imagery, with the ability to mark observations obstructed by features such as clouds, cloud shadows and snow based on existing quality range classification [2].

The Global River Widths from Landsat (GRWL) database is the first global compilation of river plan geometry at constant discharge (Figure 1). The GRWL database contains measurements in plan of more than 2.1 million km (> 58 million measurements) of rivers ≥30 m wide at an average annual discharge. It also contains over 7.6 million tagged measurements of lakes, reservoirs and canals connected to the river network [7]. The GRWL width data is most accurate and complete with a width of more than 90 m (about three Landsat pixels), and therefore this algorithm has been successfully applied to the Ile River, which has an average width of 160 m..

The satellite images were imported from the Landsat-8 collection of images. Landsat-8 is an American Earth remote sensing satellite, the eighth in the Landsat program (the seventh launched into orbit) [10]. Originally called Landsat Data Continuity Mission (LDCM), jointly developed by NASA and USGS. Launched into orbit on February 11, 2013. The main scientific tasks of Landsat-8:

- Collection and storage of multispectral images of medium resolution (30 meters per point) for at least 5 years;
- Preservation of geometry, calibration, coverage, spectral characteristics, image quality and data availability at a level similar to previous satellites of the Landsat program;
- Free distribution of Landsat-8 images [9].

Calculation of the river width includes three main stages: 1) extraction of the river mask from the satellite image; 2) obtaining the river centerline from the river mask; 3) measuring the width of the river along the center line [11].

After calculating the centerline 1 pixel wide, two steps are used to calculate the width of the river. First, for each pixel of the centerline, it calculates a direction orthogonal to the local centerline. Second, it calculates the width of the river along these orthogonal directions.

To compute the orthogonal direction, the algorithm collapses the river's centerline with a 9x9 pixel kernel filled with zero values, excluding pixels along the edge. The value of each pixel of the rim is assigned as follows: starting from the center-right position moving counterclockwise, each pixel is assigned the value of the angle (in degrees) between the horizontal line and the line connecting the pixel itself and the core center. For example, the rim pixel in the upper right corner is set to 45 °, and the rim pixel in the upper center is set to 90 °. The centerline pixel values from the thumbnail divided by two roughly correspond to their orthogonal directions. However, the resulting orthogonal direction is accurate only when exactly two pixels cross the rim of the core, one from each end of the overlapping centerline segment (Figure 1). For pixels located near the end of the river centerline that intersect with the kernel once, we estimated their orthogonal direction by adding 90 ° to the collapsed value. We then discard those orthogonal directions if more than two centerline pixels cross the nucleus - a condition that often occurs near a merge [2].
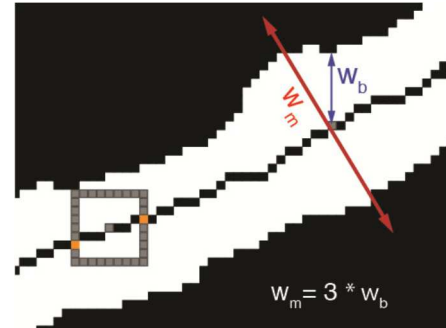

Figure 1. Calculation of the orthogonal direction and width of the river. The rectangle is a 9x9 pixel kernel that is used to compute the orthogonal direction for its center point (gray). The red line running in the orthogonal direction of the centerline is used to calculate the width of the river

After obtaining the orthogonal directions, the width of the river is calculated for each pixel of the centerline as follows. First, a line segment ω_m is created along the orthogonal direction with a length equal to three times the distance to the nearest pixel other than water; the algorithm then calculates the width of the river (ω) by multiplying ω_m by the mean of the channel mask along ω_m.

$$\omega = \omega_m * u \qquad (1)$$

$$u = \frac{\omega}{\omega_m} = \frac{\sum_{i=1}^{N} p_i}{N} \qquad (2)$$

here N is the total number of pixels crossing the segmentω_m, p_i - the value of pixel i on the channel mask. Average u value equals the ratio between wet length of ω_m and this value ω. This particular method for calculating the final width is chosen because of its simple implementation in the GEE.

The accuracy of the width measurements depends on the accuracy of the water classification, which can be affected by clouds, cloud shadows, topographic shadows, and the presence of snow/ice. Clouds cover 67% of the earth's surface at any given time [11]. Recent work has pioneered the ability to distinguish between these functions using a machine learning approach [12], and this ability is extended to allow such processing to be done automatically in the GEE, making it easy to apply globally. When using Landsat imagery as input, the automated algorithm [11] uses the Fmask classification algorithm [13] to note whether the width measurement affects clouds, cloud shadows, or snow/ice. Each of these conditions has its own flag, which is assigned the average value of the corresponding image along the cross-sectional line of ω_m. Thus, a flag with a nonzero value means that the checked condition can affect the width measurement.

Timelapse is an ambitious new GEE project that allows you to travel back in time to the Earth's past and see how its landscape has changed in a specific area over the past decades, a kind of time machine [14]. To create the Timelapse, our tool used 200 photos of the selected area on the map, taken by space satellites in the frame of the Landsat mission, starting in 1984, and only the most detailed and clear photos without clouds were selected, which were then combined into a single animation using GIF technology.

To set the timeline, the jQuery UI Slider plugin was used, which makes the selected elements sliders [15]. With this slider, you can specify a range for the last 10, 15, or 20 years.

Building a graph of the change in the width of the river is the most important and final stage of this tool, since through the graph you can see how the river changes over time.

III. RESULTS AND DISCUSSION

To test the KazRivDyn application, the Ili River, flowing across the border with China, was chosen as the study area. Obtaining the results takes 1-3 minutes, depending on the quantity and quality of images from the Landsat-8 satellite. All images pass through a filter, which selects images with sharper boundaries of wet rivers, without clouds, shadows, etc. All rivers are highlighted in yellow on the map (Fig. 2(a)). To select the desired area, simply click on the center line of the river (yellow) to calculate the time series for the width of the river. In the center on the right - an interactive frame-by-frame animation Timelaps of the river (Fig. 2 (b)) in the size of 300x300 pixels in gif format, demonstrating the change in the width of the river for the entire period. The jQuery UI Slider plugin (Fig. 2 (c)) was used to set the default data slider. After selecting the area of interest, images from 1984 to 2019 are processed, and if necessary, the user can change the time by dragging the slider. The plotting (Fig. 2 (d)) is the last process of observing the dynamics of the application in the application.
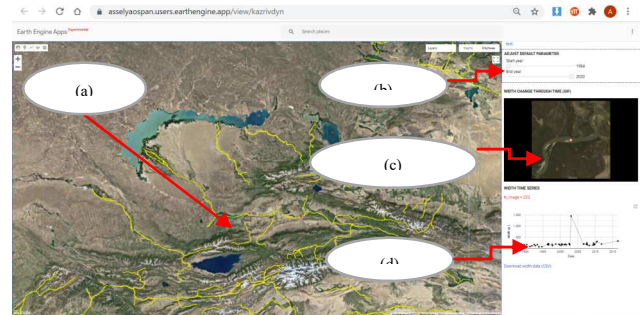

Figure 2. The Ili river highlighted in yellow in the KazRivDyn application. The area of the river that flows through the border with China is highlighted in a red square

When you hover the cursor over a point on the graph, the exact date (Fig. 3) of the image and the width of the river at that time will be shown; there is also a function for loading data in .csv format, for further data processing in Python. For construction, 385 images were selected that passed the installed filters, in which images without clouds and shadows were selected, of which the first 200 images were included in the construction process.
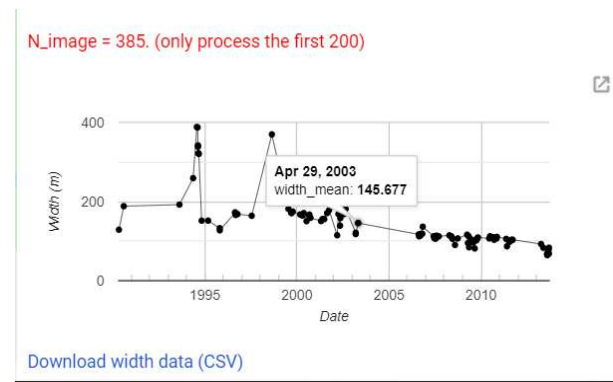

Figure 3. Plotting a chart

For a detailed study, the chart can be opened in a new window and loaded in different formats (Fig. 4).

Figure 4. Chart in PNG format

Based on the results obtained, a statistical analysis was made in Python (Table 1), and the average width of the river is 166.5 m, since according to official data [16] the average width

of the Ile River is 160 m, we can say with confidence that the application is working successfully and gives us the expected results.

TABLE I.   DATA STATISTICS OF THE ILE RIVER

Date, year	Average width, meter	Standard deviation	Maximum width, meter	Minimum width, meter
1990	149.35	23.4	190.47	123.37
1996	170.2	28.4	209.36	112.3
2000	180.77	32	257.11	83.55
2011	186.88	30	213.7	127
2012	152.89	36	204	75
2015	149.87	34	204.467	75.119
2017	166.5	46.2	276.605	104.371
2019	151.71	44.3	255.789	97.489
2020	149.74	42.4	255.78	81.148

Since many parameters affect the change in the volume of water in rivers, of all these parameters, glacier melting, precipitation, temperature, and the resulting volume of water from the Ile River on the border with China will be used in our work, in further work on this topic all these parameters will be embedded in the application. The main goal of this project was to use machine learning to predict the dynamics of river width over the next 20 years. This work is in progress, and our KazRivDyn application will be the main platform where machine learning will be applied.

IV.   CONCLUSION

When working on a research problem, the following results were obtained:

1) application n was developed on the GEE platform, it implements an automatic algorithm for identifying rivers on a map, operating in real time.

2) three tools have been developed and implemented for importing and processing the received images from remote sensing: measuring the width of the river along the orthogonal axis, launching Timelapse animation graphics, building a graph for visualizing the received data.

This completes the main part of our research work. The results showed that the application works efficiently and the measurements of the river width correspond to the real data. Due to the lack of data on the parameters (glacier volume and precipitation), on which the volume of river catchments depends, it is not yet possible to draw accurate conclusions for making a decision. In the future, the KazRivDyn application will also calculate the volume of glaciers and the amount of precipitation, as well as reveal the dependence (correlation) of the volume of water on these parameters. Based on these data, a forecasting model will be built using the regression method.

Conclusion from this article - work is going well and the automatic object recognition algorithm works fine on the Google Earth Engine platform. In the future, it is planned to continue working on the GEE platform using the above parameters and machine learning.

REFERENCES

[1] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," Remote Sens. Environ., vol. 202, pp. 18–27, Dec. 2017. https://doi.org/10.1016/j.rse.2017.06.031

[2] Yang, Xiao, George H. Allen, Tamlin M. Pavelsky, and Gennadii Donchyts. (2019). "RivWidthCloud: An Automated Google Earth Engine Algorithm for River Width Extraction from Remotely Sensed Imagery." IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 17, NO. 2, p.217-221. FEBRUARY 2020.

[3] Pariva Dobriyal, Ruchi Badola, Chongpi Tuboi, Syed Ainul Hussain. A review of methods for monitoring streamflow for sustainable water resource management. This article is published with open access at Springerlink.com. 31 October 2016. https://doi.org/10.1007/S13201-016-0488-Y

[4] Hauer FR, Lamberti GA (2007) Methods in stream ecology, 2nd edn.Academic Press, San Diego

[5] Harmel RD, Smith DR, King KW, Slade RM (2009) Estimating storm discharge and water quality data uncertainty: a software tool for monitoring and modeling applications. Environ Model Softw 24(7):832–842

[6] Gordon ND, McMahon TA, Finlayson BL, Gippel CJ, Nathan RJ (2013) Stream hydrology: an introduction for ecologists. Wiley, England

[7] Global River Widths from Landsat (GRWL) DataBase. Allen, George H., Pavelsky, Tamlin M. https://doi.org/10.1126/science.aat0636

[8] G. H. Allen and T. M. Pavelsky, "Global extent of rivers and streams,"Science, vol. 361, no. 6402, pp. 585-588, 2018. https://doi.org/10.1126/science.aat0636

[9] U.S. Geological Survey Landsat Data Continuity Mission. Rolla Publishing Service Center (July 2012). Retrieved February 12, 2013. Archived June 16, 2013.

[10] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images," Remote Sens. Environ., vol. 159, pp. 269–277, Mar. 2015. https://doi.org/10.1016/j.rse.2014.12.014

[11] M. D. King, S. Platnick, W. P. Menzel, S. A. Ackerman, and P. A. Hubanks, "Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites," IEEE Trans. Geosci. Remote Sens., vol. 51, no. 7, pp. 3826–3852, Jul. 2013. https://doi.org/10.1109/TGRS.2012.2227333

[12] F. Isikdogan, A. C. Bovik, and P. Passalacqua, "Surface water mapping by deep learning," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 11, pp. 4909–4918, Nov. 2017. https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1109%2FJSTARS.2017.2735443

[13] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images," Remote Sens. Environ., vol. 159, pp. 269–277, Mar. 2015. https://doi.org/10.1016/j.rse.2018.04.046

[14] Google Earth Timelaps https://earthengine.google.com/timelapse/

[15] Jquery Slider Widget https://api.jqueryui.com/slider/

[16] River Ile. https://ru.wikipedia.org/wiki/%D0%98%D0%BB%D0%B8_(%D1%80%D0%B5%D0%BA%D0%B0) (In Russian)