Master Thesis

Using a Random Forest Model for Water Classification to Limit Erosion Impact in Bangladesh

Quincy van Iersel, 8351252

Supervisor: Dr. Menno W. Straatsma

Master Applied Data Science

Utrecht University

In cooperation with: Deltares

Date: 31-07-2023

Abstract

Riverbank erosion is a significant issue caused by various factors such as floods, deforestation, heavy rainfall, and strong river currents. Bangladesh, a country with three converging rivers, faces the impact of erosion during the monsoon season. To understand erosion's impact, water/no-water classification maps can assist water management decision making, which can result in timely erosion mitigation efforts. The current dashboard (Bangladesh Erosion Monitor – BEM) uses JRC-GSWE water/no-water classifications. However, the classifications from JRC-GSWE are only available until 2021, which means there is no near-real-time information available, limiting the effectiveness of the BEM.

Due to the unavailability of near-real-time JRC-GSWE classifications, an alternative method needs to be created. Therefore, this study aims to develop a Random Forest water/no-water classification model using Landsat and Sentinel-2 data and compare its accuracy with JRC-GSWE classifications. Additionally, the impact of a dual-sensor option using Landsat 8 and Sentinel-2 data will be analysed to determine if there is additional value in terms of classification accuracy. The results show that the Random Forest model using Landsat 8 data achieves the highest overall accuracy and Kappa coefficient when compared to the JRC-GSWE data sources, which means it is a reliable alternative to the JRC-GSWE classifications. However, combining Landsat 8 and Sentinel-2 data yields a balanced performance in classifying both water and no-water areas. The proposed model serves as a reliable alternative to JRC-GSWE data, which could support near-real-time water management decision making and timely erosion mitigation efforts in Bangladesh. This research is conducted in collaboration with Deltares, which is an institute that helps with water management decision making. They are working together with the Bangladesh Water Development Board (BWDB) and the institute of Water Modelling (IWM) to bring more insight on the impact of erosion at national scale.

Contents

1. Introduction

Riverbank erosion is the irregular surface cut by the eroding action of flowing water (Allaby, 2008). Riverbank erosion can be caused by flood, deforestation, heavy rainfall, strong current of the river or stilt deposition (Munna, 2018). The impact of erosion can have negative impact on the land and population. Interventions are needed which require planning, and monitoring to determine where and when erosion occurs. A country that suffers from erosion is Bangladesh. The country's climate and its three converging rivers (Padma, Jamuna, and Meghna), causes a conveyance of water, especially during the monsoon season. This can result in a high discharge that flows through these rivers and causes a dynamic shift in terms of river layout from year to year, where in some areas erosion occurs, while other areas have accreted by sediment deposition. To know the impact of erosion, an over-time water/no-water difference can be made. This can be done by mapping an image of water/no-water before and after the monsoon, comparing both images with each other. The process from finding weak spots alongside riverbanks, to applying erosion mitigation works takes time. Therefore, the faster this process goes, the less likely the loss of agricultural and residential land.

The current dashboard (Bangladesh Erosion Monitor (BEM)) makes an initial estimate where erosion can be expected in the short term, using a simple extrapolation of the observed trends in recent years. This information can then be used by the BWDB (or other relevant ministries) to better (and faster) plan and counter the impact of erosion with mitigation works (Deltares, 2023). The current dashboard uses JRC Global Surface Water Explorer (JRC-GSWE). These classifications assist Deltares with water surface information to help support water management decision making (e.g., water/no-water classifications).

The JRC-GSWE is a tool developed to monitor and analyse different facets of surface water dynamics. Together the maps show where and when open water is present on the earth's surface. The data from the JRC-GSWE provides detailed information on the extent and changes in waterbodies such as lakes, reservoirs, and rivers over time. In order to map water, JRC-GSWE used thermal imagery and the constraining spatial properties of water on other features (including snow, clouds, shadows, bare rock and vegetated land) in the Landsat sensors' six visible, NIR and SWIR channels to separate pixels acquired over open water from those acquired over other surfaces (European Commision Joint Research Centre, 2016). The model that can make the water/no-water classifications utilized for the JRC-GSWE involves a decision tree that incorporates the multispectral imagery of the Landsat archive, along with additional ancillary data layers. The model's performance was judged in term of errors of omission and commission at the pixel scale. The validation design considered the small spatial extent of inland water surfaces and its intrinsic spatio-temporal variation. The validation was performed using 40.123 control points distributed both geographically (globally), across the 32 years, and across all the different Landsat sensors (TM, ETM+ and OLI). Two reference data sets were produced, one was dedicated to

the omission errors and the other data set was used to characterize errors of commission. The presence of water for each single validation point was confirmed by visually checking all the points as followed: only points confirmed as water were used in the estimation of the omission error. The outcome shows a map with omission error points, commission error points, and correctly detected water points. If no-water was present, then this is corresponded to an error of commission. The results of this research demonstrated that errors of omission for seasonal water are higher than for permanent because there are fewer opportunities to observe each water body. These errors will result in underestimation of its occurrence (Pekel, Cottam, Gorelick, & Belward, 2016).

Although the current BEM can show useful information to counter the impact of erosion, an important part that is missing is the recent information. The JRC-GSWE makes water/no-water classifications which Deltares currently use for their BEM. The downside is that the JRC-GSWE data is only available until the year 2021, and it is unknown when more recent data will be available. Because of this limitation, near-real-time classifications cannot be made. Near-real-time means to use the most recent available Landsat imagery. The limitations of not having near-real-time classifications (water/no-water classifications after 2021) can cause problems in terms of not knowing where potential riverbank weak spots are and thus not unknown where potential erosion can occur. A solution for this is to develop a model that is also able to make water/no-water classifications similar to the JRC-GSWE classifications. Landsat images can provide insight in land classifications (e.g., water/no-water).

For the last 5 decades, the Landsat program has launched several satellites, each one equipped with increasingly advanced sensor technology. Landsat collections ensure that all Landsat products contain known data quality. The Landsat collection 2 is currently in use as available data collection which represents the next generation Landsat data. In order to classify waterbodies, the surface reflectance collection is used, which improves the images by accounting for atmospheric effects such as aerosol scattering and thin clouds, which help in the detection and characterization of land surface change (U.S. Geological Survey, 2015). The sensors aboard each of the Landsat satellites were designed to acquire data in different ranges of frequencies along the electromagnetic spectrum (see table 1).

Table 1 Landsat 8 Level 2 Surface Reflectance

Bands	Wavelengths (mm)	Resolution (m)
SR-Band 1 – Coastal aerosol	0.43-0.45	30
SR-Band 2 – Blue	0.45-0.51	30
SR-Band 3 – Green	0.53-0.59	30
SR-Band 4 – Red	0.64-0.67	30
SR-Band 5 – Near Infrared (NIR)	0.85-0.88	30
SR-Band 6 – Shortwave Infrared (SWIR) 1	1.57-1.65	30
SR-Band 7 – Shortwave Infrared (SWIR) 2	2.11-2.29	30

Another way to collect multi-spectral images of the planet is using Sentinel-2 data. Sentinel-2 has 13 different bands: four bands at 10 m, six bands at 20 m and three bands at 60 m (QA60 used for cloud masking) spatial resolution (table 2) (European Space Agency, 2023). Compared to Landsat 8, Sentinel-2 consists of two satellites (S2A and S2B) that have a five-day revisit frequency at the equator compared to the single Landsat 8 satellite that only has a frequency of 16-days (Drusch, et al., 2012).

Table 2 Sentinel-2 Level 2

Bands	Wavelength (nm)	Resolution (m)
B1 - Aerosols	443.9 (S2A) / 442.3 (S2B)	60
B2 - Blue	496.6 (S2A) / 492.1 (S2B)	10
B3 - Green	560 (S2A) / 559 (S2B)	10
B4 - Red	664.5 (S2A) / 665 (S2B)	10
B5 - Red Edge 1	703.9 (S2A) / 703.8 (S2B)	20
B6 - Red Edge 2	740.2 (S2A) / 739.1 (S2B)	20
B7 - Red Edge 3	782.5 (S2A) / 779.7 (S2B)	20
B8 - NIR	835.1 (S2A) / 833 (S2B)	10
B8A - Red Edge 4	864.8 (S2A) / 864 (S2B)	20
B9 - Water Vapor	945 (S2A) / 943.2 (S2B)	60
B11 - SWIR 1	1613.7 (S2A) / 1610.4 (S2B)	20
B12 - SWIR 2	2202.4 (S2A) / 2185.7 (S2B)	20
QA60 - Cloud Mask		60

In earlier research from (Carrasco, O'Neil, & Morton, 2019), the impact of combining Sentinel-1, Sentinel-2 and/or Landsat 8 with each other to produce land use classifications was analysed, to determine if there is additional value of combining multi-sensor data. This enables the efficient use of multi-temporal data and the exploitation of cloud-gap filling techniques for land cover mapping. The research provides the accuracy between land cover maps created with Sentinel-1, Sentinel-2, and Landsat 8 data from one year and test whether a combination would improve the classification accuracy. The results in Appendix I show that combining different sensor data improves the classification accuracy compared to a single-sensor approach.

In earlier research from (Tottrup, et al., 2022), a round robin exercise was organized to conduct an intercomparison of 14 different satellite-based approaches for monitoring inland surface dynamics with Sentinel-1, Sentinel-2, and Landsat 8 imagery. The objective was to achieve a better understanding of the pros and cons of different sensors and models for surface water detection and monitoring. The results of this research indicated that while using a single sensor approach can provide comprehensive results for very specific locations, a dual sensor approach is the most effective way (combining Landsat 8 with Sentinel 2 data) to undertake largescale national regional surface water mapping. Appendix II shows the accuracy results of this research where the dual sensor approach shows an overall higher accuracy

compared to the single sensor approach in most of the cases. The results also suggest that the impact using a multi-sensor approach might differ per region.

In other research where Landsat 8 data was utilized, specifically the visible and near-infrared (NIR) bands, to generate Normalized Difference Vegetation Index (NDVI) images (Taufik, Syed Ahmad, & Ahmad, 2016). Subsequently, an analysis was conducted to distinguish between vegetation, non-vegetation, and water areas. The determination of suitable thresholds for separating these areas was accomplished with the assistance of ground-truth information obtained from the study area. The accuracy of the classification process was evaluated through the utilization of a confusion matrix, which facilitated the computation of the overall classification accuracy and Kappa coefficient. The findings revealed that the NDVI-based classification demonstrated a high level of accuracy of 95.55% with a Kappa coefficient of 0.915 in classifying Landsat 8 data.

Machine learning models are able to use input data (e.g., Landsat 7 or 8) to make classification predictions (water/no-water classifications). In recent years, random forest classification models have emerged as an effective approach for accurately identifying water bodies from Landsat satellite imagery (REFs). Random forest is an ensemble learning algorithm that combines multiple decision trees to make predictions. It is known for its ability to handle complex and high-dimensional datasets, as well as its robustness against noise and outliers. (Nguyen, Doan, & Radeloff, 2018, Breiman, 2001, Byoung, Kim, & Nam, 2015, Ko, Kim, & Nam, 2011).

The aim of this research is to develop a random forest water/no-water classification model that uses available Landsat data, a combination of Landsat 8 and Sentinel-2 data, and compare it with the JRC-GSWE classifications. The resulting classifications will determine how reliable the model is and if there is any additional value in including Sentinel-2 data. Because the JRC-GSWE only provides these classifications until 2021, no real-time data is available. The limitation of real-time data can cause disruptions in terms of not being fast enough to locate potential erosion and take the necessary actions such as applying groynes. A new alternative model that can make the same type of classifications can solve this issue. Therefore, this research has the following proposed research question:

What is the reliability and accuracy of a random forest water/no-water classification model using Landsat and/or Sentinel-2 data, compared to the JRC-GSWE classifications alongside riverbanks in Bangladesh?

The proposed method will use historic Landsat 7, Landsat 8 and Sentinel-2 surface reflectance collection images, including the JRC-GSWE classifications (prior to 2021). A Random Forest model will be developed using four different data options (Landsat 7, Landsat 8, Sentinel-2 and a combination of Landsat 8 and Sentinel 2) to produce water/no-water classifications. The output from the Random Forest classification model will be compared with the JRC-GSWE classifications. It will also be determined if there is an additional value in using a combination of Landsat 8 and Sentinel-2 in terms

of classification performance. A reliable model should have an overall prediction accuracy and kappa coefficient above 90%, when compared to the JRC-GSWE classifications. Additionally, both Random Forest and JRC-GSWE classifications will be compared to a validation polygon map with water/ no-water areas, to measure performances on an independent dataset. A recommendation will be made regarding the feasibility of implementing the proposed Random Forest model, utilizing Landsat and/or Sentinel-2 data, as a replacement for the JRC-GSWE classifications within the BEM.

2.

2. Data and Methods

2.1 Study Area

Bangladesh is a country in South-East Asia with a population of 168.4 million (United Nations, Department of Economic and Social Affairs, 2022). The impact of erosion can cause the loss of important agricultural- or residential areas. Bangladesh is an example where erosion has a negative impact. The country has three converging rivers (Padma, Jamuna, and Meghna), which causes an accumulation of water, especially during the monsoon season. Figure 1 shows the region of interest which are the areas around the major rivers. These regions are characterized by their susceptibility to regular inundation and the impact of water-related hazards on the social population, agriculture, infrastructure, and overall socioeconomic conditions. By making the model focus on these areas, a more accurate and timely water classification result may be generated.

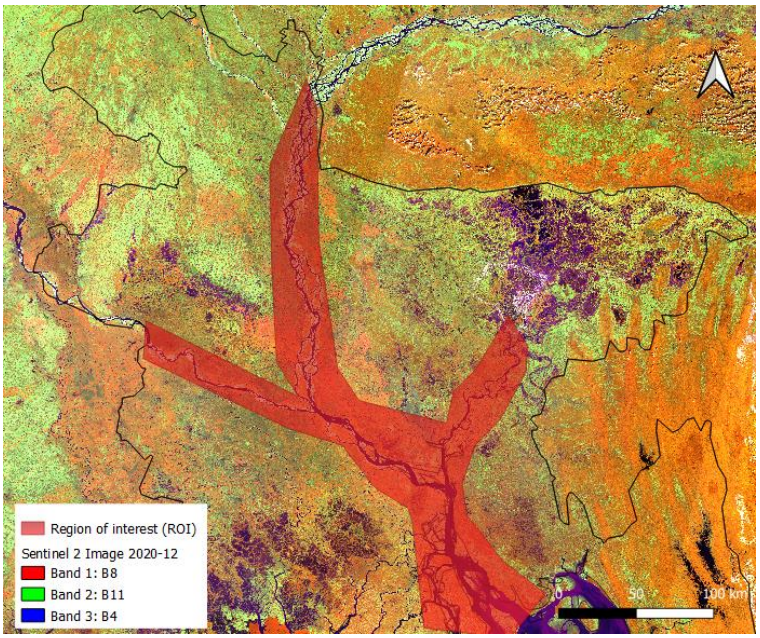


Figure 1 Region of Interest

2.2 Input Data

Before creating the Random Forest model, input data is needed to train the model. The data sources used includes the monthly Landsat 7, 8 and Sentinel-2 surface reflection collection and water classification reference data from the JRC-GSWE dataset. In the case of detecting water/no-water, the combination of bands 4-5-3 is used for Landsat 7, 5-6-4 is used for Landsat 8 and band 8-11-4 for Sentinel-2 with a scale of 30 meter. Figure 2 shows an example on how the band combinations are visualised, showing a clear distinction between water and no-water (Acharya, Dong Ha, In Tae, & Jae Kang, 2016), (Yu, Di, & Yang, 2019), (Tamouk, Lotfi, & Farmanbar, 2013), (Pekel, Cottam, Gorelick, & Belward, 2016).

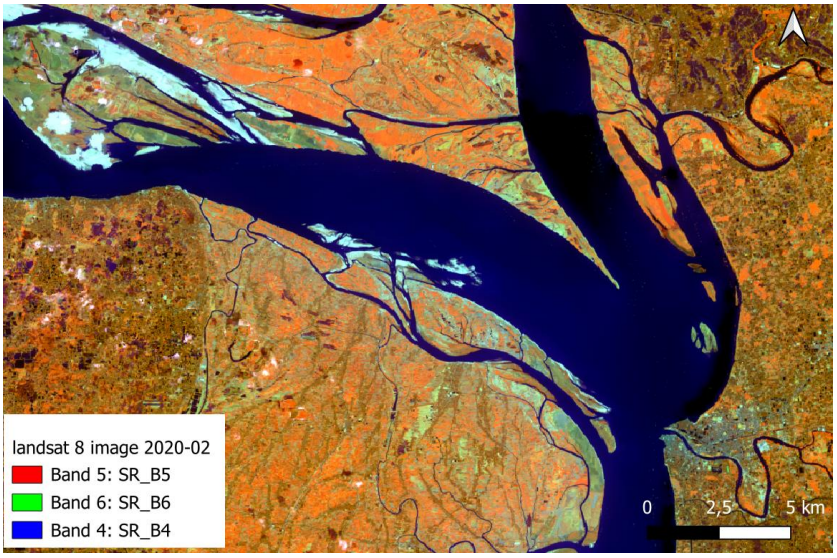


Figure 2 Landsat 8 Surface Reflectance 2020-2

The class property that needs to be predicted is the JRC-GSWE Monthly History Water classifications with three different classes named water, no-water and masked. Figure 3 shows a map with the three different classes where the blue colour is water, beige is no-water and red the masked values. The masked values represent unknown classification results caused by clouds, noise or sensor recording gaps. Because the proposed model only needs to predict water and no-water, all the masked values from the JRC-GSWE data will be excluded in the training process.

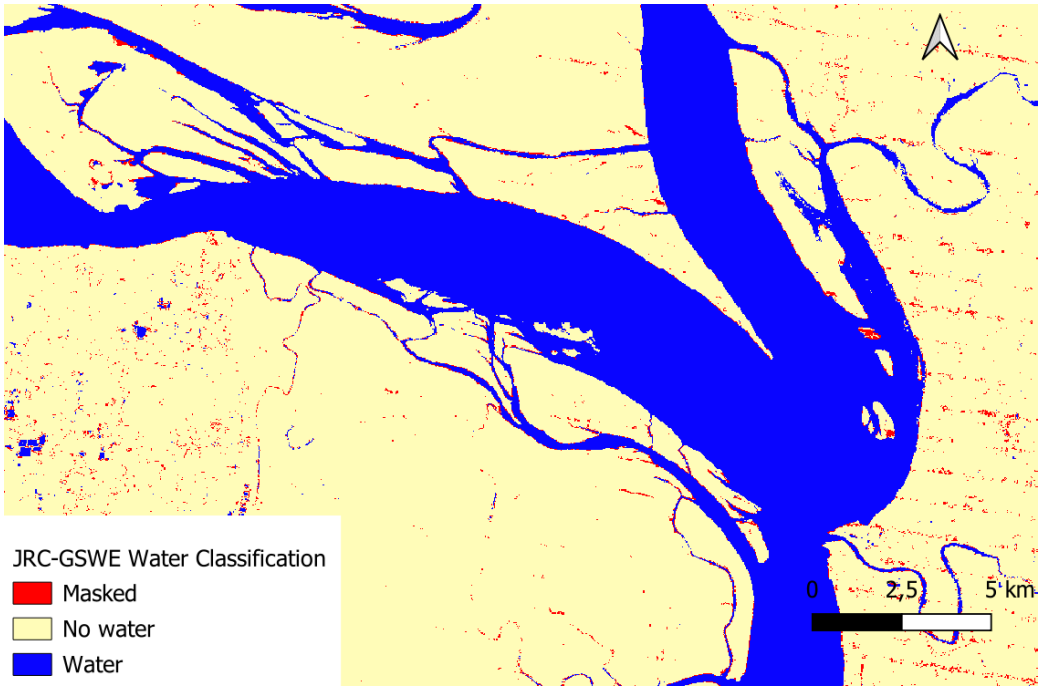


Figure 3 JRC-GSWE Water Occurrence classifications 2020-2

2.3 Pre-processing

Before the Landsat and Sentinel data can be used, the raw data needs to be pre-processed to further improve the quality. Figures 4 and 5 show an example of a raw Landsat 7 and Landsat 8 image that was recorded during the monsoon season, where you can clearly see the impact of clouds and Scan Line Corrector Failure (SLCF). Clouds can significantly degrade the quality of the Landsat and Sentinel data. Cloud masking is a technique used to remove clouds and their shadows from the input data. The process involves the detection of cloud pixels to create a cloud-free image.

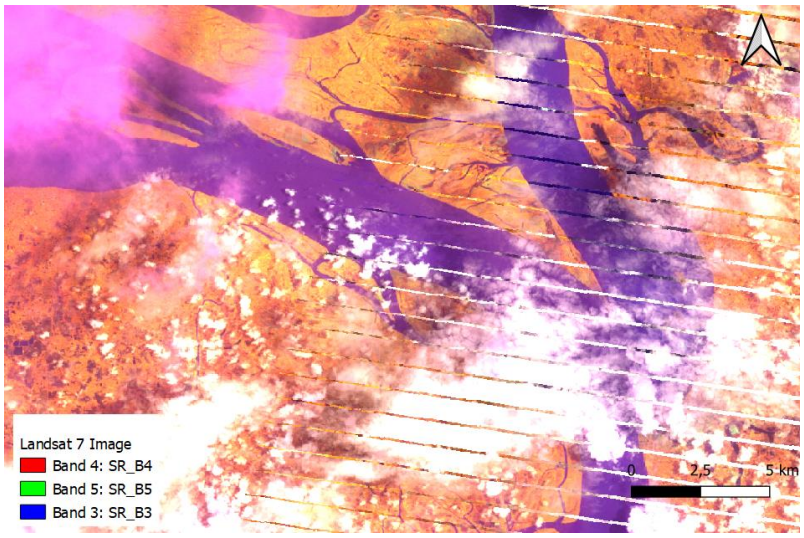


Figure 4 Raw Landsat 7 Image

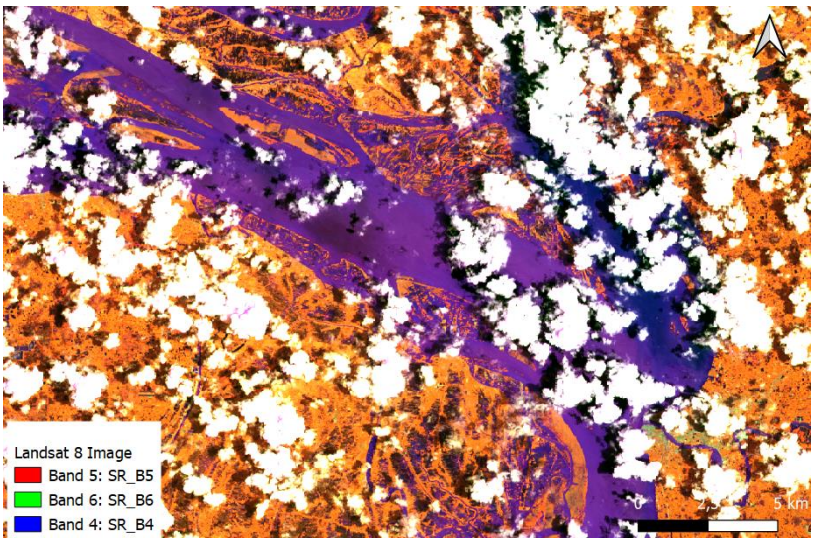


Figure 5 Raw Landsat 8 image

For Landsat and Sentinel-2 images, separate functions are needed. For Landsat 7 and 8 surface reflectance images, the QA_PIXEL band is used to identify unwanted pixels by a binary representation of '11111'. A value of 0 is assigned to these binary representations resulting masked assigned values. The band QA_RADSAT band is used to create a mask for saturated pixels, which are also set to 0. For

Sentinel-2 imagery, the quality assessment band QA60 is used to identify and remove cloud pixels. The QA60 band has a bit 10 (opaque clouds) and a bit 11 (cirrus clouds) representation which helps distinguishing between clouds and no clouds. In the case that the pixel is a cloud, bit 10 and/or bit 11 will have a value of 1, which means that there is a cloud present. The final output of these functions returns the masked image as the output. Because of the cloud masking, the image may not be complete anymore. To avoid biased results because of the number of cloud pixels in the dataset, the completeness of each test image will be included. Figure 6 and 7 shows an example of the pre-processed Landsat 7 and Landsat 8 image that were recorded during the monsoon season. Here you can see that the clouds and SLCF are masked out.

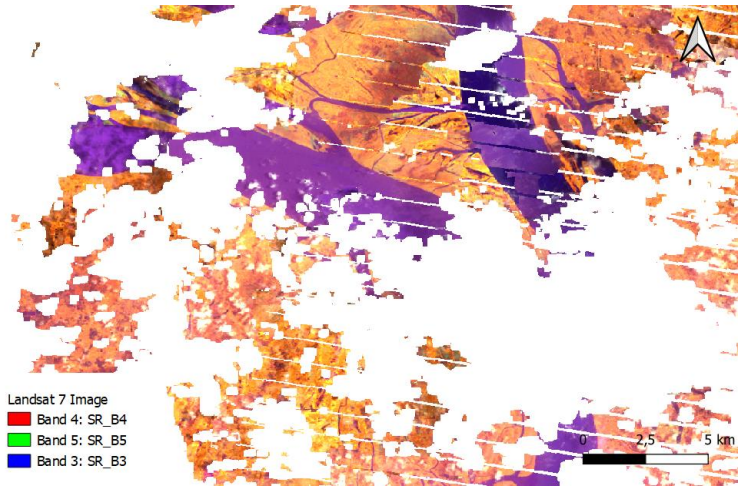


Figure 6 Landsat 7 image after pre-processing

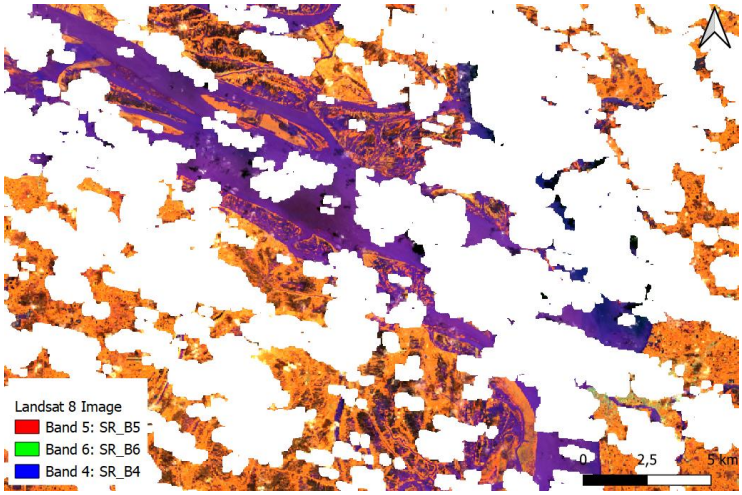


Figure 7 Landsat 8 image after pre-processing

In Appendix III, a full overview of the data completeness is shown. To further improve the quality of the Landsat images, mosaicking is applied which spatially assembles the Landsat images to produce a spatially continuous image (Google, 2021). In order to cover different types of environmental conditions throughout the year, a monthly image will be created. The final training data contains the monthly

Landsat and/or Sentinel data combined with the JRC-GSWE for the years 2018 and 2019 (24 months in total). The test data consists of Landsat and/or Sentinel data for every two months of the year 2020.

2.4Model Setup

The proposed Random Forest model will use the three band combinations from the Landsat and/or Sentinel-2 data to predict two classes, water and no-water. Because the model only needs to predict water and no-water classes, the masked class will be excluded during the training and testing process. The number of trees in a random forest model are generated in a way that attempts to split the data set at every node in half. Each tree in the forest is grown from training pixels which are randomly selected to train the random forest classification model. In Figure 8, a general architecture of a Random Forest model is presented. The Random Forest model's architecture is defined with different parameters, such as the number of trees, the allowed number of splits per tree node, the minimum number of leaf samples (output samples) required at each leaf node, and the number of levels within the tree (Segal, 2004).

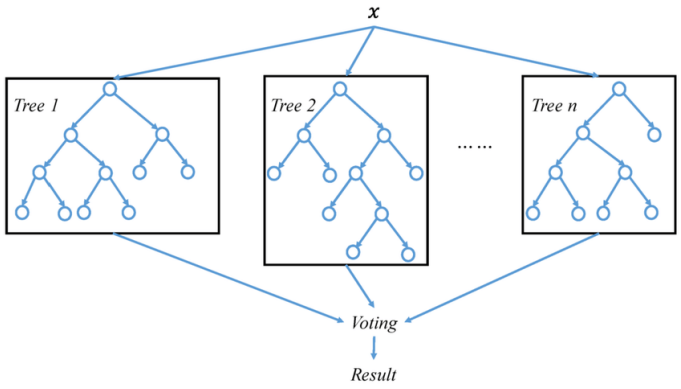


Figure 8 General architecture Random Forest model

The performance of the model is analysed using the Kappa coefficient and the overall accuracy of the model, which can be calculated using a confusing matrix. The overall accuracy can by calculated by the sum of the true positive (TP) and the true negative (TN), divided by the sum of the TP, TN, false positive (FP) and false negative (FN) (equation 1).

$$Overall\ accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Equation 2 explains the Kappa coefficient, where P_o represents the observed agreement and P_e the expected probability agreement (Kreamer, 2015). Appendix IV shows the different Random Forest architectures considered with their performance results. By comparing the performance of different architectures, where for each parameter, various values were tested, the proposed Random Forest model has 300 threes, with 3 splits per tree and a minimum leaf sample of 3.

$$k = \frac{P_0 - P_e}{1 - P_e} \quad (2)$$

2.5 Model Validation

During the training process, the monthly Landsat and/or Sentinel and JRC-GSWE data of years 2018 and 2019 are used as training data. For the test data, the year 2020 is used. The test results are generated using the Landsat data for every two months of the year 2020, so that all seasons are represented (including the monsoon season). Because the JRC-GSWE classifications also have some errors in them, and ground-truth observations are unavailable, an additional self-made map of the Landsat data is made containing 50 different polygons of water and no-water alongside riverbank areas. For the validation map, a Landsat 8 image of February 2020 is selected because of the completeness of the dataset (Appendix III table 4). Figure 9 shows an example of the self-made polygon validation map where the blue polygons represent water, and the beige polygons represent no-water. The polygons will then be converted into raster data with a scale of 30 meters to analyse spatial distribution between the Random Forest classification results, JRC-GSWE and the validation data. Four different classification results will be compared with the validation map, the JRC-GSWE classifications, the Random Forest Landsat 8 classifications, Random Forest Sentinel-2 classifications, and the Random Forest classifications which includes a combination of Landsat 8 and Sentinel-2.

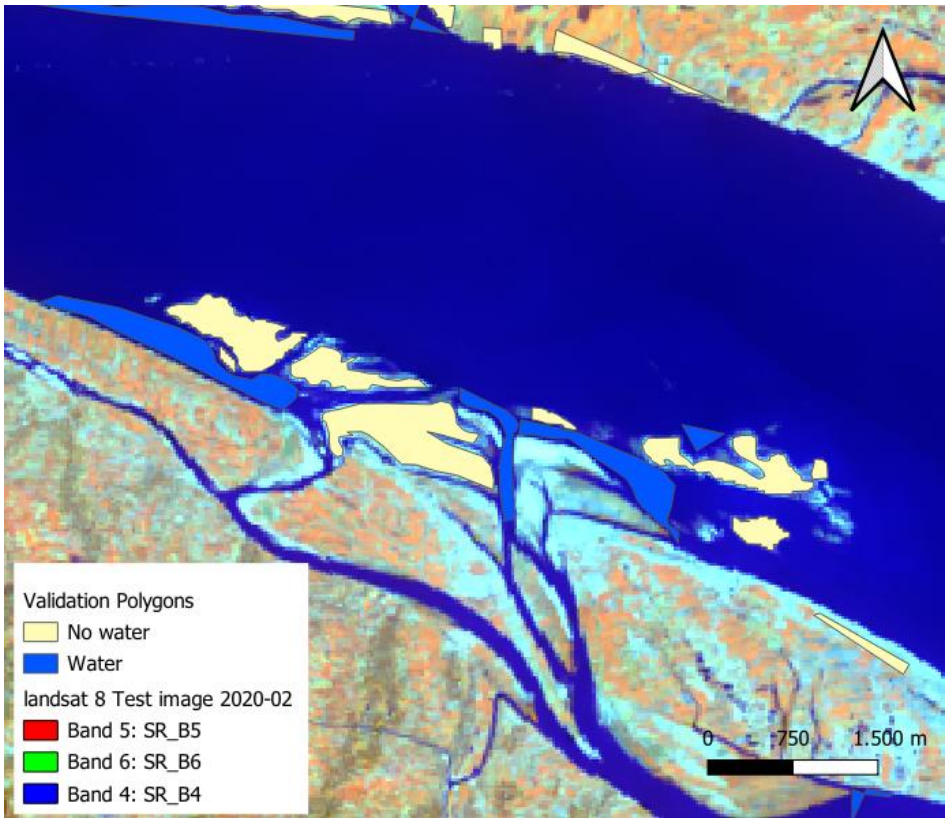


Figure 9 Example self-made polygons

3. Results

The results in this research represent the classification accuracy of a random forest model using Landsat 7, Landsat 8, and Sentinal-2 image collection. The Random Forest classification results are measured by calculating the overall accuracy and Kappa coefficient between the predicted class and actual class of the JRC-GSWE classification.

Figure 10 show the combined average test results of the year 2020. Upon a closer examination of the figures, it becomes evident that using Landsat 7 or 8 data to predict the JRC-GSWE classifications is more accurate compared to using only Sentinel-2 data.

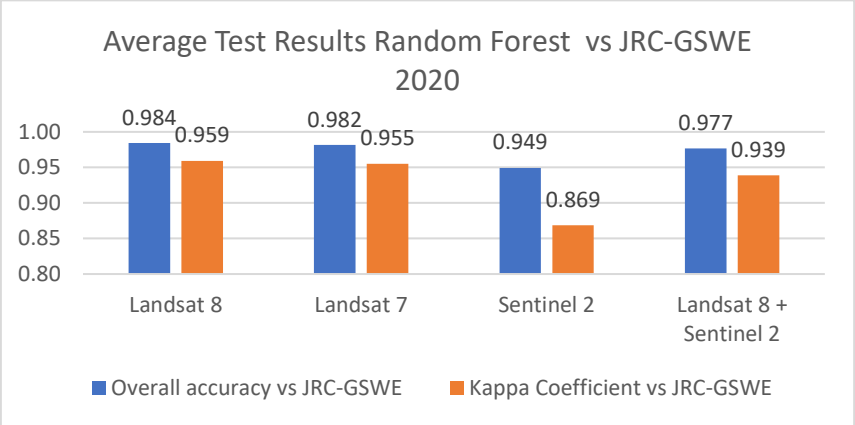


Figure 10 Average test results Random Forest vs JRC-GSWE 2020

Figures 11 and 12 show the overall test accuracy and Kappa coefficient for the Random Forest model using Landsat 7, Landsat 8, Sentinel-2 and a combination of Landsat 8 and Sentinel-2. When closer examining the monthly results, Landsat 7 shows an overall accuracy between approximately 0.97 and 0.99, with a Kappa coefficient between 0.91 and 0.99. When using Landsat 8 data, the overall accuracy is between 0.96 and 0.99, with a Kappa coefficient between 0.92 and 0.98. The use of Sentinel-2 data shows a wider range in terms of accuracy, where the overall accuracy ranges between 0.88 and 0.97, with a Kappa coefficient between 0.75 and 0.94. When combining Landsat 8 and Sentinel-2, the overall accuracy ranges between 0.96 and 0.99, with a Kappa coefficient between 0.91 and 0.95. In Appendix V, a full overview of the Random Forest test classification accuracy and Kappa coefficient results for every two months.

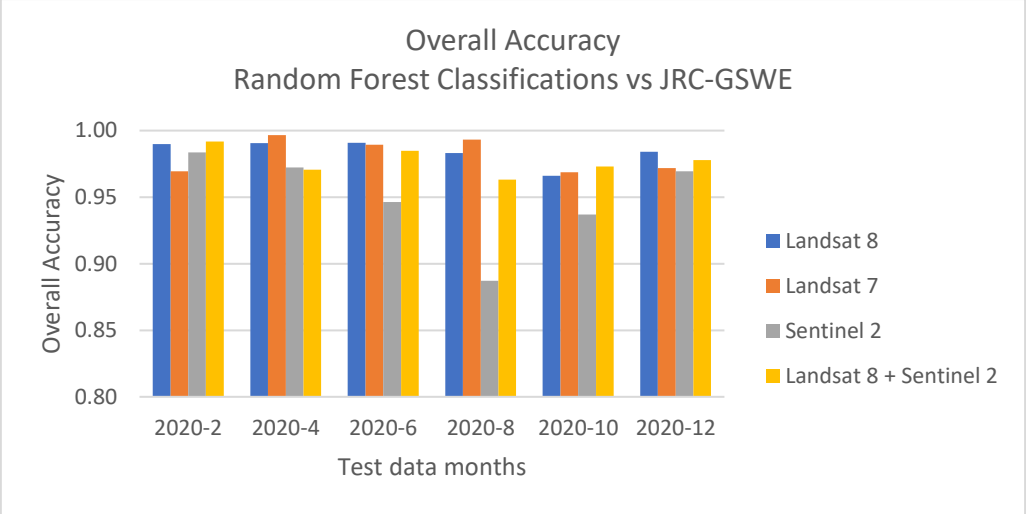


Figure 11 Overall accuracy Random Forest classifications vs JRC-GSWE

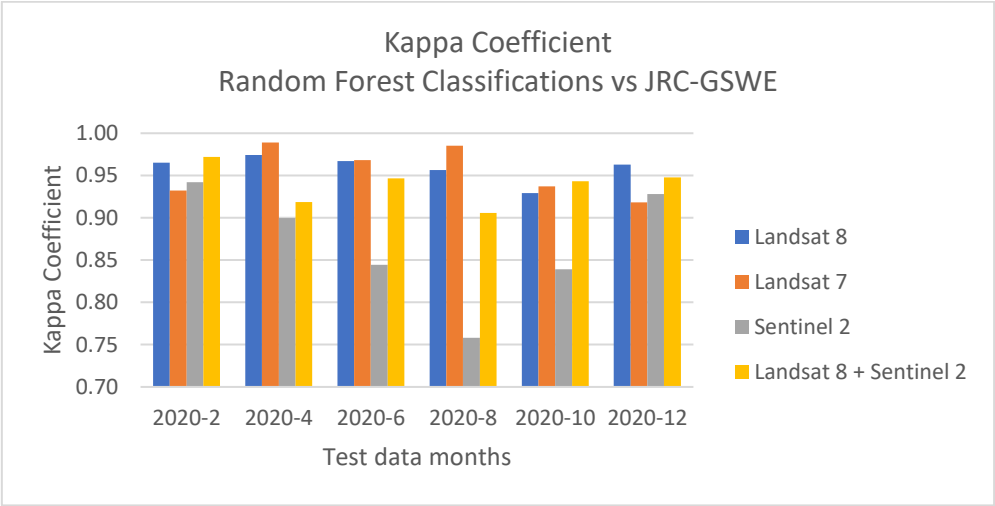


Figure 12 Kappa coefficient Random Forest classifications vs JRC-GSWE

Figures 13-16 show a visual representation of the results, where the red areas represent a no-water classification from the Random Forest model, while the JRC-GSWE classification shows water. The black colour represents a water classification from the Random Forest model, where the JRC-GSWE classification shows no-water. Comparing figures 13 and 14 demonstrate that the Random Forest model using Landsat 7 data has difficulties in classifying no-water pixels alongside the riverbanks, whereas the use of Landsat 8 shows difficulties in classifying water pixels.

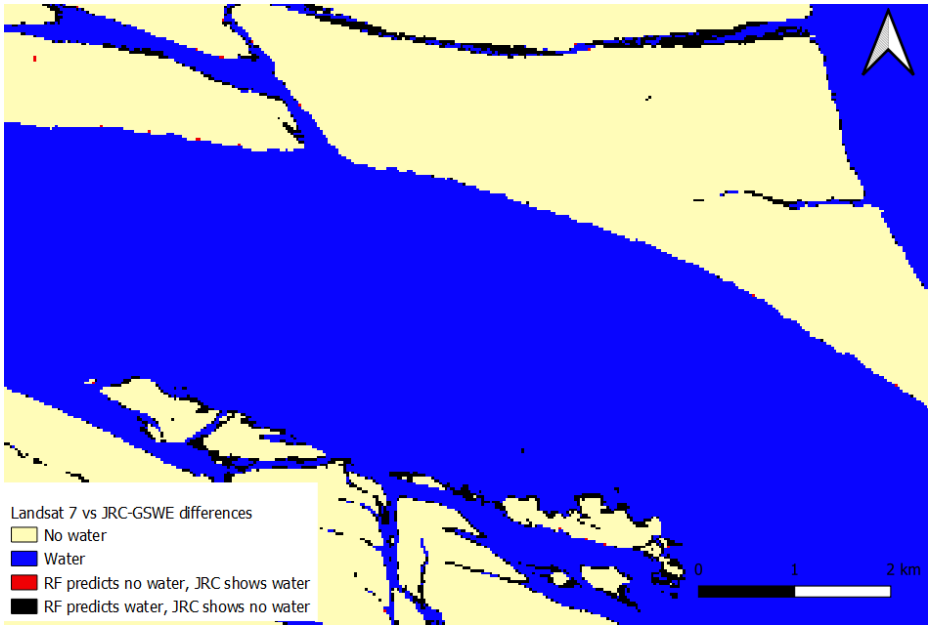


Figure 13 Landsat 7 example water/no-water classification results vs JRC-GSWE

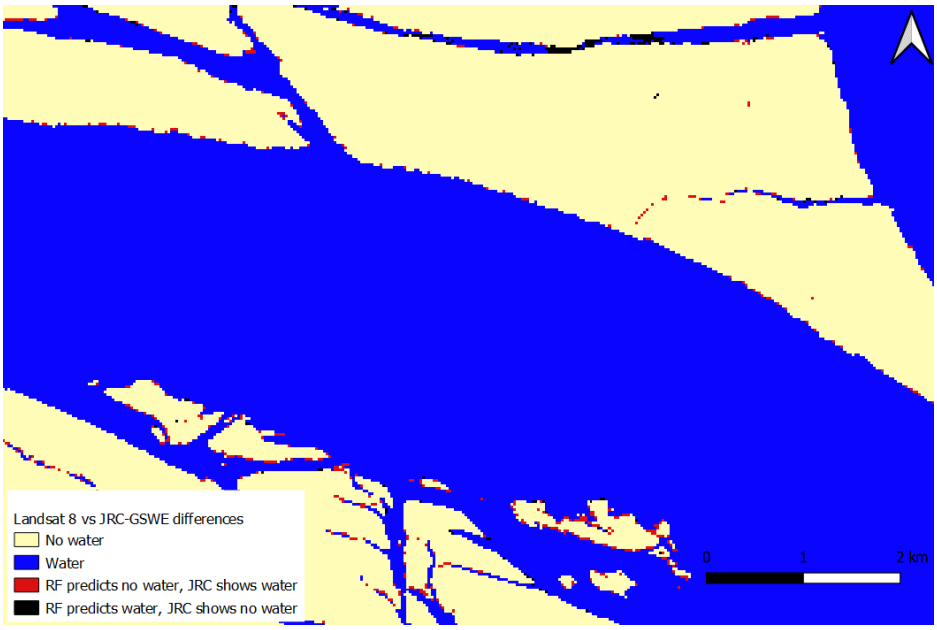


Figure 14 Landsat 8 example water/no-water classification results vs JRC-GSWE

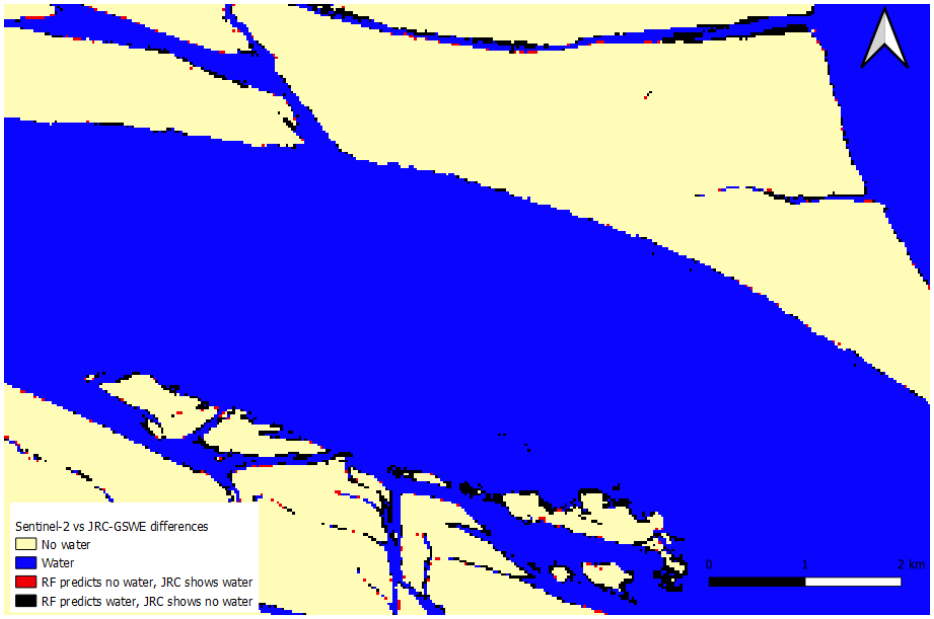


Figure 15 Sentinel-2 example water/no-water classification results vs JRC-GSWE

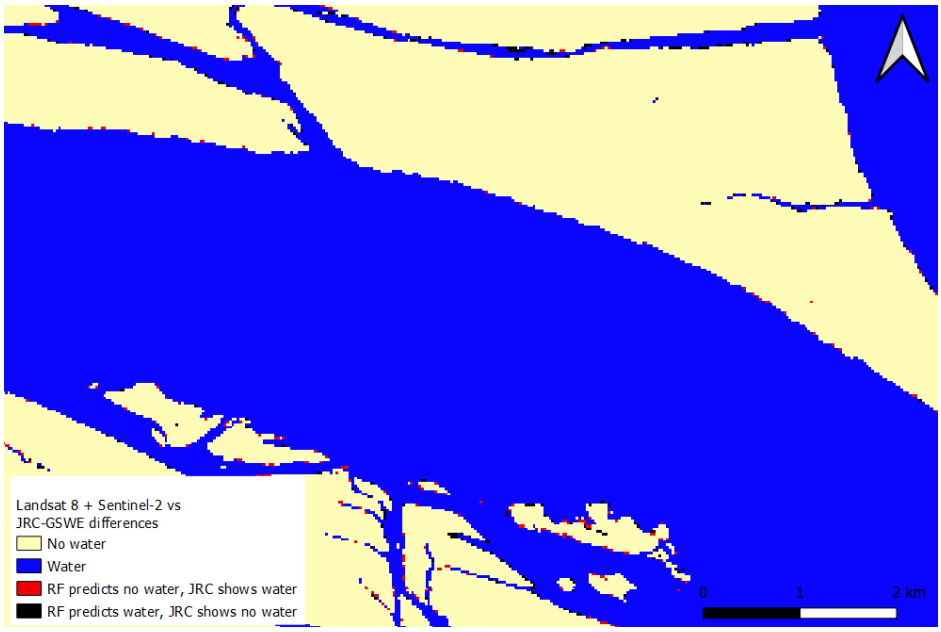


Figure 16 Landsat 8 + Sentinel-2 example water/no-water classification results vs JRC-GSWE

As mentioned in the previous section, the JRC-GSWE classifications also may have some errors, therefore a self-made polygon map was made of the month February 2020, which has the least amount of cloud cover. Figure 17 show the accuracy results between the Random Forest model (using Landsat 8, Sentinel-2, and a combination both), the JRC-GSWE classifications, compared to the self-made validation data. The results show an overall accuracy between 0.85 and 0.90.

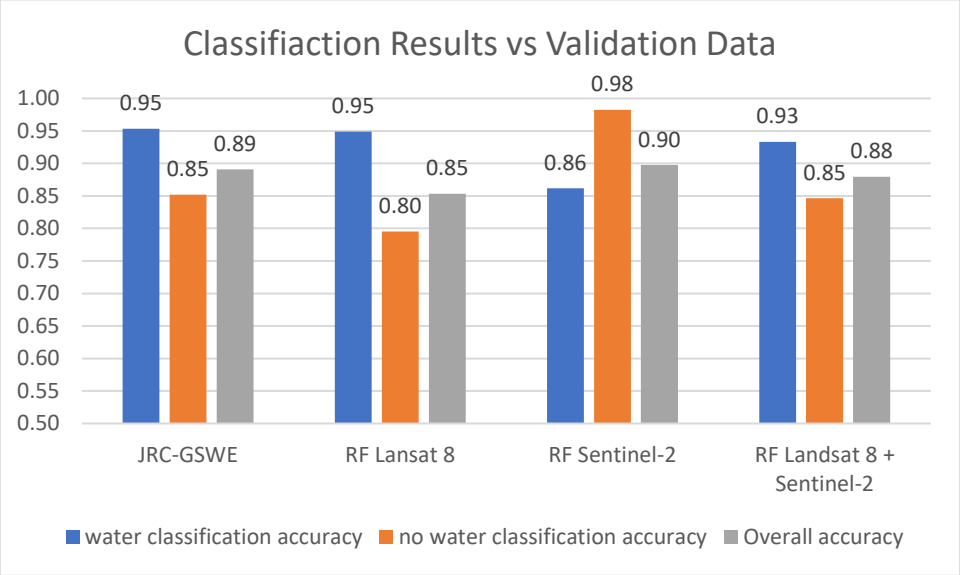


Figure 17 Classification results vs validation data

Figures 18-21 show a visual representation of the spatial distribution between the validation map and the Random Forest classification results when different data is used. Misclassifications of no-water pixels are marked in red, whereas misclassifications of water pixels are marked in black. By analysing the results, the Random Forest model using Sentinel-2 data shows to have the highest overall accuracy. JRC-GSWE shows to have the highest accuracy in detecting water whereas Sentinel-2 shows to be the least accurate. For no-water classifications, Sentinel-2 shows to have the highest accuracy whereas Landsat 8 shows to be the least accurate.

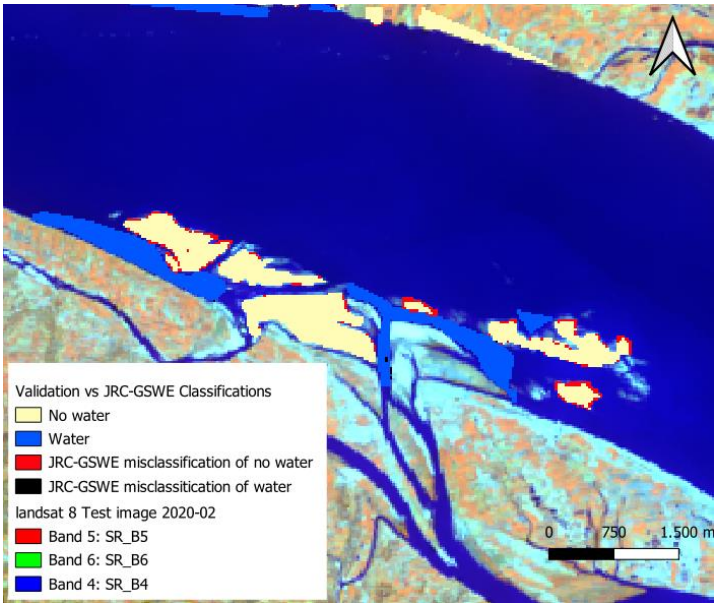


Figure 19 Validation data differences compared to the JRC-GSWE classifications.

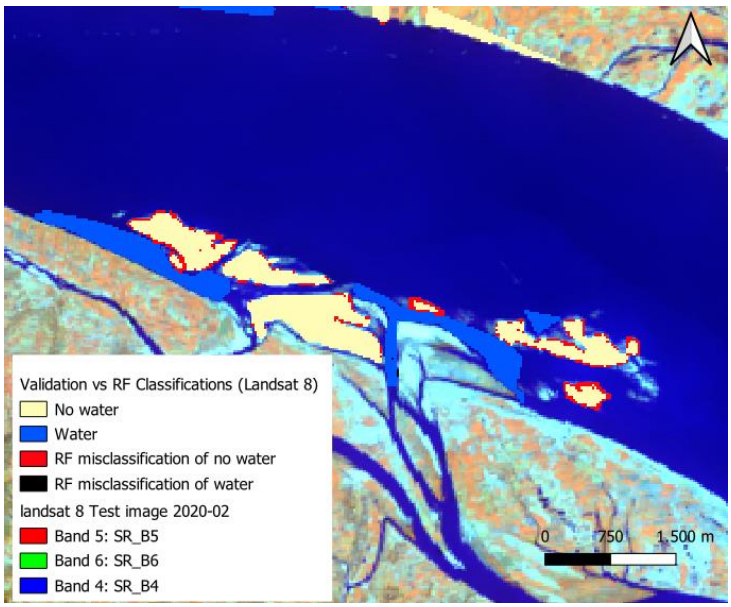


Figure 18 Validation data differences compared to the Random Forest classifications (Landsat 8)

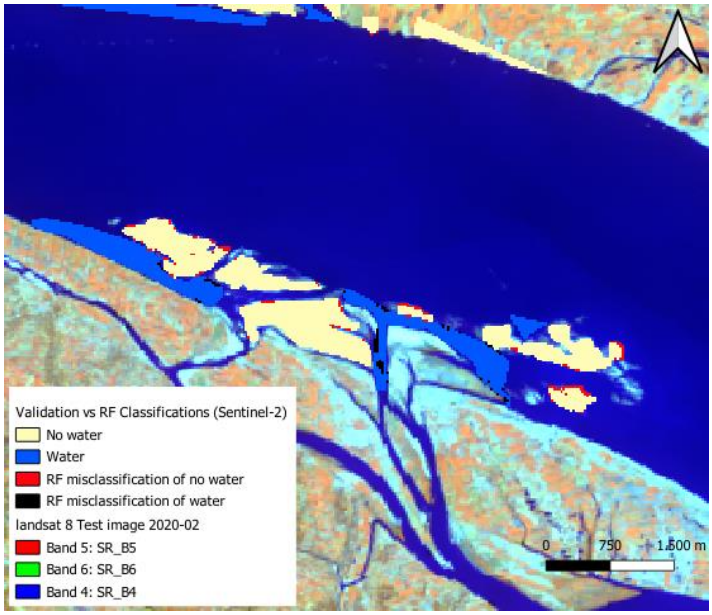


Figure 20 Validation data differences compared to the Random Forest classifications (Sentinel-2)

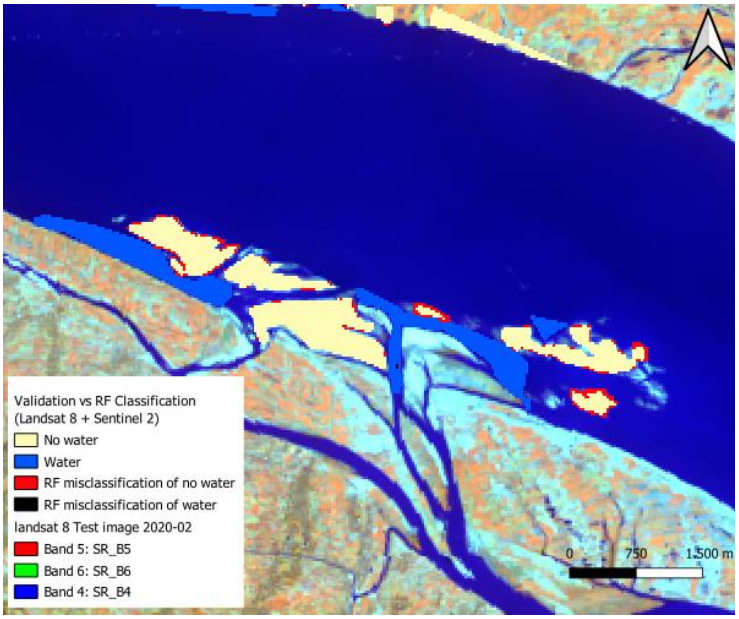


Figure 21 Validation data differences compared to the Random Forest classifications (Landsat 8 + Sentinel-2)

4. Discussion

One of the objectives of this study was to develop a random forest water/no-water classification model using Landsat and/or Sentinel-2 data and compare it with the JRC-GSWE classifications. The results of the Random Forest classification model show a reliable accuracy and Kappa coefficient when compared to the JRC-GSWE classifications when using Landsat data. This means that the Random Forest model can be used as an alternative to the current JRC-GSWE classifications. The reason that Landsat 7 outperforms the Landsat 8 model is because the JRC-GSWE classifications are performed using Landsat 7 data (European Commision Joint Research Centre, 2016, Pekel, Cottam, Gorelick, & Belward, 2016). From the perspective of using remote sensing data, such as Landsat and Sentinel-2, in order to make water/no-water classifications, the results are consistent with the literature that this type of data shows to be effective when applied to a Random Forest model (Nguyen, Doan, & Radeloff, 2018, Breiman, 2001, Byoung, Kim, & Nam, 2015, Ko, Kim, & Nam, 2011). The use of Landsat 7 and Landsat 8 data shows to have a reliable accuracy in predicting the JRC-GSWE classification test data, where the use of only Sentinel-2 data is less reliable.

Because the potential number of errors in the JRC-GSWE classifications, another objective was to compare the Random Forest classification results and the JRC-GSWE classifications with the validation polygons. Both JRC-GSWE classifications and the random forest classifications were able to correctly classify most of the validation data, with an accuracy between 0.85 and 0.9. When analysing the validation results of the Random Forest model and JRC-GSWE classifications, the use of Sentinel-2 shows a promising effect where it shows the highest overall classification accuracy. Nevertheless, the validation results also show that the use of Sentinel-2 data shows to be less accurate when detecting water compared to the Landsat- and JRC-GSWE classifications. On the other hand, JRC-GSWE and Landsat 8 show to be less reliable when detecting no-water compared to Sentinel-2. Selecting a method that only shows high accuracy for a single class (e.g., high classification accuracy for water only) could cause biased results. This suggests that the combination of Landsat 8 with Sentinel-2 offers a reliable method in for both water/ no-water classifications. This aligns with previous studies on using a dual sensor approach, further validating the suitability of the proposed dual-sensor approach. (Tottrup, et al., 2022) (Carrasco, O'Neil, & Morton, 2019).

The findings made in this research support the aim of the research question, demonstrating that the Random Forest model using Landsat and/or Sentinel-2 can provide as a reliable replacement for the JRC-GSWE classifications produce and near-real-time water/no-water classifications in Bangladesh. The model's performance for near-real-time classifications can assist the BEM in identifying vulnerable areas alongside riverbanks and timely implementing of erosion mitigation efforts, which can protect valuable agricultural and residential land.

It is important to acknowledge the limitations and uncertainties of this research. Although the classifications during the monsoon season show a high accuracy, it must be considered part of the data is unknown due to cloud cover. This shows to be evident when analysing the test results of then month 2020-8, where using Landsat show to perform better compared to Sentinel-2, but has significantly less complete (Landsat 7 is 47% complete and Landsat 8 is 16% complete) compared to Sentinel-2 (92% complete). Another limitation is the potential errors within the JRC-GSWE classification data. While efforts were made to ensure the preciseness of the validation data, the absence of having more accurate ground-truth validation data on precise locations alongside riverbanks might lead to misclassification in certain areas that are not captured precise enough.

5. Conclusion

This research aimed to develop a Random Forest water/no-water classification model using Landsat or Sentinel-2 data. Additionally, the study explored the potential of using a dual-sensor option where Landsat 8 and Sentinel-2 data are combined to make more accurate classifications. The classification results were compared with the existing JRC-GSWE classifications and a self-made validation polygon map to determine the proposed model is reliable enough as a replacement to assist limiting the impact.

The results indicate that a Random Forest classification model using Landsat 8 data shows the most reliable overall accuracy and Kappa coefficient when predicting the JRC-GSWE classifications. This outperforms the use of Landsat 7, Sentinel-2, and a combination of Landsat 8 and Sentinel-2. When comparing the model's results with the self-made validation map, the use of Sentinel-2 data produced the highest overall accuracy due to its high-performance classifying no-water polygons. However, the use of Landsat 8 demonstrated to be more accurate when classifying water polygons. Combining Landsat 8 and Sentinel-2 data, the model produced reliable outcomes, offering a balanced performance in classifying both water and no-water polygons.

The research findings support the research objective, showcasing that the proposed Random Forest model using Landsat 8 and Sentinel-2 data can serve as a reliable alternative to the JRC-GSWE data for water/no-water classifications in Bangladesh. Near-real-time classifications hold significant potential in assisting the identification of vulnerable areas alongside riverbanks, enabling the timely implementation of erosion mitigation measures to protect vulnerable agricultural and residential land.

It should be mentioned that the absence of precise ground-truth validation data alongside riverbanks may lead to some misclassifications. Further research could focus on obtaining more accurate ground-truth data to refine the model's accuracy.

In conclusion, the Random Forest classification model using a combination of Landsat 8 and Sentiniel-2 data is reliable enough for making water/no-water classifications. The model's reliability and capability of using near-real-time Landsat 8 and Sentinel-2 data can assist in identifying vulnerable areas and support water management decision making, which can result in timely erosion mitigation efforts.

6. References

Acharya, T. D., Dong Ha, L., In Tae, Y., & Jae Kang, L. (2016). Identification of Water Bodies in a Landsat 8 OLI Image Using a J48 Decision Tree. *Sensors 2016*, 1075.

Allaby, M. (2008). *A Dictionary of Earth Sciences.* New York: Oxford University Press.

Basten, K. (2016). *Classifying Landsat Terrain Images via Random Forests.* Nijmegen: Radboud University.

Breiman, L. (2001). Random Forests. *Machine Learning 45*, 5-32.

Byoung, C. K., Kim, H. H., & Nam, J. Y. (2015). Classification of Potential Water Bodies Using Landsat 8 OLI and a Combination of Two Boosted Random Forest Classifiers. *Sensors 2015*, 13763-13777.

Carrasco, L., O'Neil, A., & Morton, D. (2019). Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2 and Landsat 8 for Land Cover Mapping with Google Earth Engine. *Remote Sens. 2019*, 288.

Deltares. (2023). *Bangladesh Erosion Monitor*. Retrieved from arjenhaag.users.earthengine.app

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., & Hoersch, B. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 25–36.

European Commision Joint Research Centre. (2016). *Global Surface Water - FAQ*. Retrieved from Global Surface Water Explorer: https://global-surface-water.appspot.com/faq

European Commision: Global Surface Water. (2020). *Global Surface Water - FAQ*. Retrieved from global-surface-water.appspot.com: https://global-surface-water.appspot.com/faq

European Space Agency. (2023). *SENTINEL-2 MISSION GUIDE*. Retrieved from Sentinel.esa.int: https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Google. (2021, May 27). *Compositing, Masking, and Mosaicking*. Retrieved from Google Earth Engine: https://developers.google.com/earth-engine/tutorials/tutorial_api_05

Kartikeyan, B., Majumder, K., & Dasgupta, A. (1996). An expert system for land cover classification. *Computing & Control Engineering Journal*, 58 - 66.

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. *Lecture Notes in Computer Science*, 76-90.

Ko, B. C., Kim, S. H., & Nam, J.-Y. (2011). X-ray Image Classification Using Random Forests with Local Wavelet-Based CS-Local Binary Patterns. *Journal of Digital Imaging volume 24*, 1141-1151.

Kreamer, H. C. (2015). Kappa Coefficient. *Kraemer, H. C. (2015). Kappa Coefficient. Wiley StatsRef: Statistics Reference Online, 1–4. https://doi.org/10.1002/9781118445112.stat00365.pub2*, 1-4. doi:https://doi.org/10.1002/9781118445112.stat00365.pub2

Loveland, T. R., & Irons, R. J. (2016). Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 1-6.

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 823-870.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., & Jenstrom, D. T. (2020). Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 111968.

Munna, N. H. (2018). *5 causes and 7 effects of riverbank erosion in Bangladesh.* Barishal, Bangladesh: EarthReview.

NASA. (2012, July 23). *40 Years of Earth Observation: From the Beginning*. Retrieved from landsat.gsfc.nasa.gov: https://landsat.gsfc.nasa.gov/article/40-years-of-earth-observation-from-the-beginning/

Nguyen, H., Doan, T., & Radeloff, V. (2018). APPLYING RANDOM FOREST CLASSIFICATION TO MAP LAND USE/LAND COVER USING LANDSAT 8 OLI. *GeoInformation For Disaster Management (Gi4DM)*.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature 540*, 418-422.

Segal, M. R. (2004). *Machine Learning Benchmarks and Random Forest Regression.* San Francisco, CA, USA: University of California.

Shoshany, M. (2008). Knowledge based expert systems in remote sensing task: quantifying gains from intelligent inference. *Int. Soc. Photogramm. Remote Sens. Arch., 37,* , 1085-1088.

Smith, A. R. (1978). Color gamut transform pairs. *ACM SIGGRAPH Computer GraphicsVolume 12Issue 3*, 12-19.

Tamouk, J., Lotfi, N., & Farmanbar, M. (2013). *SATELLITE IMAGE CLASSIFICATION METHODS and LANDSAT 5TM BANDS.*

Taufik, A., Syed Ahmad, S., & Ahmad, A. (2016). Classification of Landsat 8 Satellite Data Using NDVI Tresholds. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 37-40.

Tottrup, C., Druce, D., Meyer, R. P., Christensen, M., Riffler, M., Dulleck, B., . . . Paganini, M. (2022). Surface Water Dynamics from Space: A Round Robin Intercomparison of Using Optical and SAR High-Resolution Satellite Observations for Regional Surface Water Detection. *Remote Sens. 2022*, 2410.

U.S. Geological Survey. (2015, April 21). *Landsat Surface Reflectence*. Retrieved from U.S. Geological Survey: https://www.scribbr.com/citation/generator/folders/30RREphPeaq6TGsufHkxvm/lists/6tg3KYhu5NwSuUYelvzy6/?lastAddedSourceId=45a54085-7a95-4676-b852-702ba2942058

United Nations, Department of Economic and Social Affairs. (2022). *World Population Prospects 2022.*

United States Geological Survey. (2019, February 1). *What are the band designations for the Landsat satellites?* Retrieved from USGS Science for a changing world: https://www.usgs.gov/faqs/what-are-band-designations-landsat-satellites

United States Geological Survey. (2020, September 17). *Using the USGS Landsat Level-1 Data Product*. Retrieved from USGS Science for a changing world: https://www.usgs.gov/landsat-missions/using-usgs-landsat-level-1-data-product

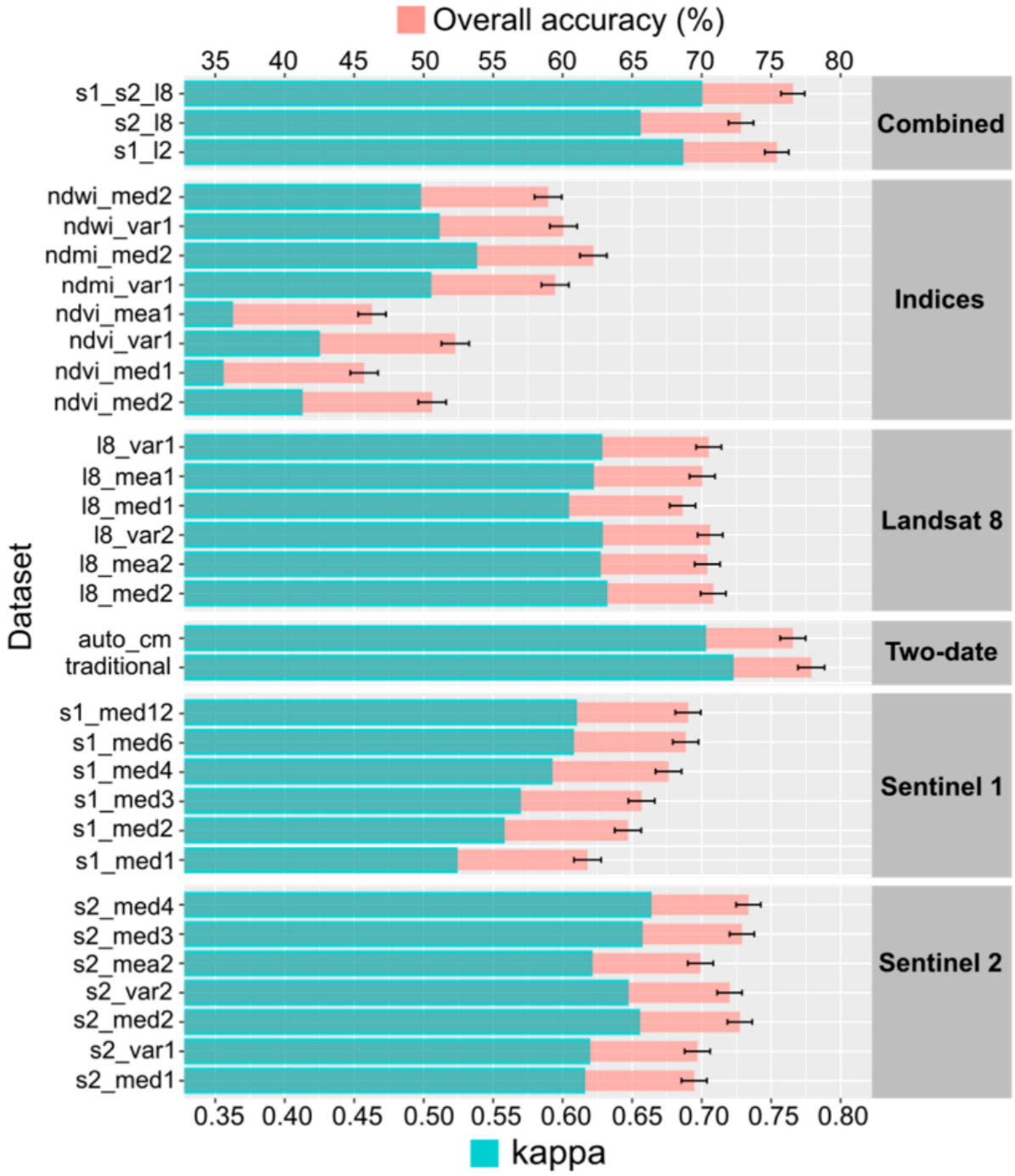
United States Geological Survey. (2023, Januari 1). *Using the USGS Landsat Level-1 Data Product*. Retrieved from USGS Science for a changing world: https://www.usgs.gov/landsat-missions/using-usgs-landsat-level-1-data-product

Van Tho, N. (2020). Coastal erosion, river bank erosion and landslides in the Mekong Delta: Causes, effects and solutions. *Geotechnics for Sustainable Infrastructure Development* , 957–962.

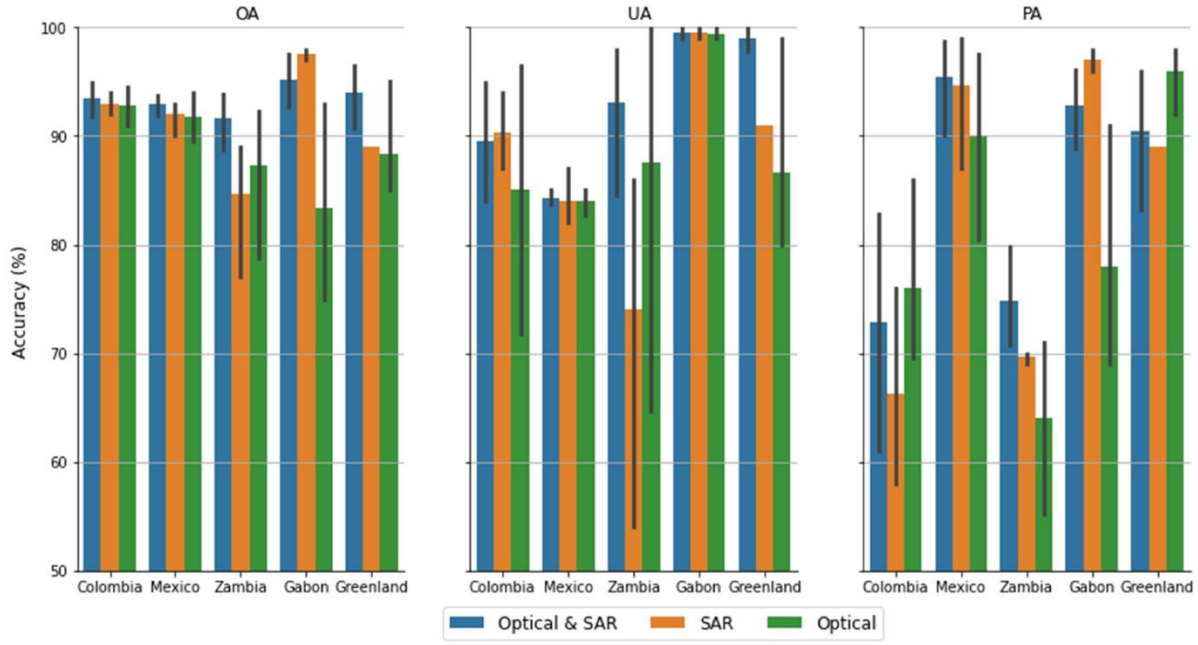
Yang, J.-B., & Xu, D.-L. (2002). On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 32, no. 3,*, 289-304.

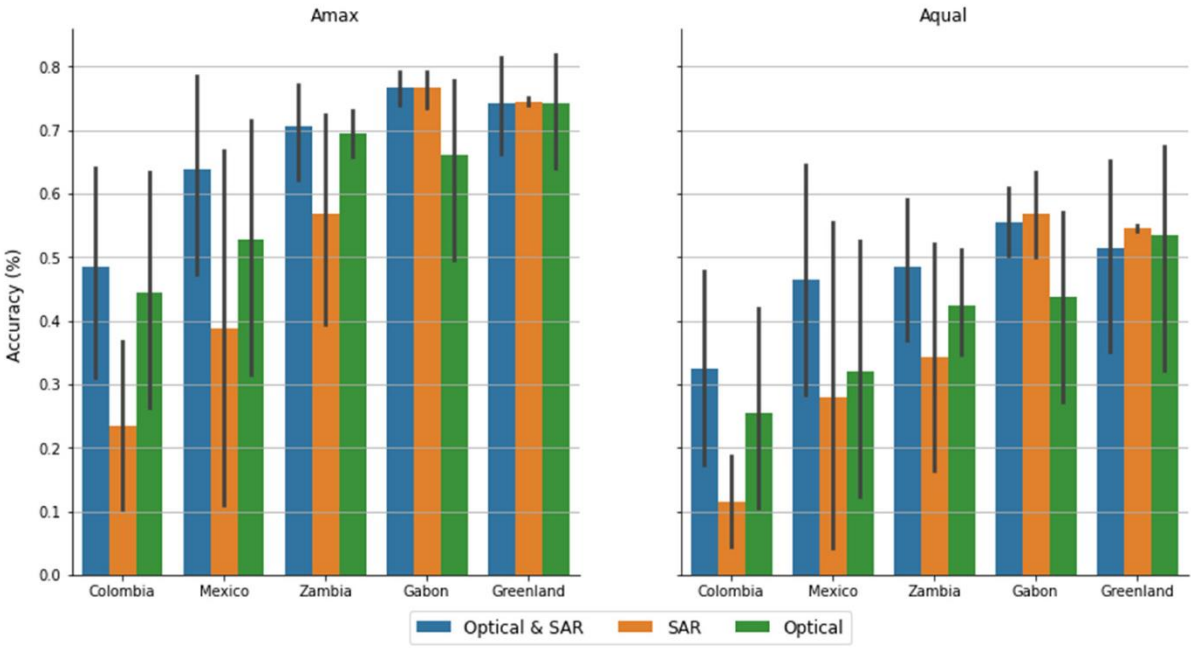
Yu, Z., Di, L., & Yang, R. (2019). Selection of Landsat 8 OLI Band Combinations for Land Use and Land Cover Classification. *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*.

Appendix I: Classification accuracies Sentinel-1, Sentinel-2, and Landsat 8 for Land Cover Mapping (Carrasco, O'Neil, & Morton, 2019)



Appendix II: Accuracy statistics from the WorldWater round robin test sites (Tottrup, et al., 2022)





Appendix III: Data completeness

Table 3 Training Data completeness

Image Completeness training data	Landsat 7	Landsat 7 Percentage	Landsat 8	Landsat 8 Percentage	Sentinel 2	Sentinel 2 Percentage
2018-1	8308031	0.883	9408074	1.000	0	0.000
2018-2	1925451	0.205	9385681	0.997	0	0.000
2018-3	8703619	0.925	8555457	0.909	0	0.000
2018-4	8306280	0.883	7327983	0.779	0	0.000
2018-5	2112280	0.224	9161532	0.974	0	0.000
2018-6	3843563	0.408	3695965	0.393	0	0.000
2018-7	2595673	0.276	749555	0.080	0	0.000
2018-8	44625	0.005	4051582	0.431	0	0.000
2018-9	3999711	0.425	8470462	0.900	0	0.000
2018-10	6214695	0.660	9289870	0.987	0	0.000
2018-11	9134952	0.971	9410254	1.000	0	0.000
2018-12	7690708	0.817	9394620	0.998	9410563	1.000
2019-1	8199545	0.871	9397562	0.999	9410563	1.000
2019-2	403518	0.043	9375868	0.996	9410563	1.000
2019-3	8072673	0.858	9273838	0.985	9410563	1.000
2019-4	7891117	0.839	7069959	0.751	9410563	1.000
2019-5	8093648	0.860	7214868	0.767	9410563	1.000
2019-6	4163871	0.442	1559213	0.166	9266244	0.985
2019-7	396230	0.042	705842	0.075	8888742	0.945
2019-8	6162712	0.655	5184122	0.551	8491666	0.902
2019-9	330856	0.035	5562870	0.591	9078127	0.965
2019-10	8278988	0.880	7295342	0.775	9410563	1.000
2019-11	8204036	0.872	9290508	0.987	9410563	1.000
2019-12	6598659	0.701	8432311	0.896	9410563	1.000

Table 4 Test data completeness

Image Completeness test data	Landsat 7	Landsat 7 Percentage	Landsat 8	Landsat 8 Percentage	Sentinel 2	Sentinel 2 Percentage
2020-12	8597992	0.91	9410441	1.00	9410563	1.00
2020-2	8584725	0.91	9410549	1.00	9410563	1.00
2020-4	8102217	0.86	1498608	0.16	9410563	1.00
2020-6	1033498	0.11	5981678	0.64	7544823	0.80
2020-8	4459580	0.47	1538539	0.16	8699204	0.92
2020-10	7495098	0.80	5922742	0.63	9410563	1.00

Appendix IV: Random Forest Parameter Tuning

Table 5 Random Forest Parameter Tuning

RF	nTrees	variablesPerSplit	minLeafPopulation	Kappa Coefficient
	100	1	1	0.95937
	100	1	2	0.95258
	100	1	3	0.95702
	100	2	1	0.96157
	100	2	2	0.95775
	100	2	3	0.96228
	100	3	1	0.96305
	100	3	3	0.96297
	100	3	2	0.96222
	300	1	1	0.96311
	300	1	2	0.95557
	300	1	3	0.95780
	300	2	1	0.96307
	300	2	2	0.96151
	300	2	3	0.96155
	300	3	1	0.96455
	300	3	2	0.96372
	300	**3**	**3**	**0.96522**
	500	1	1	0.95937
	500	1	2	0.95481
	500	1	3	0.95709
	500	2	1	0.96232
	500	2	2	0.96151
	500	2	3	0.96228
	500	3	1	0.96455
	500	3	2	0.96445
	500	3	3	0.96299
	1000	1	1	0.95941
	1000	1	2	0.95630
	1000	1	3	0.95707
	1000	2	1	0.96382
	1000	2	2	0.96001
	1000	2	3	0.96153
	1000	3	1	0.96087
	1000	3	2	0.96220
	1000	3	3	0.96301

Appendix V: Random Forest Classification Accuracy Results

Table 6 Random Forest Classification Accuracy Results vs JRC-GSWE

Landsat 8	**Overall accuracy vs JRC-GSWE**	**Kappa Coefficient vs JRC-GSWE**
2020-2	0.9898	0.9651
2020-4	0.9907	0.9743
2020-6	0.9909	0.9669
2020-8	0.9832	0.9564
2020-10	0.9661	0.9294
2020-12	0.9841	0.9631
Landsat 7	**Overall accuracy vs JRC-GSWE**	**Kappa Coefficient vs JRC-GSWE**
2020-2	0.9695	0.9324
2020-4	0.9967	0.9889
2020-6	0.9893	0.9682
2020-8	0.9932	0.9853
2020-10	0.9688	0.9370
2020-12	0.9718	0.9181
Sentinel 2	**Overall accuracy vs JRC-GSWE**	**Kappa Coefficient vs JRC-GSWE**
2020-2	0.9835	0.9422
2020-4	0.9723	0.8998
2020-6	0.9462	0.8443
2020-8	0.8871	0.7582
2020-10	0.9369	0.8393
2020-12	0.9695	0.9280
Landsat 8 + Sentinel 2	**Overall accuracy vs JRC-GSWE**	**Kappa Coefficient vs JRC-GSWE**
2020-2	0.9919	0.9720
2020-4	0.9707	0.9187
2020-6	0.9849	0.9468
2020-8	0.9631	0.9056
2020-10	0.9729	0.9433
2020-12	0.9777	0.9479

Table 7Random Forest overall average Results

Overall average	**accuracy vs JRC-GSWE**	**Kappa Coefficient vs JRC-GSWE**
Landsat 7	0.9816	0.9550
Landsat 8	0.9841	0.9592
Sentinel 2	0.9493	0.8686
Landsat 8 + Sentinel 2	0.9769	0.9390

Table 8 Polygon validation map accuracy results

Polygon Validation Map	water	no-water	Overall accuracy
JRC-GSWE	0.9532	0.8521	0.8910
Random Forest Landsat 8	0.9490	0.7953	0.8532
Random Forest Sentinel-2	0.8620	0.9825	0.8976
Random Forest Landsat 8 + Sentinel-2	0.9334	0.8465	0.8793