Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

DOI: 10.37943/19LRYW4856

Kamilla Rakhymbek
Bachelor of Science in Information and Communication Technologies, Engineer, Laboratory of Digital Technologies and Modeling
rakhymbekkamilla@gmail.com, orcid.org/0009-0008-7404-8433
Sarsen Amanzholov East Kazakhstan University, Kazakhstan

Nurassyl Zhomartkan
Bachelor of science, Junior researcher, Laboratory of Digital Technologies and Modeling
zhomartkan.kz@mail.ru, orcid.org/0009-0006-3935-2013
Sarsen Amanzholov East Kazakhstan University, Kazakhstan

Dauren Nurekenov
Master of Science, Senior researcher, Laboratory of Digital Technologies and Modeling
dnurekenov@gmail.com , orcid.org/0009-0003-5126-6319
Sarsen Amanzholov East Kazakhstan University, Kazakhstan

Zheniskul Zhantassova
Candidate of Technical Sciences, Associate Professor, Department of Computer Modeling and Information Technology
zheniskul_z@mail.ru , orcid.org/0000-0001-5550-7587
Sarsen Amanzholov East Kazakhstan University, Kazakhstan

FLOOD RISK MAPPING IN THE IRTYSH RIVER BASIN USING SATELLITE DATA

Abstract: Floods are among the most frequent and devastating natural disasters, causing significant economic damage and loss of life worldwide. Effective flood risk management relies on accurate modeling techniques that can predict vulnerable areas and assess potential impacts. In this study, flood dynamics are simulated in the Irtysh River Basin near Ust-Kamenogorsk, a city in East Kazakhstan prone to seasonal flooding using high-resolution satellite imagery and digital elevation data. The primary objective is to visually model flood risks based on terrain characteristics. The study utilizes imagery sourced from the Mapbox platform, which combines data from MODIS, Landsat 7, Maxar, and the Google Earth Engine, providing access to Sentinel-2 surface reflectance imagery at 10-meter resolution. Elevation data from the Copernicus global digital elevation model, with a 30-meter resolution, is used to simulate flood progression. The flood simulation involves calculating flood depth relative to the terrain's elevation, allowing for a pixel-by-pixel determination of submerged areas. Each simulation incrementally increases water levels to generate a sequence of images, showcasing the progression of flooding over time. The study describes hydraulic soil characteristics usage, and focuses on visualizing flood risk based on terrain data and water level changes. The simulation results indicate that flooding initially impacts riverbanks as water flow starts from the northwest of the city with critical infrastructure becoming vulnerable once water levels exceed 2 meters from the lowest elevation point. These findings highlight the potential of high-resolution satellite imagery and terrain data for flood risk assessment and improving ur-

Received: 11.09.2024 Accepted: 26.09.2024 Published: 30.09.2024

DOI: 10.37943/19LRYW4856

ban flood preparedness. The results provide valuable insights into flood progression enabling more informed decision-making for disaster mitigation.

Keywords: Flood map, Python, Mapbox, Google Earth Engine, Digital Elevation Models.

Introduction

East Kazakhstan is a region characterized by its diverse topography, ranging from expansive steppes and dense forests to rugged mountains. This geographical variety, coupled with seasonal weather patterns, makes it particularly susceptible to flooding. Understanding and predicting these flood events is crucial for the safety and development of the region. Flood animation and modeling have been the focus of extensive research due to their crucial role in flood risk management and mitigation. The methodologies employed range from traditional hydrodynamic models to more contemporary machine learning (ML) and deep learning (DL) approaches [1].

Hydrodynamic models have long been the standard for simulating flood behavior. These models rely on the physical principles of fluid dynamics to represent the movement of water through river networks and floodplains. Commonly used models include the one-dimensional (1D) HEC-RAS (Hydrologic Engineering Center's - River Analysis System) and two-dimensional (2D) models like MIKE 21 and TELEMAC. These models are praised for their accuracy in simulating flood extents, depths, and velocities but require extensive calibration data, which can be a limitation in data-scarce regions. Studies have shown that these models are highly effective in small to medium-sized catchments. For instance, Bellos and Tsakiris [2] developed a hybrid method combining hydrodynamic and hydrological techniques, which proved effective for small catchments prone to flash floods. Another study [3] utilized a 2D hydrodynamic approach to evaluate hydrological responses in small watersheds, highlighting the importance of high-resolution data in improving model performance.

The advent of Machine Learning (ML) and Deep Learning (DL) has introduced new avenues for flood modeling, offering an alternative to traditional hydrodynamic methods. These data-driven models, such as Artificial Neural Networks (ANNs) and Recurrent Neural Networks (RNNs), have shown promise in predicting flood events with minimal physical input data. Recent research emphasizes the potential of DL models, particularly for handling complex datasets involving multiple hydrological variables. Reference [4] conducted an extensive review of ML and DL applications for flood inundation modeling, concluding that DL models generally outperform traditional methods in terms of accuracy. However, it was also pointed out challenges such as the lack of standardized datasets and the limited ability of ML models to incorporate expert hydrological knowledge. Despite these drawbacks, DL models are gaining traction due to their ability to generalize across different hydrological conditions and catchment characteristics. Comparative studies indicate that while DL models can achieve higher predictive accuracy, they often lack the interpretability and physical grounding of hydrodynamic models. This has led to a growing interest in hybrid approaches that integrate both data-driven and physically based models. For instance, Bulti and Abebe [5] reviewed various flood modeling methods and advocated for the integration of empirical and physically based models to enhance flood prediction in urban areas.

Moreover, advancements in geospatial technologies have significantly improved the spatial resolution of flood models. Ben Khalfallah and Saidi [6] demonstrated the application of HEC-RAS coupled with Geographic Information System (GIS) tools for spatiotemporal floodplain mapping, which enhances the accuracy and usability of flood risk assessments. Recent advances in flood modelling have been driven by the integration of satellite imagery and advanced data processing. Tools such as FLOMPY [7] and workflows using Sentinel-1 time series data [8], [10] have greatly improved the accuracy and efficiency of flood mapping. Additionally,

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

platforms like the Google Earth Engine, combined with Sentinel-1 and Landsat data [9], allow for rapid and large-scale flood monitoring. Synthetic Aperture Radar (SAR) based assessments [11], [12] and remote sensing techniques [13], [14] are also crucial for refining flood prediction models, improving both scalability and adaptability to effectively manage flood risks in different environments.

Besides, the proposed model has significant potential by including additional hydrological parameters, which would greatly improve its accuracy and reliability. One way to achieve this is through several models developed using python, such as the Brooks-Corey model [15], which can calculate the hydraulic properties of soil based on its texture. Additionally, some literature highlights the application of more comprehensive model improvement approaches. For instance, in work [16], the need for a step-by-step approach to modeling hydrological processes is discussed due to their complexity and data uncertainty. Similarly, the authors [17] also consider mathematical modeling of watersheds with an emphasis on the gradual inclusion of various hydrological processes to improve the accuracy and reliability of models. Furthermore, the MIKE SHE system [18], which is an integrated model of surface and groundwater, uses a modular approach that allows adding new model components as needed.

Reflecting the progress in the field of geospatial technologies, the primary goal of this article is the construction of the animation of a flood map on the Irtysh River near the city of Ust-Kamenogorsk in the East Kazakhstan region. Therefore, the main objectives of the study are:

- Acquiring and filtering the satellite imagery and digital elevation models for Ust-Kamenogorsk to ensure high-quality data.
- Identification and analysis of potential flood-prone areas by the calculation of water level rises relative to elevation data.
- Analysis and visualization of the simulation model using the capabilities of Mapbox and Google Earth Engine.

The study highlights the potential of integrating high-resolution satellite imagery with elevation data to construct a flood simulation model that enhances urban flood preparedness. This approach contributes to the disaster risk management field by offering a replicable framework, adaptable to other flood-prone regions, and capable of further refinement through the incorporation of additional parameters such as meteorological and hydraulic data.

Methods and Materials

Data description

The modelling and mapping to visually represent potential flood scenarios and risks are performed utilizing high-resolution satellite imagery and terrain data, focusing on the Ust-Kamenogorsk city in the East Kazakhstan region. For the study, high-resolution satellite imagery is obtained through the Mapbox and Google Earth Engine platform. In Mapbox platform the imagery is sourced from a combination of open and proprietary datasets including MODIS, Landsat 7, Maxar's products. Google Earth Engine is used to obtain Sentinel 2 surface reflectance imagery at 10m resolution. For terrain modeling, the elevation data is collected from the Copernicus global digital elevation model which represents the Earth's surface, incorporating buildings, infrastructure and vegetation at 30 m spatial resolution. As mentioned, the study area is centered on the city, with the delineation of its boundaries represented in the following Fig. 1.

DOI: 10.37943/19LRYW4856

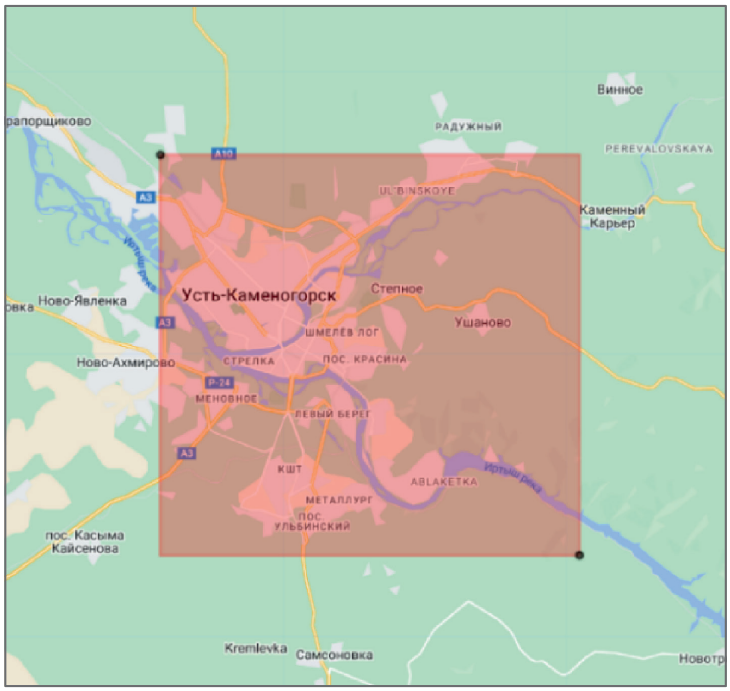
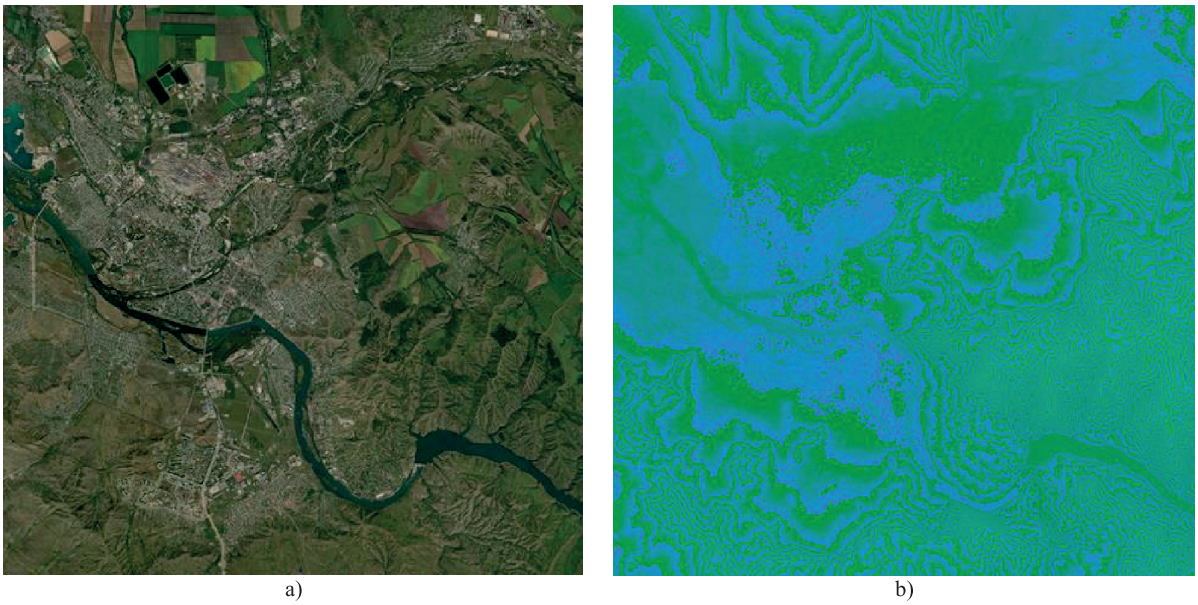


Figure 1. Boundary box of the study area.

To ensure high-quality data for the analysis, Sentinel-2 satellite images were carefully selected based on three key criteria: spatial coverage, cloud cover, and seasonal timing. First, images were filtered to include only those with a minimum of 60% spatial coverage over the region of interest. Second, images with cloud cover percentages below 10% were prioritized to minimize atmospheric interference and enhance the clarity of the surface features. Finally, images were restricted to those captured during the spring and summer months to coincide with periods of maximum vegetation growth and relevant hydrological conditions. Additionally, the preprocessing steps involved using both satellite imagery and terrain elevation data to create comprehensive composite images, allowing for detailed spatial analysis of the study area. In the Fig. 2, the satellite and elevation map are illustrated.



a) b)

Figure 2. Visual representation of the selected area:
(a) Satellite imagery depicting the geographic features of the chosen region;
(b) Elevation map illustrating the topographical variations within the same area.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

The individual satellite and terrain tiles were processed and stitched together to form large, continuous mosaics. The tiles, each representing a segment of the region are combined based on their geographic position, ensuring that each tile is placed in its correct spatial location.

Flood modeling and mapping

The flood simulation is conducted by combining high-resolution satellite imagery with detailed elevation data to visualize potential inundation scenarios across the study area. The approach involved several steps, including the calculation of depth, generation of an overlay to represent water levels, and visualization of flooded areas at incremental water. The flood simulation begins by establishing a base flood level, set initially to the minimum elevation within the study area. A series of flood levels was simulated by progressively increasing the water level in fixed increments. For each flood level L_i the relative flood depth $D(x, y)$ at each pixel is calculated as:

$$D(x, y) = L_i - E(x, y) \tag{1}$$

where L_i is the water level, and $E(x, y)$ is the terrain elevation. Positive values of $D(x, y)$ indicated that the pixel was submerged, while negative or zero values indicated that the area remained above water. To effectively visualize the flooded areas, a semi-transparent overlay was applied to the base satellite image. This overlay was colored blue to represent water and adjusted based on the depth. The transparency of the overlay was controlled by an alpha mask, with transparency scaled according to the flood depth. The alpha mask $A(x, y)$ is normalized using the maximum flood depth at each water level:

$$A(x, y) = \frac{D(x, y)}{D_{\max}} \times 255 \tag{2}$$

where $D_{\max}$ represents the maximum flood depth at a given water level. For each incremental flood level, the resulting image was generated by combining the base satellite imagery with the flood overlay. The series of images produced provided a time-sequenced visualization of flood progression across the study area, from the initial flood level to a maximum water level set by the defined range. These images were saved sequentially, forming a complete set of flood simulations across multiple water levels.

Refining flood forecasts through hydraulic characteristics of the soil

When modeling floods, it is important to consider various hydraulic parameters that affect the behaviour of the water flow and the level of flooding. Some of these parameters are the soil conductivity, which depends on its moisture content, as well as infiltration rate, which directly affect the volume of runoff and the duration of the flood. The Darcy equation, referenced in [19] as equation (3), describes the volume of water passing through a medium in a given time, with Darcy's Law (introduced in 1856) explaining water flow through porous materials like sand under atmospheric pressure.

$$Q = \frac{ks(H + e)}{e} \tag{3}$$

where Q is the water volume per unit time, s is the surface area, e is its thickness, H is the water height, and k is a coefficient reflecting the sand's properties.

The process of soil water infiltration can be described using the Green-Ampt model, as formulated in equation (4), by the authors in reference [20]. This model takes into account the initial moisture content and the pressure at the water-soil interface, which allows for a more accurate assessment of the process of water penetration into the soil.

DOI: 10.37943/19LRYW4856

$$f_p = K_S[1 + (\theta_S - \theta_i)S_r / I_p], t > t_p \quad (4)$$

Where f_p denotes the infiltration rate, K_S is the saturated hydraulic conductivity, θ_i and θ_S represent the initial and saturated soil moisture content, respectively, S_r stands for capillary pressure, t_p is the time it takes for ponding to occur after rainfall begins, R refers to the rainfall intensity, I_p is the cumulative infiltration, with $I_p = t_p R$.

Another method for calculating water absorption into the soil involves using Horton's equation (5), as referenced by Horton [21]. This equation shows how infiltration starts at a high initial rate and gradually decreases to a steady final rate.

$$f = f_c + (f_0 - f_c)e^{-kt} \quad (5)$$

Where f is the infiltration rate at any given moment (in mm/h), f_c represents the final infiltration rate (in mm/h), f_0 is the initial infiltration rate (in mm/h), k represents the decay constant, t is the time of infiltration (in hours).

Computational tools and software

The flood simulation and data processing tasks are carried out using a combination of specialized software and programming libraries. The primary tools employed include Python, Mapbox and Google Earth Engine. Mapbox, a platform for interactive mapping, is combined with Python data processing capabilities to develop animations that effectively visualize flood data. Mapbox provides efficient extraction of landscape and satellite images, allowing to configure geospatial visualizations adapted to specific regions.

Google Earth Engine, a powerful cloud geospatial platform, is also used for analysis and acquisition of data. GEE provides access to a vast archive of satellite images and geospatial datasets. It is particularly well suited for environmental monitoring and flood modelling due to its ability to handle large data volumes and an integrated set of remote sensing and spatial analysis tools. The main advantage of GEE is its ability to quickly process global datasets with huge amounts of data, making in comparison with locally processed Mapbox and Python workflows.

Results and Discussions

Flood simulation enables the prediction of trends based on rising water levels. The flood modeling for the Irtysh River basin around Ust-Kamenogorsk, conducted using Python-Mapbox and Google Earth Engine, produced reliable and visually informative results. The images, arranged as a collage, show water levels at 1, 5, 10, and 15 meters above the lowest elevation (Fig. 3), effectively illustrating the dynamic progression of flooding throughout the city. Modeling trends indicate that as water levels rise, flooding begins in the northwest and gradually spreads southeast, following the elevation map (Fig. 4).

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

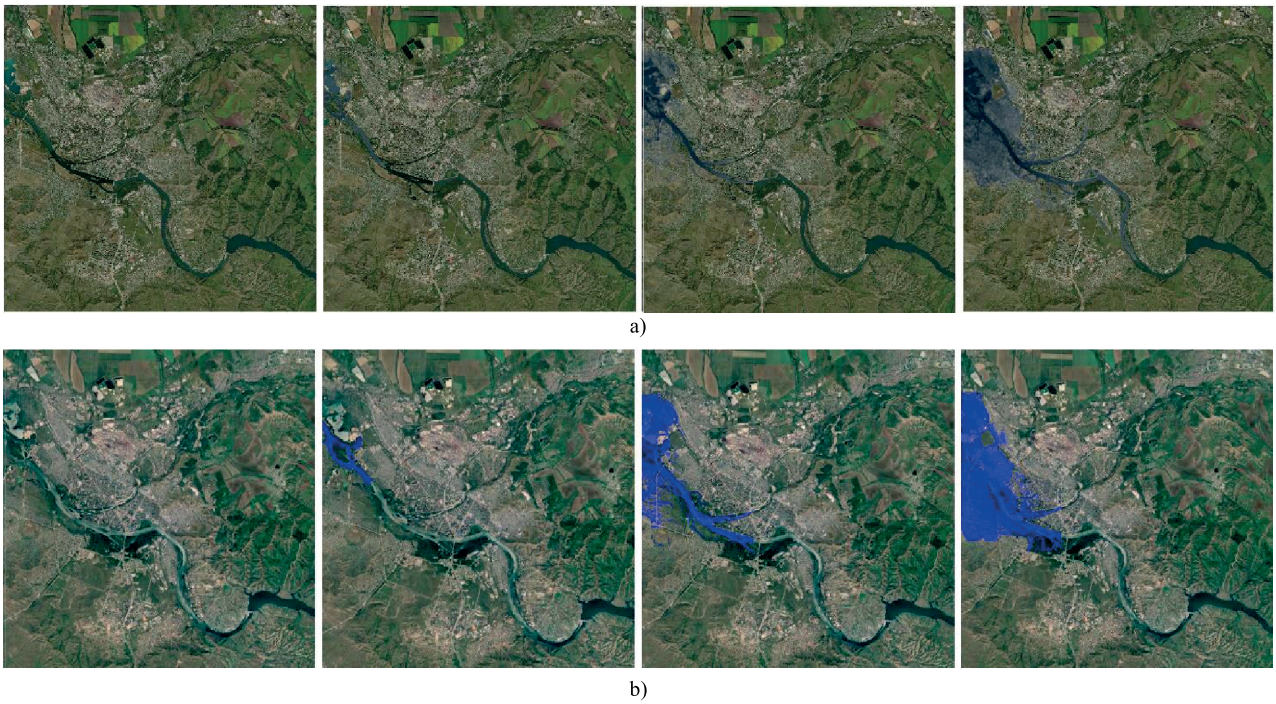


Figure 3. Flood simulation at water levels of 1, 5, 10, and 15 meters, showing the progression of flooding from the northwest to the southeast of the city. (a) Python-Mapbox, b) Google Earth Engine.

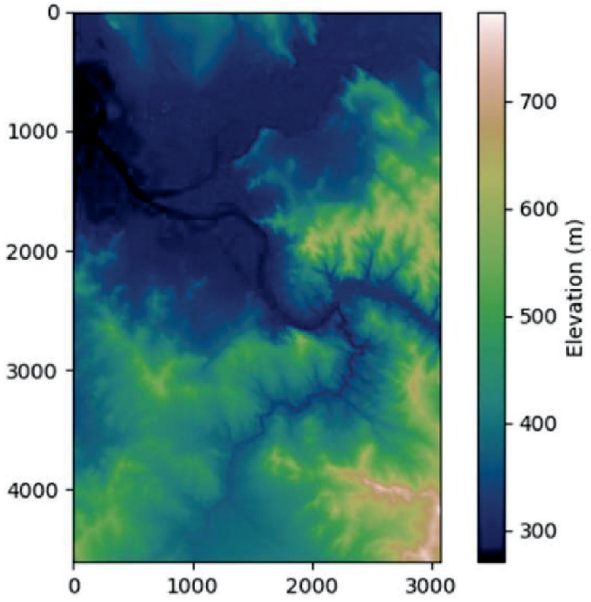


Figure 4. Elevation map depicting the landscape.

Although scenarios with water levels exceeding 4 meters are unlikely, this study focuses on simulating floods at levels up to 4 meters, providing insights into the most probable impacts. In Fig. 5, the red-highlighted areas represent those most vulnerable to flooding at a 4-meter water level. Subsequent images will provide a closer look at these areas, offering a detailed view of the flooding dynamics and its effects on specific districts.

DOI: 10.37943/19LRYW4856

Figure 5. The areas of the city most prone to flooding at a 4-meter water level are highlighted in red.

The progression of the flood was analyzed in 1-meter increments, allowing for a detailed understanding of the flooding process step-by-step. At the initial stage, when the water rises by 1 meter from the lowest point in the terrain, flooding remains minimal and is confined to areas directly adjacent to the river. This is illustrated in Fig. 6(a) (Python-Mapbox) and Fig. 7(a) (Google Earth Engine). At this level, the terrain serves as a natural barrier, preventing the water from spreading to higher elevations. When the water rises to 2 meters, as shown in Fig. 6(b) for Python-Mapbox and Fig. 7(b) for Google Earth Engine, flooding begins to affect nearby infrastructure and low-lying urban areas. This level marks a critical threshold, as key infrastructure, such as the "Industroy" concrete factory, located in coastal industrial zones, starts to be impacted.

a) b) c) d)

Figure 6. Flood maps of the study area generated with Python-Mapbox, labeled (a) 1 m, (b) 2 m, (c) 3 m, and (d) 4 m, showing water level increases and the terrain's response to rising floodwaters.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 19, SEPTEMBER 2024

As the water level reaches 3 meters, the flood zone gradually expands, moving towards the garden areas and surrounding territories, as illustrated in Fig. 6(c) and Fig. 7(c) for Python-Mapbox and Google Earth Engine, respectively. With a further rise in water to 4 meters, more significant areas will be at risk of flooding, including industrial zones and garage cooperatives. At the same time, the riverbanks, where the garden plots are located, such as the "Svetoch", "Rybnik", "Zvezda", and "Druzhba" associations, will be heavily affected. These changes are clearly visible in Fig. 6(d) (Python-Mapbox) and Fig. 7(d) (Google Earth Engine). According to the analysis, the critical threshold begins at the 2-meter mark, beyond which important infrastructure in the low-lying urban areas near the river starts to flood.

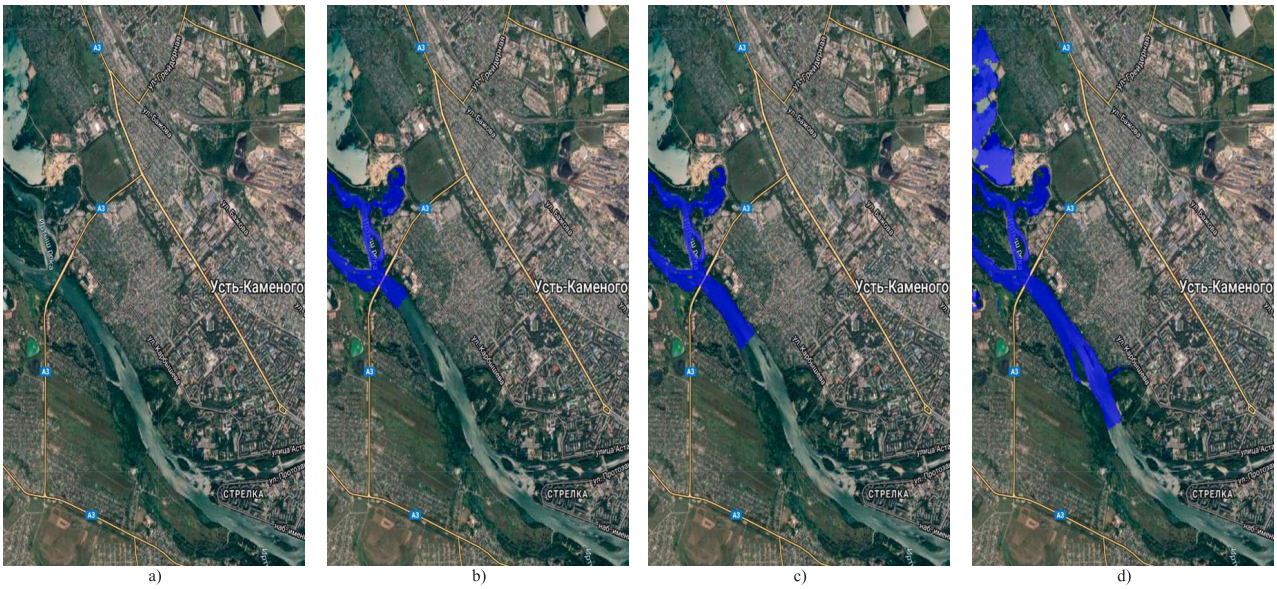


Figure 7. Flood maps of the study area generated with Google Earth Engine, labeled (a) 1 m, (b) 2 m, (c) 3 m, and (d) 4 m, showing water level increases and the terrain's response to rising floodwaters.

Conclusion

This study underscores the critical importance of advanced flood modeling for effective risk management in the Ust-Kamenogorsk region. By leveraging Python-Mapbox and Google Earth Engine, we successfully simulated potential flood scenarios, revealing that the lowlands along the Irtysh River, particularly in the city's northwestern floodplain, are the most susceptible to flooding. The consistent results obtained from both Python-Mapbox and Google Earth Engine demonstrate the robustness of the modeling framework. Python-Mapbox offers enhanced flexibility for detailed, localized visualizations, while Google Earth Engine excels in large-scale, cloud-based analyses.

Future research should focus on expanding the model by integrating the suggested environmental parameters and extending the model to other vulnerable regions. Additionally, exploring the incorporation of hydraulic and meteorological data will further refine the model, making it a more robust tool for flood forecasting and management.

Acknowledgement

This work has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No BR24992899).

References

[1] Jerome, G.F. (2021). Flood disaster hazards; causes, impacts and management: A state-of-the-art review. *Natural Hazards, IntechOpen: Rijeka.*

[2] Bellos, V., & Tsakiris, G. (2016). A hybrid method for flood simulation in small catchments combining hydrodynamic and hydrological techniques. Journal of Hydrology, 540, 331-339. https://www.doi.org/10.1016/j.jhydrol.2016.06.040
[3] Barbero, G., Costabile, P., Costanzo, C., Ferraro, D., & Petaccia, G. (2022). 2D hydrodynamic approach supporting evaluations of hydrological response in small watersheds: Implications for lag time estimation. *Journal of Hydrology*, *610*, 127870. https://www.doi.org/10.1016/j.jhydrol.2022.127870.
[4] Karim, F., Armin, M.A., Ahmedt-Aristizabal, D., Tychsen-Smith, L., & Petersson, L. (2023). A review of hydrodynamic and machine learning approaches for flood inundation modeling. *Water*, *15*(3), 566. https://www.doi.org/10.3390/w15030566 .
[5] Bulti, D.T., & Abebe, B.G. (2020). A review of flood modeling methods for urban pluvial flood application. Modeling Earth Systems and Environment, 6, 1293-1302. https://www.doi.org/10.1007/s40808-020-00803-z .
[6] Ben Khalfallah, C., & Saidi, S. (2018). Spatiotemporal floodplain mapping and prediction using HEC-RAS-GIS tools. Journal of African Earth Sciences, 142, 44-51. https://www.doi.org/10.1016/j.jafrearsci.2018.03.004 .
[7] Karamvasis, K., & Karathanassi, V. (2021). FLOMPY: An Open-Source Toolbox for Floodwater Mapping Using Sentinel-1 Intensity Time Series. Water, 13(21), 2943. https://www.doi.org/10.3390/w13212943 .
[8] Martinis, S., Plank, S., Ćwik, K. (2018) The use of Sentinel-1 time-series data to improve flood monitoring in arid areas. Remote Sens, 10, 583. https://www.doi.org/10.3390/rs10040583
[9] DeVries, B., Huang, C., Armston, J., Huang, W., Jones, J.W., Lang, M.W. (2020) Rapid and robust monitoring of flood events using Sentinel-1 and Landsat data on the Google Earth Engine. Remote Sens. Environ. 240, 111664. https://www.doi.org/10.1016/j.rse.2020.111664
[10] Filipponi, F. (2019, June). Sentinel-1 GRD preprocessing workflow. In Proceedings (Vol. 18, No. 1, p. 11). MDPI. https://www.doi.org/10.3390/ECRS-3-06201 .
[11] Papaioannou, G., Vasiliades, L., Loukas, A., Alamanos, A., Efstratiadis, A., Koukouvinos, A., ... & Kossieris, P. (2021). A flood inundation modeling approach for urban and rural areas in lake and large-scale river basins. Water, 13(9), 1264. https://www.doi.org/10.3390/w13091264 .
[12] Landuyt, L., Van Wesemael, A., Schumann, G.J.P., Hostache, R., Verhoest, N.E.C., Van Coillie, F.M.B. (2019) Flood Mapping Based on Synthetic Aperture Radar: An Assessment of Established Approaches. IEEE Trans. Geosci. Remote Sens. 57, 722–739. https://www.doi.org/10.1109/TGRS.2018.2860054 .
[13] Matgen, P., Hostache, R., Schumann, G., Pfister, L., Hoffmann, L., & Savenije, H.H.G. (2011). Towards an automated SAR-based flood monitoring system: Lessons learned from two case studies. *Physics and Chemistry of the Earth, Parts A/B/C*, *36*(7-8), 241-252. https://www.doi.org/10.1016/j.pce.2010.12.009.
[14] Zotou, I., Bellos, V., Gkouma, A., Karathanassi, V., & Tsihrintzis, V. A. (2020). Using Sentinel-1 imagery to assess predictive performance of a hydraulic model. *Water Resources Management*, *34*, 4415-4430. https://www.doi.org/10.1007/s11269-020-02592-7
[15] Memari, S.S., & Clement, T.P. (2021). PySWR-A Python code for fitting soil water retention functions. Computers & Geosciences, 156, 104897.
[16] Beven, K. (2012). Rainfall-Runoff Modelling: The Primer. Wiley-Blackwell.
[17] Singh, V.P., & Woolhiser, D.A. (2002). Mathematical modeling of watershed hydrology. Journal of Hydrologic Engineering, 7(4), 270-292.
[18] Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., & Rasmussen, J. (1986). An introduction to the European Hydrological System—Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. Journal of Hydrology, 87(1–2), 45–59.
[19] Richards, L.A. (1950). Laws of soil moisture. *Eos, Transactions American Geophysical Union*, *31*(5), 750-756.
[20] Hsu, S.Y., & Hilpert, M. (2011). Incorporation of dynamic capillary pressure into the Green–Ampt model for infiltration. Vadose Zone Journal, 10(2), 642-653.
[21] Horton, R.E. (1937). Determination of infiltration-capacity for large drainage-basins. *Eos, Transactions American Geophysical Union*, *18*(2), 371-385.