Measuring Tailings Storage Facility Bathymetry Using Sentinel-2 and Landsat-8/9 Multispectral Imagery and Machine Learning

Caio E. Stringari[1]*, Jeanine Engelbrecht[1,2], Brett Eaton[1]

[1] BGC Engineering, 980 Howe St, Vancouver, BC, Canada - (cstringari, jengelbrecht, beaton)@bgcengineering.ca
[2] Department of Geography and Environmental Studies, Stellenbosch University, Stellenbosch 7955, South Africa

Keywords: Bathymetry, Sentinel-2, Machine Learning, Deep Learning, Convolutional Neural Networks, Tailings Storage Facility, Mining.

Abstract

Tailings, a byproduct of mining, consist of fine sediment particles suspended in water that are stored in tailings storage facilities (TSFs). The discharge of untreated TSF water into the environment is typically prohibited due to its contact with mine tailings and processing chemicals. TSF failures have caused damage to communities and the environment, prompting calls for better management practices and advanced monitoring tools. For operational mine water management, boat-based bathymetric surveys have been used. However, these technologies have limitations, especially when the surveying of large facilities is required. Advances in remote sensing, particularly satellite-based earth observation (SBEO), offer cost-effective solutions for monitoring TSFs. This study explores the use of machine learning models, including XGBoost and Convolutional Neural Networks (CNNs), applied to Sentinel-2 and Landsat-8/9 data to estimate TSF bathymetry. Surveyed bathymetry datasets were used for model training, testing, and results validation. The results of the experiments revealed that high-accuracy bathymetric estimates could be obtained with mean absolute errors between 6 and 12 cm depending on the source of the data (i.e. Sentinel-2 or Landsat-8/9) and the model used (XGBoost vs CNN). Limitations include mixed pixel effects on the pond-beach interface and lower accuracies obtained in shallow areas, notably when XGBoost is used. This research underscores the potential of using satellite data and machine learning for TSF bathymetric monitoring, with implications for enhancing environmental and safety standards in mining operations.

1. Introduction

Tailings are a common by-product of the mining process and are created when mined ore is crushed, ground, and processed to extract the valuable minerals. Tailings usually consist of a slurry of fine particles of sediment suspended in water and are stored in specially designed, fit-for-purpose impoundments called tailings storage facilities (TSFs). The water in TSFs has been in contact with mine tailings and chemicals used in mineral processing. Consequently, discharge into the environment without prior treatment is typically not permitted and where treatment is not a viable option, the water must be stored. While TSFs around the world are typically well managed, 257 failures have been recorded between 1915 and 2020 resulting in 2,650 fatalities and the release of over 250 million m^3 of contaminated mine waste materials into the environment (Piciullo et al., 2022). Materials released due to TSF failures can travel hundreds of kilometres, contaminating rivers, lakes, and the surrounding land. With between 6,810 and 20,230 TSFs globally and an estimated cumulative failure rate of between 1.2% and 4.4% (Rana et al., 2022), the potential impacts of TSF failures are significant. Since TSF failures have had catastrophic impacts on lives, communities, local economies, and the environment (Cacciuttolo and Cano, 2023; Navarro et al., 2019), these incidents have led regulators, communities and mining investors to demand improved tailings management practices, increased safety standards and improved monitoring tools (Cacciuttolo and Cano, 2023).

1.1 TSF water management and the need for bathymetric data

An operational TSF is highly dynamic with a tailings surface that constantly changes as new tailings are deposited, or water removed for mining operations. The water depth of the supernatant pond in the TSF fluctuates due to water losses (as process water is reclaimed or from evaporation and seepage) and water input (with water as a component of slurry as it is deposited or from precipitation and surface runoff). The presence of large volumes of water exceeding the capacity of the TSF has been a contributor to most recent TSF failures with internal and external erosion of the TSF, seepage and overtopping being the main causes of failures (Lumbroso et al., 2021). Inadequate control of the volumes of water in a TSF can lead to an increased risk of failure and the consequences thereof. Additionally, since the main pathway for contaminants into the environment is by water, the effective management of TSFs is primarily a water management problem and starts with knowledge of the volume of water in the TSF at any given time.

A means to assess the volume of water in the supernatant pond is by periodic bathymetric surveys. Due to the importance of the bathymetric data in operational TSF management plans, bathymetric surveys are typically performed annually, quarterly, or even monthly. Historically, bathymetric surveys were conducted using crewed boats equipped with echo sounders. However, occupational health and safety concerns associated with boating accidents and exposure to potentially hazardous tailings, combined with the inability of large boats to survey

* Corresponding (presenting) author

https://doi.org/10.5194/isprs-archives-XLVIII-M-4-2024-63-2024

shallow ponds have led to the increased adoption of unmanned surface vehicles equipped with bathymetric survey gear. These systems have the advantage of being operated remotely although they tend to get trapped in viscous sediments or mud. The main limitation of bathymetric surveying approaches is the inability to produce detailed surveys of very large TSFs in a reasonable amount of time. Therefore, bathymetric survey approaches frequently rely on the surveying of transects and interpolating between the gaps, which leads to uncertainties in the depth estimates. While bathymetric surveys provide a means of estimating the pond volume, due to the time-consuming nature of the surveys, bathymetry surfaces are usually produced at monthly, quarterly, or seasonal intervals, limiting the potential for use in operational mine water management settings.

1.2 The role of SBEO for bathymetric measurement

Advances in satellite and sensor technology to capture data from remote sensing systems mean that cost-effective tools are being developed for the monitoring of water bodies in mining projects. The main advantage of satellite-based Earth observation (SBEO) is that it allows for the remote monitoring of hazardous or inaccessible areas. To further develop tools for effective water management, satellite data have been applied to detect the depth of water in lakes, rivers, oceans, and supernatant ponds of TSFs (Navarro et al., 2019).

For water depth measurement, remote sensing data (including airborne hyperspectral, satellite multispectral and bathymetric lidar data) have become viable tools for mapping water depth (Legleiter and Harrison, 2019). A variety of algorithms have been developed for remote sensing of river bathymetry, which relies on establishing a relationship between measured remote sensing image reflectance and water depth using empirical, physical or machine learning models. The physical foundation for estimating water depth from various kinds of remotely sensed data is the attenuation of electromagnetic waves with distance travelled through the water column. The attenuation and absorption of different wavelengths as a function of water depth have been investigated (Legleiter and Harrison, 2019) and suggest that shorter blue wavelengths were affected more by attenuation and absorption compared to green and red wavelengths. This was attributed to chlorophyll, and organic matter dissolved in the water column. Attenuation also increased significantly in the longer wavelength (beyond 700 nm - near infrared).

The simplest bathymetry estimation algorithms use regression to establish a relationship between the measured image reflectance to co-located depth measurements (obtained by in-situ measurements or bathymetric surveys). The main limitations of regression models are that the relationships are affected by the reflectance from the stream bed, but also by the optical properties of the water column, reflection from the water surface and atmospheric contamination (Legleiter & Harrison, 2019). To overcome these limitations, alternative approaches have evaluated band ratios as potential predictors of water depth. The optimal band ratio analysis (OBRA) approach considers all possible satellite image band combinations and seeks to identify the pair of wavelengths that provides the strongest relationship between measured reflectance and water depth (Legleiter and Harrison, 2019; Niroumand-Jadidi and Vitti, 2016). The OBRA technique was used in bathymetric estimates of a fluvial system in California, USA. The results demonstrated that OBRA yields reliable depth estimates in the presence of variable substrates, water column characteristics, and water surface textures (Legleiter and Harrison, 2019). Furthermore, the outcomes of applying the algorithms to data at different stages of radiometric calibration suggested that atmospheric correction of data did not appreciably improve depth retrieval. Similarly, using OBRA techniques on a small, shallow alpine river (Niroumand-Jadidi & Vitti, 2016) in Italy revealed that the red-edge and near-infrared bands yielded optimal results when used as denominators, while using the coastal blue, blue, green and yellow bands as numerators, yielded high correlations. Despite these successful demonstrations, the optical properties of the water column, influenced by suspended and dissolved constituents such as sediment and organic matter, were found to be the primary factor determining the feasibility of mapping river bathymetry via remote sensing.

To address the potential limitations imposed by OBRA techniques, the K-nearest neighbours (KNN) machine learning approach was applied for water depth estimation due to its ability to be applied across a broader range of environmental conditions. KNN provided a high level of depth retrieval performance superior to the OBRA models, implying that this machine-learning approach could facilitate depth retrieval, particularly for multispectral satellite images (Legleiter and Harrison, 2019). Citing the recent advancements in the radiometric resolution of satellite sensors resulting in a greater sensitivity to small variations in water-leaving radiance, an improved ability to retrieve river bathymetry from satellite data was demonstrated (Niroumand-Jadidi et al., 2022). The analytical approach leveraged neural networks (NN) to retrieve bathymetry by analysing the complex, non-linear relationships between the spectral response of satellite data and measured water depth. The NN-based retrievals were compared with retrievals from OBRA techniques with NN-models outperforming the OBRA techniques in all experiments (Niroumand-Jadidi et al., 2022). The results of the investigation demonstrated the ability to retrieve bathymetry from Landsat-9 data with high accuracy (i.e. that model predictions align more closely with ground truth data), at depths of up to 20 m.

Although experiments on satellite-bathymetry retrievals for TSF bathymetry are more limited, the use of Sentinel-2 to retrieve TSF bathymetry in Peru has been demonstrated (Navarro et al., 2019). The approach analysed the difference in reflected radiation from the water surface and from the depth by analysing the differential absorption of different wavelengths as the depth of the water increases (Navarro et al., 2019). The study compared bathymetry estimates from Sentinel-2 data with bathymetry data obtained from sonar equipment. The experiments demonstrated the ability to estimate the TSF bathymetry from Sentinel-2 with an average error of <10 % when compared to sonar-based bathymetric measurements. The main limitations were identified to be related to the turbidity of the water at the time of satellite image capture (Navarro et al., 2019). The results of these investigations demonstrated that SBEO has the potential to be applied, operationally, for the retrieval of TSF bathymetry for the management of TSF water balance.

The potential advantages of using satellite-derived bathymetry for mine water management include a reduction in the operational risks since human presence is not required, together with the ability to extract spatially continuous bathymetric surfaces even for shallow tailings beaches (Navarro et al., 2019). Satellites also provide daily, weekly, bi-weekly, or monthly revisit potential, improving the frequency of bathymetry estimates for operational decision-making. Additionally, since satellites have been imaging the Earth for many decades, the ability to estimate bathymetry from satellite data will allow for the assessment of historical deposition rates and spatial and temporal patterns of these depositions.

2. Site characteristics, bathymetric and satellite data preparation

To test the potential and limitations of SBEO bathymetry retrieval algorithms, the first phase of the project was dedicated to selecting an appropriate site for algorithm development and results validation. Two main considerations for site selection included that: i) several epochs of bathymetric surveys have been conducted and data is accessible and ii) water turbidity conditions are favourable. Sites with longer records of bathymetric surveys and lower turbidity were prioritized. Ultimately, the site selected for the experiments is an operational gold and copper mine in Northern British Columbia, Canada, with TSF dimensions of 2 km x 0.8 km. The site was selected since, during construction, one of the main considerations was the protection of natural drainages. Therefore, all water that is used or moves across the surface of the site is collected and directed to the TSF. This water is recycled from the TSF and reused in the mill. This makes the mine TSF a very dynamic surface, with pond volumes fluctuating constantly due to water collection, storage, and reuse.

Given the importance of tailings pond volume for mine water management, regular boat-based bathymetric surveys are also conducted at the mine. These surveys are conducted at roughly monthly intervals during the summer months. No bathymetric surveys are conducted in the winter when ice is present on the water surface. For this investigation, 26 bathymetric survey datasets, collected between July 2021 and October 2023 were made available. These data were organized and pre-processed (including gridding and re-projection to common coordinate reference systems).

Since TSFs are highly dynamic with new materials continuously being added or removed, satellite data acquisition dates were filtered to dates matching, within two days, the dates of the bathymetric surveys. The experiments focused on Sentinel-2 and Landsat8/9 atmospherically corrected surface reflectance data obtained from the Copernicus Browser and the USGS Earth Explorer respectively. Scenes with cloud cover affecting the TSF were discarded. A summary of the remaining datasets that were used for further analysis is provided in Table **1**.

Bathymetric survey date	Sentinel-2	Landsat 8/9
June 23, 2023	June 21, 2023	June 21, 2023
October 17, 2022	NA	October 10, 2022
September 26, 2022	September 24, 2022	NA
August 23, 2022	August 23, 2022	August 21, 2022
July 26, 2022	July 26, 2022	July 27, 2022
July 4, 2022	NA	July 4, 2022
October 2, 2021	October 2, 2021	NA

Table 1. TSF bathymetric data and corresponding satellite datasets.

3. Model implementation and accuracy assessment

Empirical models are models based on observations rather than on a pre-defined physical model or mathematical formula. Examples of empirical models include linear regression models, decision trees or neural networks. Mathematically, a model can be denoted as a function (f) that takes as input a matrix (X) with size MxN in which M is the number of training samples and N is the number of imagery bands. The goal of the model is to generate the vector $\hat{y}$ with predicted depths that minimize a chosen cost (or loss) function. This relationship can be written as:

$$\hat{y} = f(X; y) \quad (1)$$

In this study, a decision tree ensemble model, eXtreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016), as well as convolutional neural networks (CNN) were tested for their ability to estimate TSF bathymetry.

XGBoost uses a collection of decision trees to predict a target variable based on several input features. Each tree in the forest is built from a sample drawn with replacement from the training set and can be thought of as a flowchart-like structure in which each internal node represents a “test” on an attribute (e.g., whether reflectance of Band 1 is greater than 0.4), the branch represents the outcome of the test (true or false), and each leaf node represents a predicted depth ($\hat{y}$). The paths from root to leaf represent classification rules.

Convolutional neural networks (CNNs) are more complex to create and train when compared traditional machine learning models as they take into consideration the two-dimensional aspect of the input data. CNNs, however, tend to produce better depth estimations (Peng et al., 2022). The CNN approach attempted in this project will follow the methods proposed by (Annan and Wan, 2022) and (Lumban-Gaol et al., 2021). A sliding window of size S is applied to the satellite image and sub-images of size SxS are created. Each sub-image is mapped to measured bathymetric values. Next, the data is organized into input (X) and output (y) matrices of size (M, S, S, N) where M is the number of training samples and N is the number of spectral bands. As with the empirical models, the predicted depth is given by $\hat{y} = f(X; y)$, albeit with more complex inputs and outputs. In this work, each sliding window has a 4-pixel overlap with the previous windows in both (x, y) directions. The overlapping windows will result in overlapping bathymetric predictions, which are then interpolated to a predefined grid using linear interpolation to smooth the prediction results.

The architecture selected for the CNN is based on the VGG-16 model (Simonyan and Zisserman, 2014). VGG-16 uses convolutional layers to gradually extract information from the input images and is, in this case, four layers deep, uses 3x3 convolution windows and has no data normalization between layers. The input image size to the model is 32x32 pixels. The loss function for training is the root mean square error ($RMSE$) and the models are set to train for 1000 epochs (i.e., time step iterations) but may stop training earlier if the validation dataset loss does not change by more than a set value for more than 10 epochs.

To implement the models, we extracted the reflectance values for each band of each satellite image, as well as the corresponding measured bathymetry values. The data was organized into two tables, one with reflectance band values (X) and another with the measured bathymetry values (y). The data was split into 3 datasets: training, validation, and testing. The training dataset was used for model training, meaning that the model learns from this data. The validation dataset is a subset of the data that is used to fine-tune model parameters (e.g., the learning rate in a neural network or the depth in a decision tree) and to prevent overfitting. Finally, the test dataset is subset of data that was not used for training or validation. The test data are used to evaluate the final performance of the model, theoretically providing an unbiased assessment of how the model will perform on unseen data. For the XGBoost models, the data splits are 25% training data, 25% validation data and 50% test data. For Sentinel-2, there are 86,000 training, 86,001 validation, and 172,002 testing samples, and for Landsat-8/9, there were 113,716 training, 113,716 validation, and 227,432 testing samples. For the CNN model, the data splits are 56% training data, 25% test data and 19% validation data, meaning that across all dates, the Sentinel-2 dataset has 12406 training, 5514 test and 4136 validation samples and Landsat-8/9 dataset has 16576 training, 7368 test and 5526 validation samples.

After applying the models to the satellite data, the performance of all models to predict the tailings bathymetry can be evaluated using three metrics (Niroumand-Jadidi et al., 2022). These metrics are:

i) Pearson's correlation coefficient which is expressed as:

$$r_{xy} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(y_i - \bar{\hat{y}})}{\sqrt{\sum_{i=i}^{n}(y_i - \bar{y})^2}\ \sqrt{\sum_{i=i}^{n}(y_i - \bar{\hat{y}})^2}}, \tag{1}$$

ii) the root mean square error ($RMSE$) expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \tag{2}$$

iii) the mean absolute error (MAE) expressed as:

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{3}$$

where n is the sample size, y_i is the observed bathymetry, $\hat{y}_i$ is the predicted bathymetry, and $\bar{y}$ are sample means $(\frac{1}{n}\sum_i^n y_i)$. In general, models with higher r_{xy} and lower $RMSE$ and MAE are better performing models.

4. Results

The results of implementing the respective models are summarised in Table 2. The results of the XGBoost suggest that, for both Sentinel-2 and Landsat-8/9 data, high accuracy bathymetric estimates could be retrieved, with the Sentinel-2 test dataset achieving accuracies of r_{xy}= 0.98, $RMSE$ = 0.12 m and MAE = 0.08 m. The Landsat-8/9 results scored slightly lower with $RMSE$ = 0.14 m. As part of the XGBoost results, an assessment of feature importance for Sentinel-2 and Landsat-8/9 was provided as shown in Figure **1**. The results suggest that, for both the Landsat-8/9 and Sentinel-2 datasets, the coastal aerosol (ultra blue) bands (Band 1 (B01) in both datasets) have the highest importance. In Landsat-8/9, this is followed by green (B03) and red (B04) bands. For Sentinel-2, the SWIR band (B09) appears to be of higher importance than the visible bands (B02, B03, B05 and B04), likely due to the improved ability to separate the land/water interface.

The TSF bathymetry prediction obtained by XGBoost for the Sentinel-2 and Landsat-8/9 scenes captured on August 23, 2023, are shown in Figure **2**. It is observed that the difference between the measured bathymetry and the predicted bathymetry is greater towards the edges of the TSF. Further investigation is required to find a cause for lower prediction accuracies in these areas, but the low resolution of the satellite data, leading to mixed pixel effects on the edges of the TSF may be a potential cause for these effects. Furthermore, the differential attenuation depending on the wavelength of the data used may have reduced impact in areas where the water is shallow, leading to lower accuracy bathymetric measurements.

Data	**Model**	**Dataset**	**r_{xy}**	***RMSE* (m)**	***MAE* (m)**
Sentinel-2	XGBoost	Train	0.98	0.1	0.07
		Test	0.98	0.12	0.08
		Validation	0.98	0.12	0.08
Landsat-8/9	XGBoost	Train	0.98	0.13	0.08
		Test	0.98	0.14	0.08
		Validation	0.98	0.14	0.08
Sentinel-2	CNN	Train	0.99	0.01	0.01
		Test	0.98	0.09	0.06
		Validation	0.99	0.07	0.05
Landsat-8/9	CNN	Train	0.99	0.03	0.01
		Test	0.93	0.17	0.12
		Validation	0.98	0.11	0.12

Table 2. Evaluation metrics for the models applied to Sentinel-2 and Landsat-8/9 data, calculated using all available satellite image dates

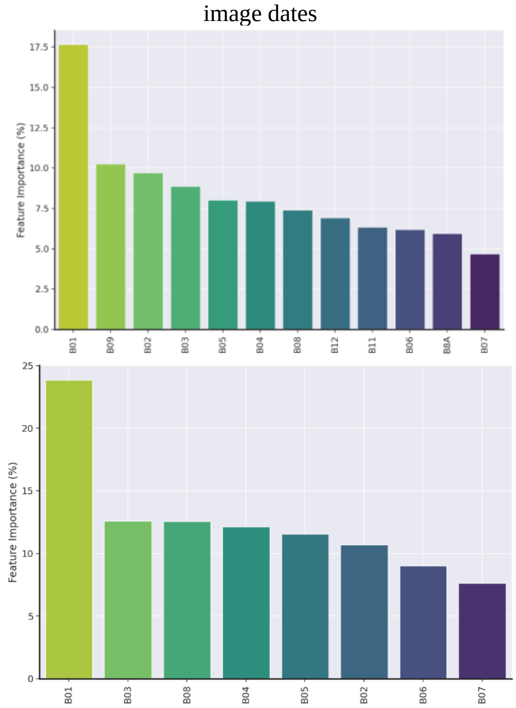

Figure 1. XGBoost feature importance for Sentinel-2 (top) and Landsat-8/9 (bottom).

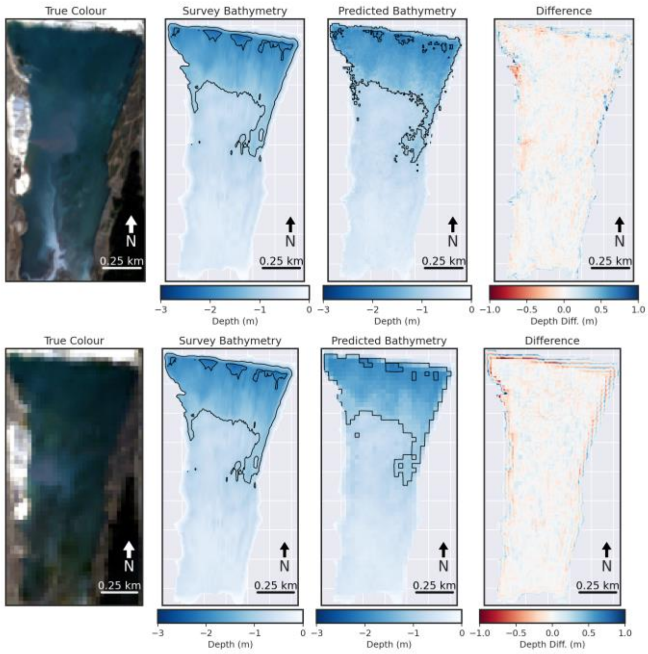

Figure 2. Results of the XGBoost model for bathymetry and Sentinel-2 data (top row) and Landsat-8/9 (bottom tow), collected on August 23, 2024. The colour scale in all plots is in meters. Contours are displayed in 1 m intervals.

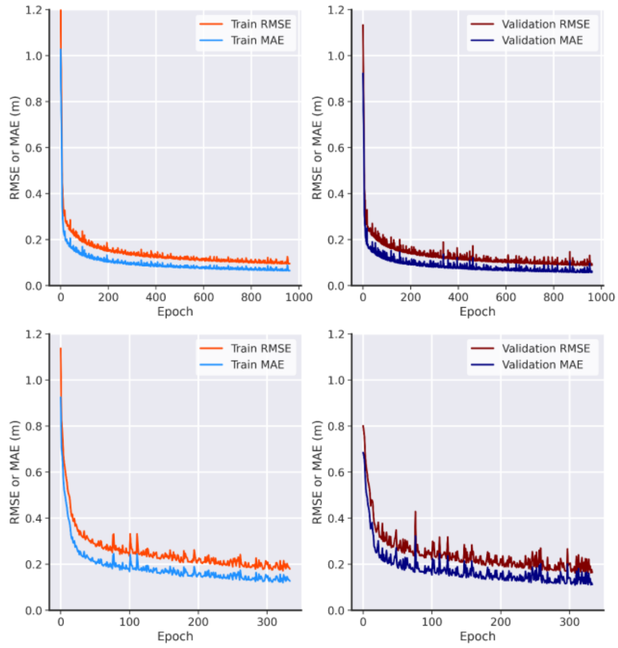

Figure 3. Training curves for the CNN trained on the Sentinel-2 data (top row) and for the Landsat-8/9 data (bottom row). These curves show how the model evolves over training time steps (epochs)

In applying the CNN models, the learning curves for the VGG-16 models were derived to evaluate the goodness of fit. These plots are shown in Figure 3. The plots indicate a good fit with both the Sentinel-2 and Landsat-8/9 training and validation loss decreasing to a point of stability and exhibiting a small gap between the training and validation loss. The comparisons between predicted and survey bathymetry and model evaluation metrics using the test datasets across all dates are shown in Figure 4. Considering the test dataset, for Sentinel-2 data, the CNN model is the best-performing model, with $RMSE$ = 0.09 and MAE = 0.06 compared to XGBoost ($RMSE$ = 0.12 and MAE = 0.08). For Landsat data, the XGBoost model proved to be superior, with $RMSE$ = 0.14 m and MAE = 0.08 m compared to the CNN with $RMSE$ = 0.17 and MAE = 0.12.

Figure 5 shows the CNN prediction for the Sentinel-2 and Landsat-8/9 data captured on August 23, 2023. The predicted bathymetry agrees well with the survey bathymetry although for the Landsat data, the poor model performance on the edges of the TSF is more prominent than on the Sentinel-2 data. It is also observed that the predicted bathymetry from Landsat-8/9 appears to be smoothed compared to the pixelated appearance of the predicted bathymetry using XGBoost due to the low resolution of the Landsat-8/9 data. This is a direct consequence of using the sliding window and interpolation approach which is an additional benefit of working with CNN on the coarse resolution data. The Landsat data also reveals errors in the prediction where vegetation on the edges of the TSF is mistaken for water and a depth estimate is provided in these areas. These artefacts are not present in the Sentinel-2 CNN, likely because the higher number of spectral bands are used by the CNN to better distinguish between water and land.

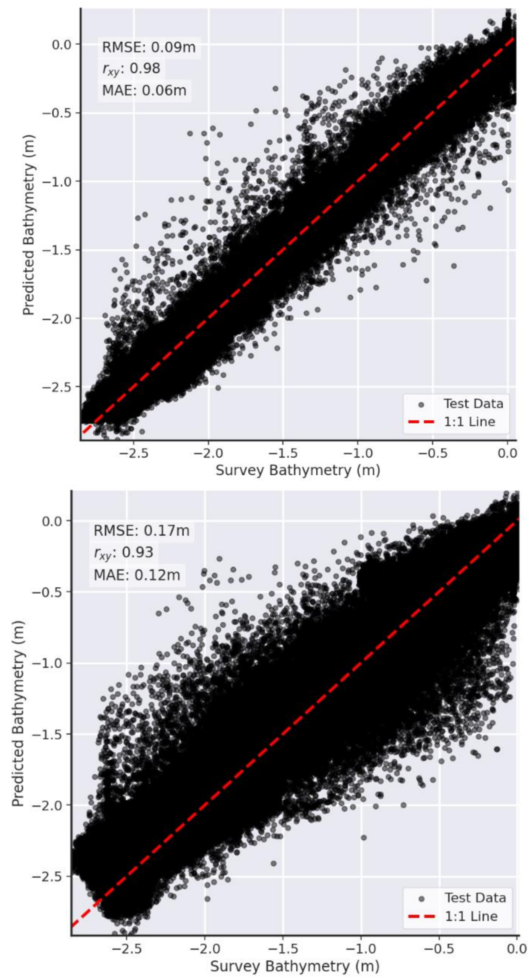

Figure 4. CNN Model performance assessment using the test dataset using data from all dates for Sentinel-2 (top) and Landsat-8/9 (bottom).

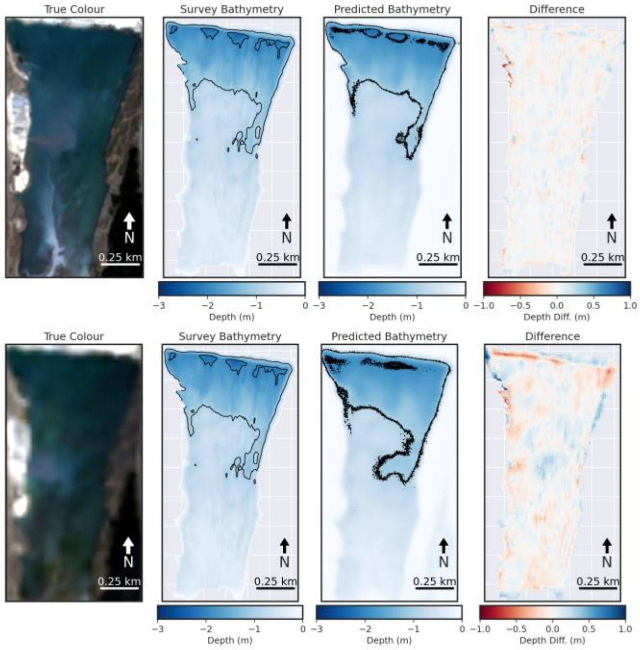

Figure 5. Results of the CNN model for bathymetry and Sentinel-2 data (top) and Landsat-8/9 (bottom), collected on August 23, 2024. The colour scale in all plots is in meters. Contours are displayed in 1 m intervals.

5. Conclusions and future work

To further explore the potential of using satellite data for monitoring TSF bathymetry, we tested the performance of the XGBoost and CNN models in estimating TSF bathymetry from publicly available data and validated the results against surveyed bathymetric measurements obtained at an operational mine. The current work allowed us to test the operational limitations and opportunities of using satellite data for bathymetric measurements for operational mine water monitoring applications. The results revealed that high-accuracy bathymetric estimates could be obtained with mean absolute errors between 6 and 12 cm depending on the source of the data (i.e. Sentinel-2 or Landsat-8/9) and the model used (XGBoost vs CNN). The accuracies obtained from the Sentinel-2 data were higher than the accuracies obtained from Landsat-8/9. This is likely due to the coarser resolution of the Landsat-8/9 data (30 m) compared to Sentinel-2 (10 m for the visible bands and 60 m for the coastal blue). The higher resolution data enable bathymetric estimates with greater detail, leading to higher accuracy results. The results for both Sentinel-2 and Landsat-8/9 suggest that, when using XGBoost, lower accuracy measurements could be obtained in shallow water environments. However, when using CNN on Sentinel-1 data, the accuracies in the shallow water environments were improved. Lower accuracies on the edge of the TSF remained present in all datasets, suggesting that mixed pixel effects due to the low resolution of the data would affect the results. However, these effects were less prominent in the Sentinel-2 data than in Landsat-8/9 due to higher spatial resolution.

Given the improvements provided by higher spatial resolution data, future research would aim to augment the lower resolution publicly available satellite data with high-resolution data to improve the accuracy of the bathymetric retrievals in shallow TSFs. While satellites such as SPOT-6/7 at 6 m resolution can provide improved resolution compared to Landsat-8/9 and Sentinel-2, SPOT-6/7 has a lower spectral resolution, notably missing the coastal blue and SWIR bands that were the wavelengths of highest feature importance according to the results from XGBoost. Additional investigations are needed to determine the impact of spatial resolution vs. spectral resolution on the accuracy of bathymetric estimates to determine if SPOT-6/7 would be a suitable alternative to lower-resolution retrievals. Very high-resolution data obtained by satellites such as Worldview-3 likely provide the greatest opportunity for operational TSF bathymetry retrieval from high-resolution sensors. Not only does the sensor provide data with a spatial resolution of between 1.2 and 3.7 m, but the imaged wavelengths include the coastal blue, visible, NIR and several SWIR bands. The improvements in both spatial and spectral resolution provided by sensors such as Worldview-3 will likely lead to improvements in the accuracy of the bathymetric estimates and are recommended for investigation in the future.

Although our initial results are promising, the main limitation of this work is that the ability to measure TSF bathymetry from satellite data was only tested at a single mine site and using a limited training dataset. Since different mines are expected to be highly diverse, with different bottom substrate composition, water chemistry and water-depth variations depending on the site, it is unlikely that models trained at one mine and over a limited temporal range would be directly transferable to different mines, in different geographic locations and different conditions. Therefore, to demonstrate that satellite bathymetric data are robust and reliable alternatives to conventional boat-based surveys, additional work is needed to develop, train and implement the models for TSF bathymetric measurement on different mines. It should also be noted that, although the bathymetric survey data was used in these experiments as a measure of true bathymetry, the bathymetric surfaces produced from surveys are known to be prone to errors and data noise. Therefore, future research should also validate the results against alternative measures such as the results of water balance equations and hydrological models. The main opportunity identified in this work is the ability to retrieve TSF bathymetric estimates from satellite data over large TSFs using a single image, effectively reducing the time associated with bathymetric surveys. Coupled with the ability to collect data every 5 to 10 days (or more often if multiple satellites or satellite constellations can be used), the end products will provide more regular bathymetric measurements for improved insights into TSF water management.

Acknowledgements

The authors gratefully acknowledge the funding of this work that was provided by the Canadian Space Agency (CSA) as part of the smartEarth initiative: Accelerating EO Innovations. The authors further acknowledge the expert advice provided by Daryl Dufault and Brieanna Shaw in the execution of the project.

References

Annan, R.F., Wan, X., 2022. Recovering Bathymetry of the Gulf of Guinea Using Altimetry-Derived Gravity Field Products Combined via Convolutional Neural Network. Surv Geophys 43, 1541–1561. https://doi.org/10.1007/s10712-022-09720-5

Cacciuttolo, C., Cano, D., 2023. Spatial and Temporal Study of Supernatant Process Water Pond in Tailings Storage Facilities: Use of Remote Sensing Techniques for Preventing Mine Tailings Dam Failures. Sustainability (Switzerland) 15. https://doi.org/10.3390/su15064984

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining https://doi.org/10.1145/2939672.2939785

Legleiter, C.J., Harrison, L.R., 2019. Remote Sensing of River Bathymetry: Evaluating a Range of Sensors, Platforms, and Algorithms on the Upper Sacramento River, California, USA. Water Resour Res 55, 2142–2169. https://doi.org/10.1029/2018WR023586

Lumban-Gaol, Y.A., Ohori, K.A., Peters, R.Y., 2021. Satellite-derived bathymetry using Convolutional Neural Networks and multispectral Sentinel-2 images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B3-2021, 201–207. https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-201-2021

Lumbroso, D., Collell, M.R., Petkovsek, G., Davison, M., Liu, Y., Goff, C., Wetton, M., 2021. DAMSAT: An Eye in the Sky for Monitoring Tailings Dams. Mine Water Environ 40, 113–127. https://doi.org/10.1007/s10230-020-00727-1

Navarro, L., Aravena, G., Engels, J., Turner, J., 2019. Satellite bathymetry for the monitoring of supernatant water volumes within tailings storage facilities, in: Proceedings of the 22nd International Conference on Paste, Thickened and Filtered Tailings. Australian Centre for Geomechanics, Perth, pp. 205–217. https://doi.org/10.36487/acg_rep/1910_13_engels

Niroumand-Jadidi, M., Legleiter, C.J., Bovolo, F., 2022. River Bathymetry Retrieval from Landsat-9 Images Based on Neural Networks and Comparison to SuperDove and Sentinel-2. IEEE J Sel Top Appl Earth Obs Remote Sens 15, 5250–5260. https://doi.org/10.1109/JSTARS.2022.3187179

Niroumand-Jadidi, M., Vitti, A., 2016. Optimal band ratio analysis of WorldView-3 imagery for bathymetry of shallow rivers (case study: Sarca River, Italy), in: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. International Society for Photogrammetry and Remote Sensing, pp. 361–364. https://doi.org/10.5194/isprsarchives-XLI-B8-361-2016

Peng, K., Xie, H., Xu, Q., Huang, P., Liu, Z., 2022. A Physics-Assisted Convolutional Neural Network for Bathymetric Mapping Using ICESat-2 and Sentinel-2 Data. IEEE Transactions on Geoscience and Remote Sensing 60, 1–13. https://doi.org/10.1109/TGRS.2022.3213248

Piciullo, L., Storrøsten, E.B., Liu, Z., Nadim, F., Lacasse, S., 2022. A new look at the statistics of tailings dam failures. Eng Geol 303. https://doi.org/10.1016/j.enggeo.2022.106657

Rana, N.M., Ghahramani, N., Evans, S.G., Small, A., Skermer, N., McDougall, S., Take, W.A., 2022. Global magnitude-frequency statistics of the failures and impacts of large water-retention dams and mine tailings impoundments. Earth Sci Rev. https://doi.org/10.1016/j.earscirev.2022.104144

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition.