Risk-Guided Flood Segmentation from Optical Satellite Imagery Using NDWI Threshold Optimization and Segment Anything Model

Andrea Reid [1], Shabnam Jabari [1], Heather McGrath [2]

[1] Dept. of Geodesy & Geomatics Engineering, University of New Brunswick, Fredericton, NB – (andrea.reid, sh.jabari)@unb.ca
[2] Natural Resources Canada, Government of Canada, Ottawa, ON – heather.mcgrath@nrcan-rncan.gc.ca

Keywords: Flood Extent Mapping, Segment Anything Model, NDWI, Flood Risk Priors.

Abstract

Accurate and timely flood extent mapping is essential for emergency response. Optical satellite imagery is widely used for rapid flood mapping due to its global coverage and free availability. A common approach for delineating surface water from optical imagery involves the Normalized Difference Water Index (NDWI), which detects water features using green and near-infrared spectral bands. However, NDWI-based flood mapping requires the selection of a threshold value, and small variations in this threshold can lead to substantial differences in the estimated flood extent. At the same time, recent foundation segmentation models such as the Segment Anything Model (SAM) can identify object boundaries without task-specific training but requires manual prompting. This study proposes a risk-guided flood segmentation framework that integrates spectral thresholding with SAM refinement. First, NDWI thresholds are optimized using a risk score derived from flood hazard maps, allowing the threshold selection process to prioritize water detections in areas where flooding is more likely to occur. Then, the resulting NDWI-based flood mask is refined using SAM to improve boundary delineation and recover missed flood pixels. The method is evaluated using imagery from the 2018 spring flood along the Saint John (Wolastoq) River in New Brunswick, Canada, across five study regions using both Sentinel-2 and Landsat-8 scenes. Results show that the proposed risk-guided NDWI threshold selection with SAM refinement improves recall while maintaining stable precision. The framework requires no model training and provides a reproducible workflow for automated flood mapping from optical satellite imagery.

1. Introduction

1.1 Background & Motivation

1.1.1 Flood Mapping with Satellite Imagery: Flooding is increasing in frequency and intensity across the globe, causing significant impacts on communities and the infrastructure that they depend on (IPCC, 2022). In Canada, the financial and societal risks associated with flooding have increased due to a combination of climate change, population growth, and intensified development in flood-prone areas (Task Force on Flood Insurance and Relocation, 2022). To help prepare for, mitigate, and respond to flooding, there is a need for improved monitoring, mapping, and management of flood events.

Accurate and timely flood extent mapping is especially critical to the disaster response and is also essential for managing risk. Flood maps derived during and immediately after an event provide information that can be used by emergency services to identify affected areas, coordinate evacuations, and prioritize response efforts. They are also widely used for damage assessment and for updating flood hazard maps used in land-use planning and insurance risk assessments (Task Force on Flood Insurance and Relocation, 2022).

Satellite remote sensing is an effective tool for mapping flood extents over large spatial areas, providing consistent global coverage and frequent revisit intervals. Imagery captured from Sentinel-2 and Landsat-8 optical sensors are particularly valuable due to their free availability, large spatial coverage, and relatively short revisit times. Improvements in spatial resolution and revisit frequency have made satellite imagery increasingly suitable for monitoring flood events.

1.1.2 Normalized Difference Water Index (NDWI): One of the most widely used indices for delineating surface water in multi-spectral imagery is the Normalized Difference Water Index (NDWI), which exploits the strong absorption of near-infrared radiation by water relative to the green band (McFeeters, 1996). However, despite its simplicity and computational efficiency, NDWI-based flood mapping has several limitations. Selecting an appropriate threshold value can be particularly challenging, as small changes in the threshold may lead to large differences in the estimated flood extent. Optimal thresholds can vary across sensors, atmospheric conditions, land cover types, and illumination conditions. In addition, NDWI may produce false detections in complex environments, particularly in urban areas where reflective surfaces such as rooftops or bare soil can exhibit spectral characteristics similar to water.

1.1.3 Image Segmentation: On the other hand, the use of machine learning and deep learning approaches for flood detection from satellite imagery has been shown to be an effective alternative to NDWI thresholding. Traditional machine learning methods, such as random forest classifiers, can incorporate multiple data sources—including spectral information, elevation, and land cover—to improve flood prediction accuracy. For example, recent studies have demonstrated that random forest models can effectively integrate satellite imagery with terrain and land cover data to predict flood areas (Esfandiari et al., 2020; Composto et al., 2025).

More recently, deep learning approaches have shown strong performance for image segmentation tasks and offer significant potential for automated flood boundary extraction. Convolutional neural networks (CNNs) have been widely applied to classify water and non-water pixels in satellite imagery. For instance, a CNN-based model was developed to fuse optical and synthetic

aperture radar (SAR) imagery from Sentinel-1 and Sentinel-2 for segmenting permanent water, temporary water, and surface water using the Sen1Flood11 dataset (Bai et al., 2021). While deep learning approaches can significantly improve segmentation accuracy, they require large volumes of labelled training data, which can be difficult and time-consuming to produce, especially for urban areas.

1.1.4 Segment Anything Model (SAM): Recent advances in foundation models have opened new possibilities for image segmentation with limited training data. The Segment Anything Model (SAM) is a large-scale segmentation model that can identify object boundaries in images without task-specific training (Kirillov et al., 2023). Recent work by Shokati et al. (2026) showed that SAM outperformed a UNet model for detecting floodwater pixels in aerial imagery when provided with appropriate prompts. Their results also indicated that point-based prompts were more effective than bounding boxes for guiding the segmentation. However, their method still requires manual selection of prompts to guide the segmentation process, which limits its applicability for large-scale or near-real-time flood mapping applications. Furthermore, without semantic understanding, SAM may generate irrelevant segments in complex environments such as urban areas or densely vegetated landscapes.

1.2 Goals & Objectives

In this study, we aim to address these limitations by proposing a risk-guided segmentation framework that combines NDWI-thresholding with SAM segmentation to improve automated flood boundary delineation in optical satellite imagery.

First, we introduce a risk-weighted optimization procedure to determine the NDWI threshold used for water detection. Instead of selecting a threshold arbitrarily or through manual tuning, we incorporate flood hazard information to guide the optimization process, prioritizing water detections in areas where flooding is more likely to occur. Second, we refine the resulting NDWI-based water mask using SAM, which helps capture visually consistent flood boundaries.

By combining spectral indices and flood hazard maps with a foundation segmentation model, this process incorporates knowledge about where a flood is more likely to occur (flood hazard maps), how it is represented physically in the satellite data (NDWI), and how it appears visually (SAM). Moreover, the proposed method does not require training large deep learning models or generating extensive labelled datasets, making it suitable for flood mapping applications that are limited by data scarcity.

We evaluate the proposed framework using multi-spectral satellite imagery of the 2018 Spring flooding along the Saint John (Wolastoq) River in New Brunswick, Canada, one of the most significant flood events recorded in the region. The results demonstrate how integrating risk-guided NDWI thresholding with SAM-based refinement can improve the accuracy and robustness of flood extent delineation compared to NDWI- or SAM-based segmentation alone.

2. Methodology

2.1 Overview

This section describes the proposed framework for automated flood delineation from optical satellite imagery. The methodology combines NDWI water detection with a risk-guided threshold optimization and segmentation refinement using SAM. Refer to Figure 1 for an overview of the methodology.

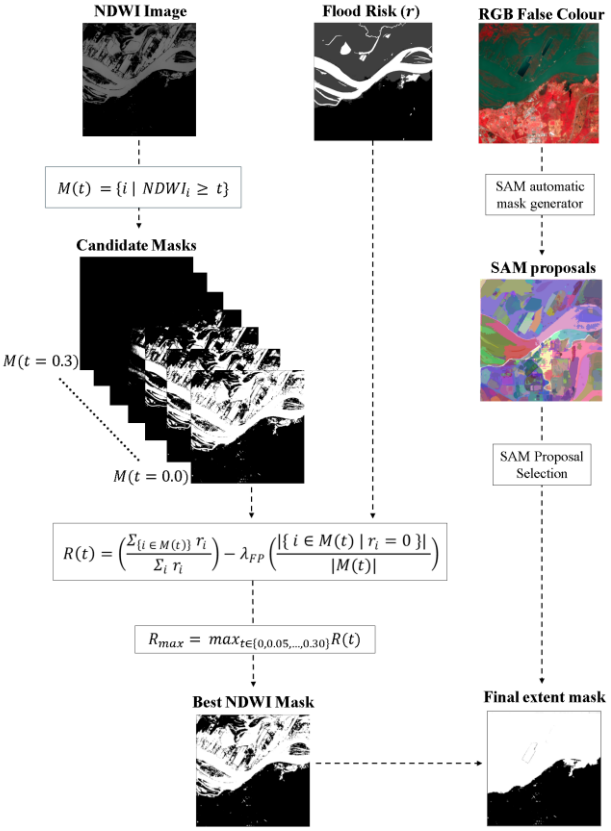

Figure 1. Overview of Methodology.

2.2 Study Area

To evaluate the methodology, five study areas (patch size = 7.7 x 7.7 km) affected by Spring flooding along the Saint John River in New Brunswick, Canada, were selected (Figure 2). For each study region, we analysed Sentinel-2 (10 m resolution) and Landsat-8 (30 m resolution) multi-spectral imagery acquired on May 2, 2018, corresponding to peak (or near-peak) flood conditions for the selected areas. Comparing results across different satellite optical sensors with different resolutions allows us to assess the robustness of the method.

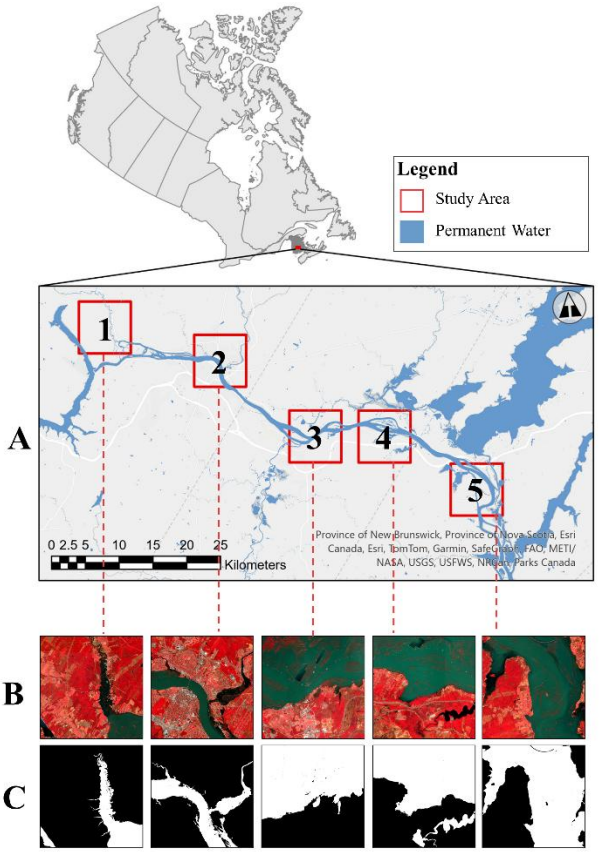

Figure 2. The location of five study areas along the lower Saint John (Wolastoq) River in New Brunswick, Canada (A), along with RGB False-Colour composites from Sentinel-2 (B) and the corresponding 2018 peak flood extent derived from provincial hydrological/hydraulic modelling for the event (C).

2.3 NDWI flood mask generation

For each of the five study regions and for both satellite sensors, NDWI was computed using atmospherically corrected surface reflectance from Near-infrared (NIR) and Green spectral bands using the following expression:

$$NDWI = \frac{Green - NIR}{Green + NIR}. \quad (1)$$

Water is typically associated with NDWI > 0 (McFeeters, 1996), but the optimal threshold (t) used to predict water pixels varies by scene and sensor. For each region, we therefore generated candidate flood masks using NDWI thresholds ranging from $t = 0.0$ to $t = 0.3$ in increments of 0.05. Despite its superior performance for water detection, we did not apply the modified NDWI (MNDWI), because replacing NIR with Short wave infrared (SWIR) would reduce spatial resolution for Sentinel-2 (SWIR spatial resolution = 20 m; NIR and Green spatial resolution = 10 m).

2.4 Risk-based threshold scoring

2.4.1 Calculating Risk Score: To guide threshold selection with geographic context, we constructed a flood risk raster that integrates provincial flood hazard maps to identify potential 'risk' areas and hydrological data to identify where the permanent water is located (Service New Brunswick, 2025; Figure 3). The flood hazard maps are based on an Annual Exceedance Probability (AEP) of 5% (i.e., a 1 in 20-year flood) and an AEP of 1% (i.e., 1 in 100-year flood). Risk (r) was determined for each pixel (i) in the flood risk raster based on the probability of a flood occurring: 100% for permanent water bodies, 5% for 1 in 20-year flood areas, 1% for 1 in 100-year flood areas, and zero chance for no risk areas. Note that in the study areas, the 1% and 5% AEP maps are largely overlapping.

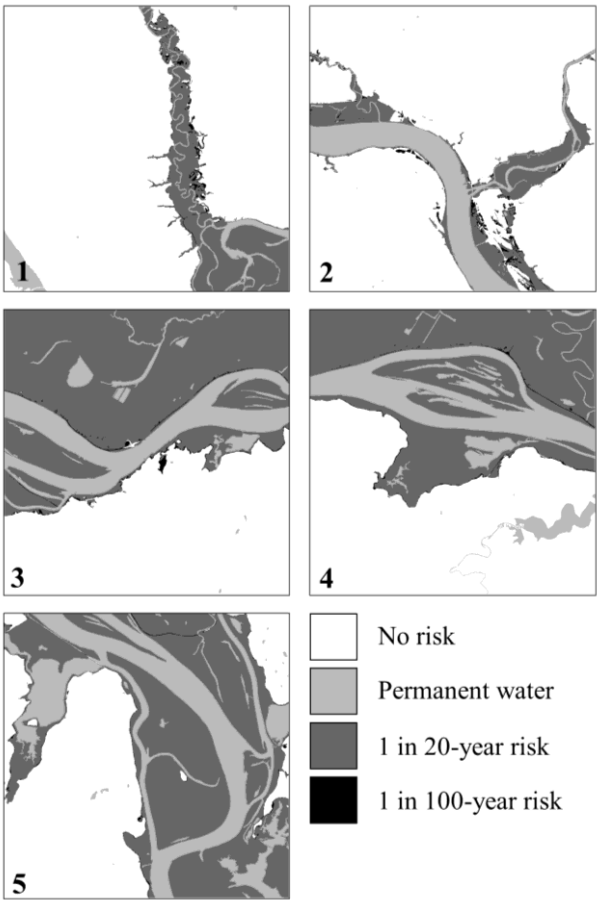

Figure 3. Flood risk maps for each of the five study regions derived from provincial flood hazard maps and hydrological data. Note the large overlap between the 1 in 20-year and 1 in 100-year flood risk.

For each value of t, a NDWI binary mask, $M(t)$, is generated using Equation (2).

$$M(t) = \{ i \mid NDWI_i \geq t \}. \quad (2)$$

Then, for each threshold-derived NDWI mask, a risk score $R(t)$ is computed as:

$$R(t) = \left(\frac{\Sigma_{\{i \in M(t)\}} r_i}{\Sigma_i r_i}\right) - \lambda_{FP}\left(\frac{|\{ i \in M(t) \mid r_i = 0 \}|}{|M(t)|}\right), \quad (3)$$

where the first term rewards capturing pixels in higher-risk areas (reduces potential false negatives), the second penalizes pixels outside risk areas (reduces potential false positives), and λ_{FP} is an adjustable scaling factor. Finally, using Equation (3) the best risk score is calculated as:

$$R_{max} = max_{t\in\{0,0.05,\ldots,0.3\}} R(t), \quad (4)$$

and the NDWI mask with the maximum risk score is selected as the optimal threshold.

2.4.2 False-positive weighting factor (λ_{FP}): We expect the behaviour of the risk score to vary with the spatial scale and anticipated extent of a flood event. To account for this, a weighting factor, λ_{FP}, is applied to the false-positive penalty term to control the extent to which detections in zero-risk areas reduce the final score.

For small or spatially localized flood events, the true inundation extent is likely to occupy only a limited portion of the mapped hazard zone. Under these conditions, lowering the NDWI threshold may increase the number of detected pixels within flood-prone areas and therefore improve the first term of the score, even if many of those detections do not represent actual floodwater. At the same time, thresholds below the optimum often increase the detection of isolated or spurious water pixels outside likely flood areas due to spectral confusion with urban rooftops, shadows, wet soil, and other non-water surfaces. In such cases, assigning a higher value of λ_{FP} increases the penalty for these false positives and helps prevent the risk score from favouring non-optimal thresholds.

In contrast, during very large flood events, inundation may extend across a large fraction of the mapped hazard zone and, in some cases, even beyond the current extent of the risk raster. In this case, some detections in zero-risk areas may still correspond to floodwater rather than classification error. Therefore, assigning a lower value of λ_{FP} reduces the influence of the false-positive penalty and prevents the score from being overly conservative when floodwaters exceed the known flood hazard areas.

In practice, λ_{FP} can be selected based on expected flood magnitude, hydrologic context, or empirical calibration using historical flood events. This provides flexibility to adapt the risk-guided threshold optimization to different flood scenarios while maintaining a balance between (i) capturing likely inundated areas and (ii) limiting spurious detections outside the expected flood domain.

To estimate the likely magnitude of the 2018 flood along the Saint John (Wolastoq) River, water levels recorded at three hydrometric stations located within or near the five study areas were obtained for May 2, 2018 (i.e., the same day the satellite imagery was acquired) from historical hydrometric records (Environment and Climate Change Canada, 2026). The daily average water levels were then compared with the water levels predicted for the 1-in-20-year and 1-in-100-year flood hazard maps for each station location. Because the flood hazard maps use Canadian Geodetic Vertical Datum of 2013 (CGVD2013) as the vertical reference datum, while the hydrometric station records are referenced to the Canadian Geodetic Datum of 1928 (CGVD28), conversion factors were applied to the station readings using published conversion values obtained from data made available by the Government of Canada (Environment and Climate Change Canada, 2026).

The recorded water levels were close to, but did not exceed, the peak water levels predicted for the 1-in-20-year flood hazard at all three hydrometric stations (Table 1). We therefore expect the inundated areas visible in the satellite imagery to fall predominantly within the boundaries of the mapped risk zones and accordingly set $\lambda_{FP} = 1$, such that the full penalty is applied to detections in zero-risk areas.

Hydrologic Station	Daily Average (May 2, 2018)	1/20 risk	1/100 risk
01AK003	7.54 m	8.32 m	9.11 m
01AO002	6.41 m	7.05 m	7.81 m
01AO011	5.99 m	6.35 m	7.16 m

Table 1. Water levels recorded at the three hydrometric stations in the study region, including the daily average observed water levels on May 2, 2018, and the corresponding water levels predicted from flood hazard maps at each station location. All water levels are referenced to CGVD2013.

2.5 SAM refinement

NDWI thresholding alone may produce noise in urban areas (e.g., due to bright rooftops falsely detected as water) or it may fail to capture flooded vegetation. To improve flood delineation boundaries, we apply SAM's zero-shot, automatic mask generator to the false-colour RGB composite image (R: NIR, G: Red, B: Green) resulting in a set of proposal segments. This refinement step allows SAM to use the NDWI mask and risk maps as spatial priors, while contributing clean object boundaries derived from RGB features.

SAM segments were retained for NDWI refinement based on the following rules: (1) segments with > 80% overlap with the selected best NDWI mask were accepted directly; (2) segments with > 70% overlap with the NDWI mask and > 80% overlap with non-zero risk pixels were also retained; and (3) segments with > 60% overlap with the NDWI mask and > 90% overlap with non-zero risk pixels were likewise retained. The first rule prioritised segments that strongly matched the spectral flood signal, while the second and third rules allowed inclusion of segments with slightly lower NDWI agreement where there was strong supporting evidence from the flood risk surface. The final refined mask was produced by taking the logical union of all retained SAM segments and merging them with the original best NDWI mask.

2.6 Accuracy Assessment

Segmentation performance was evaluated using a peak flood inundation map derived from remote sensing and hydrologic/hydraulic data for the Spring 2018 event. This map is freely available from the Province of New Brunswick's open data portal (Service New Brunswick, 2025). Water-surface elevations were applied across 83 model transects, and the resulting inundation extents were merged to produce a continuous flood-extent surface for the Lower Saint John River representing the maximum water levels of the 2018 flood. This map represents a very accurate depiction of peak flood extents and water depths for the region and was thus used as our 'ground truth' flood extent

reference masks (Figure 2C). For each study region and satellite sensor, we computed the typical segmentation metrics, including Intersection over Union (IoU), F1-Score, Precision, and Recall.

3. Results & Discussion

3.1 Risk-based NDWI scoring

The results show that risk score strongly correlates with segmentation accuracy (Figure 4), demonstrating that flood hazard maps can objectively guide NDWI threshold selection. The trend was similar for both the Landsat and Sentinel imagery. This supports the use of a data-driven framework for selecting NDWI thresholds based on the spatial likelihood of flooding in a region. Moreover, this risk-guided approach reduces the uncertainty associated with traditional NDWI-based segmentation and provides a reproducible strategy for flood mapping across different scenes and environmental conditions, as well as different sensors.

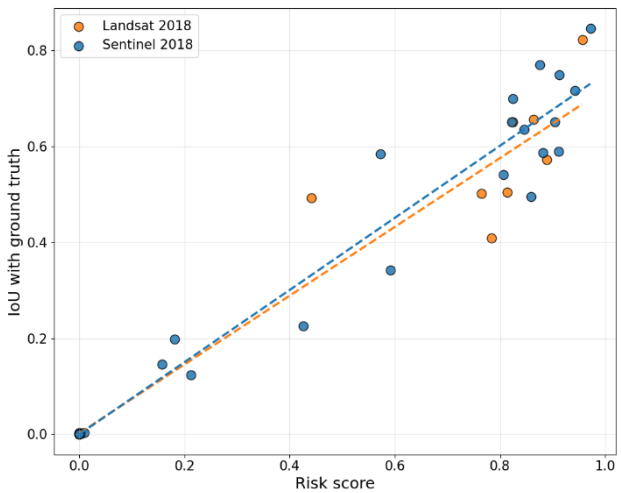

Figure 4. Scatter plot showing the relationship between risk score and IoU (computed from a ground truth flood extent mask) for each NDWI threshold-derived flood mask.

3.2 SAM refinement

Refinement of the NDWI-derived flood mask using the Segment Anything Model (SAM) further improves segmentation performance. Across the study regions, applying SAM to the NDWI mask associated with the highest risk score consistently enhanced flood boundary delineation (Table 2, Figure 5). It is noteworthy that precision remains relatively stable for both Landsat and Sentinel imagery. This pattern suggests that SAM mainly reduces false negatives (i.e., it adds flood pixels that were missed by the NDWI thresholding step) without introducing a significant number of false positives. Thus, the visual consistency of object boundaries added by SAM helps to produce a more complete flood mask.

Satellite	Method	IoU	F1-score	Precision	Recall
Landsat	NDWI	0.610	0.751	0.992	0.613
	NDWI+ autoSAM	0.751	0.851	0.989	0.758
Sentinel	NDWI	0.738	0.848	0.996	0.740
	NDWI+ autoSAM	0.866	0.925	0.992	0.872

Table 2. Summary of segmentation performance averaged across the five study regions for the best risk-guided NDWI mask (selected based on max risk score, R_{max}) alone and refined with the SAM automatic mask generator.

3.3 Landsat vs. Sentinel

Sensor characteristics also influenced segmentation performance. Sentinel-2 imagery yielded the highest overall accuracy, likely due to its finer spatial resolution (10 m) compared to Landsat-8 (30 m). Higher spatial resolution reduces mixed pixels along water–land boundaries and allows for better representation of narrow floodplain features such as channels, levees, and inundated vegetation zones. Nevertheless, the proposed workflow demonstrated an improvement in flood delineation for both sensors, indicating that the approach can be generalized across satellite imaging systems. Moreover, because the NDWI relies only on green and near-infrared spectral bands, there is potential to expand the methodology to other optical sensors and platforms.

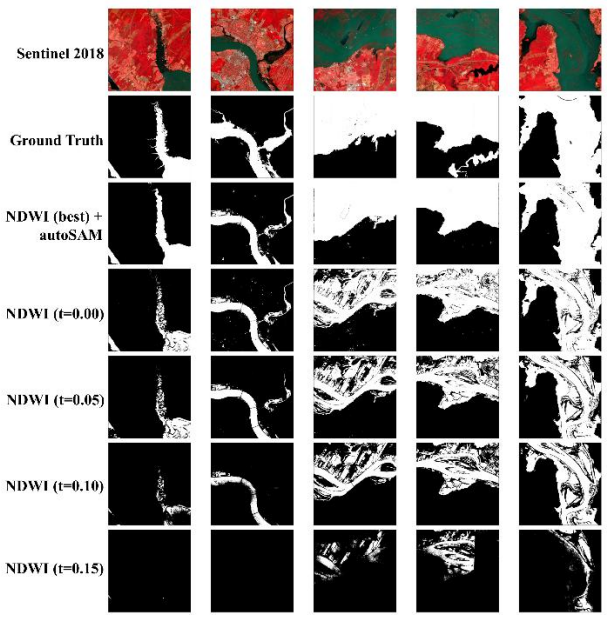

Figure 5. Comparison of flood segmentation outputs across the five study regions using Sentinel-2 false-colour composites (NIR–Red–Green). For each region, the panels include the ground-truth flood extent, the highest-ranked NDWI mask selected according to the risk score, the corresponding NDWI + SAM refined segmentation, and the next best NDWI threshold-based masks. The threshold used to generate each NDWI mask is denoted by t. Masks derived from thresholds > 0.15 are excluded for visual clarity.

4. Conclusions

4.1 Overview

Accurate flood extent mapping is essential for emergency response, infrastructure planning, and flood risk management. This research demonstrates that combining risk-guided spectral thresholding with foundation models can improve flood delineation in optical satellite imagery. The NDWI thresholding stage identifies candidate water pixels based on spectral properties and flood risk context, while SAM refines these detections by introducing visual cues present in the imagery. The workflow requires no supervised model training, making it suitable for rapid deployment in regions where labelled datasets are unavailable or where new flood events must be mapped quickly.

4.2 Limitations & Future Work

Recent improvements to SAM (i.e., SAM 3) provide further opportunities to extend this work. SAM 3 enables segmentation through text-based prompts in addition to manual point and bounding-box prompts (Carion et al., 2025). Using simple expressions such as "water," SAM 3 can leverage semantic information to identify objects within an image. Although this model became available after the completion of the present study, it represents a promising avenue for future research. Comparing SAM 3 with the risk-guided NDWI–SAM workflow may help us understand whether semantic prompting alone can effectively identify floodwater in complex landscapes or whether integrating physical risk constraints remains beneficial. Future work could also explore hybrid approaches that combine semantic prompting with our method.

The fact that our method relies on optical satellite imagery, which is sensitive to atmospheric and environmental conditions, may limit the application of our study. Flood events are frequently associated with heavy cloud cover and precipitation, which can obscure the surface and limit the availability of usable observations during the flood period. Synthetic aperture radar (SAR) sensors offer an alternative because they can penetrate clouds and acquire imagery regardless of environmental conditions. Integrating SAR observations into the proposed framework could improve the temporal reliability of flood mapping during extreme weather events.

Despite the effect of cloud cover on optical imagery, the spectral bands required for NDWI (including visible green and near-infrared wavelengths) are widely available across many remote sensing platforms. The proposed methodology could thus potentially be extended beyond satellite imagery to higher-resolution airborne and unmanned aerial vehicle (UAV) systems. These platforms provide flexible data acquisition during flood events (often beneath the cloud canopy) and can capture fine-scale inundation patterns in urban landscapes. Future research will investigate the application of the risk-guided segmentation framework to airborne and UAV imagery.

Finally, our method uses an adjustable scaling factor to adjust the importance of risk areas depending on the size of the flood. In this work, nearby water gauge readings are used to determine the magnitude of the flood, thus allowing us to choose appropriate values for the scaling factors. However, the effect of adjusting scaling factor values on the results requires a more rigorous evaluation and both small and large flood events will be tested in future work.

4.3 Key Contributions

This study makes the following contributions to the field of remote sensing-based flood mapping:

1. A novel data-driven approach for selecting NDWI thresholds using flood hazard information, providing an objective and physically informed alternative to manual or heuristic threshold selection.
2. An automated segmentation framework that combines NDWI-based water detection with refinement using a foundation segmentation model (SAM), improving flood boundary delineation without requiring model training.
3. The method generalizes across different optical satellite sensors (Sentinel-2 and Landsat-8) and relies solely on publicly available datasets.
4. A reproducible workflow that reduces manual intervention in flood mapping and can support rapid assessment during flood events.

5. Acknowledgements

We would like to thank our colleagues and collaborators for their helpful discussions and feedback during the development of this research.

We also acknowledge the space agencies and organizations responsible for the Sentinel-2 and Landsat-8 satellite missions for providing open-access earth observation data that make large-scale flood monitoring possible.

Finally, we thank the Government of New Brunswick for providing publicly available and highly accurate flood extent maps, which were used as reference data in this study. Their commitment to open data greatly supports research and improved flood risk understanding.

6. References

Bai, Y., Wu, W., Yang, Z., Yu, J., Zhao, B., Liu, X., Yang, H., Mas, E., Koshimura, S., 2021. Enhancement of detecting permanent water and temporary water in flood disasters by fusing Sentinel-1 and Sentinel-2 imagery using deep learning algorithms: Demonstration of Sen1Floods11 benchmark datasets. Remote Sensing, 13(11), 2220. https://doi.org/10.3390/rs13112220.

Carion, N., Gustafson, L., Hu, Y.-T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., Lei, J., Ma, T., Guo, B., Kalla, A., Marks, M., Greer, J., Wang, M., Sun, P., Rädle, R., Afouras, T., Mavroudi, E., Xu, K., Wu, T.-H., Zhou, Y., Momeni, L., Hazra, R., Ding, S., Vaze, S., Porcher, F., Li, F., Li, S., Kamath, A., Cheng, H.K., Dollár, P., Ravi, N., Saenko, K., Zhang, P., Feichtenhofer, C., 2025. SAM 3: Segment anything with concepts. arXiv preprint. https://arxiv.org/abs/2511.16719.

Composto, R.W., Tulbure, M.G., Tiwari, V., et al., 2025. Quantifying urban flood extent using satellite imagery and machine learning. Natural Hazards, 121, 175–199. https://doi.org/10.1007/s11069-024-06817-5.

Environment and Climate Change Canada, 2026. Water Level and Flow. Available online: https://wateroffice.ec.gc.ca/index_e.html (accessed 6 April 2026).

Esfandiari, M., Jabari, S., McGrath, H., Coleman, D., 2020. Flood mapping using random forest and identifying the essential conditioning factors: A case study in Fredericton, New Brunswick, Canada. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-3-2020, 609–615. https://doi.org/10.5194/isprs-annals-V-3-2020-609-2020.

IPCC. (2022). Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the IPCC.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 4015–4026. https://doi.org/10.48550/arXiv.2304.02643.

McFeeters, S.K., 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425–1432. https://doi.org/10.1080/01431169608948714.

Service New Brunswick, 2025. *New Brunswick Flood Hazard Maps*. https://geonb.snb.ca/flood_hazard_maps/index.html (accessed August 1, 2025).

Shokati, H., Seufferheld, K.D., Fiener, P., Scholten, T., 2026. Rapid flood mapping from aerial imagery using fine-tuned SAM and ResNet-backboned U-Net. Hydrology and Earth System Sciences, 30, 743–756. https://doi.org/10.5194/hess-30-743-2026.

Task Force on Flood Insurance and Relocation, 2022. Adapting to Rising Flood Risk: An Analysis of Insurance Solutions for Canada. Public Safety Canada, Ottawa, Canada. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/dptng-rsng-fld-rsk-2022/dptng-rsng-fld-rsk-2022-en.pdf