

JamUNet: predicting the morphological changes of braided sand-bed rivers with deep learning

Antonio Magherini

JamUNet: predicting the morphological changes of braided sand-bed rivers with deep learning

by

Antonio Magherini

in partial fulfillment of the requirements for the degree of

Master of Science in Civil Engineering

at Delft University of Technology, Faculty of Civil Engineering and Geosciences

to be defended publicly on Thursday October 31st 2024 at 09:30 AM in CEG Lecture Hall E.

Committee composition:

Chairman:	Dr. Riccardo Taormina,	TU Delft
Main supervisor:	Dr. ir. Erik Mosselman,	TU Delft & Deltares
Supervisor:	Dr. Víctor Chavarrías,	Deltares

Cover: Landsat 5 satellite image retrieved from Google Earth Engine. It represents the Brahmaputra-Jamuna River at the border between India and Bangladesh on January 25, 1994.

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

The code repository is available at <https://github.com/antoniomagherini>.



*Ai miei genitori, a mio fratello
e soprattutto alle mie nonne*

Abstract

Braided rivers are among the most dynamic natural Earth systems, with a rapid and complex morphological evolution. Limited understanding and inadequate algorithm implementation of specific processes affect the accuracy of physics-based models. This leads to uncertainties that complicate the effective design of interventions and protection measures. In recent years, artificial intelligence techniques rapidly advanced and the availability of satellite imagery products increased. This research sets a novel attempt to predict the planform evolution of braided rivers with deep learning using satellite images. The study focuses on the middle and lower reaches of the Brahmaputra-Jamuna River in India and Bangladesh. We developed JamUNet, a U-Net-based convolutional neural network (CNN). The model is trained with the Global Surface Water Dataset (GSWD) to classify each pixel as "Non-water" or "Water". Four images from the same month over four consecutive years were used as input. The fifth year image served as target. JamUNet demonstrates a general capability in capturing the planform evolution. Morphological processes like meander migration, channel abandonment, and confluence and bifurcation development are often well captured. However, temporal patterns are lacking. More complex phenomena, like channel formation and channel shifting, remain unpredicted. JamUNet also tends to underpredict the total areas of erosion and deposition. Overall, JamUNet achieves a 5 – 6% improvement compared to the benchmark method for which no morphological change occurs in metrics such as precision, recall, F1-score, and critical success index (CSI). Among these, recall is the most meaningful metric for evaluating the model performance. JamUNet can serve as a preliminary tool for water management authorities in India and Bangladesh. It can assist in prioritising bank protection in erosion-prone areas and support land reclamation projects and inland navigation. Caution is always advised due to the model tendency to underpredict erosion. More research is required to improve the current model. Nonetheless, deep-learning modelling seems a promising field of research. Testing alternative model architectures and incorporating additional data, such as water levels or river discharge, could improve the model performance.

Acknowledgements

This thesis marks the end of my studies at Delft University of Technology and, at least for the moment, also the end of my academic journey. The last two years were a roller-coaster of emotions, happy and unforgettable moments as well as challenges and difficulties. Many people have crossed my way. Some of them just for a while, while others have brought to me much more than I could ever imagine. The list of people I wish to thank is so vast that hopefully I don't miss anyone.

I start thanking my supervisors Erik Mosselman, Riccardo Taormina and Víctor Chavarrías. Your suggestions, guidance and help during the last months were indispensable for the development and completion of this project and, above all, I'm grateful for the opportunity you gave me.

Thank you Erik, I'm incredibly glad to have had you as main supervisor of my thesis. Your immense knowledge of rivers, among the rest, has always fed my curiosity and always made me eager to learn new things. Thank you for all the conversations, documents, images, and everything that you shared with me. Above all, thank you for your willingness to answer all the possible questions that came to my mind. Thank you Riccardo for the patience and kindness you had when I felt lost at the beginning of this project. Your positive approach was necessary to cheer me up. You always had interesting and challenging ideas which I hope will be part of follow-up studies. Thank you Víctor for being the first to welcome me in Deltares. Thank you for the support, the suggestions, the availability to answer my questions, and the interesting discussions you always brought to the table.

I'm extremely grateful to Arjen Haag for the help provided at the beginning of the thesis regarding the use of *Google Earth Engine* and for sharing a lot of knowledge on satellite images. Without you, all of this would have taken a much longer time and bigger effort.

Another big thank you to Bert Jagers, for providing me the files you used for your research, and for being available for questions whenever I had some.

On top of that, I wish to thank all Deltares employees and interns I met in the last months. It was a pleasure working around all of you.

Many thanks to Chit Yan Toe, Laura Stanganelli, Davide Wüthrich and Wim Uijttewaal for giving me the opportunity to perform some experiments and then present the results at the NCR DAYS 2024. Thank you Chit for the valuable and interesting discussions we had at the time, your perseverance and determination are remarkable and I wish you the best for your career. Thank you Laura and Davide for all the times you helped me before, during and after the experiments, for the preparation to the presentation and for all the chats and suggestions you gave me during these months. A huge thank you also goes to Wim for the discussions and suggestions, your approach to research is admirable.

I never thought I could say this, especially after the first two overwhelming quarters of the first year, but a big thanks goes to the whole MUDE Team and specifically to Robert Lanzafame. You allowed me to see how things work from the other side of the table (or the other side of the river, probably a better metaphor in this case), gave me the opportunity to become a small part of such module and appreciate the hard work that is put behind it. Thank you also to Nina Piccoli and Kieran Dunne for the supervision during the serious game development. It was fun and enriching to work for such project.

I owe everything to my parents Andrea and Antonella and my brother Francesco. You allowed me to move away from home and study abroad. You supported me and gave me the strength to hold on when I couldn't find it myself. I missed you a lot, but I know that you were, are and will always be there by my side.

Thank you to all my family, cousins, aunts and uncles, for your never-ending love and closeness across these years I was far from home. A special thought goes to my nonne Maria and Palma, I hope I made you and our beloved nonni Marcello and Enzo proud.

I thank my longstanding friends Lorenzo and Lorenzo. After more than ten years we still share the same experiences around the world. Unfortunately, the destiny wants us apart, but God bless who invented the video calls! I wish you the best for your future, wherever it's going to be. Thank you also to all my friends from

Borgo, you are so many to mention. I really missed every one of you. It's a pity we cannot see each other more often, but it's always a pleasure when we gather back together.

Special thanks to my *goliardici* friends Agatha, Edoardo and Francesca (in alphabetical order, otherwise someone might say *Zamuner et al., 2024...*) for sharing this Dutch experience together. From the flag-photo on top of the library roof to MUDE introduction, passing through Italian lunches, Sundays at the library and especially *quel maledetto frico!* I wish the best of luck for your future as well, and I hope we will meet again soon.

Thanks Lucas for sharing our time in Deltares after many projects and assignments and, obviously, for teaching me Python too! I wish you the best for your future, I'm sure it will be bright. Thank you also to Luis, Daan, and Dani for all the moments spent together and for always being *good eggs*.

I could continue for long. Many other people have entered my life in the last years. From my roommates to all group project fellows, from the other Italians in the Netherlands to all the friends I made in Civil Engineering and beyond that. Maybe we will never see each other again, or maybe we will. It doesn't matter, you will all be part of my memories.

To you all, a wholeheartedly thank you!

*Antonio Magherini
Delft, 31st October 2024*

List of abbreviations

AI	artificial intelligence	LSTM	long short-term memory
ANNs	artificial neural networks	MAE	mean absolute error
AUC	area under curve	MNDWI	modified normalized difference water index
BCE	binary cross entropy	MSE	mean squared error
BN	batch normalization	MSL	mean sea level
CCE	categorical cross entropy	NDVI	normalized difference water index
ConvLSTM	convolutional long short-term memory	NN	neural network
CNN	convolutional neural network	PINN	physics-informed neural network
CSI	critical success index	PR	precision-recall
DNNs	deep neural networks	PWD	Public Works Datum
DTM	digital terrain model	ReLU	rectified linear unit
EVI	enhanced vegetation index	RNN	recurrent neural networks
FN	false negative	ROC	receiver operating characteristic
FP	false positive	RMSE	root mean squared error
FPR	false positive rate	SAR	synthetic-aperture radar
GEE	Google Earth Engine	SWE	shallow water equations
GNN	graph neural network	TN	true negative
GSWD	Global Surface Water Dataset	TP	true positive
HLS	Harmonized Landsat and Sentinel-2	TPR	true positive rate
LES	large eddy simulation	XAI	explainable artificial intelligence

List of symbols

α_1	Scaling weight of total area of erosion (-)
α_2	Scaling weight of total area of deposition (-)
δ	Threshold parameter for Huber loss (-)
ρ	Pearson correlation coefficient (-)
ρ_s	Spearman's rank correlation coefficient (-)
$\sigma(x)$	Sigmoid activation function (-)
σ_X	Standard deviation of variable X (-)
σ_Y	Standard deviation of variable Y (-)
$\text{cov}(X, Y)$	Covariance between variables X and Y (-)
d_i	Difference between the ranks of the i -th sample of variables X and Y (-)
$f(x)$	Rectified Linear Unit (ReLU) activation function (-)
L_{BCE}	Binary cross entropy (BCE) classification loss (-)
L_{Huber}	Huber regression loss (-)
L_{MAE}	Mean absolute error (MAE) regression loss (-)
L_{MSE}	Mean squared error (MSE) regression loss (-)
L_{total}	Total training loss, including the BCE and Huber losses (-)
N	Total number of samples (-)
$R(X_i)$	Rank of the i -th sample of the variable X (-)
$R(Y_i)$	Rank of the i -th sample of the variable Y (-)
w_{thr}	Water threshold for binary classification (-)
X	First variable (-)
x	Model output before the non-linear activation (-)
Y	Second variable (-)
y_i	Target pixel class (-)
\hat{y}_i	Model output predicted probability after the non-linear activation (-)

List of Figures

2.1 Sketch of straight, meandering, braided and anastomosing channels. Image from https://tucanada.org/2022/10/14/wandering-waterways/	4
2.2 Geographical context. The map shows the full course of the Brahmaputra (yellow reach) from the Himalayas to the confluence with the Ganges River. The zoom image shows the stretch of the river considered for this case study (blue reach).	7
2.3 River cross section at Bahadurabad, in Bangladesh. The horizontal lines represents the water level at different discharges. The survey was conducted in 1992 for the River Survey Project within the Flood Action Plan 24 (1996). Water levels and discharge data are retrieved from Figure A.2.	8
2.4 Brahmaputra River close to Dibrugarh, India. Image courtesy of Erik Mosselman.	8
2.5 Bank erosion along the Brahmaputra. Left image from Sarkar (2017). Right image courtesy of Sanjay Giri.	10
2.6 Flood extent in Rangpur district, Bangladesh, July 2024. Image from Agence France-Presse (2024).	10
2.7 Convolution of an input layer. It is possible to see how the weights ω_{ij} are shared and do not change while the convolution slides through the image. Images c) and d) show the effect of padding=1, also known as same padding. This preserves the original image dimensions by surrounding the image edges by zeros. Image from Prince (2023).	13
2.8 Comparison between fully connected layers with weights weighting each input (left) and a convolutional layer with only neighbouring weights (right). Image from Prince (2023).	13
2.9 Comparison between Landsat-8 (2.9a) and Sentinel-1 (2.9b) images representing the flood conditions of the Jamuna reach upstream of the confluence with the Ganges River of 2015 monsoon season. Landsat images are affected by cloud coverage, whereas Sentinel images are not.	15
3.1 Adopted workflow for this research.	17
3.2 Example of an original image. Red, yellow, and blue pixels represent “No data”, “Non-water”, and “Water” areas, respectively. The image represents the most downstream reach in January 1990. The white arrow indicates the flow direction.	18
3.3 Example of a preprocessed binary image after the “No data” replacement. Yellow and blue pixels represent “Non-water” and “Water” areas. The image represents the most downstream reach in January 1990. The white arrow indicates the flow direction.	19
3.4 Spatial dataset reaches division. Blue, yellow, and red areas represent training, validation, and testing reaches, respectively.	20
3.5 Sketch of the JamUNet architecture and operations performed.	21
4.1 The left plot shows the evolution of the training and validation loss. The right plot shows the evolution of the validation metrics computed setting $w_{thr} = 0.5$	25
4.2 Loss and metrics across the testing datasets samples.	26
4.3 Boxplots of the loss and metrics for the test dataset. It shows the quartiles of the distribution, including outliers. The quartiles represent those values of the distribution that divide it into four equal parts, namely the 25%, 50%, 75%, and 100%.	26
4.4 Correlation matrix between the loss and metrics. The matrix is symmetric. The values present in the above-diagonal part of the matrix are the Pearson correlation coefficients between the two variables, computed with Equation 4.1.	27
4.5 ROC and PR curves for the min loss model.	28

4.6	The first row of images contains the model inputs of testing sample 2. The first image of the second row represents the model target, the second image is the model prediction using the min loss model. Yellow and blue areas represent the “Non-water” and “Water” areas, respectively. The third image is the misclassification map, i.e. the difference between model prediction and target. Green and red areas represent FP (predicted “Water” but actually “Non-water”) and FN (predicted “Non-water” but actually “Water”), respectively. The last image represents the real and predicted total areas of erosion and deposition (crossed diagonal and diagonal hatches, respectively)	30
4.7	Predicted and real erosion and deposition locations of testing sample 2. Point ①, ②, and ④ represent wrong predictions. Point ③ and ⑤ represent good predictions.	31
4.8	The left image shows the revetment location along the Kojuri channel, in Sirajganj District, Bangladesh. The right image shows the 2012 and 2013 predictions of erosion locations (left column) and the real erosion locations (right column). As it is possible to observe, the year following the implementation of the measure, JamUNet predicts erosion to occur, whereas in fact it did not happen ①. However, the year after it already correctly predicts almost no erosion in the same location ②.	31
4.9	Distribution of the predicted and real areas of erosion and deposition.	32
4.10	Correlation matrix between the predicted and real erosion and deposition areas. The matrix is symmetric. The values present in the above-diagonal part of the matrix are the Pearson correlation coefficients between the two variables, computed with Equation 4.1.	33
4.11	Example of the reach evolution using the no-change method for sample 2 to be compared with Figure 4.6. The legend is the same.	34
5.1	Example of model output sensitivity to last input image. The absence of the tributary channel on the left of the image, point ①, is reflected in the output image, which results in underpredicted erosion (although the channel is stable across all years). The figure shows the evolution of sample 12.	36
A.1	Average daily recorded water levels at Bahadurabad, Bangladesh, between 1964 and 1994. The red line shows the daily average across the years of record.	50
A.2	Average biweekly recorded water level and discharge at Bahadurabad, Bangladesh, between 1989 and 2016.	51
A.3	Correlation matrix between recorded water levels and discharge of Figure A.2. In the above-diagonal element of the matrix the Pearson and Spearman’s rank correlation coefficients are included.	51
A.4	Average daily recorded water levels at Bahadurabad, Bangladesh, between 2012 and 2023.	52
B.1	Original and rotated images representing the training reach 22 in December 1987.	53
B.2	1988 low-flow season images (first four from left) and seasonal average image of testing reach.	54
B.3	Sketch of neighbours averaging step. Red, yellow, and blue pixels represent “No data”, “Non-water”, and “Water” areas, respectively.	55
B.4	1988 neighbours average image and final average image after the final replacement for the testing reach.	55
B.5	Image representing the March 1988 conditions of the testing reach after “No data” replacement.	56
C.1	Example of a discarded input-target combination containing at least one fully “Non-water” image.	57
C.2	Sketch of an input-target combination.	58
D.1	Sketch of a 2D convolution operating on multiple input channels. The red arrows indicate the direction across which the convolution operates. The output pixels contain random values.	59
D.2	Sketch of a 3D convolution operating on multiple input channels. The red and green arrows indicate the direction across which the convolution operates. As the input dataset has depth equal to 1, the final output of the 3D convolution resembles that of a 2D convolution. The output pixels contain random values.	60
D.3	Sketch of a max pooling operation. The input pixels contain random values.	60

D.4 Sketch of a transposed convolution. The input pixels and the kernel weights contain random values.	60
D.5 Sketch of rectified linear unit (ReLU) and Sigmoid activation functions.	61
F.1 Inputs, target, prediction, misclassification map and total areas of erosion and deposition of sample 2. The prediction is generated using the $w_{thr} = 0.402$	64
F.2 Inputs, target, prediction, misclassification map and total areas of erosion and deposition of sample 19.	65
F.3 Location of the Chauhali revetment.	66
F.4 Predicted and real erosion locations of testing samples 16 and 17. Point ① and ② indicate the location of the Chauhali revetment and the predicted and real erosion.	66
F.5 Inputs, target, prediction, misclassification map and erosion and deposition areas of sample 15. Note the evolution of the Chauhali channel during the input years (first row of images) before the implementation of the revetment.	67
G.1 ROC and PR curves for the max recall model.	68
G.2 Loss and metrics boxplots of the test dataset using the max recall model.	69
G.3 Model prediction of testing sample 2 with max recall model. The image should be compared to Figure 4.6 to observe the differences between the min loss and the max recall model.	69
H.1 Evolution throughout epochs of the training and validation losses (left plot) and validation metrics for the model trained with the temporal dataset.	71
H.2 ROC and PR curves for the min loss model trained with the temporal dataset.	71
H.3 ROC and PR curves for the max recall model trained with the temporal dataset.	72
H.4 Model prediction of testing sample 1 with min loss model trained with the temporal dataset. . .	72

List of Tables

3.1	Original water band values for the chosen collection.	18
3.2	Converted water band values after binarization.	19
3.3	2D and 3D convolutions comparison.	22
3.4	Model hyperparameters and values.	23
4.1	Testing loss and performance metrics for JamUNet model. All quantities are adimensional.	26
4.2	Comparison between testing metrics for JamUNet and the no-change method. All quantities are adimensional.	32
C.1	Link between testing sample ID, input years, and target year.	58
G.1	Testing loss and metrics achieved by the max recall model. All quantities are adimensional.	68
H.1	Testing loss and metrics achieved by the min loss and max recall models trained with the temporal dataset. All quantities are adimensional.	70

Contents

Abstract	ii
Acknowledgements	iii
List of abbreviations	v
List of symbols	vi
List of Figures	vii
List of Tables	x
1 Introduction	1
1.1 Background	1
1.2 Problem definition	1
1.3 Objective and research questions	3
1.4 Thesis outline	3
2 Background	4
2.1 Braided rivers	4
2.1.1 Morphological characteristics and processes	4
2.1.2 Existing models of braided rivers	6
2.2 Case study	6
2.2.1 Context	6
2.2.2 Hydrology and hydrodynamic	7
2.2.3 Geomorphology	9
2.2.4 Morphodynamics	9
2.2.5 Social and international context	9
2.2.6 Existing models of the Brahmaputra River	11
2.3 Artificial intelligence	11
2.3.1 Machine learning and neural networks	11
2.3.1.1 Loss functions and metrics	12
2.3.2 Deep-learning modelling	12
2.3.2.1 Related research	14
2.4 Remote sensing	14
2.4.1 Satellite imagery	14
2.4.2 Available collections	15
2.4.3 Preprocessing techniques	16
3 Methods and materials	17
3.1 Satellite images	17
3.1.1 Image collection	17
3.1.2 Image preprocessing	18
3.1.3 Input-target combinations	19
3.2 JamUNet	20
3.2.1 Model architecture	20
3.2.2 Hyperparameters	22
3.2.3 Performance metrics	22
3.3 Benchmark models	24
3.3.1 Neural Network	24
3.3.2 No-change method	24

4	Results	25
4.1	JamUNet predictions	25
4.1.1	Loss and metrics	25
4.1.2	ROC and PR curves	28
4.1.3	Visual assessment	29
4.2	Benchmark models comparison	32
5	Discussion and limitations	35
5.1	Interpretation and limitations	35
5.2	Limitations of the research	37
5.2.1	Training dataset	37
5.2.2	“No data” replacement	37
5.2.2.1	Validation and testing datasets	38
5.2.3	Image quality	38
5.2.4	Hyperparameter optimisation	38
6	Conclusions and recommendations	39
6.1	Answers to research questions	39
6.2	Recommendations	41
6.2.1	Model application	41
6.2.2	Future research	41
6.2.2.1	Main recommendations	41
6.2.2.2	Further possibilities	42
6.2.3	Final consideration	43
	Bibliography	44
A	Hydraulic data	50
B	Satellite image preprocessing	53
B.1	Image standardization	53
B.2	“No data” replacement	54
B.2.1	Seasonal averaging	54
B.2.2	Neighbours averaging	54
C	Dataset generation	57
C.1	Quality filtering	57
C.2	Final datasets	57
D	JamUNet operations	59
E	Additional physics-induced loss	62
F	Additional figures	64
G	Max recall model	68
H	Temporal dataset	70
H.1	Dataset generation procedure	70
H.2	Model results	70
H.3	Motivation	71

1

Introduction

1.1. Background

Braiding is one of the three main channel patterns described in literature to categorise the plan view of rivers, along with meandering and straight channels ([Leopold and Wolman, 1957](#)). As described by [Jagers \(2003\)](#), braided rivers are “systems of channels splitting and rejoining around islands”. Although different classifications exist and outline more channel types, such as anabranching and anostomising rivers ([Alabyan and Chalov, 1998](#)), braided rivers represent one of the most distinctive categories because of their unique characteristics ([Ferguson, 1993](#)).

Braided rivers are generally identified by the combined presence of wide shallow channels, an alluvial bed of non-cohesive sediments, and limited presence of vegetation ([Jagers, 2003](#)), as well as the channel division around alluvial islands and bars ([Leopold and Wolman, 1957](#)). Braided rivers are often found on steeper slopes than meandering rivers ([Jagers, 2003](#)). For these reasons, in braided rivers the flow velocity generally reaches higher values compared to other rivers. Consequently, relevant erosion and deposition processes take place in braided rivers, which are considered the most dynamic type of rivers ([Stecca et al., 2019](#)).

Different morphodynamic changes are observed in braided rivers, ranging from micro- (bedforms and dune migration) to macro- (bar development) and especially mega-scale (formation of islands, confluences and bifurcations) ([Jagers, 2003](#)). All scales are relevant for the overall evolution of the river morphology and mutually influence each other. Consider for instance the effect of scour holes on bend migration and channel attraction ([Mosselman and Sloff, 2002](#)). However, mega-scale changes have greater impacts on the river and the adjacent areas. Specifically, bank erosion and lateral channel migration, as well as new channel formation, are the most concerning ones for the safety of the surrounding population. When a bank retreats or a new channel is formed, existing (hydraulic) structures, embankments, and bank protection measures can be damaged. The functionality of these structures is therefore harmed and eventually failures can occur. The surrounding population becomes exposed to higher flood and geomorphic risk ([Noh et al., 2024](#)), which can lead to severe damages and loss of lives. Moreover, the failure of critical infrastructure, such as bridges and inland port facilities, can negatively influence trade, with implications on the well-being and economic and social development of the nearby communities ([Best et al., 2022](#)).

1.2. Problem definition

Being able to predict where and how the river evolves is crucial for supporting spatial-related decision-making processes in the vicinity of such rivers. To safely and efficiently design interventions, protection measures, evacuation plans as well as infrastructure and all other buildings, engineers, urban planners, and policy makers need to be aware of the spatio-temporal morphological developments of the rivers. In the past, laboratory scale experiments were employed for understanding specific processes and consequently predict the evolution of the concerned physical systems. Nowadays, large-scale experiments are generally replaced by numerical models ([Chavarriás, 2019](#)). Several different models were developed for braided rivers. To mention some, [Murray and Paola \(1994\)](#) formulated a physics-based cellular model. [Klaassen et al. \(1993\)](#) developed a probabilistic framework based on scenario-occurrence probabilities. [Jagers \(2003\)](#) implemented

their functional model design into a working branches model.

However, a reliable physics-based modelling of such rivers is a complex task (Jagers, 2003), especially when compared to other types of rivers. This is due to the rapidly changing morphological patterns, along with a lack of process understanding and the non-implementation in the model algorithms of specific morphodynamic processes (Jagers, 2003). Consider for instance scour hole evolution, bank erosion, and the effects of riparian vegetation, wood, sediment gradation, and secondary flow (Mosselman and Sloff, 2002; Jagers, 2003; Siviglia and Crosato, 2016; Hu et al., 2023). Achieving accurate short-term predictions (ranging from few months to one year ahead) is already challenging. Longer-term predictions (beyond one year), often more relevant for the design of large projects and land-use plans, pose even greater difficulties.

Consequently, prediction uncertainty hinders the design process and the effectiveness of interventions, with negative repercussions on the local population.

In recent years, artificial intelligence (AI) techniques emerged and gained popularity for several different applications and contexts (Goodfellow et al., 2016; Blake et al., 2021). As defined by Gignac and Szodorai (2024), AI is “the maximal capacity of an artificial system to successfully achieve a novel goal through computational algorithms”. Despite raising new ethical and moral questions (Lo Piano, 2020; Prince, 2023), the performance reached by AI models to solve complex problems is promising. Among other applications, deep learning is the key subset of machine learning for computer vision tasks that involve images. Unlike other traditional machine-learning models, the deep-learning ones leverage the inclusion of specific inductive biases. These represent the inherent tendency of the model to prioritize one solution over another by extrapolating patterns between data points (Krizhevsky et al., 2012; Prince, 2023). This is crucial to overcome the so-called curse of dimensionality, the issue of training data sparsity. The amount of training data needed increases exponentially as the number of dimensions in the data grows. Without inductive biases, vast amounts of data would be required to ensure sufficient coverage of the input space and limit unexplored regions of this. For that, an almost infinite amount of data would be required for large datasets. Instead, by reducing the complexity of the model solution space, inductive biases guide the model learning and help the generalization to unseen datasets. Representation learning also enhances deep-learning models performance by automatically extracting meaningful patterns from raw data. As a result, complex patterns are expressed in simpler representations necessary for the modelling task (LeCun et al., 2015; Goodfellow et al., 2016).

Several deep-learning models were developed as surrogate models, which are trained on the outputs of physics-based model simulations. These models are useful as they accelerate computations by eliminating the need to run new simulations. Consider the convolutional neural network (CNN) by Melo et al. (2022) for emulating estuarine morphology or the graph neural network (GNN) by Bentivoglio et al. (2023) for flood propagation, both relying on Delft3D simulations (Deltares, 2024). However, physics-based models are complex to implement for braided rivers and their accuracy may not always be satisfactory. It is not meaningful to develop surrogate models for predicting braided rivers morphology, because the deep-learning model uncertainties would sum with the physics-based ones. As a result, the predictions would be even less reliable. For this reason, deep-learning morphological modelling should be based on monitored data.

Remote sensing, particularly through satellite imagery, offers valuable tools for predicting the planform evolution of braided rivers. Satellite images were already used for rivers and other Earth systems studies (Hagenars et al., 2018; Munasinghe et al., 2021; Spinoza et al., 2021). Recently, more satellite imagery products were generated and different global collections became accessible (Woodcock et al., 2008; Wulder et al., 2012). The release of the Google Earth Engine (GEE) platform also helped users to more easily retrieve, process and export the available geographic collections for different applications (Gorelick et al., 2017; Cardille et al., 2023).

Some studies applied AI techniques using satellite images to focus on specific morphological processes, such as bank erosion (Jagers, 2003; Ali and Zobeyer, 2021), or for quantifying the total accretion, erosion and unchanged areas (Deng et al., 2022) of braided rivers. However, no research is yet performed to predict the overall planform evolution of (large) braided sand-bed rivers, including other morphological processes such as confluence and bifurcation development, bar and channel migration, channel shifting, and so on.

For all of the above reasons, this thesis represents an innovative attempt to predict the morphological evolution of a braided sand-bed river using deep learning. A U-Net-based CNN model for binary image segmentation was implemented. The model was trained with the Global Surface Water Dataset (GSWD) collection, which contains pre-classified satellite images (Pekel et al., 2016). The suggested model for this research is

inspired by the neural network (NN) developed by Jagers (2003). The case study of this project is the Brahmaputra River. The focus was posed on the middle and lower reaches, in India and Bangladesh. These reaches were split into several smaller segments used for training, validating and eventually testing the model. The model performance were assessed considering a loss function, which represents the difference between the model predictions and the real targets. Other classification metrics were included as well. The model outputs were visually inspected as well. A comparison with two benchmark morphological models was eventually performed.

1.3. Objective and research questions

The main objective of this thesis is to *develop an original deep-learning model for predicting the planform changes of a braided sand-bed river*. Additionally, another expected outcome is the development and implementation of a novel framework and methodology that can be applied to similar Earth systems, such as coastal shoreline changes for instance.

Hence, the main research question of this work is:

How can deep learning help in predicting the planform changes of braided sand-bed rivers?

To answer this question, the following subquestions are formulated:

1. What is a suitable model architecture for this problem?
2. What are the most suitable assessment metrics for this task?
3. Based on the available data, for which prediction timescale - short-term (i.e., less than one year ahead, in the order of weeks or few months) or long-term changes (one year ahead or longer) - is the model most suited?
4. Depending on the satellite image collection chosen, what image preprocessing techniques are required for improving the model performance, if there is any?
5. What relevant information (quantities, locations) on morphological processes can be inferred from the outputs of the model?
6. How does the model compare to (i) the NN developed by Jagers (2003) and (ii) the benchmark method for which no morphological change occurs?
7. What kind of physical information can be exploited to improve the model performance?

1.4. Thesis outline

This document is organised as follows. Chapter 2 presents the background of this project, with relevant literature, the related findings, and the case study. In Chapter 3 the overall methodology, the data, the pre-processing steps and the model architecture are described, along with an introduction of the benchmark models. Chapter 4 contains the model results in terms of metrics and visual inspection, as well as a comparison with the benchmark models. Chapter 5 contains a discussion and interpretation of the results, along with limitations. Eventually, Chapter 6 contains the conclusions and answers to the research questions, with recommendations for the model application and future improvements.

2

Background

This chapter provides the background required for this research. It mostly includes the literature review performed to find existing knowledge, related findings and research gaps regarding the three main topics of this project. Section 2.1 describes braided rivers, their morphological processes and some relevant models developed. Section 2.2 provides the technical and social context of the case study. It also includes information on previous research and the models specifically developed for it. Section 2.3 introduces artificial intelligence, the most suited assessment metrics, and deep-learning modelling with some applications to similar problems. Section 2.4 focuses on remote sensing and satellite images, the available collections, and the most commonly applied preprocessing techniques.

2.1. Braided rivers

2.1.1. Morphological characteristics and processes

Braiding is one of the primary channel patterns used to describe the plan view of rivers (Leopold and Wolman, 1957). According to Paola (2001), braiding is the fundamental pattern of alluvial rivers, representing the instability of streams flowing on non-cohesive material. In addition to the definition given by Jagers (2003) presented in Section 1.1, a broader description is provided by Lane (1957). According to Lane, braided rivers are characterised by “a number of alluvial channels with bars and islands between meeting and dividing again, and presenting from the air the intertwining effect of a braid”.

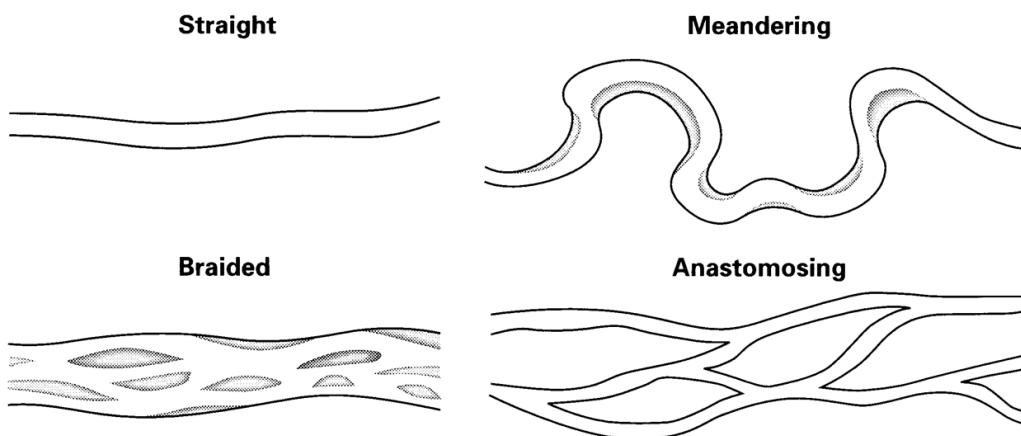


Figure 2.1: Sketch of straight, meandering, braided and anastomosing channels. Image from <https://tucanada.org/2022/10/14/wandering-waterways/>.

There is some debate about the way braided rivers develop. Lane (1957) suggests that braided rivers evolve from anastomosing rivers after coarse sediments begin to deposit. These sediments form central bars that

trap finer material and eventually lead to a disequilibrium in sediment transport rates. As mentioned by Jagers (2003), various instabilities in water and sediment balance equations can cause rivers to develop a braiding pattern rather than a meandering one. The formation of braided rivers depends on a set of controlling variables, which differentiate them from meandering channels. Compared to these, braided rivers are generally characterised by steeper bed slopes, larger discharges and larger width/depth ratios (Leopold and Wolman, 1957). However, Paola (2001) argues that only abundant bedload is the most fundamental characteristic of braided streams. Factors such as steep slopes, coarse grain sizes, and flashy discharges are deemed to be secondary, and play accessory roles in the evolution of these rivers. Consequently, braiding patterns develop because of an exaggerated response of sediment flux to variations in the water flux. Such pattern does not appropriately develop only if vegetation and cohesive sediments are present, which constraint its evolution. Potentially, without vegetation the braiding pattern would be laterally unconstrained (Paola, 2001). According to Ferguson (1993), there is not a single universal mode of braid development. Braiding patterns are not solely the result of depositional processes, like central bar deposition and the formation of chutes and lobes. Erosional processes, including chute cutoffs, dissection of multiple lobes, and avulsions initiated by bank erosion, can also lead to the development of braiding. Typically, braided rivers are characterised by "high values of valley slope, stream power, shear stress, width/depth ratio, and bedload transport rate".

The interaction between water and sediment plays a crucial role in determining the river(bed) morphological evolution (Jagers, 2003). The combination of high flow velocities and non-cohesive bed and bank material results in significant erosion and deposition processes in braided rivers (Jagers, 2003). Hence, the presence of an unstable network of multiple channels causes the river morphology to rapidly change during high-flow stages (Ashmore, 2013).

Several morphological features characterise braided rivers. The main ones are braid bars and islands, around which channels bifurcate and converge. Braid bars typically develop in the middle of the river, detached from the banks, and are subject to reorganization and evolution after floods due to lateral accretion processes. In contrast, islands are generally larger and consequently more stable, also thanks to the presence of vegetation.

Channel confluences and bifurcations are other significant elements and play a crucial role in the high-level dynamics of braided rivers, influencing the overall morphological evolution of the system (Paola, 2001). Confluences tend to accelerate the water flow, causing erosion, whereas bifurcations slow it, enhancing deposition and bar formation. The process of reversible entrainment is relevant in braided rivers, as erosion and deposition processes typically move similar volumes of material (Paola, 2001).

Confluences are often considered more dynamic than bifurcations due to variations in discharge partitioning. However, bifurcations and upstream channel alignment play an important role in distributing water and sediment among the downstream channels (Jagers, 2003). Depending on several factors, the geometry of the bifurcation can lead to channel abandonment. This occurs when sediment supply exceeds transport capacity and causes a silting-up and consequent abandonment of the channel (Klaassen et al., 1993; Jagers, 2003). Klaassen et al. (1993) showed that the frequency of channel abandonment upstream of bifurcations is correlated with the deflection angle. The larger the angle, the larger the probability of channel abandonment. Opposed to channel abandonment, channel formation through cutoff, often across bars, allows the river to maintain a braided pattern. Additionally, channel offtakes contribute to the development of avulsions, where channels abruptly change course (Best et al., 2022).

Other relevant processes in braided rivers are channel migration (Klaassen et al., 1993) and channel shifting. These processes are driven by bank erosion and the accretion of point bars. When riverbanks erode, the sediment is deposited downstream in areas with reduced transport capacity (Jagers, 2003). This interplay between erosion and deposition often leads to either a gradual narrowing or widening of the river. Bank erosion is also correlated to bar development and the presence of confluences or bifurcations (Surian, 2015). This process is considered as one of the most concerning for the nearby population because it can cause the river banks to retreat by exceptional amounts, harming the safety of people, buildings and infrastructure (Klaassen et al., 1993; Best et al., 2022).

Scour holes, particularly at confluences and meander bends, are another important feature of braided rivers, as they affect the macro-scale dynamics with implications on channel attraction, channel narrowing and generation of confluences (Mosselman and Sloff, 2002; Jagers, 2003).

Additionally, planform changes are sensitive to flood discharge variations. The extent of bank erosion is influenced by the magnitude and duration of the floods, although these do not necessarily affect the locations of erosion (Jagers, 2003). Channel formation and bar migration along the channel also depend on the flood

extent. This has consequence on the future morphological evolution. Consider the bar arrival at a bifurcation. This can be anticipated or postponed in time, depending on the flood magnitude. The siltation of one of the bifurcated channels with bar deposited material is anticipated or postponed depending on the time needed for the bar to reach the bifurcation (Klaassen et al., 1993).

2.1.2. Existing models of braided rivers

Field studies and laboratory experiments were common approaches to study the dynamics of braided rivers (Jagers, 2003). These were later replaced by physics-based numerical models (Chavarriás, 2019), which allow to achieve results in a shorter time whereas not being affected by laboratory scale effects. However, the fast and chaotic morphological changes of braided rivers complicate the development of reliable and accurate models. In addition to this, some morphodynamic processes are either not fully understood or lack a proper implementation in physics-based models. Scour hole development, bank erosion, sediment transport characteristics and bedforms development, the effects of riparian vegetation, wood, sediment gradation, and secondary flow are some of these processes (Mosselman and Sloff, 2002; Jagers, 2003; Siviglia and Crosato, 2016; Hu et al., 2023). For these reasons, the models predictions can be uncertain and not always provide a solid base for the design of infrastructure, interventions or other measures. Nevertheless, solving a system of equations, comprising the Saint Venant shallow water equations (SWE), coupled with a sediment-flux equation and Exner equation for mass balance, is particularly complex for braided rivers, as the channel network constantly reorganises (Paola, 2001). The computation time of these models can also be a further concern, especially when modelling large systems.

To partially overcome these issues, Murray and Paola (1994) developed a simple two-dimensional cellular model without modelling the flow in detail. The model does not solve explicitly the above-mentioned set of equations but simplifies them into a set of local rules that govern how water and sediment move on a discrete grid. The model turned out to be able to produce braiding patterns only using non-linear sediment transport laws. This is in line with the more general knowledge that physically realistic sediment transport laws are non-linear.

Besides some physics-based models, such as Delft3D (Deltares, 2024), were developed for specific situations (Giri et al., 2021; Barua, 2023), other original models were implemented for braided rivers. Jagers (2003) mentions two conceptual models. The first is a one-dimensional Markov Chain, that statistically represents the alongstream variations of the number of channels in a cross section and allows to generate bifurcations and confluences. Additionally, Random Walk models based on a regular two-dimensional lattice grid were implemented. In these models, initially single channels can randomly split or not into two branches and progress independently until they rejoin together or split again.

Other models, theoretically suitable for all types of rivers, exist (Jagers, 2003). One of these is the time series extrapolation. Such autoregressive models rely on the presence of maps, surveys, or satellite images. By overlaying the bankline positions across different years and including random variations, the average bank migration rate can be extrapolated from past years behaviour, for instance.

Eventually, to deal with planform changes of braided rivers, Jagers (2003) suggests to either treat the branches of braided rivers as separate channels and model these as single meandering rivers or, alternatively, model all branches with a two-dimensional model for the entire braidplain.

2.2. Case study

2.2.1. Context

The case study of this project is the Brahmaputra River. The name of the river actually changes several times along its course. Brahmaputra is the name of the Indian reach, but it also represents the most adopted name when referring to the whole river. In Bangladesh, for instance, the river splits in two branches. The main branch is referred to as the Jamuna River, and the Old Brahmaputra is the name used for the secondary branch. When dealing with both the Indian and the main Bangladeshi reach, as for in this research, the river is often called Brahmaputra-Jamuna. However, for a better readability and adherence to the international most recognised name, the river is referred to only as the Brahmaputra River. The name Jamuna is used when specifically referring to the Bangladeshi reach only.

The Brahmaputra springs in the Chinese province of Tibet, from the Chemayungdung glacier near Mount

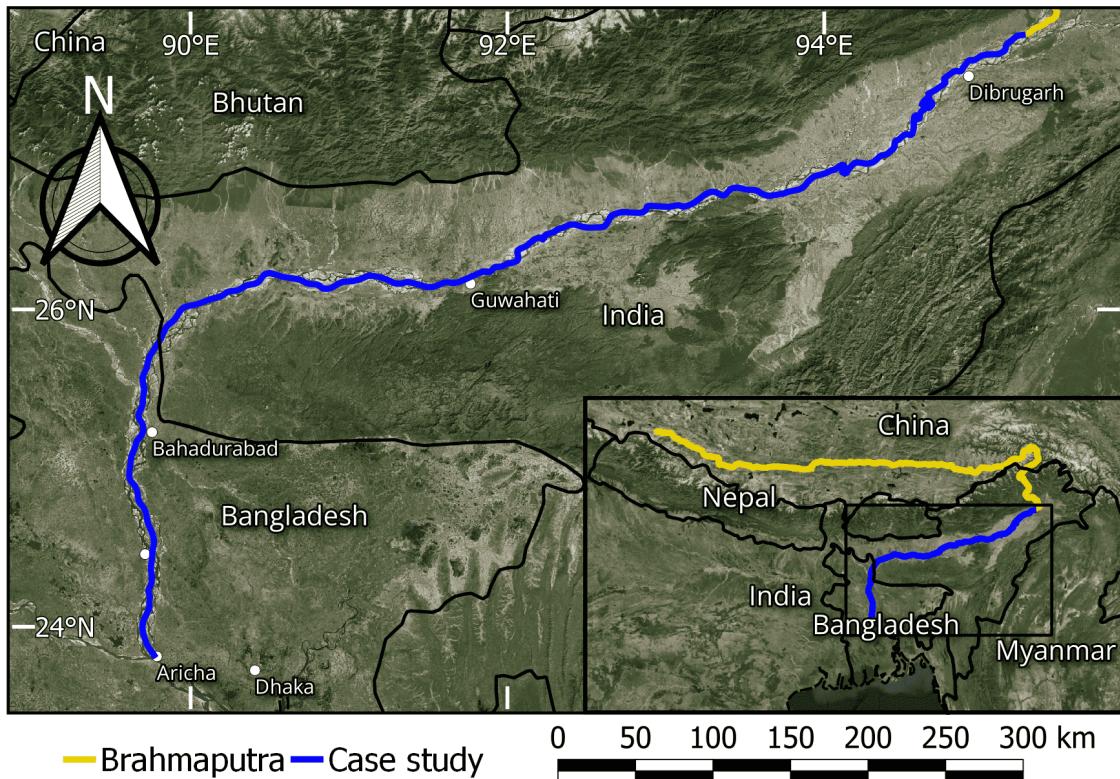


Figure 2.2: Geographical context. The map shows the full course of the Brahmaputra (yellow reach) from the Himalayas to the confluence with the Ganges River. The zoom image shows the stretch of the river considered for this case study (blue reach).

Kailash in the Himalayas, at an altitude of 5100 m (Jagers, 2003). It flows along the northern board of the Himalayas for about 1400 km. It then crosses the mountains, entering the Indian province of Assam. Close to Sadiya it joins the Lohit and Dibang River. After approximately 600 km, the river sharply bends left and enters Bangladesh, where it flows for approximately 250 km. Near Bahadurabad, in Bangladesh, it also splits in two branches, the Jamuna, currently the main branch, and the Old Brahmaputra (more information on this is given in Section 2.2.3). At Aricha, in Bangladesh, it joins the Ganges River, forming the Padma River. Further downstream, the Padma River joins the Upper Meghna River, forming the Lower Meghna River, which eventually discharges in the Bay of Bengal, in the Indian Ocean (Jagers, 2003; Best et al., 2022). The bottom right image of Figure 2.2 shows the planview of the Brahmaputra from its origin to the confluence with the Ganges River.

The stretch of the river considered for this project is approximately 875 km long. The upstream boundary is set close to Sadiya, just downstream of the confluence with the Lohit and Dibang River, in Indian territory. The downstream boundary is set close to Aricha, just upstream of the confluence with the Ganges River, in Bangladesh. Figure 2.2 shows the stretch of the river considered for this project. It also includes the river from its origin to the confluence with the Ganges River.

2.2.2. Hydrology and hydrodynamic

The Brahmaputra River has a catchment area of approximately 560 000 km², which is on average subject to 1.9 m/yr of rainfall. Due to its geographical location, the river experiences large floods during summer months. The first contribution to the flood discharges is the snow melt originated from the Himalayas, which occurs around May. However, tropical monsoon rainfall also largely contributes to the hydrograph increase and, actually, dominates it between July and September, when the hydrograph peak is reached. The monsoon season generally lasts from May until October. For these reasons, the discharge largely varies throughout the year. During the low-flow season between January and April, the river discharge is between 4000 and 12 000 m³/s, with a minimum record of 2037 m³/s. The mean annual discharge at Bahadurabad is 20 200 m³/s. The average annual flood is about 60 000 m³/s, but larger values were recorded, with a peak of 102 535 m³/s

reached in 1998. The water levels recorded at Bahadurabad also increase by 6–7 m, from around 13 m Public Works Datum (PWD)¹ during the low-flow season to 19–20 m PWD during floods (Jagers, 2003; Best et al., 2022). More information on recorded water levels and discharges are given in Appendix A.

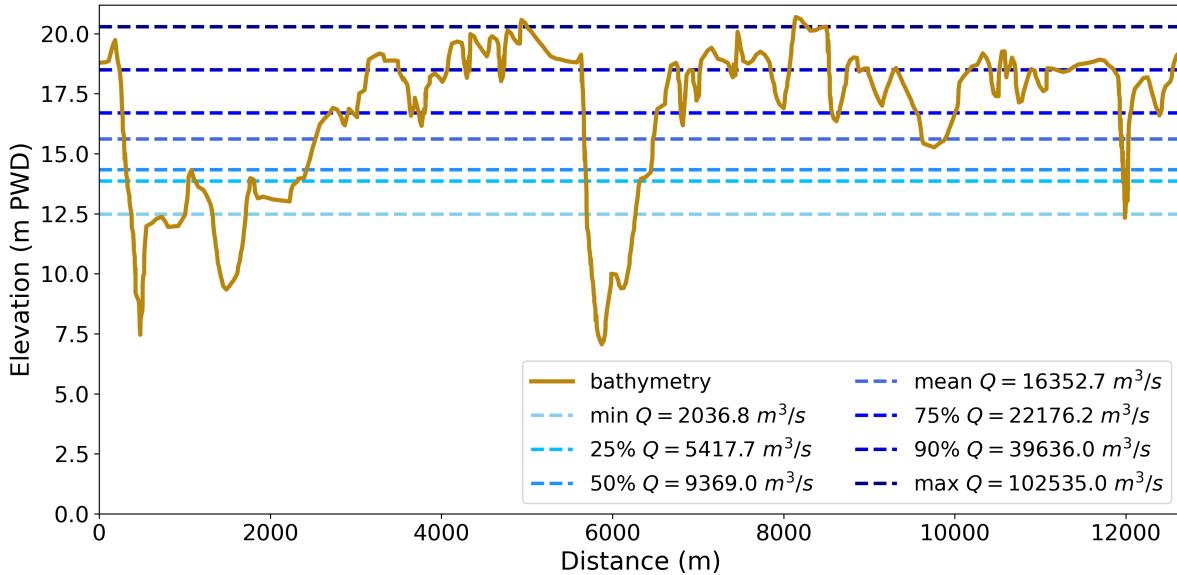


Figure 2.3: River cross section at Bahadurabad, in Bangladesh. The horizontal lines represents the water level at different discharges. The survey was conducted in 1992 for the River Survey Project within the [Flood Action Plan 24 \(1996\)](#). Water levels and discharge data are retrieved from Figure A.2.

Figure 2.3 shows a cross section of the Jamuna reach at Bahadurabad. It also includes the water levels associated to the minimum, average, and maximum discharges as well as the 25%, 50% (median), 75%, and 90% percentiles of the discharge distribution. The data were recorded within 1990 and 2016 (Figure A.2). The bathymetry was surveyed in 1992.



Figure 2.4: Brahmaputra River close to Dibrugarh, India. Image courtesy of Erik Mosselman.

The braidplain river width varies between 5–17 km, with an average of 11 km, whereas individual channels can be up to 2 km wide (Klaassen et al., 1993; Best et al., 2022). The flow depth varies between 2–20 m and

¹The PWD was established by the Bangladeshi Department of Public Works. It is a horizontal datum originally believed to have zero level at the mean sea level (MSL) at Calcutta, India. This datum is 0.46 m below the MSL.

up to 30 m in the largest channels, although local bed scour holes of 40 – 45 m were recorded (Jagers, 2003; Best et al., 2022). Depth-averaged velocities can reach over 3.5 m/s during floods.

2.2.3. Geomorphology

The Brahmaputra flows across one of the most tectonically active zones in the world, subject to a 5 cm/year north migration (Jagers, 2003). The Himalayas mountains experience uplifting processes, whereas the Bengal region is subducting. For these reasons, large earthquakes occurred in the last centuries (Best et al., 2022) and might occur again in the future. Such earthquakes can highly influence the local morphology, as they cause landslides that feed the river with sediments and mobilise the alluvial riverbed fine sediments, causing an increase in sediment transport. Additionally, small faults are also considered to influence local channel migration and width changes (Klaassen et al., 1993; Best et al., 2022).

In Bangladeshi territory the Brahmaputra splits in two branches, the Jamuna and the Old Brahmaputra. The latter used to be the main branch until the end of the 18th century, when a gradual avulsion started developing. Tectonic activity, increased discharges and catastrophic floods are deemed to be the most likely causes for the initiation of such avulsion (Best et al., 2022). Moreover, westward migrating patterns are currently observed in the river (Jagers, 2003). Between 1973 and 2003, river bank erosion rates were estimated to be 207 and 83 m/year on the right and left bank, respectively (Best et al., 2022).

Additionally, in some locations the course of the Brahmaputra is affected by the presence of inselbergs, also known as monadnocks. These geological features are isolated mountains generally composed of granitic or gneiss matrix. Such features represent a constraint for the river development, and eventually confine the river braidplain. The main inselbergs of the region are the Shillong and the Mikir Hills (Lahiri, 2022). Vegetation also plays a crucial role in stabilizing the river banks. Locations characterised by the presence of vegetation are generally more stable, compared to non-vegetated areas, although Best et al. (2022) argues that deep roots are insufficient to halt bank erosion in the deepest channels (20 – 30 m deep) of the Brahmaputra.

2.2.4. Morphodynamics

The Brahmaputra is widely classified as a braided river. This is confirmed by the Brice braiding index, included between 4 – 6 (Best et al., 2022), and by the physics-based formula proposed by Crosato and Moselman (2009). In the braid belt, the upstream reaches typically have three major channels, whereas the downstream ones are characterised by two major channels. Most of the Brahmaputra bed and bank material is fine sands (85%) and silt/clay (15%). For this reason, the Brahmaputra is also referred to as a braided sand-bed river. Due to abrasion processes, the grain size gets finer moving downstream, with an average value of 0.22 mm at Bahadurabad. During floods, a huge amount of sediment is transported. Some studies estimate a sediment yield between 555 – 1157 Mt/year, with a 10 – 30% contribution of bedload (Best et al., 2022). For an average transporting width of 3 km, 5 Mt/day of total load is estimated to be transported by the river. This large amount of sediments causes the river to reorganise itself after every flood. Most of this material originates from erosion of the bed and the banks of the river and the sediment supply from the Himalayas.

Among the various morphodynamic processes that shape braided rivers, including the Brahmaputra, bank erosion likely received the most attention due to its significant impacts on the nearby communities. According to Klaassen et al. (1993), the bank erosion rates of the Brahmaputra can vary between 0 – 500 m/year, with local exceptional situations reaching up to 750 – 1000 m/year. Additionally, according to recent estimates the Brahmaputra erodes around 2000 – 5000 ha/year of mainland floodplain (Best et al., 2022).

2.2.5. Social and international context

The devastating floods of the Brahmaputra yearly affect tens of millions of people in India and Bangladesh, with estimated damages in the order of billions of dollars. Crops, buildings, infrastructure, and economical activities largely suffer from these floods, which can also claim hundreds of human lives (Best et al., 2022). Several projects were developed with international stakeholders and NGOs to protect the surrounding areas and the local population. Most of these works focused on river training and riverbanks protection and stabilization (Best et al., 2022). As a direct consequence of the 1987 and 1988 large floods, for instance, the Bangladeshi Government, in collaboration with the World Bank, developed the Flood Action Plan (FAP). Within this plan, several projects with different goals were implemented. FAP1 aimed at stabilising the banks of the Brahmaputra for protecting the city of Sirajganj and the Fulchari railway ferry terminal. It also aimed at preventing avulsions into the Bengali River and future outflanking of the Bangabandhu Bridge. FAP21 tested



(a) Erosion of the bank.

(b) Erosion of an embankment.

Figure 2.5: Bank erosion along the Brahmaputra. Left image from [Sarkar \(2017\)](#). Right image courtesy of Sanjay Giri.

permeable groynes at Kamarjani and revetments at Bahadurabad and Guthail ([Best et al., 2022](#)).

Figure 2.6: Flood extent in Rangpur district, Bangladesh, July 2024. Image from [Agence France-Presse \(2024\)](#).

One of the most recent initiatives is the Flood and Riverbank Erosion Risk Management Investment Program (FRERMIP), funded by the Asian Development Bank and the Governments of Bangladesh and the Netherlands. The project involved a joint venture of consultancy firms such as Northwest Hydraulic Consultant and Euroconsult Mott MacDonald. Deltares and Bangladeshi national authorities, such as the Ministry of Water Resources (MoWR) and the Bangladesh Water Development Board (BWDB), collaborated as well. The program started in 2016 aiming at “providing structural and non-structural flood and riverbank erosion risk management measures” in three main priority Bangladeshi areas in order to “provide a more secure and improved livelihood for people living along the main rivers of Bangladesh” ([McLean et al., 2020](#)).

Despite significant efforts and the success of many projects, several interventions failed between 1998 and 2004 ([Best et al., 2022](#)). These failures are partly due to the occurrence of large floods but also to the difficulties in predicting the morphological development of the river.

Other international initiatives focused on the construction of critical infrastructure, such as the Bangabandhu Bridge ([Mottaleb and Rahut, 2018](#)), or on flood adaptation measures, like the construction of plinths ([Begum, 2023; Koepen Stichting, 2024](#)).

2.2.6. Existing models of the Brahmaputra River

Some physics-based models of the Jamuna exist. However, most of these models were implemented for understanding the morphological response of the river to specific local interventions. Some examples are the bank stabilization of the avulsion channel in Kamarjani ([Barua, 2023](#)) or the prediction the behaviour of key bifurcations under different conditions to support adaptive river management strategies ([Giri et al., 2021](#)). In this section other original models are reported.

[Klaassen et al. \(1993\)](#) developed a probabilistic framework to predict the planform changes of the Jamuna reach. Branches were treated individually as meandering channels. Elementary planform changes were modelled either with kinematic or stochastic processes, such as Monte Carlo simulations. They showed it was possible to reach quite accurate result for short term changes. Longer-term changes remained a challenge as the number of possible developments increases.

Building on this framework, [Jagers \(2003\)](#) implemented the so-called branches model. Simulations were performed with different stochastic influences. Every result represented the realisation of one ensemble of possible planforms. By overlaying all planforms one over the other, the frequency representing the predicted occurrence of channels was interpreted as the probability of erosion.

2.3. Artificial intelligence

2.3.1. Machine learning and neural networks

The concept of [AI](#) was initially introduced in 1955 by [McCarthy et al.](#), who stated that “every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it”. Whereas different definitions exist, [AI](#) can be considered as the set of computational algorithms used to solve complex tasks by simulating the human intelligent behaviour ([Prince, 2023](#)).

Machine learning is the subset of [AI](#) systems that acquires knowledge by extracting patterns from preprocessed raw data ([LeCun et al., 2015; Goodfellow et al., 2016](#)). This set of techniques allows to make subjective decisions based on real world knowledge. Two main categories of methods exist: supervised and unsupervised learning. Supervised learning involves the presence of labeled data, which implies the existence of a defined relationship between inputs and outputs ([Prince, 2023](#)). Within supervised learning, the two most common tasks are classification, which involves predicting two or more distinct categories, and regression, where the goal is to predict the continuous value of one or more variables. For our problem both approaches could be appropriate. The task choice depends on the data used, i.e. binary satellite images or bankline coordinates, and the research objectives.

Among various algorithms, artificial neural networks ([ANNs](#)) are one of the most widely used models. These are inspired by the functioning of biological brains ([Goodfellow et al., 2016](#)), which consist of interconnected neurons that transmit signals to each other through synapses. Similarly, [ANNs](#) consist of multiple fully connected neurons, also known as hidden units, organised into several distinct hidden layers. [Bishop \(1994\)](#) describes [ANNs](#) as “a series of functional transformations”. Initially, linear combinations of the inputs are computed using adaptive parameters known as weights and biases. These combinations are then followed by a non-linear transformation through differentiable activation functions, which vary depending on the task. These functions introduce non-linearity in the network, allowing to learn more complex patterns.

To propagate the information through the network, the output of a single (hidden) layer serves as input to the subsequent one. Finally, before reaching the output layer, the results are transformed by a further activation function suited to the specific task. This entire process of mapping inputs to outputs, with information flowing through the hidden units of the various layers, is called *forward propagation*.

During the training process, the learning and the performance of the model are evaluated using loss functions. These functions measure the difference between the real targets and the model predictions. The objective is to minimise this loss, making the model predictions as close as possible to the actual targets. This is achieved through *error backpropagation* ([Bishop, 1994; LeCun et al., 2015](#)). This method computes the gradient of the loss function with respect to the model parameters and then propagates this error gradient through the network. Specific optimization algorithms use these gradients to update the model parameters, aiming to reduce the loss and improve the model performance.

One key issue of these models is overfitting. A model is said to overfit when it optimises its parameters to fit not only the relevant patterns in the data, but also the noise present in those data. The model can achieve high accuracy on the training data, however it often fails to generalise well to unseen data, leading to poor

performance on new examples. To overcome this issue, the available data is split in three distinct datasets: training, validation and testing datasets. The training set is used to train the model, which enables it to learn the relevant patterns and update its parameters. The validation dataset is used during the training step to monitor the model performance in order to avoid overfitting. The testing dataset is used for a final evaluation of the model on unprocessed data and is relevant to assess how the model can generalise (LeCun et al., 2015).

2.3.1.1. Loss functions and metrics

The loss functions and metrics depend on the problem and task. For regression tasks, the most common functions are the root mean squared error (**RMSE**) and the mean absolute error (**MAE**) losses, which measure the distance between the predicted and the target value. Similar functions are the mean squared error (**MSE**) and Huber losses. Classification losses focus instead on comparing the predicted probability to the actual target class. For binary classification tasks, binary cross entropy (**BCE**) is the most common function, whereas for multiclass classification categorical cross entropy (**CCE**) is used.

Additionally, for classification problems other assessment metrics are available. Accuracy measures the percentage of correct predictions out of the total. Precision measures the percentage of correctly predicted positive pixels with respect to all positive predictions. Recall measures the percentage of correctly positive predictions out of the real positive targets. F1-score is the harmonised mean of precision and recall. In addition to recall, Bentivoglio et al. (2023) suggest to include the critical success index (**CSI**), specifically when dealing with imbalanced datasets.

2.3.2. Deep-learning modelling

ANNs typically consist of few hidden layers, each containing a limited number of hidden units. Due to their relatively simple structure, these models are often referred to as "shallow" **ANNs**. However, when it comes to extracting high-level abstract features, their performance get worse. Furthermore, when dealing with high-dimensional structured data², like images, they would require an impractically large number of parameters, making computation inefficient (Goodfellow et al., 2016; Prince, 2023). This inefficiency is worsened by the so-called *curse of dimensionality*. This refers to the issue of training data sparsity, where the amount of training data needed increases exponentially as the number of dimensions in the data grows.

To overcome these challenges, deep learning was introduced. Compared to shallow **ANNs**, deep neural networks (**DNNs**) are formed by stacking many hidden layers, which in turn increase the number of hidden units.³ Besides requiring a larger number of parameters, **DNNs** are able to capture complex patterns more effectively due to the presence of specific *inductive biases*. These biases, implemented through model architectural choices, represent built-in assumptions that allow these models to prioritise certain solutions by extrapolating patterns between data points (Krizhevsky et al., 2012; Prince, 2023). Deep-learning models performance is also enhanced by representation learning. This comprises a set of methods that extracts meaningful and informative patterns from raw data, without requiring preprocessing of these. Deep-learning models automatically implement this feature extraction and express complex patterns in terms of simpler representations, like corners and contours (Goodfellow et al., 2016). These representations are necessary for the model task (LeCun et al., 2015), such as feature detection or classification.

Among other applications, deep learning is particularly well-suited for computer vision tasks. This is largely due to the spatial inductive bias, which assumes that nearby pixels within an image are statistically related. **DNNs** can leverage this by considering the spatial organization of images. The spatial inductive bias is achieved through convolutions, linear operations that apply filters to local regions independently (Figure 2.7). By sharing the parameters across the entire image, convolutions process different local regions of an image in a similar way and better exploit spatial dimensions (Goodfellow et al., 2016). This greatly reduces the number of parameters compared to **ANNs**, which treat the relationship between every input equally and have to learn the interpretation of the pixels at every position (Prince, 2023). Figure 2.8 compares a fully connected layer with a convolutional layer, for which less parameters are required. After applying a non-linear activation function to the convolution output we obtain a convolutional layer. The output of a convolution is called feature map and it represents abstract hidden variables extracted from the input data.

²The term *high-dimensional data* refers to data that have a large number of input variables. Each pixel of an image represents one variable. As a result, even relatively small images with 64x64 pixels would already contain 4096 input variables.

³The number of hidden units is often referred to as *width* of the network, whereas the number of hidden layers as *depth* of the network. To avoid confusion with the width and depth concepts of rivers, from here onward, hydraulic variables will be defined as *channel width* and *channel depth*, whereas the number of hidden units and hidden layers will be referred to as *network width* and *network depth*, respectively.

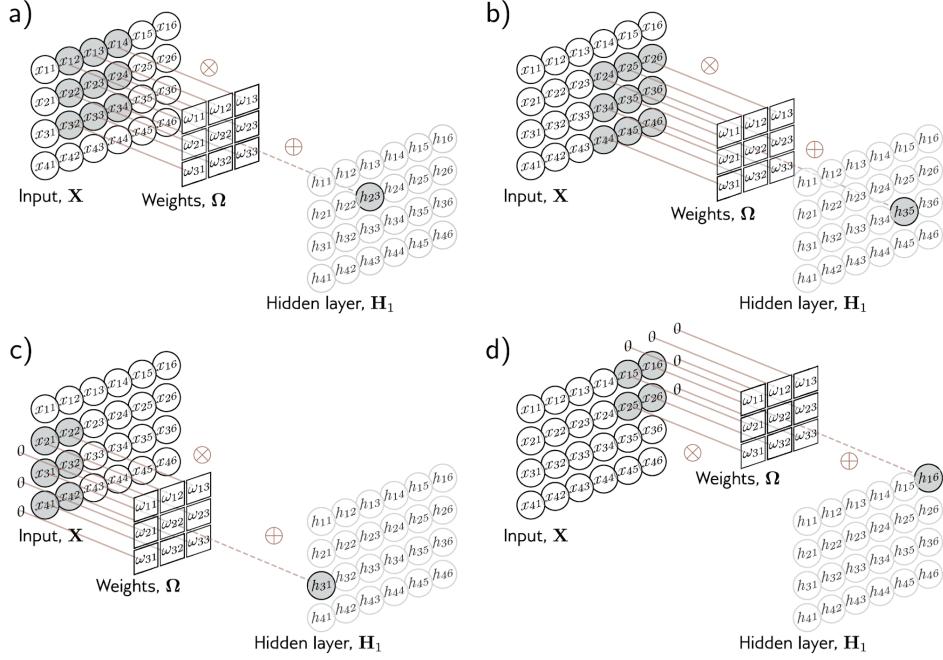


Figure 2.7: Convolution of an input layer. It is possible to see how the weights ω_{ij} are shared and do not change while the convolution slides through the image. Images c) and d) show the effect of padding=1, also known as same padding. This preserves the original image dimensions by surrounding the image edges by zeros. Image from Prince (2023).

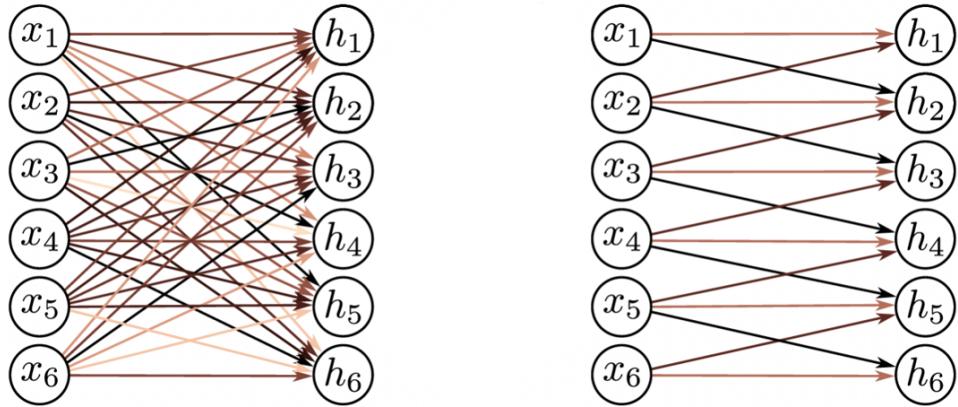


Figure 2.8: Comparison between fully connected layers with weights weighting each input (left) and a convolutional layer with only neighbouring weights (right). Image from Prince (2023).

Pooling layers are often used after convolutional layers to reduce the spatial dimensions of these feature maps, improving the computational efficiency. Finally, stacking several convolutional and pooling layers forms a **CNN**. These models can process multiple input channels⁴ simultaneously by applying different filters to each input channel and generating multiple feature maps (O’Shea and Nash, 2015).

Similarly to the way a **CNN** deals with two-dimensional input data like images, recurrent neural networks (**RNN**) are suitable for sequential data, like time series (Schmidt, 2019). These models are specifically designed to capture temporal patterns thanks to the presence of an internal state, which enables them to retain the information from previous time steps. By introducing memory cells and gating mechanisms to **RNN** we get a long short-term memory (**LSTM**) unit (Hochreiter and Schmidhuber, 1997), which is able to track both short- and long-term time dependencies in the data. Eventually, combining a **CNN** with an **LSTM** results

⁴The term *input channels* refers to the number of different layers characterizing the input data, such as images. For grayscale images, there is one channel, whereas RGB images have three channels (red, green, and blue). To avoid confusion with river channels, we will refer to this concept simply as *channels* or directly as *river channels*, reserving the term *input channels* for the number of layers in the input.

in a convolutional long short-term memory ([ConvLSTM](#)) (Shi et al., 2015), which is particularly suited when dealing with spatio-temporal phenomena.

2.3.2.1. Related research

One of the main applications of deep learning is surrogate modelling. Surrogate models are trained with the simulations of numerical models, with the aim to speed these computations up without having to run the whole simulation again. For this reason, these models are particularly useful for almost real-time applications, such as flood warning systems.

For instance, [Bentivoglio et al. \(2023\)](#) implemented a novel [GNN](#) for spatio-temporal flood inundations. [Nazari et al. \(2022\)](#) developed a physics-informed neural network ([PINN](#)) for modelling the water flows in a river during floods. [Melo et al. \(2022\)](#) developed a surrogate [CNN](#) for emulating the results of a hydrodynamic model such as Delft3D ([Deltares, 2024](#)) for estuarine morphological evolution.

However, within the framework of this research surrogate modelling is not considered a valid option, given the uncertainties of physics-based models for braided rivers. Consequently, the predictions of the morphological evolution should be based on monitored data, such as flow discharge, sediment transport rate, and river bathymetry, for instance. [Khankhoje and Choudhury \(2024\)](#) developed a dynamic [NN](#) for sediment discharge forecasting in the Mississippi River and some of its tributaries using measured data from gauging stations. [Zhang et al. \(2024\)](#) implemented a [CNN](#) for predicting the shear stress distribution and river bed topography in meandering rivers coupling measured bathymetry data with large eddy simulation ([LES](#)) results. [Deng et al. \(2022\)](#) implemented a regression lightweight [NN](#) with only few layers and neurons to predict the overall accretion, erosion, and unchanged areas of a braided river. Satellite images and records of flow rate, annual average sediment transport rate, and sediment concentration were the inputs of the model. This turned out to predict quite accurately the areas of accretion and erosion, indicating a relevant depositional trend of the studied reach. [Ren et al. \(2024\)](#) also implemented a novel algorithm for riverbank erosion risk assessment based on a multi-source dataset. They categorised the risk in five levels, from low to extreme. Their research integrates methods related to the emerging field of explainable artificial intelligence ([XAI](#)). These methods enhance the model interpretability to gain more insights in the factors and processes influencing riverbank erosion. In this way, they aim to help decision-makers and involved stakeholders to prioritise the locations for protective projects.

Specifically for the Brahmaputra, two [AI](#) models exist. Using classified satellite images, [Jagers \(2003\)](#) developed an [NN](#), to predict the locations where bank erosion was likely to occur one year into the future. Five physics variables, such as the point distance to the closest channel or the local width of the nearest channel, were parameterised and given as inputs. Despite some errors and missing some morphological features, the model was able to provide reasonable predictions regarding bank erosion at several locations. This is the model that actually inspired this research and more information on this is given in Section 3.3.1.

[Ali and Zobeyer \(2021\)](#) developed a [ConvLSTM](#) for predicting the bank erosion of the Jamuna reach as well. However, contrary to the [NN](#) model by [Jagers \(2003\)](#), they did not make use of classified images in two distinct land classes. They developed a regression model which used the relative coordinates of the delineated banklines. Compared to the model by [Jagers \(2003\)](#), which did not take into account the temporal evolution of the river across different years, this one included specific features that enable to represent the temporal components as well. The model turned out to underpredict the erosion magnitude, whereas being able to accurately identify the erosion prone areas.

2.4. Remote sensing

2.4.1. Satellite imagery

Remote sensing is a set of instrument-based techniques used for acquiring and measuring spatially distributed data by means of devices which are not in contact with the surveyed items ([Salomonson, 2014](#)). For this research, only satellite-derived data were considered. This choice was motivated by the wide availability of satellite image collections through online platforms. Satellite data also provide extensive temporal and spatial coverage. This is crucial for analysing of how wide geographic areas change in time. This contrasts with other remote sensing techniques like aerial photography or unmanned aerial vehicle (UAV) imagery, which typically cover smaller areas and have shorter temporal records.

In 1972, NASA designed and launched the first Landsat mission, with the aim to provide unclassified remotely sensed data to facilitate assessment of land and water resources ([Lauer et al., 1997](#)). Since then, more satel-

lites were launched. These generated several collections with different characteristics and fields of application. The release of the [GEE](#) cloud-based platform helps users to access, process and export the available geographic collections for different Earth systems applications ([Gorelick et al., 2017; Cardille et al., 2023](#)).

Passive and active are the two types of existing remote sensing instruments. Passive satellites are also known as optical or multispectral sensors and they rely on the presence of an external energy source, specifically the sunlight. Optical sensors measure the visible, near-infrared, and shortwave infrared spectral regions, and measure the emergent radiation from the Earth. However, as they rely on the sun radiation, the presence of cloud coverage negatively affects the generated images. Active instruments, instead, emit an artificial radiation that generates a scatter back signal from the ground. synthetic-aperture radar ([SAR](#)) is the most common active system for satellite imagery. Since these systems emit micro-waves, these can penetrate the clouds and consequently return images regardless of the weather conditions ([Richards, 2022](#)).

Apart from the sensor type, the spatial nominal resolution (the square area covered by one single pixel), the temporal coverage (the total years of record), and the temporal granularity or satellite overpass frequency (the average temporal interval between consecutive images) are key factors in the collection choice. Other relevant properties are the emitted radiation amplitude (the brightness of the emitted electromagnetic radiation), the radiation frequency (which relates to the spectral resolution), and the satellite spatial coverage (the total area that a satellite can cover in one image).

2.4.2. Available collections

Landsat and Sentinel are the two most common satellite collections. Landsat 1 satellite was launched in 1972, although the most used collections trace back to 1982, when Landsat 4 was launched ([Wulder et al., 2012](#)). All Landsat satellites contain optical sensors, hence their collections are hindered by the cloud presence. The image resolution increased from 80 to 30 m, whereas the overpass frequency is 16 days. Sentinel collections are developed within the Copernicus Program of the European Union. It comprises two satellites. Sentinel-1, radar, was launched in 2014 and has an overpass frequency of 6 days. Sentinel-2, optical, was launched in 2015 and has an overpass frequency of 10 days. Both satellites have a spatial resolution of 10 m. Figure 2.9 shows how Landsat collections are hindered by cloud coverage, whereas Sentinel-1 is not.

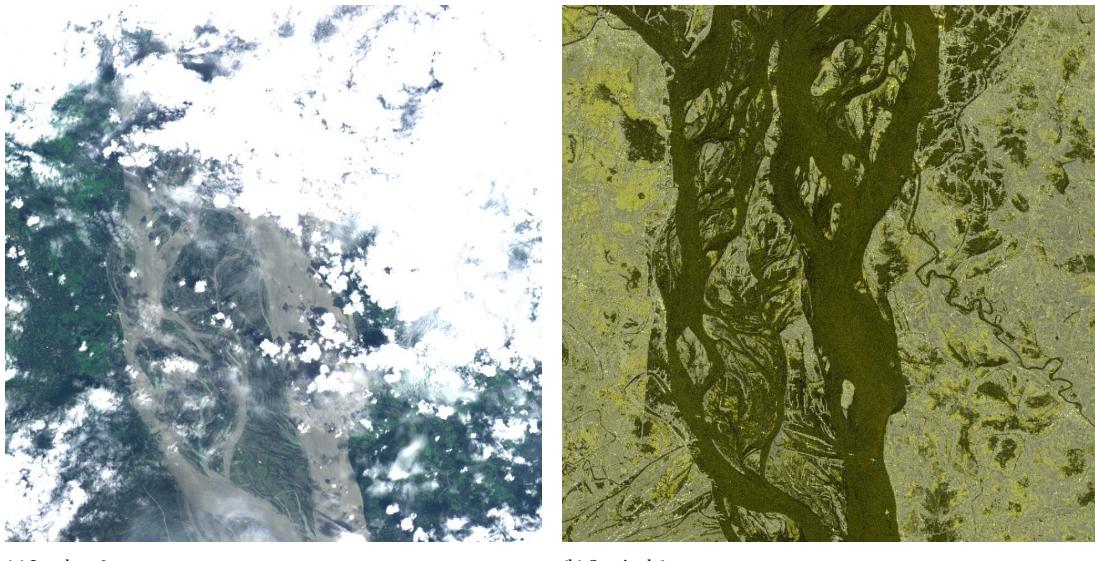


Figure 2.9: Comparison between Landsat-8 (2.9a) and Sentinel-1 (2.9b) images representing the flood conditions of the Jamuna reach upstream of the confluence with the Ganges River of 2015 monsoon season. Landsat images are affected by cloud coverage, whereas Sentinel images are not.

NASA and IBM developed the Harmonized Landsat and Sentinel-2 ([HLS](#)) collection, which combines Landsat-8 and 9 Operational Land Imager (OLI) with Sentinel-2A and 2B MultiSpectral Instrument (MSI). The harmonization between these collections is made possible due to the spectral similarities between these sensors. This enabled the generation of higher frequency imagery products ([Claverie et al., 2018](#)).

Another collection worth mentioning is the [GSWD](#), developed by the European Commission's Joint Research

Centre (JRC). This collection provides globally classified images of surface water dynamics, dating back to 1984. As it is based on Landsat satellite imagery, it is also hindered by cloud presence ([Pekel et al., 2016](#)).

2.4.3. Preprocessing techniques

The image preprocessing steps depend on the collection used. [Munasinghe et al. \(2021\)](#) describe several techniques for pixel-based classification. Most of these are one- or two-step algorithms, such as Band Ratioing, Random Forest classifiers, and Support Vector Machines. The choice among these algorithms depends on the available data type and task. Besides these, other relevant methods are sub-pixel based algorithms, to accommodate different land-use classes in a single pixel, change detection methods, and ensemble classification, which combines different approaches to improve the results. [Khurshid and Khan \(2012\)](#) implemented a two-step algorithm for river channel extraction, combining an approximate and a detailed segmentation process. [Basnayaka et al. \(2022\)](#) suggest to use multispectral bands for pixel classification. The modified normalized difference water index ([MNDWI](#)) is the most recommended index for water bodies extraction, but to minimise misclassification normalized difference water index ([NDVI](#)) and enhanced vegetation index ([EVI](#)) should be included as well. [Deng et al. \(2022\)](#) also use a combination of the above-mentioned water-body indices with Landsat images. Additionally, they argue that their proposed algorithm is more accurate for water-bodies extraction than the [GSWD](#) annual collection for braided rivers.

3

Methods and materials

This chapter aims to describe the workflow applied in this project, the data, and the deep-learning model that was trained. Section 3.1 describes the data used, specifically the satellite images, the preprocessing steps, and the final datasets. In Section 3.2 the implemented deep-learning model is described, along with the loss function and assessment metrics used to monitor and assess the model performance throughout the training process and for future evaluations. Section 3.3 introduces the benchmarks models used for comparing and assessing the performance of our proposed model.

Figure 3.1 shows the adopted workflow for the current research.

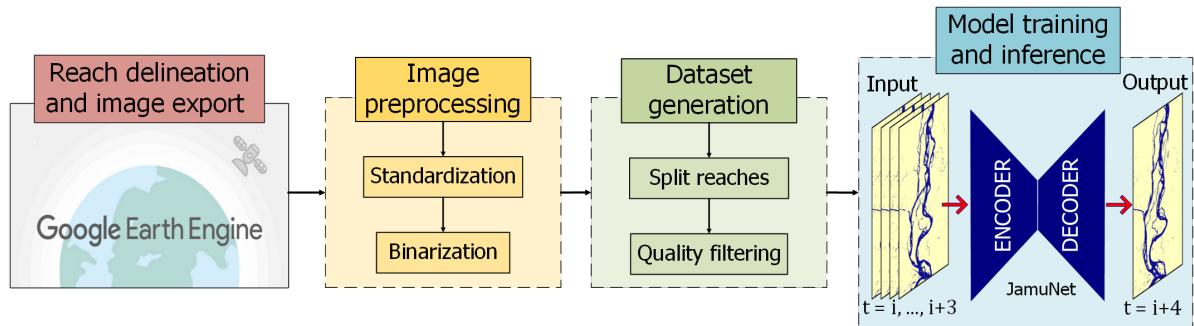


Figure 3.1: Adopted workflow for this research.

3.1. Satellite images

3.1.1. Image collection

Due to cloud coverage, during the monsoon season the number of suitable images from optical satellites is significantly reduced, if not absent at all, compared to the rest of the year. Additionally, in the months of November and December the flood season is at its falling stage. During these months, the water levels are indeed lower compared to the summer monsoon months, but still quite high. Consequently, several morphological features such as bars, secondary channels and bifurcations are not clearly distinguishable or sometimes not visible at all. On top of that, construction works for bank protections and any other intervention can only be done during the low-flow months. Hence, for the involved stakeholders it is more important to be aware of the post-floods morphological conditions.

For these reasons, in this research only images taken during the low-flow stage were considered. As introduced in Section 2.2.2, for the Brahmaputra River the low-flow stage occurs between January and April.

The available collections were accessed and analysed using the GEE platform. Apart from the parameters introduced in Section 2.4.1, the final collection was chosen also based on the number of cloud-free images and the expected preprocessing steps needed.

The *Monthly History GSWD* collection was chosen (Pekel et al., 2016). This collection was developed by the European Commission's Joint Research Centre (JRC) in the framework of the Copernicus Programme. It maps the location and temporal distribution of water surfaces at global scale (Global Surface Water Explorer, 2024). An expert system classifier was trained with Landsat 5, 7, and 8 images. Hence, the collection is affected by clouds. For the validation dataset the classification achieves less than 1% of false water detection (error of commission) and misses less than 5% of water areas (error of omission). The collection contains monthly images from March 1984 to December 2021, although for the case study area images are available starting from December 1987. The images have a spatial resolution of 30 m and are characterised by a single band named **water**. This band can take three values, each representing a specific land-use category, as included in Table 3.1. Figure 3.2 shows an original image containing the three distinctive classes.

Value	Class
0	No data
1	Non-water
2	Water

Table 3.1: Original **water** band values for the chosen collection.

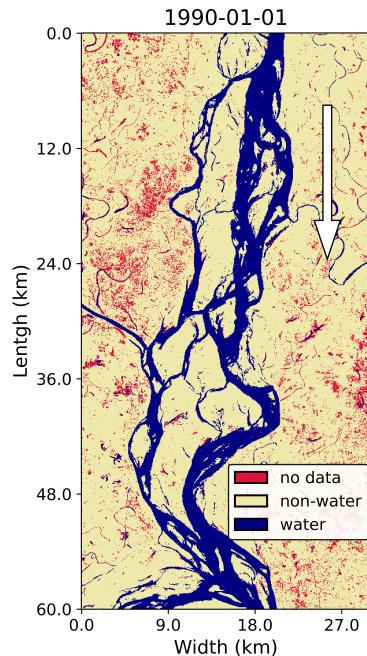


Figure 3.2: Example of an original image. Red, yellow, and blue pixels represent “No data”, “Non-water”, and “Water” areas, respectively. The image represents the most downstream reach in January 1990. The white arrow indicates the flow direction.

Images were accessed using an automated script in the GEE code editor. The specific dataset ID is the New Asset ID 1984-2021 **JRC/GSW1_4/MonthlyHistory**. The image resolution was downgraded to 60 m (i.e., with a decreasing factor of 2). Besides a potential loss of detail and feature representation, this downgrade reduces the image size, hence the export time, and most importantly the memory usage and computation time during the model training and inference. The case study stretch introduced in Section 2.2 was divided in thirty smaller reaches. Using QGIS 3.28.15, approximately 60×30 km rectangles were drawn to demarcate each reach area. 60 and 30 km represent the longitudinal and lateral dimensions, respectively. The shape files were imported in GEE as *Assets*. These were used to crop the satellite image areas, and consequently to export the images for all reaches individually.

3.1.2. Image preprocessing

To ensure consistent characteristics across all reaches and years, the images were preprocessed before generating the training, validation, and testing datasets. The flow characteristics of the most downstream of the thirty reaches were used as a reference for this standardization step. In this reach, the flow is directed southward, i.e. it goes from the top to the bottom of the image (see Figure 3.2). In the first step, it was ensured that the flow direction was directed in the same way for all reaches. This was done through a rigid rotation for all those reaches that did not respect this requirement. Subsequently, images were reshaped to match a 1000×500 pixels dimension, representing an exact 60×30 km reach.

“No data” pixels were individually replaced with the corresponding single pixels of an average generated image. This image was obtained performing a two-steps averaging process. In the first step, the low-flow season average was computed separately for all reaches and years. The remaining “No data” areas were replaced by averaging the neighbouring pixels. A hard replacement with the “Non-water” class was applied to all remaining “No data” pixels. In this way, images were binarized and the modelling task was partially simplified (see Section 3.2). Figure 3.3 shows the same image previously presented (Figure 3.2) after the binarization step. For coherence with literature and for algorithm simplification, the two class values were converted as given in Table 3.2. Appendix B contains more information on the preprocessing and “No data” replacement.

Value	Class
0	Non-water
1	Water

Table 3.2: Converted **water** band values after binarization.

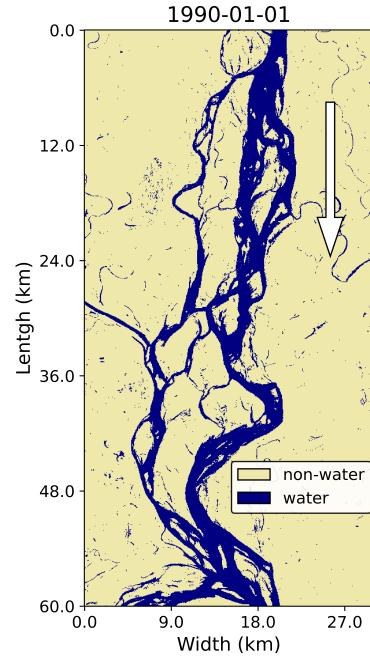


Figure 3.3: Example of a preprocessed binary image after the “No data” replacement. Yellow and blue pixels represent “Non-water” and “Water” areas. The image represents the most downstream reach in January 1990. The white arrow indicates the flow direction.

3.1.3. Input-target combinations

From a temporal perspective, the problem was modelled as a *sequence-to-one* task (Goodfellow et al., 2016). Four images, representative of the same month across four consecutive years, were used as inputs. The fifth consecutive year image represented the target. For this reason, the input and target tensors have shape (4, 1000, 500) and (1, 1000, 500), respectively.

Even after replacing the “No data” pixels, some completely unsuitable images remained. The reason is that throughout some years, mainly during the 90s, all images were only composed of “No data” pixels. As a result, the pixel replacement strategy previously adopted would return fully “Non-water” images. To avoid these images to become part of any dataset, a quality filtering operation was applied to all input-target combinations. If any image included within any input-target combination contained more than 480 000 “Non-water” pixels, the full combination was discarded. More details on the quality filtering procedure are given in Appendix C.

The training, validation and testing datasets were generated considering only images representative of March. The month choice was based on the number and quality, i.e. least number of “No data” pixels to be replaced, of the available input-target combinations.

The datasets were obtained by spatially splitting the reaches. The most downstream reach was reserved for testing the model and the one just upstream for the validation step. The remaining 28 reaches located further upstream, the majority of which in Indian territory, were utilised for training the model. This dataset is composed of 518, 19 and 19 samples for the training, validation, and testing steps, respectively. Figure 3.4 shows the planview of the river and the reaches used for the different modelling steps.

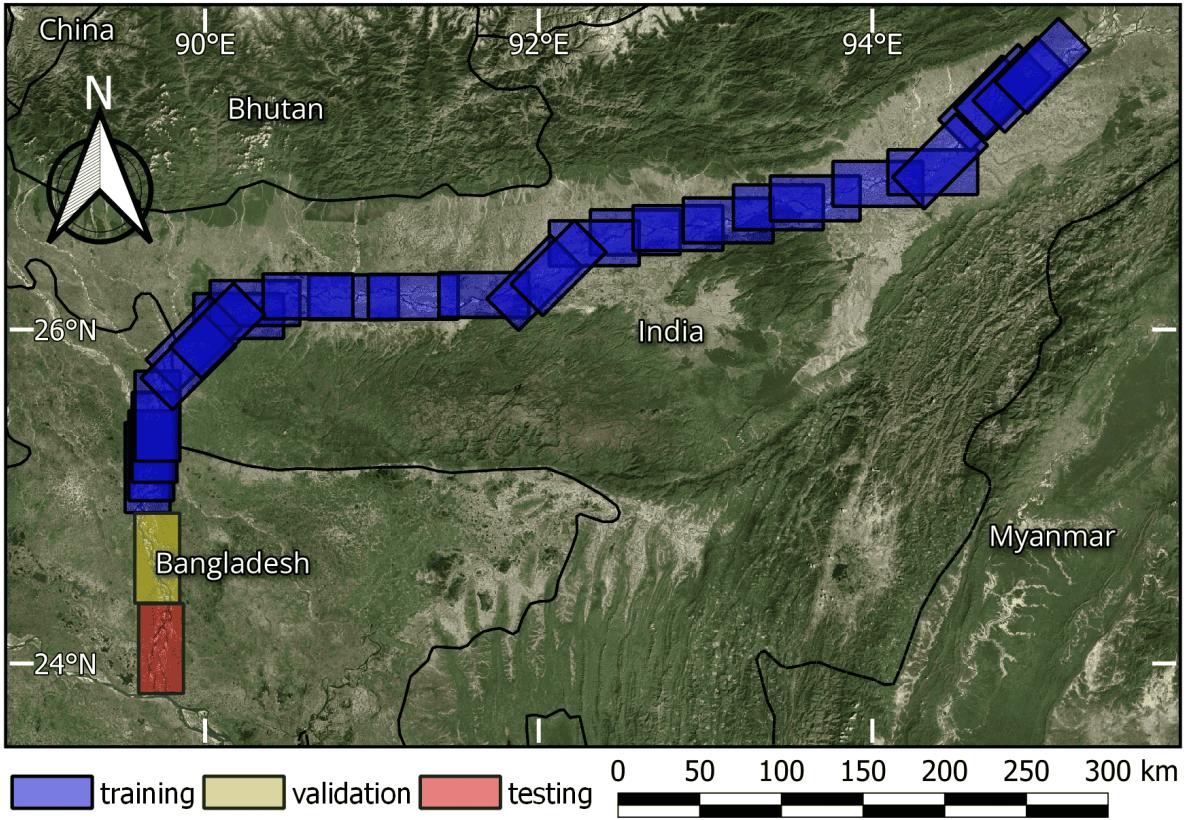


Figure 3.4: Spatial dataset reaches division. Blue, yellow, and red areas represent training, validation, and testing reaches, respectively.

A different splitting procedure was investigated. This is described in Appendix H, along with the most relevant results obtained and a brief discussion on why this dataset was not selected for the main analysis .

3.2. JamUNet

3.2.1. Model architecture

From an algorithm perspective, the modelling problem is defined a *binary semantic segmentation* task. This means that the model is trained to classify each pixel of a whole image as either “Non-water” or “Water”.

A CNN model based on the U-Net architecture was developed. Such model was originally introduced by Ronneberger et al. (2015) for biomedical image segmentation. The U-Net is a fully convolutional neural network which comprises two primary components: an encoding path and a corresponding symmetrical decoding path. The encoder is responsible for extracting relevant features from the input images, whereas the decoder uses these features to reconstruct the output image. A bottleneck layer connects the encoder and decoder, facilitating the flow of information between these two components. The model here introduced is named JamUNet, as the combination of the words “Jamuna” and “U-Net”.

CNN models are generally well-suited for spatially-oriented tasks. From a temporal perspective, these problems are often modelled as *one-to-one* problems, where a single input image is used to predict a single output image. In such cases, temporal features are generally not relevant or not present at all. The suitability of CNN models for these tasks is attributed to their spatial inductive bias, which is effectively captured through convolutional operations. However, our problem involves not only a clear two-dimensional spatial component but also significant temporal patterns. The image in year i is somehow correlated to the images in the previous years $i - 1, i - 2, \dots, i - n$, with n being an undefined number of years. It is assumed that current morphological conditions are influenced by the preceding years conditions.

To address the challenge of incorporating a temporal inductive bias into the model without using LSTM cells, an alternative approach was employed. Traditionally, the convolutional block of U-Net models consists of

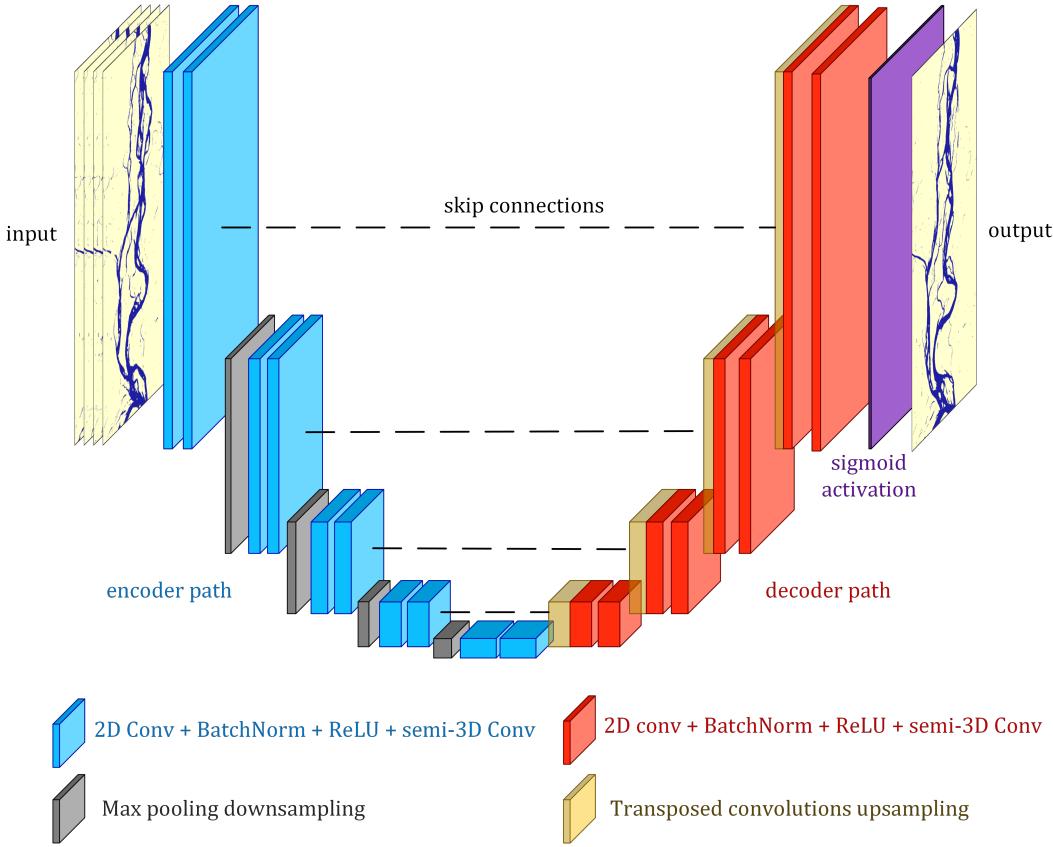


Figure 3.5: Sketch of the JamUNet architecture and operations performed.

a sequence of operations: a 2D convolution, followed by batch normalization (BN)¹, an activation function, typically rectified linear unit (ReLU) (Equation 3.1, left image of Figure D.5), and at least another 2D convolution.

$$f(x) = \max(0, x) \quad (3.1)$$

where x represents the model output after the BN step.

To better capture the temporal evolution of the images, here the second 2D convolution was replaced with a semi-3D convolution, which actually operates as a 2D convolution due to the dataset dimensions². This modification slightly improved the model outputs, both in terms of metrics and of image generation. In this setup, the initial 2D convolution is expected to extract spatial features from the feature maps independently, whereas the semi-3D convolution captures the temporal dynamics of the spatial features across the sequence map by map. This approach is inspired by Stańczyk and Mehrkanoon (2021), who applied a similar strategy for wind speed prediction.

The model is composed of 4 symmetrical hidden layers in both the encoding and decoding path, and a bottleneck temporal convolutional layer. In the encoder, every convolutional block is followed by a downsampling operation. This halves the feature map spatial dimensions, whereas the number of feature maps doubles. For instance, after the first downsample, the tensor size goes from (8, 1000, 500) to (16, 500, 250). The max

¹BN reparameterises the activations of each (hidden) layer by normalizing them to have zero mean and unit variance, based on mini-batch statistics. This helps to mitigate the internal covariate shift, that is the variation in the distribution of the network inputs. BN improves training stability, enhances the parameters optimization and also introduces a regularizing effect. Additionally, it enables to use larger learning rates, which accelerates the learning process and enhances overall training efficiency (Prince, 2023).

²The datasets have dimension (4, 1000, 500), with 4 the number of input channels, 1000, 500 image height and width. For a (semi-)3D convolution to operate the image depth is required as additional dimension. Hence, the datasets should have dimension (4, *image depth*, 1000, 500). By temporarily setting the image depth to 1, semi-3D convolutions can be performed.

pooling method with a 2x2 kernel size and a 2x2 stride was employed for this operation. Table 3.3 compares the 2D and semi-3D convolutions used in this study.

Hyperparameter	2D convolutions	Semi-3D convolutions
Number of feature maps per layer	8, 16, 32, 64, 128	8, 16, 32, 64, 128
Kernel size	(3, 3)	(1, 3, 3)
Stride	(1, 1)	(1, 1)
Padding	(1, 1)	(0, 1, 1)

Table 3.3: 2D and 3D convolutions comparison.

At the end of the encoder, a bottleneck with an additional temporal convolution is included. This is supposed to further enhances the learning of temporal features.

As mentioned, the decoder path is symmetrical to the encoder. Instead of simple upsampling techniques such as nearest-neighbor or bilinear interpolation, transposed convolutions are performed before every block. Unlike upsampling methods, which do not have learnable parameters, transposed convolutions include these, potentially enabling the model to learn more complex patterns during the upsampling operations. Transposed convolutions double the feature map spatial dimensions whereas reducing the number of feature maps. Considering the previous example, a tensor with dimensions (16, 500, 250) gets upsampled to (8, 1000, 500) using a 2x2 kernel size and 2x2 stride.

Skip connections are established between the final semi-3D convolutional layer of each encoder level and the corresponding upsampled output in the decoder. In this way, the high-resolution features learnt by the encoder are stacked with the upsampled features from the decoder. As a result, the model is supposed to better recover the spatial details lost during the downsampling step in the encoder and improve the predictive skills.

The final decoder layer reduces the number of feature maps from 8 to just 1, generating a grayscale image. This output is passed through a Sigmoid activation function (Equation 3.2, right image of Figure D.5) to produce the output segmentation map, with pixel values normalised in the range 0 – 1.

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (3.2)$$

where x represent the logits, the raw output of the network.

3.2.2. Hyperparameters

Table 3.4 contains the hyperparameters, constant parameters of the training process, and their values set for this model. The number of epochs represents the number of complete passes through the entire training dataset. One epoch is completed after the model has processed all batches in the training dataset. The batch size refers to the number of training samples processed together in one iteration, before the model updates its weights.

The optimizer is the algorithm responsible of updating the model weights and biases to minimise the loss function. It does so by adjusting the parameters after every batch of data during training, and eventually after every epoch. The learning rate controls how much the model parameters are adjusted in response to the computed gradients during backpropagation and determines the size of the steps taken by the optimizer in minimising the loss. The scheduler step is the interval between two epochs at which the learning rate is adjusted with the decreasing factor, a fraction of the current learning rate that specifies how much the learning rate is reduced.

The code was adapted from <https://github.com/tha-santacruz/BayesianUNet>. The model was implemented in Python 3.9.17 using the Pytorch library 2.0.1 (Paszke et al., 2019). It is composed of $5.23 \cdot 10^5$ parameters and has a size of 1.99 MB. It was trained on a NVIDIA GeForce RTX 3050 Laptop GPU, with 4 GB of dedicated RAM and 2048 cores. Appendix D contains further details on the various operations performed by the model.

3.2.3. Performance metrics

As we face a pixel-wise binary classification task, the appropriate loss function to evaluate model performance is the BCE loss. This function measures the difference between predicted probabilities, in the range

Hyperparameter	Value
Number of epochs	100
Batch size	16
Optimizer	<i>Adam</i>
Learning rate	0.05
Scheduler step	15
Scheduler decreasing factor	0.75

Table 3.4: Model hyperparameters and values.

0 – 1, and the actual target values, either 0 or 1. The function reads:

$$BCE_{loss} = -\frac{1}{N} \sum_{i=1}^N y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (3.3)$$

where y_i represents the pixel true label, \hat{y}_i the predicted pixel probability to be positive class, and N the total number of samples.

An attempt to include two physics-induced loss terms was performed. These terms represented the total areas of erosion and of deposition. However, the experiment failed. The results presented in the following Sections are achieved considering the BCE loss only. More details on the physics-induced terms are provided in Appendix E.

The training and validation losses were tracked during the model training process. The parameters of the model that achieves the minimum validation loss were stored. From here on, this model is referred to as the *min loss* model.

Throughout the model training, the following validation performance metrics were monitored as well.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (3.4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.5)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.6)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (3.7)$$

$$\text{CSI} = \frac{TP}{TP + FP + FN} \quad (3.8)$$

where true positive (**TP**) represents the number of correctly predicted positive pixels, false positive (**FP**) the predicted positive but actually negative pixels, false negative (**FN**) the predicted negative but actually positive pixels, and true negative (**TN**) the correctly predicted negative pixels. Positive represents the class with value equal to 1, whereas negative the one with value equal to 0. For our problem “Water” represents the positive class and “Non-water” the negative one. To compute these metrics, binary predictions are required. Binary predictions are computed from the model predicted probabilities, which were rounded according to Equation 3.9.

$$\hat{y}_i = \begin{cases} 1, & \text{if } \hat{y}_i \geq w_{thr} \\ 0, & \text{if } \hat{y}_i < w_{thr} \end{cases} \quad (3.9)$$

with $w_{thr} = 0.5$ water threshold.

Besides predicting the overall morphological evolution of the river, a significant goal of the model is to predict the location of the “Water” areas. Consequently, it is crucial to correctly forecast the riverbank locations, as it is of more concern to know where these “Water” areas are going to be found in a year ahead from now. For this reason, after minimising the **BCE** loss, it is of interest to maximise the validation Recall. When this metric is at its maximum, the number of **FN** pixels is minimised. It is crucial to achieve this as **FN** pixels might provide the misconception that the river will not be present in those specific areas. Such predictions can eventually mislead engineers and designers, who might propose an intervention or to build a structure in areas that are wrongly considered to be safe. For this reason, in addition to the min loss model, the one with the maximum validation Recall was stored and used for the testing phase as well. From here on, this model is referred to as the *max recall* model.

On top of that, the current classes are highly imbalanced, with the “Non-water” class being overrepresented compared to the “Water” class. To this end, **CSI** is recommended when interested in predicting the positive class while dealing with such imbalanced datasets.

Additionally, two points have to be remarked. Firstly, every statistical measure, such as the **BCE** loss and the other metrics, condenses all data in one single number. As a result, this number emphasises “certain aspects of the error characteristics of the model performance” (Chai and Draxler, 2014). Moreover, one should be aware of the “double penalty effect” (Bosboom, 2019). This phenomenon refers to the fact that, when dealing with grid-point based metrics in morphological predictions, correct predictions, in terms of magnitude and extension, but misplaced in space are penalised twice. In other words, there is a double negative contribution to the loss and metrics of the morphological features. In turn, the loss gets larger and the other metrics decrease, although the model predictions are misplaced in space but conceptually correct when visually inspecting them. Assessing the goodness of a model solely on such metrics might lead to a biased evaluation, and a visual inspection of the results is always necessary.

3.3. Benchmark models

3.3.1. Neural Network

This model was introduced in Section 2.3.2.1. It was developed by Jagers (2003) for predicting the bank erosion of the Jamuna reach. Several network configurations were tested. Eventually, the model with one hidden layer composed of five hidden units was chosen. Binary satellite images containing “Land” and “Water” pixels were employed. Only the seven nearest pixels to each channel were used to train the network. This corresponds to a distance of 350 m, as the image resolution was 50 m. From these pixels, five physical variables were parameterised and used as inputs of the network. Namely, these inputs were the distance to the nearest channel, the sine and cosine of the angle between the direction to the nearest channel and the flood flow direction, the local width of the nearest channel and the fraction of the 11×11 pixels neighbourhood. The output layer was eventually activated with Sigmoid activation (Equation 3.2).

3.3.2. No-change method

The benchmark for which no morphological change occurs is the second method considered for comparison. This method assumes that no change occurs among two consecutive years. As a result, the morphological conditions in year i , which we aim to predict, are the same as those in year $i - 1$. For simplicity, from here onwards this benchmark is referred to as the *no-change method*.

4

Results

This chapter contains the results of the JamUNet model. In Section 4.1 an overview of the losses and metrics is given, with graphical evaluation tools and a visual assessment of the results. In Section 4.2 JamUNet performance is compared to the benchmarks methods previously introduced.

4.1. JamUNet predictions

4.1.1. Loss and metrics

Here we only present the results achieved by the min loss model. From a visual inspection, the results of the max recall model look similar compared to those of the min loss model. The difference lies in the amount of predicted “Water” areas, that is larger for the max recall one. While predicting more “Water” areas might seem advantageous because of the class imbalance, overly conservative predictions are not ideal. Appendix G contains the most relevant results of the max recall model and a brief discussion on its applicability.

Figure 4.1 shows the evolution throughout the epochs of the training and validation losses (left plot) and the validation metrics (right plot), computed applying a water threshold of 0.5.

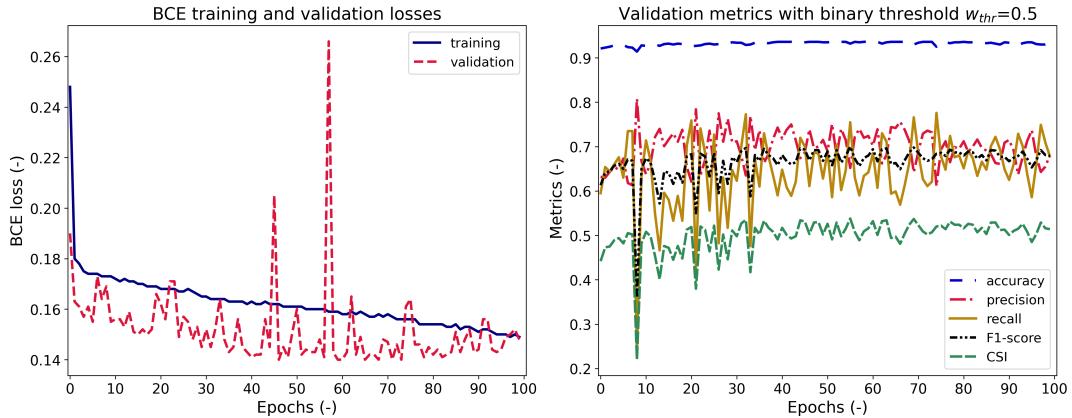


Figure 4.1: The left plot shows the evolution of the training and validation loss. The right plot shows the evolution of the validation metrics computed setting $w_{thr} = 0.5$.

The sharpest decrease in training loss occurs during the first epoch, after which the loss continues to decrease but at a slower, constant rate. In contrast, the validation one, besides several spikes and fluctuations, shows an initially decreasing trend. It reaches the minimum value of $1.40 \cdot 10^{-1}$ around half of the simulation (epoch 47), after which an average increasing trend is observed. This condition represents overfitting. The maximum Recall of 0.776 occurred at epoch 33. This behavior is common in deep learning but may be affected by the limited size and lack of diversity of the training dataset. Due to class imbalance, the model is likely to capture the minor variations of the “Non-water” areas, particularly the consistently overrepresented lateral regions,

after the initial epoch. The slower subsequent loss reduction is expected to be primarily associated with the prediction of the “Water” class.

Model	BCE loss	Accuracy	Precision	Recall	F1-score	CSI
JamUNet	$1.426 \cdot 10^{-1}$	0.935	0.720	0.710	0.713	0.555

Table 4.1: Testing loss and performance metrics for JamUNet model. All quantities are adimensional.

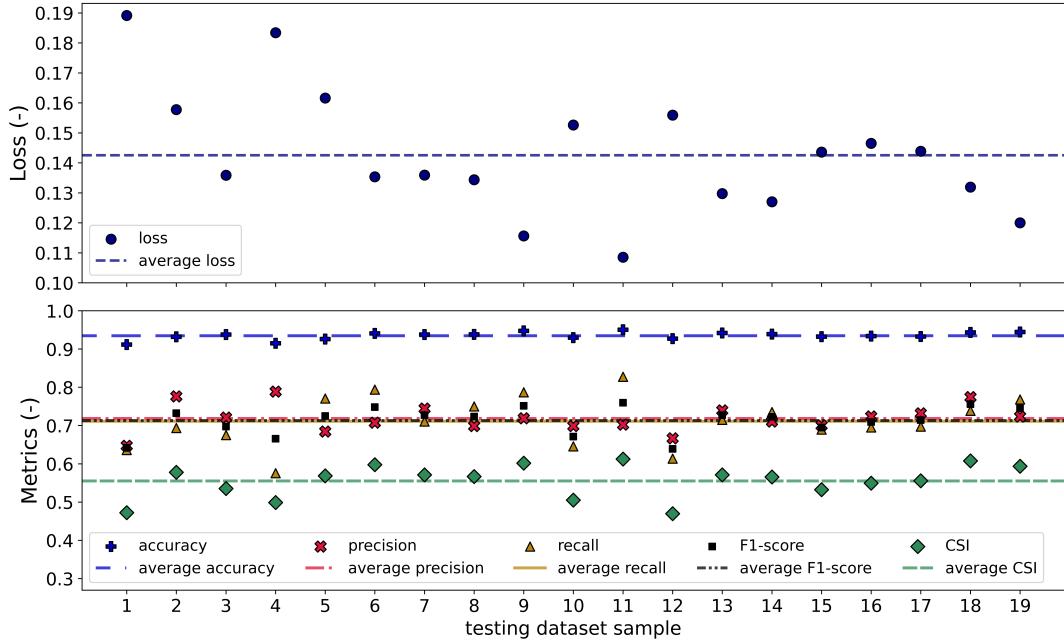


Figure 4.2: Loss and metrics across the testing datasets samples.

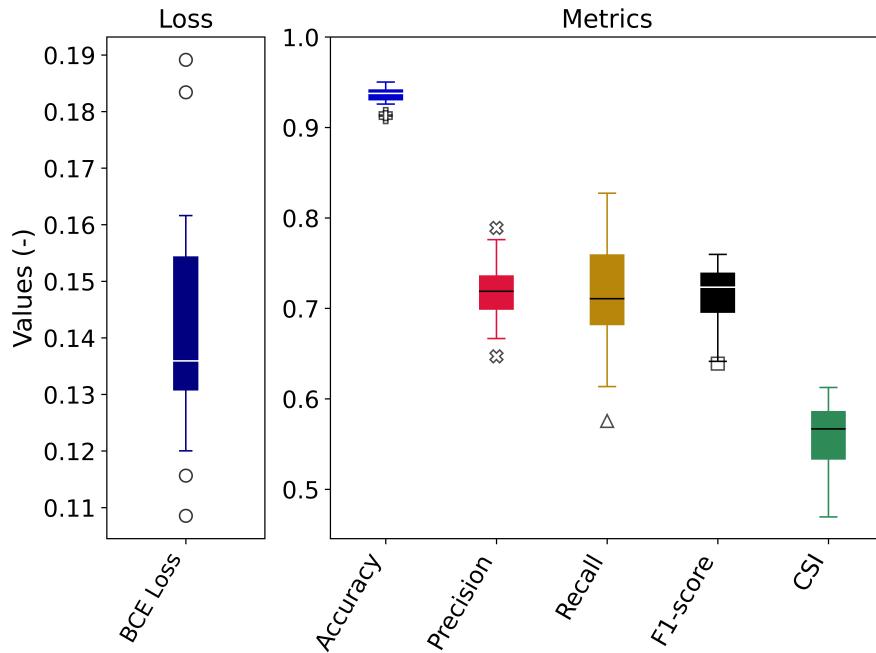


Figure 4.3: Boxplots of the loss and metrics for the test dataset. It shows the quartiles of the distribution, including outliers. The quartiles represent those values of the distribution that divide it into four equal parts, namely the 25%, 50%, 75%, and 100%.

Table 4.1 contains the average values of the loss and metrics for the testing dataset. Figure 4.2 shows the loss

and metrics across all testing samples, whereas Figure 4.3 shows the boxplots of the loss and metrics across the testing dataset.

Despite the small scale, the loss is subject to significant variation among the dataset samples. This relates to different levels of difficulty of predictions. When minor morphological changes occur, the model shows better performance. On the other hand, the quality of predictions decreases when larger changes occur. Accuracy shows a limited spread of the distribution, which again remarks the evident class imbalance. Several outliers are found for Precision, which in turns seems having a small spread. Recall shows some variability, and the sample with the highest Recall and CSI is also the one with the minimum loss (sample 11), which was expected. A similar reasoning holds for F1-score and CSI.

The sample with the minimum Recall (number 4) does not correspond with the one with minimum CSI (number 1), which is also the one with the largest loss. This has to be attributed to the lower value of FP predictions, expressed by a larger Precision as well.

Some inherent variation is present across the whole dataset. It is possible to observe that for the samples with smaller losses, all metrics are above average apart from Precision, the highest value of which is achieved by the sample with the second largest loss.

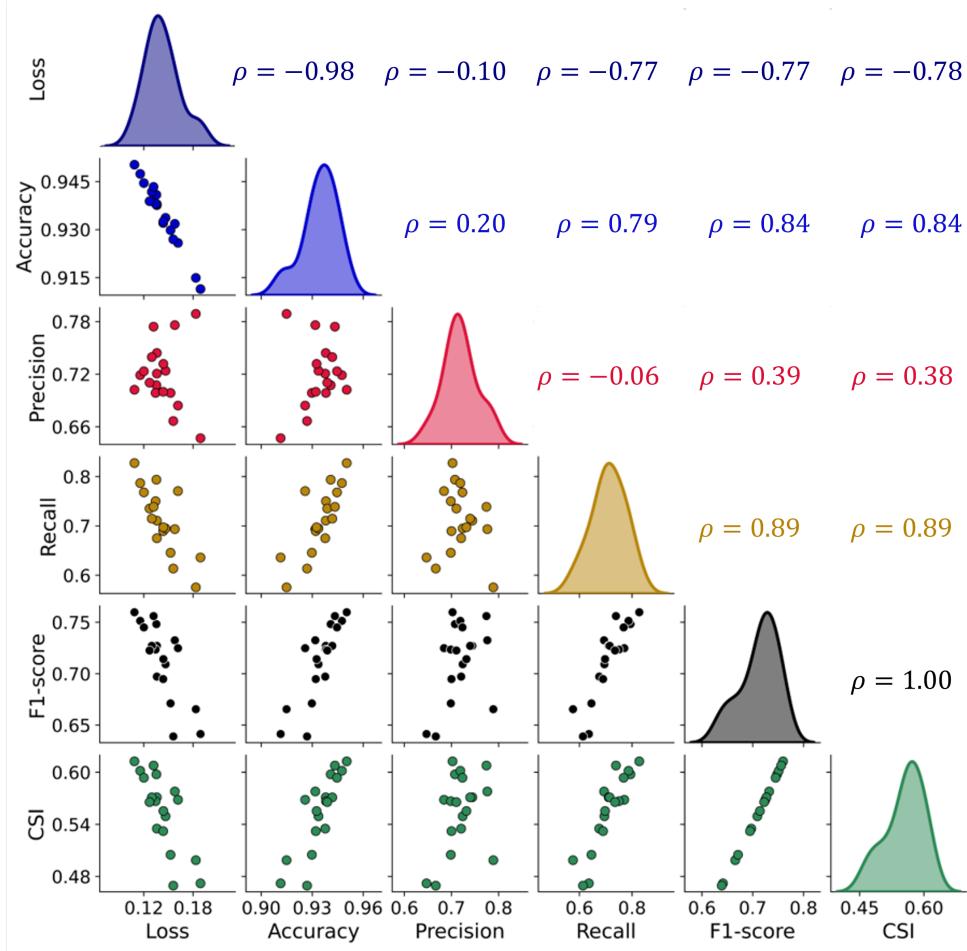


Figure 4.4: Correlation matrix between the loss and metrics. The matrix is symmetric. The values present in the above-diagonal part of the matrix are the Pearson correlation coefficients between the two variables, computed with Equation 4.1.

Additional information on the correlation between loss and metrics for the testing dataset is provided by Figure 4.4, which shows the pair plots between these. It includes also the values of the various Pearson correlation coefficients, computed as follows:

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (4.1)$$

where X and Y are the two variables, $\text{cov}(X, Y)$ is their covariance, and σ_X and σ_Y are their standard deviations. Nearly all metrics exhibit a negative correlation with the loss, with correlation values reaching up to $\rho = -0.78$. All metrics are strongly positively correlated with one another, exhibiting $\rho > 0.79$, with the exception of Precision. This metric stands out as the only one that does not demonstrate a relevant correlation with the other metrics, including the loss, although it shows slightly larger values with respect to F1-score and **CSI**. Notably, the correlation between F1-score and **CSI** is $\rho = 1$. This happens because the only difference between these metrics is the factor $\frac{1}{2}$ in the denominator of the F1-score (see Equations 3.7 and 3.8).

The goodness-of-fit of the metrics depends on the problem and task. As a rule of thumb, the classifier performance is good when the metric scores reach $0.7 - 0.8$, very good for metrics up to $0.8 - 0.9$ and excellent for scores larger than 0.9 . This holds for all metrics, with the exception of **CSI**, for which good performance are achieved with scores above 0.5 . Ideally, all metrics should be larger than 0.5 , as this threshold indicates the performance of a random classifier. In this case, accuracy is the only metric consistently larger than 0.9 , but that is also due to the relevant class imbalance. All metrics reach promising but moderate values (≈ 0.7). Larger values are desirable to enhance the model reliability and effectiveness.

4.1.2. ROC and PR curves

In binary classification tasks, the receiver operating characteristic (ROC) curve is an additional visualisation tool used to evaluate the model performance across different classification thresholds (Fawcett, 2005). It plots the model Recall, also known as true positive rate (TPR), against the false positive rate (FPR):

$$FPR = \frac{FP}{FP + TN} \quad (4.2)$$

Furthermore, the area under curve (AUC), representing the integral of the ROC curve, was computed. The AUC measures the probability that the model will correctly rank a randomly chosen positive sample higher than a randomly chosen negative sample (Google, 2024). An AUC of 0.5 indicates a model with no discriminative power (i.e., a random classifier), whereas a higher AUC signifies better model performance, with a value closer to 1 indicating excellent performance.

Additionally, the precision-recall (PR) curve is another visualization that shows how Precision and Recall vary with different classification thresholds. This curve is particularly useful for imbalanced datasets, as it highlights the model effectiveness in predicting the positive class when this is rare (Fawcett, 2005; Google, 2024). The PR curve allows to compute the binary water threshold that maximises the F1-score, i.e. that balances Recall and Precision. An optimal threshold of 0.402 was found. This is close to the value of 0.5 that was used to compute the validation metrics during the training process. An inspection of the model outputs with both thresholds was performed. Eventually, the original value of 0.5 will be kept for further discussion. This was chosen because a lower value leads to a general overprediction of “Water” areas, which might seem conceptually correct but turns out to provide too conservative results.

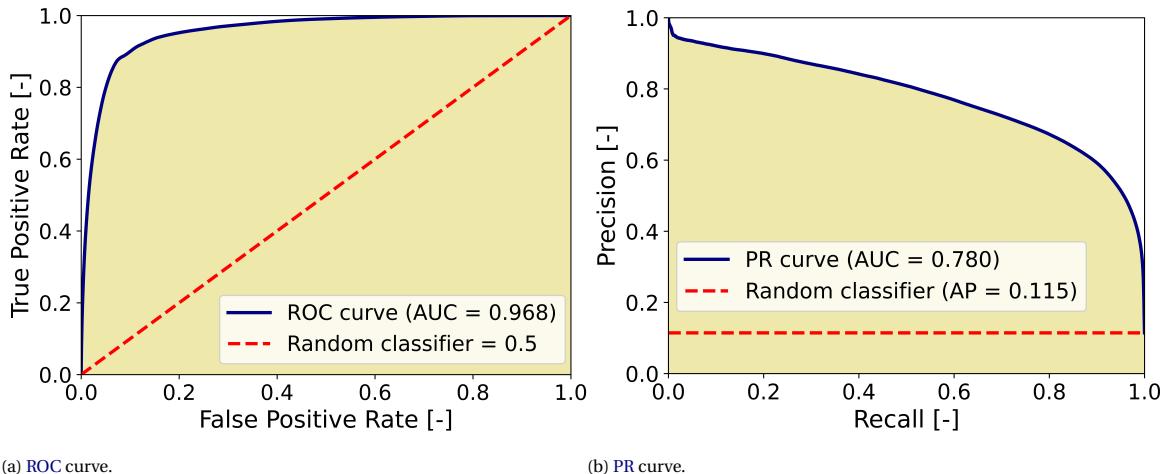


Figure 4.5: ROC and PR curves for the min loss model.

Figure 4.5a shows the ROC curve and Figure 4.5b shows the PR curve. The AUC value of these curves is 96.8% and 78.0%, respectively. The classification skills of the model can be considered adequate. However, one should not be misled by the large value of the ROC AUC. As we deal with a highly imbalanced dataset, the model can easily adapt its parameters to predict the most represented “Non-water” class. Nonetheless, the relevance of the model is to predict the positive class. The PR curve provides a more reliable evaluation. The model still seems to provide overall good performance, clearly outperforming the baseline of a random classifier. However, the ROC AUC is larger than the PR AUC. Hence, the model performs well in identifying the negative class but may struggle more when predicting the positive class, which is of greater interest.

4.1.3. Visual assessment

Figure 4.6 shows the four inputs (first row), target (second row, first image), and predicted image (second row, second image). It also shows the misclassification map (second row, third image), computed as prediction minus target, and the total real and predicted areas of erosion and deposition (second row, last image) of the testing sample 2, representing 1993 conditions. Other relevant figures of the testing dataset predictions, representative of different target years, are included in Appendix F. The link between sample number and target year is given in Table C.1.

The overall image reproduction is realistic, as the general structure of the river is recreated. Most of the expected features such as the main channels, secondary channels, meandering bends, confluences and bifurcations, and even some small islands and bars, are present. However, when looking at specific areas and single pixels, it is possible to observe that the model misses some predictions. This can be noticed by the amount of green and red areas in the misclassification map. The green areas represent conservative locations, where the model predicts “Water” but that are actually “Non-water” (i.e., FP predictions). In contrast, red areas represent dangerous locations, where the model predicts “Non-water” but that are actually “Water” (i.e., FN predictions).

Figure 4.7 also shows the predicted and real erosion and deposition locations for the same testing sample 2. This allows to look more closely on the model performance in predicting erosion and deposition sites. Point ① shows how JamUNet misses the formation of a new channel. Point ② shows the overprediction of the migration of a meandering channel. Point ③ shows the correct generation of a confluence between two close “Water” areas. Point ④ shows the overprediction of channel abandonment in the centre of the braidplain, whereas point ⑤ shows the correct prediction of channel abandonment in the outer side of the braidplain.

Relevant patterns were observed from inspecting all samples from the test dataset. JamUNet performs generally well in predicting the “Water” areas of the main channels. In some occasions the model does not predict the main channel shifting. In two cases, sample 1 and 4, the development of a bifurcation is involved. In the first case, the model predicts one branch to stabilise, whereas it actually closed. In the second case it is the opposite, where a small channel gets larger and becomes more dominant, whereas the model predicts it to remain stable. In general, channel shifting represents a challenging task, as it is complex to predict an abrupt and significant displacement of (large) channels. Besides that, the main channels are always well captured. Figure 4.6 shows that this is valid especially for the most central part of these channels, whereas the closer the pixels are to the banks and the less accurate the predictions are.

Channel abandonment is another feature usually well predicted. This holds especially, but not only, for secondary channels found on the outer side of the braidplain, i.e. not between two larger channels. However, within the braidplain JamUNet seems to slightly overpredict (secondary) channel closure. Hence, when in such locations the model output does not show many “Water” areas one should be critical, and those results should not always necessarily be interpreted as channel closure.

In contrast, and similarly to channel shifting, predicting the formation of a new channel is a particularly challenging task for the model, both in the inner and the outer side of the braidplain. Most of the areas of erosion that show channel formation are missed. The only situation when channel formation is better predicted occurs when two “Water” areas, especially main channels, get close to each other, eventually generating a confluence (see Figure 4.7).

The downstream migration of meandering channels is generally captured too. The overall process is often underpredicted. This leads the model to predict less erosion on the outer bend and less deposition, or sometimes no deposition at all, on the inner bend. Nonetheless, in other occasions the predictions are more ac-

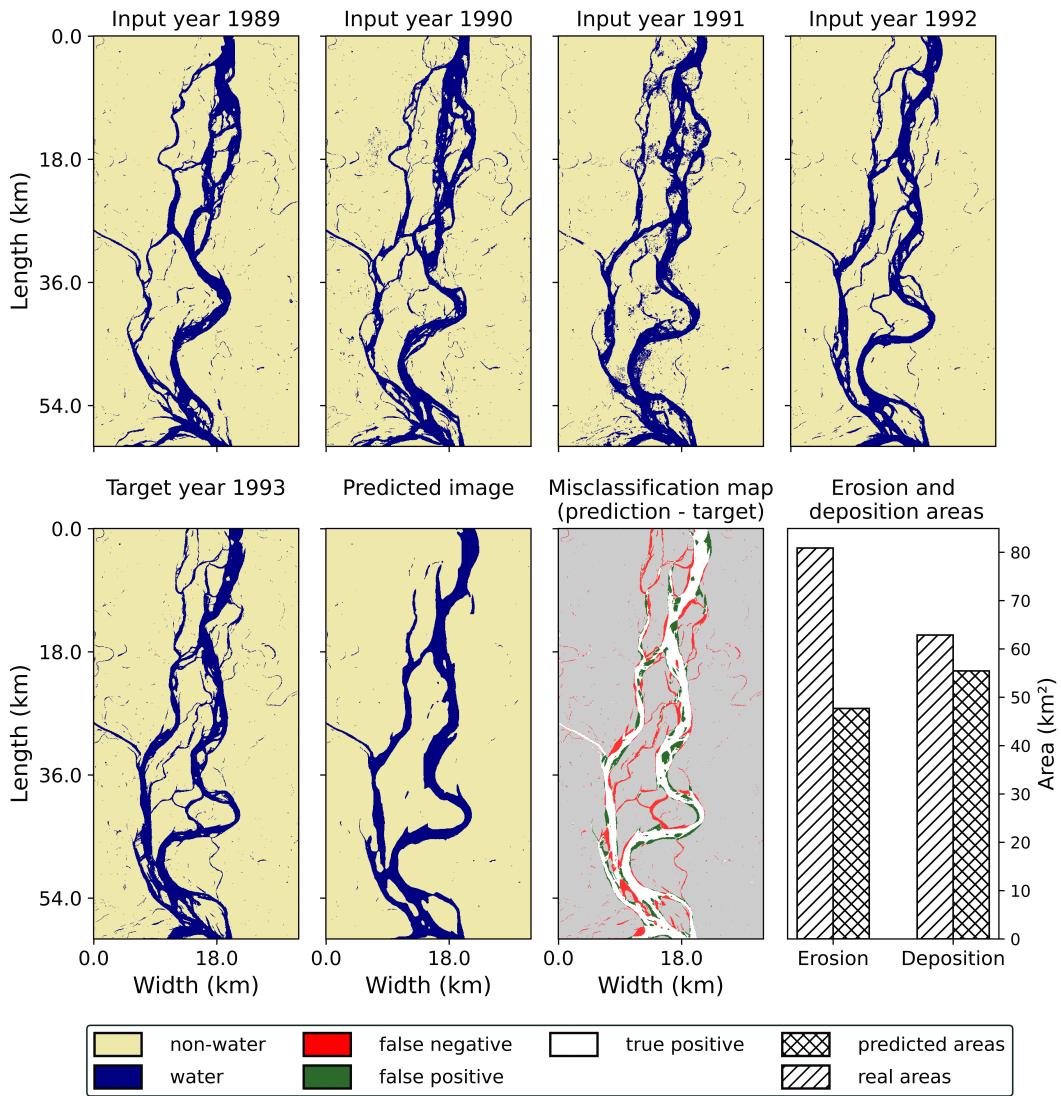


Figure 4.6: The first row of images contains the model inputs of testing sample 2. The first image of the second row represents the model target, the second image is the model prediction using the min loss model. Yellow and blue areas represent the "Non-water" and "Water" areas, respectively. The third image is the misclassification map, i.e. the difference between model prediction and target. Green and red areas represent FP (predicted "Water" but actually "Non-water") and FN (predicted "Non-water" but actually "Water"), respectively. The last image represents the real and predicted total areas of erosion and deposition (crossed diagonal and diagonal hatches, respectively).

curate, and regions where overprediction occurs were observed too. In general, caution is required as a clear recognition of situations of under- and overprediction is not determined yet.

The evolution of confluences and bifurcations of large channels is almost always accurately reproduced. However, especially when the overall morphology shows significant changes, a spatio-temporal evolution and downstream migration seems to lack. This results in bifurcations remaining close to their previous year location, with the diverging channels showing little downstream progression. Hence, erosion is underpredicted in these locations. A similar issue occurs with confluences, predicted further upstream than their actual positions. This causes converging channels to remain more static and leading to larger predicted "Water" areas. Consequently, deposition is underpredicted here. However, several observations of correct bifurcation and confluence development were found, and one particular occasion with an upstream migration confluence point was also predicted.

Finer morphological details, such as bars and small islands, are sometimes reproduced too. The location and extent of these features is not always correct. However, good predictions are present as well, with few cases of bar accretion, although being located in close but wrong locations.

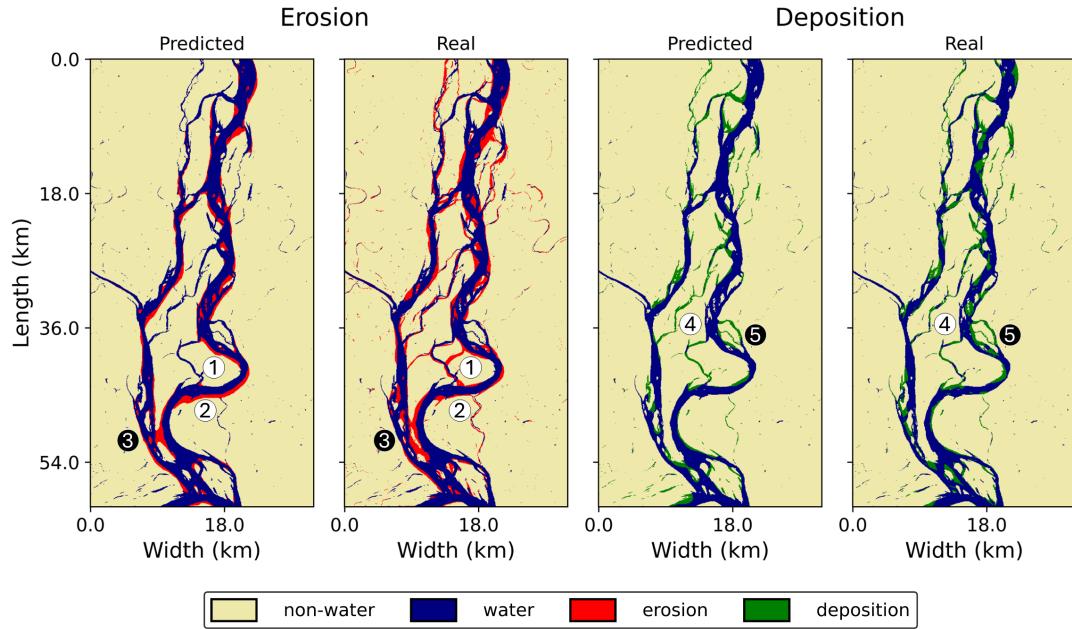


Figure 4.7: Predicted and real erosion and deposition locations of testing sample 2. Point ①, ②, and ④ represent wrong predictions. Point ③ and ⑤ represent good predictions.

Interestingly, JamUNet also captures the effects of bank protection measures. The Kojuri revetment was implemented in 2011 in Sirajganj District, in Bangladesh. Point ① in Figure 4.8 shows that in 2012, the year following the implementation of these measures, the model predicts some erosion to occur along these bends (left image), whereas in fact erosion did not happen (right image). However, for the consecutive years (from 2013 onwards), the model correctly predicted no erosion in these locations (point ② in both left and right images).

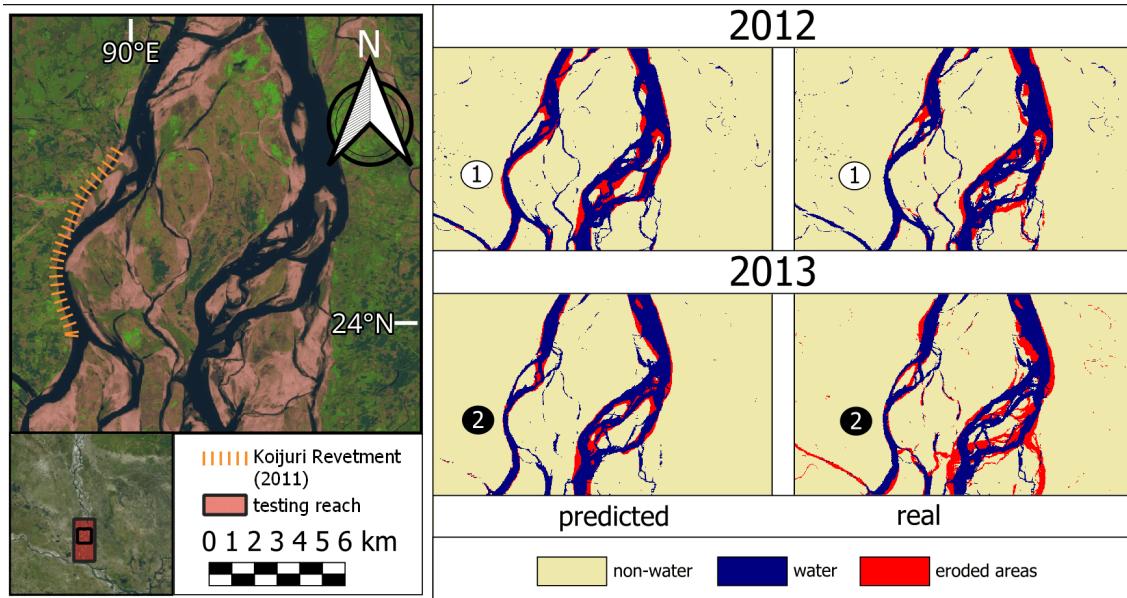


Figure 4.8: The left image shows the revetment location along the Kojuri channel, in Sirajganj District, Bangladesh. The right image shows the 2012 and 2013 predictions of erosion locations (left column) and the real erosion locations (right column). As it is possible to observe, the year following the implementation of the measure, JamUNet predicts erosion to occur, whereas in fact it did not happen (①). However, the year after it already correctly predicts almost no erosion in the same location (②).

The model generally does not seem to provide accurate predictions of the total area of erosion. The predicted values are always found within 30 and 50 km², whereas the real areas of erosion span from 40 to 120 km². The left image of Figure 4.9 shows the large difference between the two distributions, both in terms of density of probability and magnitude. Figure 4.10 also shows the low correlation between predicted and real areas of erosion, with a low correlation factor of 0.20. In contrast, the predicted values of the areas of deposition align more closely to the real ones, besides being slightly underpredicted as well. The total predicted areas span between 30 and 80 km², whereas the real values are included between 50 and 120 km². This is visible in terms of frequency and magnitude, slightly shifted to smaller values (right image of Figure 4.9), and is also confirmed by the larger correlation factor of 0.76 (Figure 4.10).

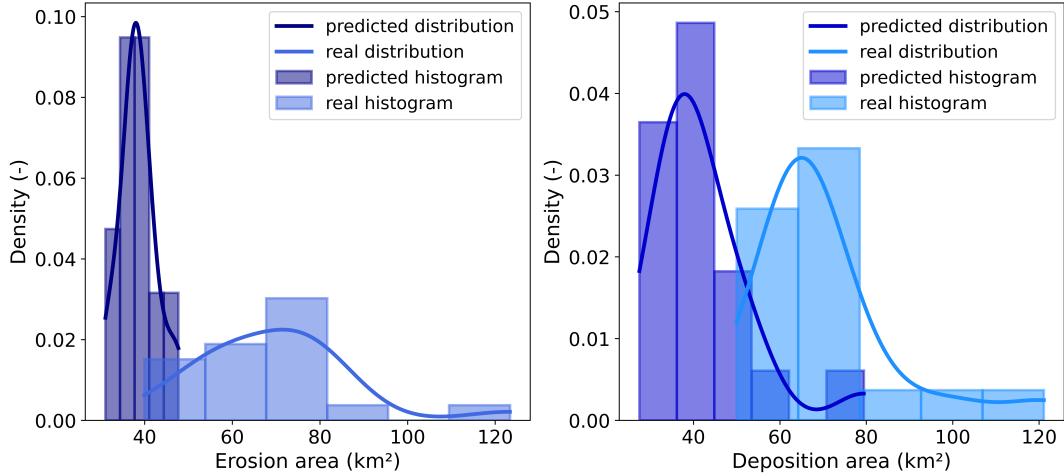


Figure 4.9: Distribution of the predicted and real areas of erosion and deposition.

4.2. Benchmark models comparison

Currently, no binary semantic segmentation model exists for predicting the overall planform changes of braided (sand-bed) rivers. As a result, the comparison with the following benchmark models might result biased in favor of JamUNet.

Additionally, the NN developed by Jagers was trained and tested using a different dataset. Since it was not possible to retrain this model within the framework of this study, only a qualitative comparison is performed.

Model	Accuracy	Precision	Recall	F1-score	CSI
JamUNet	0.935	0.720	0.710	0.713	0.555
No-change method	0.922	0.660	0.662	0.661	0.494

Table 4.2: Comparison between testing metrics for JamUNet and the no-change method. All quantities are adimensional.

The performance metrics of the no-change method are provided with those of JamUNet in Table 4.2. Despite accuracy still being high, again inflated by the significant class imbalance, all metrics reach lower values than those of JamUNet. For Precision and Recall this means that the number of both FP and FN increases, reducing the performance reliability of this method.

The computation of the BCE loss for the no-change method is meaningless. The target pixels contain binary values, either 0 or 1, instead of probabilities. The comparison would be too biased in favor of JamUNet, hence it was not included.

From a visual inspection of the predictions and misclassification maps, the no-change method misses the development of nearly all morphological features. Figure 4.11 shows the same prediction of Figure 4.6 achieved by the no-change method. This methods can indeed provide the location of the main channels, especially considering their centrelines, apart from when major shifts occur. However, it always misses all other morphological processes, such as meander migration, channel abandonment and the prediction of the river bank locations. In addition, this model could not predict any location of erosion or deposition. Presuming that no planform change happens in such dynamic rivers is a particularly strong assumption, which is proven to be

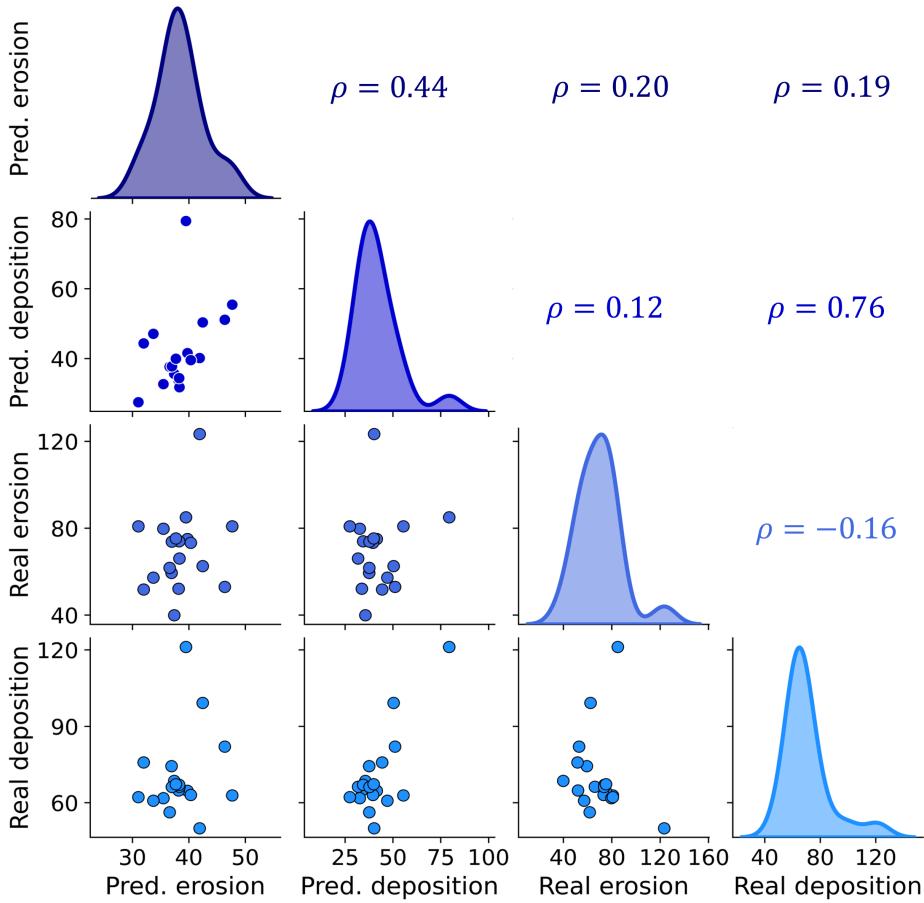


Figure 4.10: Correlation matrix between the predicted and real erosion and deposition areas. The matrix is symmetric. The values present in the above-diagonal part of the matrix are the Pearson correlation coefficients between the two variables, computed with Equation 4.1.

non-representative of the real conditions of the river.

No visual comparison is possible between JamUNet and the NN predictions. The only image that was generated by this NN represents one of the reaches used for training JamUNet. Additionally, this image represents the 1995 morphological conditions. The input-target combination with the 1995 image as target was discarded from JamUNet training dataset because of the fully “Non-water” image of 1994. As no visual assessment can be performed either, the only possible comparison is done with regard to the understanding of morphological concepts by the NN. This model was reported to provide reasonable predictions of bank erosion at several locations, primarily on the east bank of the river. Secondary channels and bar complex represented the largest sources of error. Formation of new channel was not predicted by the network, likely because such developments were not present in the training data. As no input information of river interventions and training works was included, the model could not assess the impact of human interventions. That holds both on the location of the interventions and upstream and downstream of it. Similar issues were encountered with JamUNet predictions.

Eventually, the NN was reported to learn the following main rules:

- The closer a point is to a channel, the higher the probability of erosion.
- Channels tend to migrate in downstream direction.
- The wider the channel, the faster it migrates and the more land it erodes.
- Erosion is more likely to occur for land pixels surrounded by a large amount of water than those surrounded by less water.

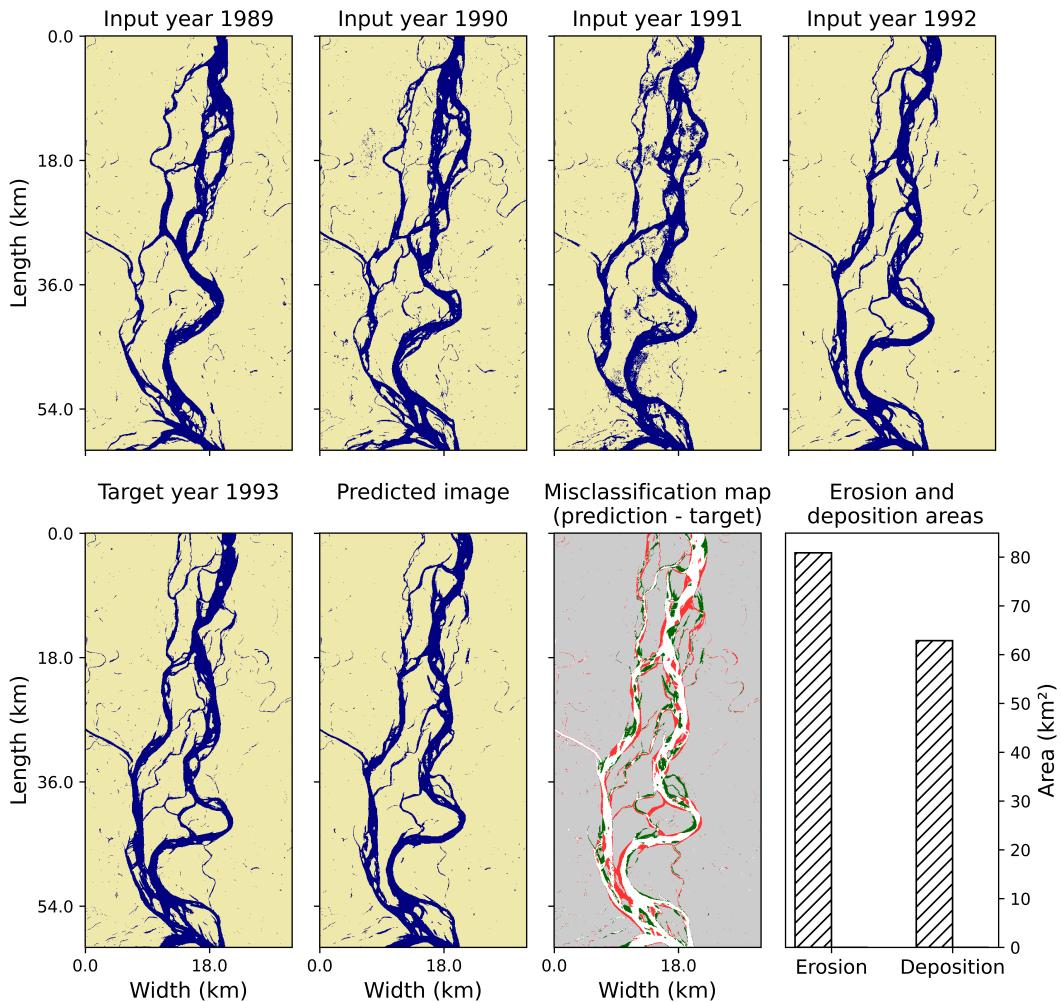


Figure 4.11: Example of the reach evolution using the no-change method for sample 2 to be compared with Figure 4.6. The legend is the same.

Thanks to the larger values of all metrics and the visual assessment, JamUNet is more recommended for predicting the morphological evolution of braided sand-bed rivers. Furthermore, although JamUNet underpredicts the total areas of erosion and those of deposition, the no-change method can not provide these values. Despite the uncertainties and failing in predicting some morphological processes, JamUNet outperforms the no-change method and can be considered a first valid alternative for morphological predictions of braided sand-bed rivers.

5

Discussion and limitations

This chapter aims to discuss the strengths and weaknesses of the proposed model, as well as the key understanding and interpretation of the outputs. Section 5.1 contains an interpretation of the model performance and limitations are assessed. Section 5.2 contains an overview of the limitations of this research.

5.1. Interpretation and limitations

The proposed model is able to represent the overall structure of the river, with almost all morphological features. Despite predicting accurately the locations of erosion and deposition, the former is generally underpredicted in terms of magnitude. Temporal patterns are often missing. This is confirmed by the evolution of confluences and bifurcations, especially considering their downstream migration, and by the migration of meandering channels. This issue is likely to be related to the model architecture. Despite the presence of 3D convolutions, these still operate as 2D convolutions given the dataset dimensions. Hence, the model still lacks a proper implementation of the temporal inductive bias that makes it suitable to process temporal sequences.

Channel formation and the main channel shifting remain challenging processes. These are highly influenced by local conditions, such as the elevation of the nearby areas as well as the flood evolution in time and space. Without these information, the model will always predict the average morphological behaviour, but can not forecast the effect of extreme events. Predicting these processes could be more complex compared to other processes also because of their minor occurrence in the training dataset. As also reported by Jagers (2003), data-driven models can not predict patterns that they did not process during the training.

The last input year flood conditions seem to influence the quality of predictions. The predictions miss the main channel shifting or the formation of new channels that occurred after larger-than-usual floods. Some of these floods occurred in 1991, 2004, 2013, 2016, 2017, and 2018. Larger-than-usual floods are not necessarily the largest floods recorded, like in 1998, when a maximum discharge of 102 535 m³/s occurred. Given the available data reported in Appendix A, channel shifting and channel formation are observed there is a combination of high water level and large discharge. These processes occur when the water level at Bahadurabad exceeds 19.5 m PWD and simultaneously the discharge at Bahadurabad surpasses the average flood discharge of 60 000 m³/s. However, within the testing years, there were some cases for which this combination of both high water level and large discharge did not cause relevant channel shifting nor channel formation. Additionally, in only one sample (number 10), channel shifting occurred despite the maximum discharge of the previous flood being 45 775 m³/s, lower than the average flood discharge. Further analysis of the relation between flood discharge and water level along with the prediction quality and the morphological processes observed are required. On the other hand, when the target year shows low-flow water levels lower than usual, the areas and locations of deposition are underpredicted. In this case, deposition does not relate to the silting-up and abandonment of channels, but rather to the exposure or drying of the riverbed.

In addition, the model seems to be sensitive to the last input year image. This is proven considering some particular predictions, like the stable secondary channels confluencing from west. The course of this channel does not change across the years. However, as Figure 5.1 shows, when in the last input year image the chan-

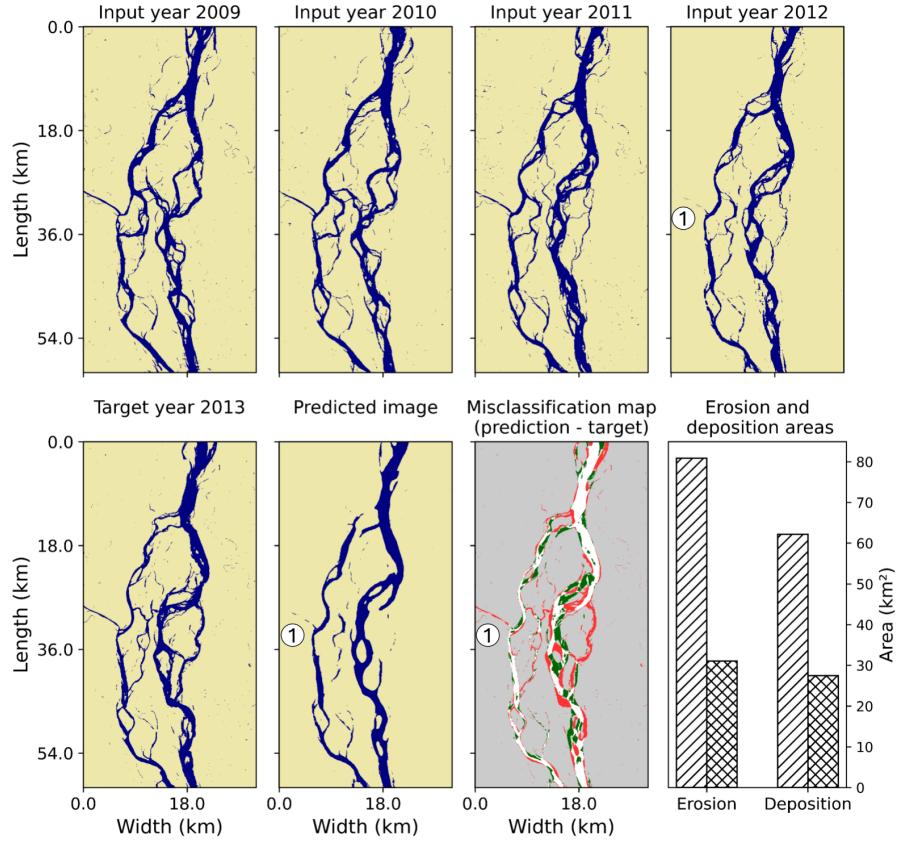


Figure 5.1: Example of model output sensitivity to last input image. The absence of the tributary channel on the left of the image, point ①, is reflected in the output image, which results in underpredicted erosion (although the channel is stable across all years). The figure shows the evolution of sample 12.

nel is partly missing, the model predicts the exact same area to be missing from that channel. A close look to confluences, bifurcations, and meander bends suggests that, at least for particularly dynamic samples, the model produces an output which resembles the last input year image.

This was also proven considering the optimal threshold found with the PR curve in Section 4.1.2. By decreasing the threshold to 0.402, more “Water” areas would obviously appear (Figure F.1). These areas are generally the locations where the river was present in the last input year. Hence, when producing an output, the model seems to assign a much larger weight to the pixel-wise “Water” occurrence rather than on the temporal features development. This also links with the previously mentioned partial lack of temporal evolution.

The model struggles the most in predicting the morphological evolution of those years affected by abrupt and fast changes. Furthermore, another limitation of the predictions is the presence of secondary channels disconnected from the main channels, which is for obvious reason not possible.

As for the prediction of bank protection measures, a similar reasoning holds for the Chauhali revetment. This was implemented in 2016 on the opposite side of the braidplain from the Kojuri revetment (Figure F.3). In the first year after the bank stabilization in Chauhali, the model predicted a major erosion spot (Figure F.4a). In the second year following the intervention, minor erosion was predicted as well (Figure F.4b). This might be caused by the relevant meandering migration that the channel experienced within those input years, as can be seen from Figure F.5. However, the extent of the predicted erosion was significantly limited compared to the previous year prediction. This fact remarks again the significant impact of the last input year influence on the overall model output.

The distributions of both the areas of erosion and deposition show a clear under-prediction compared to the real values. This is particularly enhanced for erosion and is attributed to the class imbalance. Some experiments were performed including these variables as further physics-induced loss terms to enhance the model performance. The experiments failed without showing relevant improvement of the training process.

Nonetheless, correctly predicting the total areas of erosion and deposition does not involve a clear understanding of the morphological processes and is not necessarily informative for the model. This is a further reason for which the attempt to include a physics-inductive bias might have failed. More details on this are given in Appendix E.

Eventually, it is possible to state that the model is able to represent the following concepts:

- In meandering channels, erosion and deposition occurs in the outer and inner bend, respectively.
- Two main “Water” areas that get closer to each other are likely to merge, representing the development of a confluence between two channels.
- In specific local regions, when “Non-water” areas start increasing throughout the input year images while “Water” areas decrease, the latter are likely to disappear. This would represent deposition phenomena, silting-up and consequent closure of (small) secondary channels. It also links to bar accretion, which was predicted in few occasions.
- In contrast, the “Non-water” areas located where revetments or bank protection measures are implemented are predicted to remain stable.

Furthermore, highly braided patterns hamper the model performance. In these cases, the model mainly reproduces the larger “Water” areas, i.e the main channels, but is not able to predict smaller features. It often creates disconnected secondary channels as well. Rapidly changing input images, which show large morphological variations between consecutive input years, also decrease the model performance.

5.2. Limitations of the research

5.2.1. Training dataset

As we deal with a purely data-driven model, the total amount and quality of the input data significantly influence the model output. For the current application and model, a training size of 518 input-target combination is enough to reach a testing loss equal to $1.426 \cdot 10^{-1}$ and Recall equal to 0.71. To reach smaller losses and larger metrics, more (diverse) data are required. However, the availability of satellite images is a hard constraint. For the current collection and river areas, no images are available prior to December 1987. Few other collections are available since the beginning of the 70s. However, the spatial resolution of these images is worse. Moreover, a different collection would require some modifications in the overall methodology, as different detailed preprocessing steps should be needed and a consequent classification algorithm should be implemented. Cloud coverage is another hard limitation that reduces the total number of suitable input-target combinations.

5.2.2. “No data” replacement

Besides reducing the amount of learnable morphological features, hence reducing the model performance and predictive capabilities, the presence of “No data” pixels might also induce instability in the model training and compromise the results interpretation. Replacing “No data” pixels also simplifies the modelling task by transforming it from multiclass to binary semantic segmentation, as the prediction of “No data” pixels is of no interest. These areas do not represent any relevant land-use class. However, the replacement strategy brings some uncertainty and could introduce some bias in the datasets.

The double-averaging process ensures that the replacement is done with average image conditions. The first step (seasonal average) can be deemed to be more accurate, as the assumption of no relevant morphological changes during the low-flow stage is based on recorded water level data. On the other hand, the second step (neighbours average) is more interpretable. The riverbank pixels and those located close to it are more likely to be misclassified, since they might be surrounded by similar amounts of both “Non-water” and “Water” areas. The replaced class is determined by the average, therefore having just one additional pixel from the incorrect class can cause the “No data” pixel to be assigned to the wrong class. As a result, small morphological features such as bars and islands located in the middle of larger channels, as well as small secondary channels completely surrounded by “Non-water” areas, would be misclassified.

Finally, if the two-steps replacement is not sufficient, the pixels are categorically assigned to the “Non-water” class. This was done because from a visual inspection of the images the vast majority of “No data” pixels lied in “Non-water” areas. However, for specific images this might not hold, leading to a wrong class assignment.

A deeper analysis on the correlation between “No data” and the eventual hard replacement class replacement is recommended. This should be done to check whether the assumption that most of these pixels lie in the “Non-water” areas is well grounded or not.

5.2.2.1. Validation and testing datasets

It is relevant to point out that the validation and testing datasets should be as much representative as possible of the real-world situation. Consequently, when applying the data replacement strategy, the losses and performance metrics are likely to be “artificially” inflated, hence not representing the real conditions of the river. Some bias might be introduced in the data, which causes the model to learn and consequently predict unreal patterns and features. This can also lead to overly optimistic predictions skills.

The model should be validated and tested on the original datasets, without replacing missing data. The “No data” pixels should be masked and the predictions, losses and metrics, should be computed only considering the classes of interest. This would also allow to assess the overall generalization capacities of the model, for instance checking whether the model is able to predict a specific class in the areas where it is expected to be present. However, when proceeding with this approach the prediction-target comparison is influenced by the presence of clouds. The potential applications of the model are also expected to reduce, as more uncertainties would be present in the prediction.

5.2.3. Image quality

Apart from the “No data” replacement, no further preprocessing step was applied. However, the image quality could still be improved. Consider for instance the presence of randomly scattered “Water” pixels (see Figure 3.3). Most of these pixels are isolated from the river and are found in “Non-water” areas. These pixels could represent secondary (or temporary) water bodies, such as small lakes, or could be noisy pixels that were wrongly classified. The “No data” replacement strategy might have increased the presence of some of these isolated “Water” pixels too. Their presence might confuse the model and negatively influence the learning process, which consequently degrades the predictions quality.

Additionally, it is often possible to notice the presence of small disconnected channels, which do not seem to originate from larger channels or water bodies. Their presence in “Non-water” areas might also trigger the model to predict disconnected channels. Since we are considering low-flow stage images, these channels might be partly abandoned or contain stagnant water, whereas they transport water during the high-flow stage. In this case, it is subjective whether it is better to completely remove these channels or connect them with the larger ones. As a personal consideration, those channels present in the braidplain should be connected, whereas the others can be removed, especially those further away from the braidplain.

On top of these considerations, the image resolution downgrade influence was explained in Section 3.1.1. Finer details would be lost, but the morphological features were still visible. This strategy allowed to reduce the memory usage and most importantly the computation time.

5.2.4. Hyperparameter optimisation

Manual attempts were done varying the learning rate, the kernel size, the pooling method, and the number of feature maps in the hidden layers. However, a proper hyperparameter optimisation study was not performed. Nevertheless, it is probably not relevant to implement it for the current model architecture. No significant differences were noted when manually varying these hyperparameters. This fact indicates that a hyperparameter optimisation study might not be required to improve the model performance, whereas other modifications of the methodology might have more influence.

6

Conclusions and recommendations

This chapter represents the conclusion of this research. Section 6.1 contains the answers to the research question and subquestions introduced in Section 1.3. Section 6.2 contains recommendations for the model application and future research.

6.1. Answers to research questions

This research aimed at determining to what extent deep-learning modelling can help in predicting the planform changes of braided sand-bed rivers using satellite images. To reach this goal, the research was divided in the following subquestions:

1. ***What is a suitable model architecture for this problem?***

For 2D spatial problems the most suitable deep-learning models are U-Net-based [CNN](#). That is due to the spatial inductive bias implemented through convolution operations. The U-Net architecture enables the extraction of spatial information with the encoder and its reconstruction with the decoder. We additionally implemented semi-3D convolutions to partially capture the temporal patterns as well.

2. ***What are the most suitable assessment metrics for this task?***

For binary semantic segmentation tasks the most suitable loss function is [BCE](#). It measures the distance between the predicted probability and the actual target class. Other assessment metrics exist. Recall measures the correctly predicted positive pixels out of all true positive pixels. In combination with the loss, recall is the most relevant metric to assess the model performance for practical applications. [CSI](#) also measures the correctly predicted positive pixels and is recommended especially when dealing with imbalanced datasets. Precision and F1-score are two further metrics, often used in combination with recall. Precision measures the amount of correctly predicted positive pixels out of all predicted positive. F1-score is the harmonised mean between recall and precision. Another common metric is accuracy. It measures how many correct pixels are predicted out of the total. However, when dealing with class imbalance datasets, this metric is inflated and not representative of the model performance. [ROC](#) and [PR](#) curves are additional graphical tools to assess the classification skills of the model.

3. ***Based on the available data, for which prediction timescale - short-term (i.e., less than one-year ahead, in the order of weeks or few months) or long-term (one-year ahead or longer) changes - is the model most suited?***

The timescale choice was based on the available collections, their characteristics and the expected applications of this model. A balance between a meaningful physical interpretation and the need for a large training dataset had to be found. For this reason, collections such as Sentinel-1 and Sentinel-2 were not considered due to the shorter time of record. Landsat products were initially considered, however eventually [GSWD](#) was found to be the most appropriate collection as it contains pre-classified images. This dataset is obtained from Landsat images. It contains one image per month, although during the monsoon season most images are not available because of the cloud coverage. The same prob-

lem would be present for satellites of similar type such as Landsat and Sentinel-2 collections, whereas Sentinel-1 is not affected. Consequently, low-flow season images were found to be the most suitable, also considering that protection works can be implemented during these months. Hence, one-year ahead predictions were chosen. Using the current year as the target, one representative image of the same month for each of the previous four years was used as input.

4. *Depending on the satellite image collection chosen, what image preprocessing techniques are required for improving the model performance, if there is any?*

The chosen [GSWD](#) collection contains pre-classified images with the following three classes: “No data”, “Non water”, and “Water”. The selection of this collection simplified the methodology by eliminating the need for a pixel-wise classification algorithm. Despite this, to reduce the problem to a binary semantic segmentation, the “No data” class was pixel-wise replaced through a double-averaging process. It was assumed that the water levels during the low-flow season are almost constant, hence the images look like the same (neglecting any relevant morphological change in this period). In this way, a pixel-wise seasonal average image was computed for all reaches and throughout all years, not considering the “No data” pixels. To the remaining “No data” pixels, a neighbouring average was applied. Eventually, if any “No data” pixels was left, a hard replacement was performed. “Non-water” class was assigned, as it was visually observed that most of the “No data” pixels were representative of this class. Eventually, in all datasets only the “No data” pixels were replaced by the corresponding pixels of the average image representative of that specific year. Prior to this step, images were standardized in terms of dimensions and flow characteristics.

5. *What relevant information (quantities, locations) on morphological processes can be inferred from the outputs of the model?*

JamUNet correctly predicts the main channels areas, only struggling mainly when shifting occurs. Meander migration is often captured as well, but sometimes underpredicted. As well, bifurcation and confluence development is generally captured but often underpredicted, resulting in smaller areas of erosion and deposition, respectively. Channel abandonment is almost always represented, especially in the outer side of the braidplain. In contrast, channel formation is a complex task which is almost constantly missed. Finer details such as bar complexes are less commonly predicted, although in few occasions forecast of bar accretion and downstream migration are correctly generated. The generation of disconnected (secondary) channels is another shortcoming of our model.

The total areas of erosion and deposition are both underpredicted. Predicted and real deposition values are more positively correlated, whereas a much lower correlation is found for erosion areas which is also more largely underpredicted.

6. *How does the model compare to (i) the NN developed by Jagers (2003) and (ii) the benchmark method for which no morphological change occurs?*

A quantitative comparison could be performed only with the benchmark method for which no morphological change occurs. Our proposed model outperforms it in terms of metrics and visual assessment. This benchmark is clearly less appropriate for dynamic systems as braided rivers. A qualitative comparison with the [NN](#) developed by [Jagers \(2003\)](#) was done. Similar issues were reported regarding the understanding of the [NN](#) of morphological patterns. Processes as channel formation, bar migration and secondary channels are common sources of error to both models.

However, as both benchmarks do not predict the overall morphological changes of the river, the assessment is positively biased towards JamUNet. Nonetheless, JamUNet proved to provide more promising results for predicting the overall planform changes of braided sand-bed rivers.

7. *What kind of physical information can be exploited to improve the model performance?*

An attempt was made by including the total areas of erosion and deposition in the training loss. Two additional terms were added, with the aim to enhance the feature learning process. A multiplication factor was applied to such terms in order to scale their weight on the overall loss. Larger values of such parameter greatly hindered the model learning, with the validation loss and metrics stagnating. Smaller values did not bring any contribution instead, and the additional terms were almost neglected in the overall loss. An appropriate value for such parameters was not found. Eventually, this physics-induced information did not improve the model learning.

With the above subquestions being answered, it is now possible to address the main research question:

To what extent can deep learning help in predicting the planform changes of braided sand-bed rivers?

A U-Net-based CNN model with combined 2D and semi-3D convolutions is able to reproduce nearly every morphological process. Meander migration, channel abandonment, and confluence and bifurcation development are correctly represented but their temporal evolution is often not correctly captured. More challenging processes, such as channel shifting and channel formation, are complex tasks for the model. These are almost never predicted, as they also depend on the flood magnitude, like peak flood discharges. It is also possible that these features were underrepresented in the training data. As a result, the model did not learn enough representations of these processes and consequently was not able to reproduce them. Additionally, the model also tends to underpredict the total areas of erosion and deposition.

Our model can reach a testing BCE loss up to $1.426 \cdot 10^{-1}$. Other classification metrics, such as Precision, Recall, and F1-score reach promising levels, between 0.71–0.72. CSI reaches lower values, around 0.55. The AUC value of the PR curve reaches a value of 0.78, marking the promising performance. However, improvements are required to achieve more reliable predictions. Larger metric scores, above 0.80 – 0.90, are desirable.

6.2. Recommendations

6.2.1. Model application

Given the uncertainties and the large morphological scale considered, JamUNet predictions are not recommended for small-scale planning. Besides that, the model correctly predicts the main channels location and provides information on erosion of meandering channels and channel closure. The min loss model correctly identifies the locations of bank erosion but often underpredicts their extent. To overcome this issue, a possible solution is to combine, or directly replace, its predictions with those of the max recall model.

Local (water) authorities, as well as local and international stakeholders, could use the proposed model for preliminary analysis and as additional source of information for the planning of larger interventions, such as bank protection measures along the banks more prone to erosion. Knowledge of channel abandonment could be relevant for the preservation of the river ecological status, fishing activities or for land reclamation plans. Caution must be used, though, since JamUNet performance tend to underpredict both erosion and deposition processes, and fail at predicting channel shifting and channel formation. Given the relation between larger planform changes, in terms of extent of erosion and displacement of sediment, and larger floods, even more caution is advised.

Although the model does not predict the channel depth, it is acceptable to assume that the deepest channels are also the largest ones. Consequently, inland water shipping and dredging companies responsible for maintaining the channels navigable could benefit from the predictions by knowing the evolution of the main channels.

Despite correctly showing positive performance in not predicting erosion along protected banks, JamUNet cannot be used for assessing the impact of future human interventions. This holds especially considering the upstream and downstream effects that these interventions may have on the overall morphological evolution.

6.2.2. Future research

6.2.2.1. Main recommendations

The following recommendations are of equal importance and should all be implemented. They are presented in the order in which they appear, or are expected to appear, in the methodology.

As we deal with a purely data-driven model, increasing the training dataset size is crucial. For this case study the amount of available areas almost reached its maximum, and few more areas could be included. As a result, other braided sand-bed rivers should be included. One possible example is the Congo River near Brazzaville ([Cournède and Mosselman, 2014](#)) or the Indus River in the Punjab region, in Pakistan. Some tributaries of the Brahmaputra, such as the Teesta River, the Kameng River and the Subansiri River might be included too. Caution is advised because these rivers are smaller compared to the Brahmaputra. As a result, the class imbalance would increase even more.

Another strategy involves reducing the image height. This reduction should ensure that even the largest morphological processes, like the migration of large meandering channels, are at least partially captured within

single images. In this way the possible reduction of the morphological understanding is minimised. This issue is expected to be overcome also by the larger amount of training data. Since the braidplain width can reach up to 17 km, more caution is needed for the image width reduction. 30×30 km or 40×30 km are the recommended dimensions for the images. 40 and 30 km represent the height and width of the images. Indiscriminate data augmentation is not expected to significantly improve the model performance and could, in fact, worsen the existing class imbalance.

The inclusion of more variables rather than only satellite images is also valid. Water level or discharge records are the main variables that should be included. These data could be given as time series, like daily, biweekly, or monthly records, or as single values representative of the time scale. For instance, the largest flood discharge or the higher water levels of the whole monsoon season could be chosen. Additionally, other variables such as the sediment discharge or the nominal diameter of the bed material could be included too, especially if other rivers are included in any of the datasets.

Nonetheless, one should be aware that most of these data are rarely available for long time scales and large spatial coverage, i.e. different locations along the river.

A different deep-learning model architecture should be tested as well. [ConvLSTM](#) is the first suggested model. This architecture implements both spatial and temporal inductive biases and it is likely to better capture the temporal evolution of the morphological features. A [GNN](#) inspired by the work of [Bentivoglio et al. \(2023\)](#) is another option. Such networks are composed of nodes and edges, across which information is propagated. For braided rivers modelling, nodes could represent confluences and bifurcations, while edges could represent channels. Moreover, directionality can be given to these graphs. This is relevant for hydraulic and morphological processes, such as downstream meander migration.

The Unified Focal loss ([Yeung et al., 2022](#)) or an Adaptive Region-Specific loss ([Chen et al., 2023](#)) are suggested loss functions to replace the [BCE](#) loss to partially overcome the issue of class imbalance.

Eventually, the locations of erosion and deposition could replace the total areas of erosion and deposition as additional physics-induced loss terms. Three classes would be required. “No change” would represent both “Non-water” and “Water” areas that remained stable. “Erosion” and “Deposition” would be the two other classes. For these predictions, a multiclass loss function such as [CCE](#) would be required.

6.2.2.2. Further possibilities

In this section further possibilities of longer-term research are included. These recommendations are considered secondary compared to those introduced in the previous section.

If a [GNN](#) is developed, conservation equations based on water and sediment discharge could be included to guide the model learning. This enables a more physical-based interpretation. This is crucial for communicating the results to potential users and stakeholders, also considering the increasing interest of [XAI](#) ([Mamalakis et al., 2022; Ren et al., 2024](#)).

Instead of developing a new model, another potential strategy involves leveraging foundation models like Prithvi-100M for flood detection ([Jakubik et al., 2023a,b](#)). This model is pre-trained with the [HLS](#) collection. After fine-tuning it, it is expected to be effective in tasks where labeled data is scarce and class imbalance hinders the training step.

[Bristow and Best \(1993\)](#) consider that the bed grain size does not influence significantly the large-scale braiding processes. Hence, the inclusion of more river could be suitable. This would increase the dataset size and enable a generalisation of the model to all braided rivers. The upper reach of the Tongtian River in China, the Piave River in Italy or the Waitaki River in New Zealand are some suggested rivers ([Stecca et al., 2019](#)). The same remarks on the smaller dimensions of these rivers compared to the Brahmaputra holds here.

Apart from employing a different loss function to reduce the effect of class imbalance, [Bosboom \(2019\)](#) proposed a warping method to overcome the worsening of the tracked metrics due to the “double penalty effect” introduced in Section 4.1. This method measures the closeness between images or spatial fields. It finds the smooth displacement field, an optimal transformation that maps the observed targets and the predicted model outputs, that minimises the point-wise error.

Another potential strategy to reduce class imbalance is the splitting of the “Non-water” class into “Sand” and “Vegetation” land-use classes. Vegetated areas are less likely to erode compared to sand areas. The inclusion of these classes is also expected to help the model in identifying which areas are less subject to erosion and which are more likely to be eroded. This strategy would modify the modelling task from binary semantic

segmentation to multiclass semantic segmentation. However, a different satellite image collection should be used. Given the longer records, Landsat collections are suggested. A classification algorithm should be developed as well.

Cloud coverage is another limitation that reduces the total number of suitable input-target combinations. [Zhang et al. \(2023\)](#) and [Ma et al. \(2024\)](#) propose two different approaches to combine optical and radar satellite images for cloud removal. In this way, apart from removing the clouds from the low-flow season images, images during the flood season could be retrieved as well. These could then be included in the input dataset, which consequently would contain more than one image per year. The dataset could then be composed of three images per year, for instance. The first image represents the low-flow stage. The second image represents the flood peak conditions. The last image represents the flood falling stage. By including the flood propagation in time and space, the post-flood conditions could be better predicted by the model. However, the shorter records of radar satellite images might limit the temporal extent of available images.

Different spatial and temporal spans could be of interest too. Specifically, smaller reaches and shorter time spans could have potential applications. Consider for instance the need for predicting the riverbank erosion at vulnerable locations, like schools or hospitals, before the flood peak arrives. Detailed situations as this might also require more input variables, such as a digital terrain model ([DTM](#)) of the surrounding areas and precipitation data.

Additionally, it can be of interest to research the possibility of predicting more years ahead, to allow an even longer-term planning. This could apply to large infrastructure, as well as critical (flood) protection measures and other hydraulic structures, such as dams for water intake and hydropower energy generation.

River bathymetry could be included as both input and target feature. This would modify the problem from purely planform predictions to a more thorough 3D river morphological predictions. [Hassan et al. \(1999\)](#) developed a method to derive the char elevation based on the char age. Research is conducted nowadays to extrapolate the river bathymetry from satellite images. [Valsangkar et al. \(2024\)](#) and [Bhattarai \(2024\)](#) retrieved the periodically submerged topography of the Jamuna River by integrating satellite images and in-situ observations. Additionally, [Bhattarai \(2024\)](#) unsuccessfully attempted to estimate the always submerged topography as well. High turbidity and the large depth of the Jamuna River were considered the key factors of this failure. The same limitation is reported by [Jiang and Rutherford \(2024\)](#).

6.2.3. Final consideration

Besides these suggestions, as a personal opinion, the range of applications and contexts of artificial intelligence models should be restricted to those cases where other approaches did not prove to be accurate or for which it is complex to retrieve (good) data. The concept of complex data to be retrieved is general, but can have several meanings. In the case of river morphological modelling, it mainly refers to the hydraulic and hydrological variables such as river discharge, water level, sediment discharge and sediment grading. These data require extensive and complex monitoring from both a temporal and a spatial scale. Satellite images, instead, are more easily accessible and despite some limitations can be considered a valid alternative to those data. Further research in this direction to retrieve the water levels, already going on with SWOT satellite ([Fu et al., 2009](#)), could bring potential benefits to a currently low-middle income and Least Develop Country as Bangladesh ([United Nations, 2023](#); [World Bank, 2024](#)), which could in turn save resources that can be allocated for other more relevant projects.

Bibliography

- Agence France-Presse (2024). Bangladesh floods leave at least eight dead amid fears situation could worsen. *The Guardian*. <https://www.theguardian.com/world/article/2024/jul/06/eight-dead-and-2m-affected-by-bangladesh-floods>. Retrieved 06/09/2024.
- Alabyan, A. M. and Chalov, R. S. (1998). Types of river channel patterns and their natural controls. *Earth Surface Processes and Landforms*, 23(5):467–474. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/28SICI%291096-9837%28199805%2923%3A5%3C467%3A%3AAID-ESP861%3E3.0.CO%3B2-T>.
- Ali, M. and Zobeyer, H. (2021). Research on river bank erosion dynamics using numerical modeling and deep learning techniques. Technical report, Water Resources Planning Organization (WARPO). https://warpo.portal.gov.bd/sites/default/files/files/warpo.portal.gov.bd/page/9b4814df_f835_4b0d_8c60_f2b35e0343b9/2022-09-25-04-11-fddb11ae602ab523dfa77fcda3fa8f8.pdf.
- Ashmore, P. (2013). Morphology and dynamics of braided rivers. *Treatise on Geomorphology*, 9:289–312. doi:[10.1016/B978-0-12-374739-6.00242-6](https://doi.org/10.1016/B978-0-12-374739-6.00242-6).
- Barua, A. K. (2023). The Morphological Behavior of The Kamarjani Channel of the Braided-Anabranchled Brahmaputra-Jamuna River. Master's thesis, IHE Delft. doi:[10.13140/RG.2.2.24964.09606](https://doi.org/10.13140/RG.2.2.24964.09606).
- Basnayaka, V., Samarasinghe, J. T., Gunathilake, M. B., Muttil, N., Hettiarachchi, D. C., Abeynayaka, A., and Rathnayake, U. (2022). Analysis of Meandering River Morphodynamics Using Satellite Remote Sensing Data—An application in the Lower Deduru Oya (River), Sri Lanka. *Land*, 11(7):1091. doi:[10.3390/land11071091](https://doi.org/10.3390/land11071091).
- Begum, T. (2023). “Peace of mind at last”: the Bangladeshi villagers digging their way out of the floods. *The Guardian*. <https://www.theguardian.com/global-development/2023/may/04/bangladeshi-villagers-digging-their-way-out-of-the-floods-plinth-settlements>. Retrieved 06/09/2024.
- Bentivoglio, R., Isufi, E., Jonkman, S., and Taormina, R. (2023). Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. *Hydrology and Earth System Sciences*, 27(23):4227–4246. doi:[10.5194/hess-27-4227-2023](https://doi.org/10.5194/hess-27-4227-2023).
- Best, J. L., Ashworth, P. J., Mosselman, E., Sarker, M. H., and Roden, J. E. (2022). *The Jamuna-Brahmaputra River, Bangladesh*, pages 579–640. Wiley-Blackwell, United States, 2nd edition. doi:[10.1002/9781119412632.ch20](https://doi.org/10.1002/9781119412632.ch20).
- Bhattarai, T. N. (2024). Morphological changes in large rivers and reservoirs: development and application of earth observation methods in combination with in-situ data. Unpublished Mater's thesis, Universitat Politècnica de Catalunya - BarcelonaTech. <https://upcommons.upc.edu/>.
- Bishop, C. (1994). Neural networks and their applications. *Review of scientific instruments*, 65(6):1803–1832. doi:[10.1063/1.1144830](https://doi.org/10.1063/1.1144830).
- Blake, R. W., Mathew, R., George, A., and Papakostas, N. (2021). Impact of artificial intelligence on Engineering: Past, present and future. *Procedia CIRP*, 104:1728–1733. doi:[10.1016/j.procir.2021.11.291](https://doi.org/10.1016/j.procir.2021.11.291).
- Bosboom, J. (2019). *Quantifying the quality of coastal morphological predictions*. PhD Thesis, Delft University of Technology, Netherlands. doi:[10.4233/uuid:e4dc2dfc-6c9c-4849-8aa9-befa3001e2a3](https://doi.org/10.4233/uuid:e4dc2dfc-6c9c-4849-8aa9-befa3001e2a3).
- Bristow, C. S. and Best, J. L. (1993). Braided rivers: perspectives and problems. *Geological Society London Special Publications*, 75(1):1–11. doi:[10.1144/gsl.sp.1993.075.01.01](https://doi.org/10.1144/gsl.sp.1993.075.01.01).
- Cardille, J. A., Crowley, M. A., Saah, D., and Clinton, N. E. (2023). *Cloud-Based Remote Sensing with Google Earth Engine*. Springer Cham, 1 edition. doi:[10.1007/978-3-031-26588-4](https://doi.org/10.1007/978-3-031-26588-4).

- Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3):1247–1250. doi:[10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014).
- Chavarriás, V. (2019). *Obtaining well-posedness in mathematical modelling of fluvial morphodynamics*. PhD Thesis, Delft University of Technology, Netherlands. doi:[10.4233/uuid:f741b590-8cd7-4e52-9641-b471954db5b2](https://doi.org/10.4233/uuid:f741b590-8cd7-4e52-9641-b471954db5b2).
- Chen, Y., Yu, L., Wang, J.-Y., Panjwani, N., Obeid, J.-P., Liu, W., Liu, L., Kovalchuk, N., Gensheimer, M. F., Vitzthum, L. K., Beadle, B. M., Chang, D. T., Le, Q.-T., Han, B., and Xing, L. (2023). Adaptive Region-Specific loss for improved medical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11):13408–13421. doi:[10.1109/tpami.2023.3289667](https://doi.org/10.1109/tpami.2023.3289667).
- Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., and Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219:145–161. doi:[10.1016/j.rse.2018.09.002](https://doi.org/10.1016/j.rse.2018.09.002).
- Cournède, F. and Mosselman, E. (2014). Morphological prediction of braided Congo River at Brazzaville. In *NCR-Days 2014, Book of Abstracts*, pages 63–64. <http://www.ncr-web.org/ncr-days/ncr-book-of-abstract>.
- Crosato, A. and Mosselman, E. (2009). Simple physics-based predictor for the number of river bars and the transition between meandering and braiding. *Water Resources Research*, 45(3). doi:<https://doi.org/10.1029/2008WR007242>.
- Deltares (2024). Delft3D-FLOW. Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments - User Manual, version 4.05. https://content.oss.deltares.nl/delft3d4/Delft3D-FLOW_User_Manual.pdf. Retrieved 02/09/2024.
- Deng, B., Xiong, K., Huang, Z., Jiang, C., Liu, J., Luo, W., and Xiang, Y. (2022). Monitoring and predicting channel morphology of the Tongtian River, headwater of the Yangtze River using landsat images and lightweight neural network. *Remote sensing*, 14(13):3107. doi:[10.3390/rs14133107](https://doi.org/10.3390/rs14133107).
- Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874. doi:[10.1016/j.patrec.2005.10.010](https://doi.org/10.1016/j.patrec.2005.10.010).
- Ferguson, R. I. (1993). Understanding braiding processes in gravel-bed rivers: progress and unsolved problems. *Geological Society London Special Publications*, 75(1):73–87. doi:[10.1144/gsl.sp.1993.075.01.03](https://doi.org/10.1144/gsl.sp.1993.075.01.03).
- Flood Action Plan 24 (1996). FAP24 River Survey Project, Special Report 4, Stage-Discharge Relationship for The Jamuna at Bahadurabad. Technical report, Delft Hydraulics and Danish Hydraulic Institute. Dhaka, Bangladesh.
- Fu, L.-L., Alsdorf, D., Rodriguez, E., Morrow, R., Mognard, N., Lambin, J., Vaze, P., and Lafon, T. (2009). The SWOT (Surface Water and Ocean Topography) Mission: Spaceborne Radar Interferometry for Oceanographic and Hydrological Applications. https://swot.jpl.nasa.gov/system/publications/23_23_oceanobs09_swot.pdf.
- Gignac, G. E. and Szodorai, E. T. (2024). Defining intelligence: Bridging the gap between human and artificial perspectives. *Intelligence*, 104:101832. doi:[10.1016/j.intell.2024.101832](https://doi.org/10.1016/j.intell.2024.101832).
- Giri, S., Thompson, A., Donchyts, G., Oberhagemann, K., Mosselman, E., and Alam, J. (2021). Stabilization of the Lower Jamuna River in Bangladesh—Hydraulic and morphological assessment. *Geosciences*, 11(9):389. doi:[10.3390/geosciences11090389](https://doi.org/10.3390/geosciences11090389).
- Global Surface Water Explorer (2024). Global Surface Water Explorer. <https://global-surface-water.appspot.com/>. Retrieved 15/06/2024.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.

- Google (2024). Classification: ROC and AUC. <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc>. Retrieved 02/09/2024.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. doi:[10.1016/j.rse.2017.06.031](https://doi.org/10.1016/j.rse.2017.06.031).
- Hagenaars, G., De Vries, S., Luijendijk, A. P., De Boer, W. P., and Reniers, A. J. (2018). On the accuracy of automated shoreline detection derived from satellite imagery: A case study of the sand motor mega-scale nourishment. *Coastal Engineering*, 133:113–125. doi:[10.1016/j.coastaleng.2017.12.011](https://doi.org/10.1016/j.coastaleng.2017.12.011).
- Hassan, A., Martin, T. C., and Mosselman, E. (1999). Island topography mapping for the Brahmaputra-Jamuna River using remote sensing and GIS. *Geological Society London Special Publications*, 163(1):153–161. doi:[10.1144/gsl.sp.1999.163.01.13](https://doi.org/10.1144/gsl.sp.1999.163.01.13).
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9:1735–80. doi:[10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- Hu, Y., Yang, H., Zhou, H., and Lv, Q. (2023). A review of Numerical Modelling of morphodynamics in braided Rivers: Mechanisms, Insights and challenges. *Water*, 15(3):595. doi:[10.3390/w15030595](https://doi.org/10.3390/w15030595).
- Jagers, H. (2003). *Modelling planform changes of braided rivers*. PhD Thesis, University of Twente, Netherlands. <https://research.utwente.nl/en/publications/modelling-planform-changes-of-braided-rivers>.
- Jakubik, J., Fraccaro, P., Oliveira Borges, D., Muszynski, M., Weldemariam, K., Zadrozny, B., Ganti, R., and Mukkavilli, K. (2023a). Prithvi 100M flood mapping. doi:[10.57967/hf/0973](https://doi.org/10.57967/hf/0973). Repository available at: <https://huggingface.co/ibm-nasa-geospatial/Prithvi-100M-sen1floods11>. Retrieved 28/08/2024.
- Jakubik, J., Roy, S., Phillips, C. E., Fraccaro, P., Godwin, D., Zadrozny, B., Szwarcman, D., Gomes, C., Nyirjesy, G., Edwards, B., Kimura, D., Simumba, N., Chu, L., Mukkavilli, S. K., Lambhate, D., Das, K., Bangalore, R., Oliveira, D., Muszynski, M., Ankur, K., Ramasubramanian, M., Gurung, I., Khallaghi, S., Li, H. S., Cecil, M., Ahmadi, M., Kordi, F., Alemohammad, H., Maskey, M., Ganti, R., Weldemariam, K., and Ramachandran, R. (2023b). Foundation Models for Generalist Geospatial Artificial Intelligence. *arXiv (Cornell University)*. doi:[10.48550/arXiv.2310.18660](https://doi.org/10.48550/arXiv.2310.18660).
- Jiang, H. and Rutherford, I. (2024). Potential to use free satellite imagery to retrieve the past bathymetry of large rivers. *Remote Sensing Applications Society and Environment*, 33:101133. doi:[10.1016/j.rsase.2023.101133](https://doi.org/10.1016/j.rsase.2023.101133).
- Khankhoje, T. and Choudhury, P. (2024). River system sediment flow modeling using artificial neural networks. *International Journal of Sediment Research*, 39(2):222–229. doi:<https://doi.org/10.1016/j.ijsrc.2023.11.006>.
- Khurshid, M. H. and Khan, M. F. (2012). River extraction from high resolution satellite images. In *2012 5th International Congress on Image and Signal Processing*. IEEE. doi:[10.1109/CISP.2012.6469675](https://doi.org/10.1109/CISP.2012.6469675).
- Klaassen, G., Mosselman, E., Masselink, G., Bruhl, H., Huisink, M., Koomen, E., and Seymonsbergen, A. (1993). *Planform changes in large braided sand-bed rivers*. Number Publ. 480 in Delft Hydraulics. Delft Hydraulics. <https://research.vu.nl/en/publications/planform-changes-in-large-braided-sand-bed-rivers>.
- Koepon Stichting (2024). Plinth Bangladesh – Friendship. <https://www.koeponstichting.nl/2024/05/19/plinth-bangladesh-friendship/>. Retrieved 06/09/2024.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.
- Lahiri, S. K. (2022). *The Brahmaputra River in Assam: Geomorphology, Hazards, and Natural Resources*. CRC Press, 1 edition. doi:[10.1201/9781003302353](https://doi.org/10.1201/9781003302353).

- Lane, E. W. (1957). A Study of the shape of channels formed by natural streams flowing in erodible material. *U.S. Army Engineer Division, Missouri River, 1957.* <https://erdc-library.erdc.dren.mil/jspui/bitstream/11681/8944/1/Shape-of-Channels.pdf>.
- Lauer, D. T., Morain, S. A., and Salomonson, V. V. (1997). The Landsat Program: Its Origins, Evolution, and Impacts. *Photogrammetric Engineering and Remote Sensing*, 63(7):831–838. <https://pubs.er.usgs.gov/publication/70019209>.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444. doi:[10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Leopold, L. and Wolman, M. (1957). River channel patterns: Braided, meandering, and straight. *U.S. Geological Survey professional papers*. doi:[10.3133/pp282b](https://doi.org/10.3133/pp282b).
- Lo Piano, S. (2020). Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanities social sciences communications*, 7(1). doi:[10.1057/s41599-020-0501-9](https://doi.org/10.1057/s41599-020-0501-9).
- Ma, J., Chen, Y., Pan, J., Xu, J., Li, Z., Xu, R., and Chen, R. (2024). SCT-CR: A synergistic convolution-transformer modeling method using SAR-optical data fusion for cloud removal. *International Journal of Applied Earth Observation and Geoinformation*, 130:103909. doi:[10.1016/j.jag.2024.103909](https://doi.org/10.1016/j.jag.2024.103909).
- Mamalakis, A., Barnes, E. A., and Ebert-Uphoff, I. (2022). Investigating the fidelity of explainable artificial intelligence methods for applications of convolutional neural networks in geoscience. *Artificial Intelligence for the Earth Systems*, 1(4):e220012. doi:[10.1175/AIES-D-22-0012.1](https://doi.org/10.1175/AIES-D-22-0012.1).
- McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. E. (2006). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, 27(4):12. doi:[10.1609/aimag.v27i4.1904](https://doi.org/10.1609/aimag.v27i4.1904).
- McLean, D., Oberhagemann, K., Wiebe, H., and Bennett, S. (2020). Flood and Riverbank Erosion Risk Management Investment Program (FRERMIP) - River Stabilization And Development: Jamuna-Padma and Dependent Areas, Main Report. Technical report, Institutional Strengthening and Project Management Consultant (ISPMC) and Joint Venture Northwest Hydraulic Consultants LTD/Euroconsult Mott MacDonald. http://pipeline.nhcweb.com/website/index.php/mm/report_download/RSP.pdf. Retrieved 06/09/2024.
- Melo, W., Pinho, J. L. S., and Iglesias, I. (2022). Emulating the estuarine morphology evolution using a deep convolutional neural network emulator based on hydrodynamic results of a numerical model. *Journal of hydroinformatics*, 24(6):1254–1268. doi:[10.2166/hydro.2022.068](https://doi.org/10.2166/hydro.2022.068).
- Mosselman, E. and Sloff, K. (2002). Effect of local scour holes on macroscale river morphology. In Bousmar, D. and Zech, Y., editors, *Proceedings of the international conference on fluvial hydraulics, River Flow 2002*, pages 767–772. CRC Press / Balkema - Taylor Francis Group.
- Mottaleb, K. A. and Rahut, D. B. (2018). Impacts of improved infrastructure on labor allocation and livelihoods: the case of the Jamuna Multipurpose Bridge, Bangladesh. *European Journal of Development Research*, 31(4):750–778. doi:[10.1057/s41287-018-0186-8](https://doi.org/10.1057/s41287-018-0186-8).
- Munasinghe, D., Cohen, S., and Gadiraju, K. K. (2021). A review of satellite remote sensing techniques of River Delta Morphology Change. *Remote Sensing in Earth Systems Sciences*, 4(1-2):44–75. doi:[10.1007/s41976-021-00044-3](https://doi.org/10.1007/s41976-021-00044-3).
- Murray, A. B. and Paola, C. (1994). A cellular model of braided rivers. *Nature*, 371(6492):54–57. doi:[10.1038/371054a0](https://doi.org/10.1038/371054a0).
- Nazari, L. F., Camponogara, E., and Seman, L. O. (2022). Physics-Informed neural networks for modeling water flows in a river channel. *IEEE Transactions on Artificial Intelligence*, pages 1–15. doi:[10.1109/tai.2022.3200028](https://doi.org/10.1109/tai.2022.3200028).
- Noh, B., Wani, O., Dunne, K. B. J., and Lamb, M. P. (2024). Geomorphic risk maps for river migration using probabilistic modeling – a framework. *Earth Surface Dynamics*, 12(3):691–708. doi:[10.5194/esurf-12-691-2024](https://doi.org/10.5194/esurf-12-691-2024).

- O'Shea, K. and Nash, R. (2015). An introduction to convolutional neural networks. *arXiv (Cornell University)*. doi:[10.48550/arXiv.1511.08458](https://doi.org/10.48550/arXiv.1511.08458).
- Paola, C. (2001). Modelling stream braiding over a range of scales. Retrieved from the University Digital Conservancy at: <https://hdl.handle.net/11299/164368>.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *arXiv (Cornell University)*. doi:[10.48550/arxiv.1912.01703](https://doi.org/10.48550/arxiv.1912.01703).
- Pekel, J., Cottam, A., Gorelick, N., and Belward, A. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633):418–422. doi:[10.1038/nature20584](https://doi.org/10.1038/nature20584).
- Prince, S. J. (2023). *Understanding Deep Learning*. The MIT Press. <http://udlbook.com>.
- Ren, Z., Liu, C., Zhao, X., Jin, Y., Ou, Y., Liu, R., Fan, H., Yang, Q., Lim, A., and Cheng, H. (2024). A large-scale riverbank erosion risk assessment model integrating multi-source data and explainable artificial intelligence (XAI). *Ecological Indicators*, page 112575. doi:[10.1016/j.ecolind.2024.112575](https://doi.org/10.1016/j.ecolind.2024.112575).
- Richards, J. A. (2022). *Remote Sensing Digital Image Analysis*. Springer Cham, 6 edition. doi:[10.1007/978-3-030-82327-6](https://doi.org/10.1007/978-3-030-82327-6).
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv (Cornell University)*. doi:[10.48550/arxiv.1505.04597](https://doi.org/10.48550/arxiv.1505.04597).
- Salomonson, V. V. (2014). *Remote Sensing, Historical Perspective*. Springer New York. doi:[10.1007/978-0-387-36699-9_158](https://doi.org/10.1007/978-0-387-36699-9_158).
- Sarkar, A. (2017). Brahmaputra River Bank Failures - Causes and Impact on River Dynamics. In Mikoš, M., Vilímek, V., Yin, Y., and Sassa, K., editors, *Advancing Culture of Living with Landslides*, pages 273–280, Cham. Springer International Publishing. doi:[10.1007/978-3-319-53483-1_32](https://doi.org/10.1007/978-3-319-53483-1_32).
- Schmidt, R. M. (2019). Recurrent Neural Networks (RNNs): A gentle Introduction and Overview. *arXiv (Cornell University)*. doi:[10.48550/arXiv.1912.05911](https://doi.org/10.48550/arXiv.1912.05911).
- Shi, X., Chen, Z., Wang, H., Yeung, D., Wong, W., and Woo, W.-C. (2015). Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *arXiv (Cornell University)*. doi:[10.48550/arxiv.1506.04214](https://doi.org/10.48550/arxiv.1506.04214).
- Saviglia, A. and Crosato, A. (2016). Numerical modelling of river morphodynamics: Latest developments and remaining challenges. *Advances in Water Resources*, 93:1–3. doi:[10.1016/j.advwatres.2016.01.005](https://doi.org/10.1016/j.advwatres.2016.01.005).
- Spinoza, A., Ziembra, A., Saponieri, A., Damiani, L., and Serafy, G. E. (2021). Remote Sensing-Based Automatic Detection of Shoreline Position: a case study in Apulia region. *Journal of Marine Science and Engineering*, 9(6):575. doi:[10.3390/jmse9060575](https://doi.org/10.3390/jmse9060575).
- Stańczyk, T. and Mehrkanoon, S. (2021). Deep Graph convolutional networks for wind speed prediction. *ESANN 2021 proceedings*. doi:[10.14428/esann/2021.es2021-25](https://doi.org/10.14428/esann/2021.es2021-25).
- Stecca, G., Zolezzi, G., Hicks, D. M., and Surian, N. (2019). Reduced braiding of rivers in human-modified landscapes: Converging trajectories and diversity of causes. *Earth-Science Reviews*, 188:291–311. doi:[10.1016/j.earscirev.2018.10.016](https://doi.org/10.1016/j.earscirev.2018.10.016).
- Surian, N. (2015). *Fluvial Processes in Braided Rivers*, pages 403–425. Springer International Publishing, Cham. doi:[10.1007/978-3-319-17719-9_15](https://doi.org/10.1007/978-3-319-17719-9_15).
- United Nations (2023). LDCs at a glance | Department of Economic and Social Affairs. <https://www.un.org/development/desa/dpad/least-developed-country-category/ldc-at-a-glance.html>. Retrieved 05/10/2024.

- Valsangkar, N., Nelson, A., and Hasan, M. F. (2024). Combining earth observations with ground data to assess river topography and morphologic change: Case study of the lower Jamuna River. *International Journal of Applied Earth Observation and Geoinformation*, 134:104146. doi:[10.1016/j.jag.2024.104146](https://doi.org/10.1016/j.jag.2024.104146).
- Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., and Wynne, R. (2008). Free access to Landsat imagery. *Science*, 320(5879):1011. doi:[10.1126/science.320.5879.1011a](https://doi.org/10.1126/science.320.5879.1011a).
- World Bank (2024). Overview - The World Bank in Bangladesh. <https://www.worldbank.org/en/country/bangladesh/overview>. Retrieved 05/10/2024.
- Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., and Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122:2–10. doi:[10.1016/j.rse.2012.01.010](https://doi.org/10.1016/j.rse.2012.01.010).
- Yeung, M., Sala, E., Schönlieb, C.-B., and Rundo, L. (2022). Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Computerized Medical Imaging and Graphics*, 95:102026. doi:[10.1016/j.compmedimag.2021.102026](https://doi.org/10.1016/j.compmedimag.2021.102026).
- Zhang, S., Li, X., Zhou, X., Wang, Y., and Hu, Y. (2023). Cloud removal using SAR and optical images via attention mechanism-based GAN. *Pattern Recognition Letters*, 175:8–15. doi:[10.1016/j.patrec.2023.09.014](https://doi.org/10.1016/j.patrec.2023.09.014).
- Zhang, Z., Anjiraki, M. G., Seyedzadeh, H., Sotiropoulos, F., and Khosronejad, A. (2024). Toward ultra-efficient high-fidelity prediction of bed morphodynamics of large-scale meandering rivers using a novel LES-trained machine learning approach. *arXiv (Cornell University)*. doi:[10.48550/arXiv.2407.18359](https://doi.org/10.48550/arXiv.2407.18359).

A

Hydraulic data

This appendix contains some records of water level and discharge of the Brahmaputra River. All data presented here were recorded at Bahadurabad, in Bangladesh, close to the Indian border. The water levels refer to the PWD level (see footnote 1 at Page 8).

As introduced in Section 2.2.2, the Brahmaputra River experiences a significant discharge variation within a year. Figure A.1 shows the daily water levels recorded between 1964 and 1994. It also includes the daily average across those years (red line).

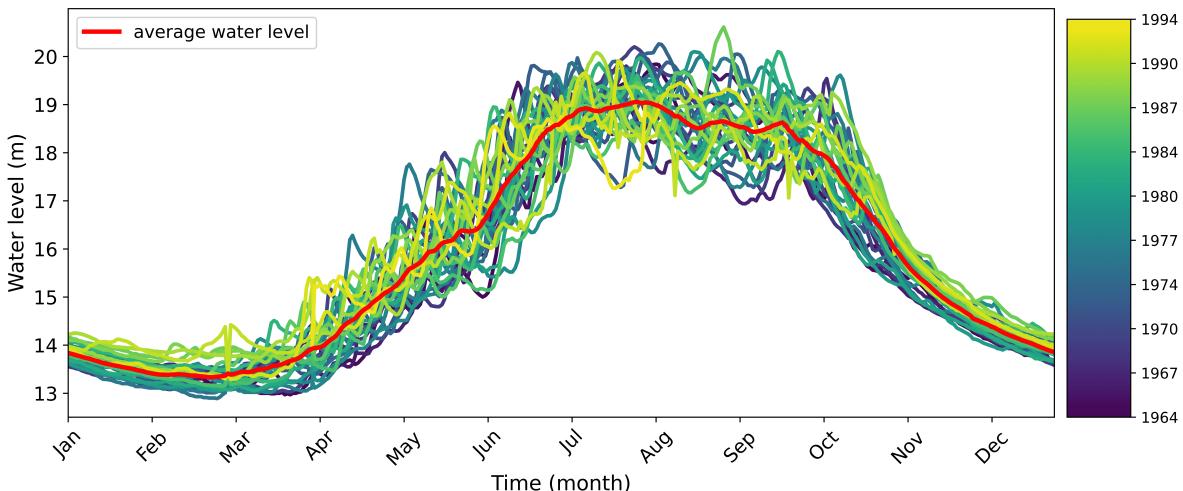


Figure A.1: Average daily recorded water levels at Bahadurabad, Bangladesh, between 1964 and 1994. The red line shows the daily average across the years of record.

This figure was used for two key assumptions:

- The low-flow season lasts on average from January to April, although December could be included as well.
- Given the relatively small variation in water level during the low-flow season, the images representative of the months between January and April can be considered as similar to each other. This assumption is the basis of the “season averaging” step for the “No data” replacement strategy (see Appendix B.2).

Figure A.2 shows the combined biweekly records of water level and discharge between 1990 and 2017. Despite missing some data, the all floods were recorded. The discharge distribution of this record was used to find the associated water levels presented in Figure 2.3.

The distribution of the discharges of Figure A.2 is plotted in Figure A.3. Given the monotonic non-linear

relationship between water levels and discharges, the Spearman's rank correlation coefficient was included. This coefficient is calculated as:

$$\rho_s = 1 - \frac{6 \sum d_i^2}{N(N^2 - 1)} \quad (\text{A.1})$$

where $d_i = R(X_i) - R(Y_i)$ is the difference between the ranks $R(X_i)$ and $R(Y_i)$ of the i -th sample of the variables X_i and Y_i , and N is the total number of samples.

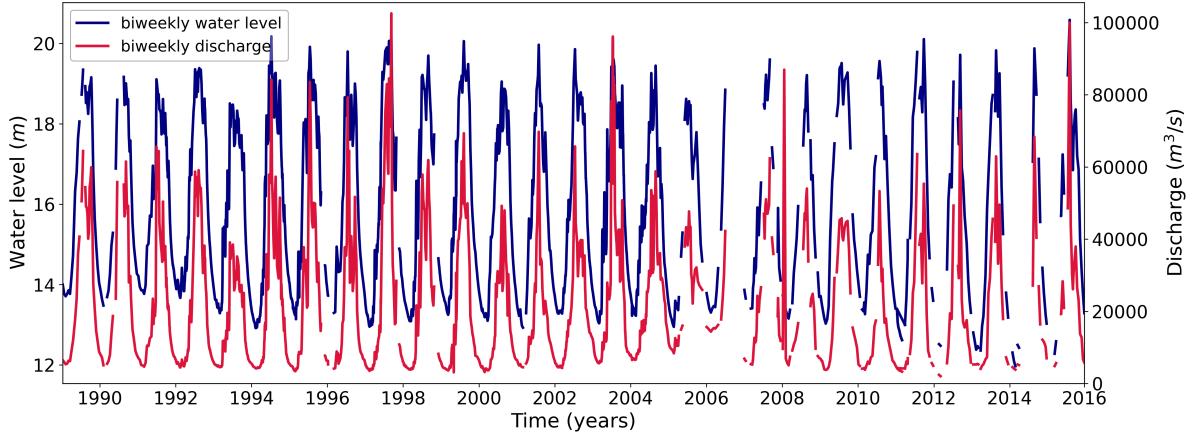


Figure A.2: Average biweekly recorded water level and discharge at Bahadurabad, Bangladesh, between 1989 and 2016.

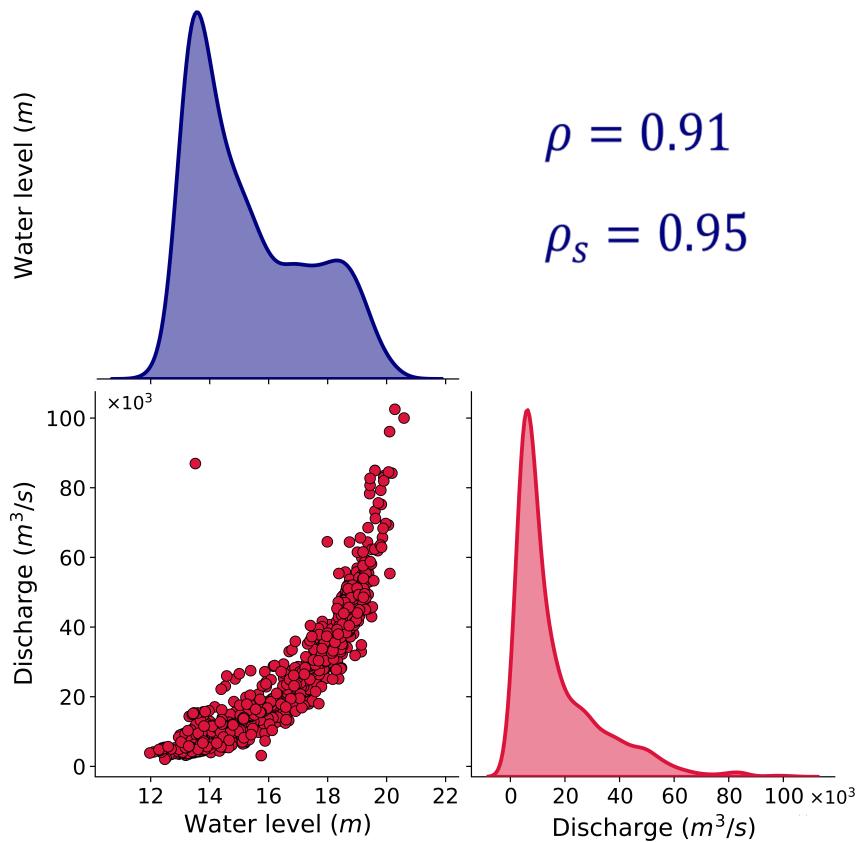


Figure A.3: Correlation matrix between recorded water levels and discharge of Figure A.2. In the above-diagonal element of the matrix the Pearson and Spearman's rank correlation coefficients are included.

Figure A.4 shows more recent daily water levels, recorded between 2012-2023.

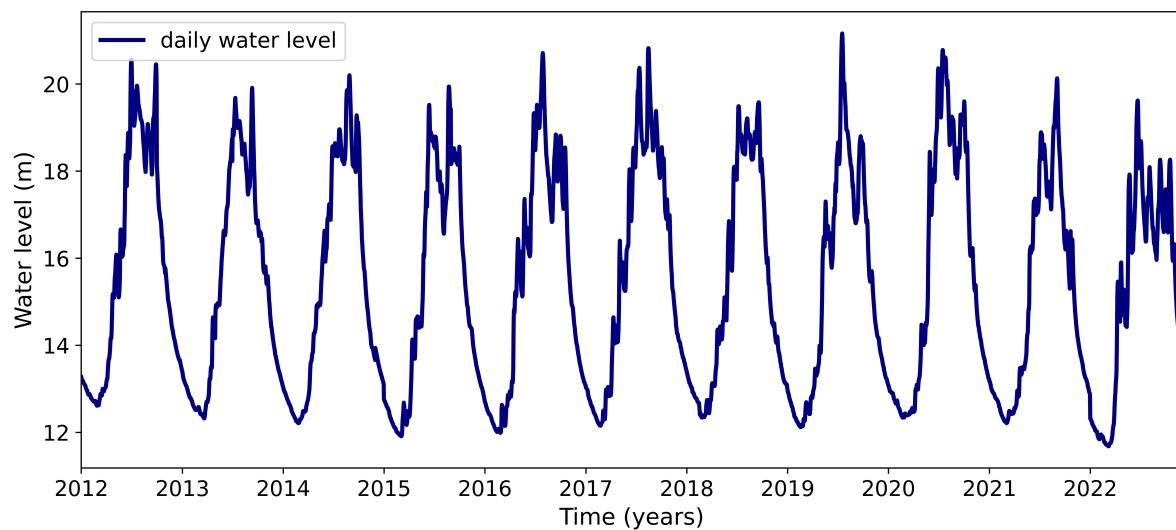


Figure A.4: Average daily recorded water levels at Bahadurabad, Bangladesh, between 2012 and 2023.

B

Satellite image preprocessing

This appendix contains a detailed description of the satellite image preprocessing, with figures and additional information.

B.1. Image standardization

Before generating the dataset, the images were preprocessed to standardize their characteristics, as described in Section 3.1.2. The most downstream reach, the one used during the testing phase, was selected as the reference. For this reach the flow is southward-directed, from the top to the bottom of the images. In several upstream reaches the flow is primarily westward- or southwestward-directed. In these reaches the flow goes from right to left or from the top right to the bottom left of the images, respectively. To ensure a consistent flow direction across all reaches, a rigid rotation was applied to the images. Reaches with westward-directed flow were rotated by 90°, while those with a southwestward direction were rotated by 45°. Figure B.1a and B.1b show the original image with westward-directed flow and the rotated image, respectively.

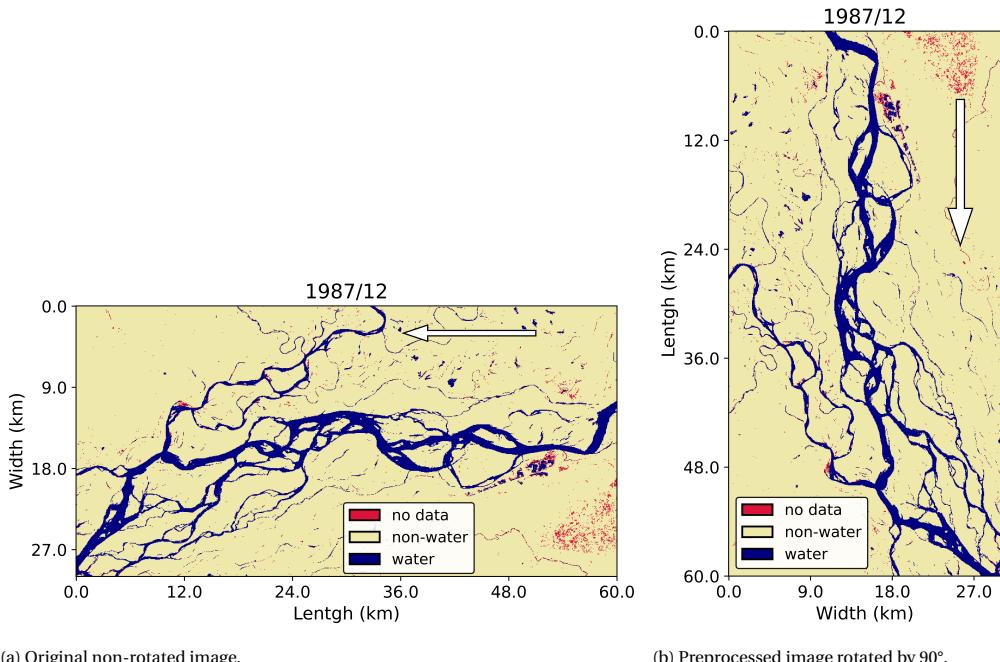


Figure B.1: Original and rotated images representing the training reach 22 in December 1987.

After the rigid rotation, images were cropped to have standard dimensions of 1000 × 500 pixels, representing the image height and width, respectively.

B.2. "No data" replacement

B.2.1. Seasonal averaging

For each combination of reach and year, a pixel-wise average image was generated using only the low-flow images. The low-flow season was considered from January to April of the same year. It was assumed that in these months the river water levels are approximately constant (see Appendix A). The four images of a given low-flow season were assumed to be almost equal to each other. The low-flow *seasonal average* image was generated using these four images per season. "No data" pixels were excluded in the computation. Figure B.2 shows the four original images and the computed seasonal average, which could still contain some "No data" pixels.

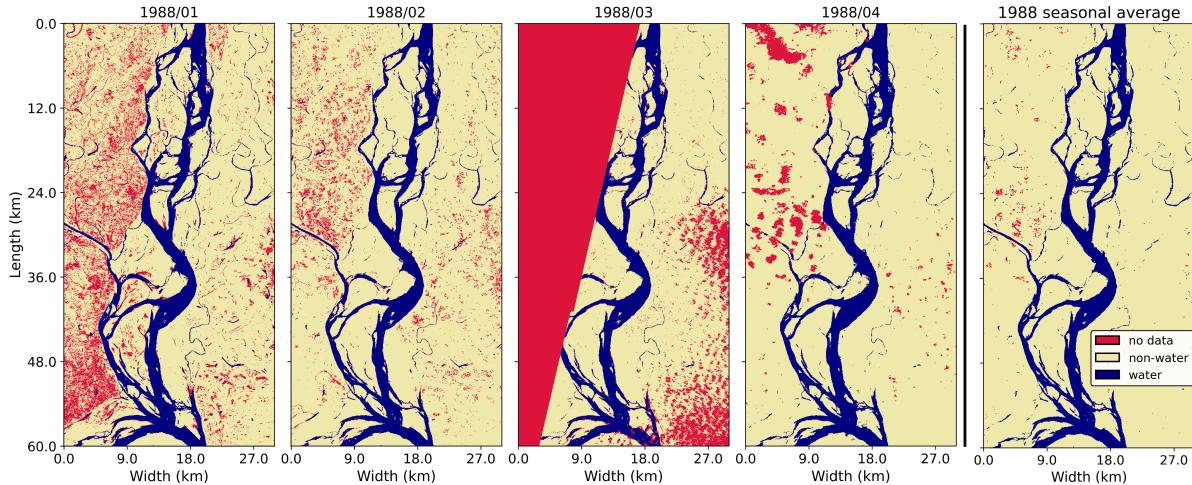


Figure B.2: 1988 low-flow season images (first four from left) and seasonal average image of testing reach.

B.2.2. Neighbours averaging

As in some cases some "No data" pixels were still present, a further averaging step was performed. This was defined the *neighbours averaging*. For each remaining "No data" pixel, a 15×15 pixels square around it was considered (Figure B.3). The values of these neighbour pixels were averaged and the result was used to replace the "No data" pixel. Again, "No data" pixels present in this neighborhood were excluded from the averaging process.

The square size was chosen to be 15×15 pixels because, given the pixel resolution of 60 m, it represented a 900×900 m square. Consider that the average width of secondary channels in the Brahmaputra is around 1.2 km (Jagers, 2003). This square size could enable to classify a single "No data" pixel in the middle of a channel to be of "Water" class. However, pixels at the interface between "Non-water" and "Water" areas risk being incorrectly assigned, depending on whether they are surrounded by more than half "Non-water" or "Water" pixels. The presence of one more pixel of one specific class is sufficient to classify the "No data" pixel in that class.

Figure B.4a shows the 1988 image after the neighbours averaging step. Few "No data" pixels are still present on the top left corner of the image.

In case also all pixels in the neighborhood were "No data", not allowing to compute an average value, by default the pixel was assigned to the "Non-water" class. This was done because the majority of "No data" pixels were visually observed to be primarily in "Non-water" areas. Hence, classify these pixels as "Water" was considered too conservative, despite this class being the most important to be correctly predicted by the model. Figure B.4b shows the final average image after the final hard replacement.

These average-generated images were used to replace each single "No data" pixels across the input and target images. Figure B.5 shows the 1988 March image representative of the testing reach after the "No data" replacement. The original image is visible in Figure B.2 (third image from the left).

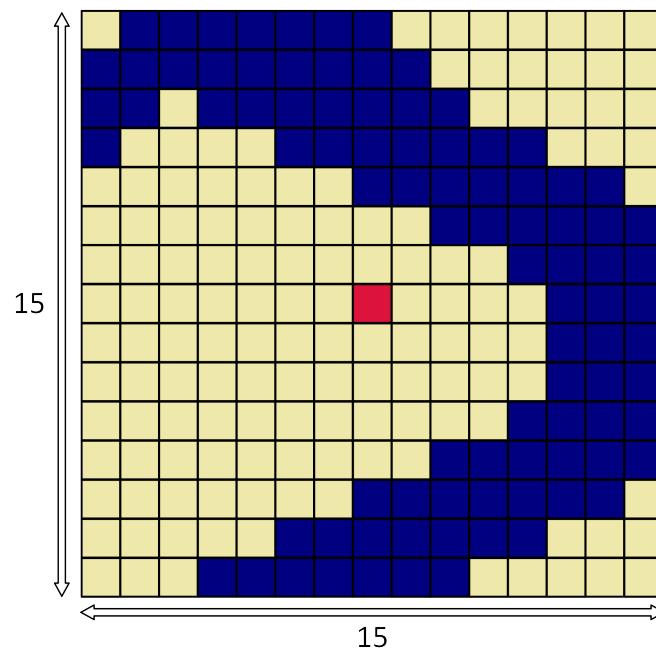


Figure B.3: Sketch of neighbours averaging step. Red, yellow, and blue pixels represent "No data", "Non-water", and "Water" areas, respectively.

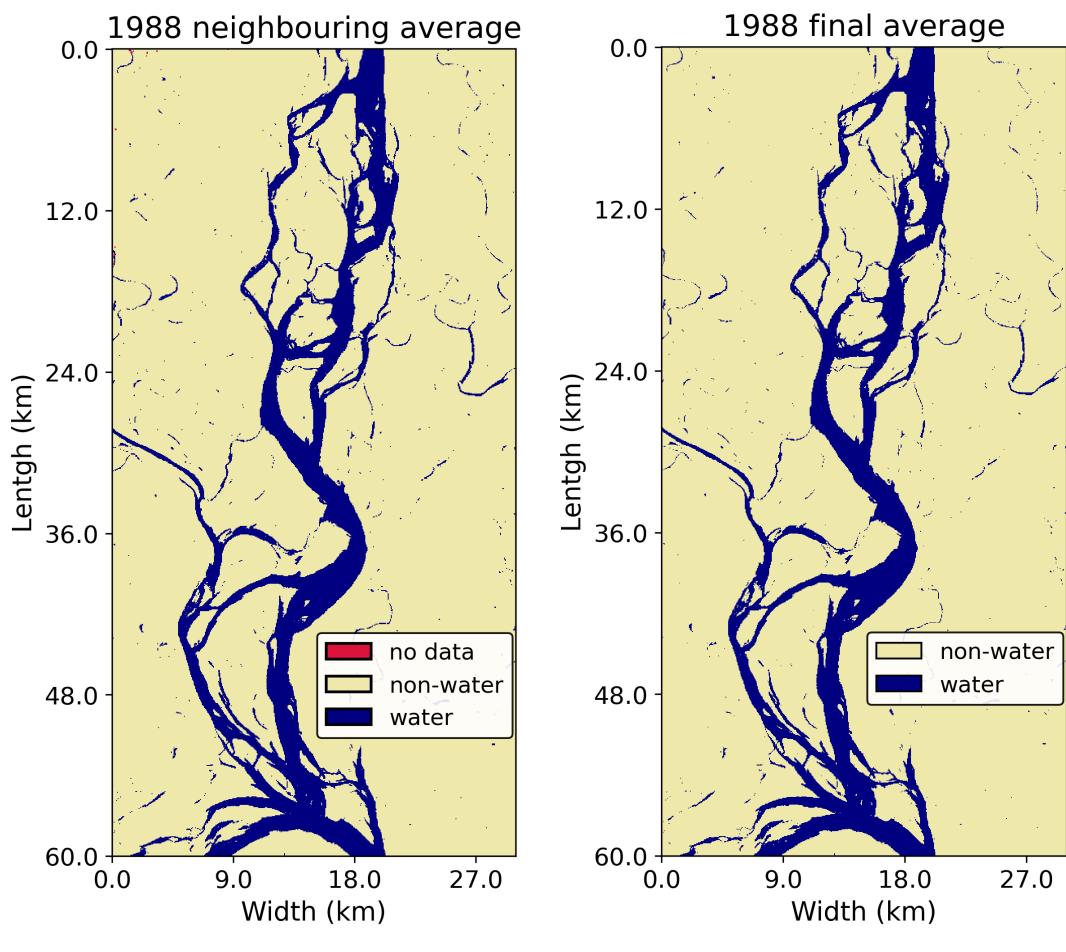


Figure B.4: 1988 neighbours average image and final average image after the final replacement for the testing reach.

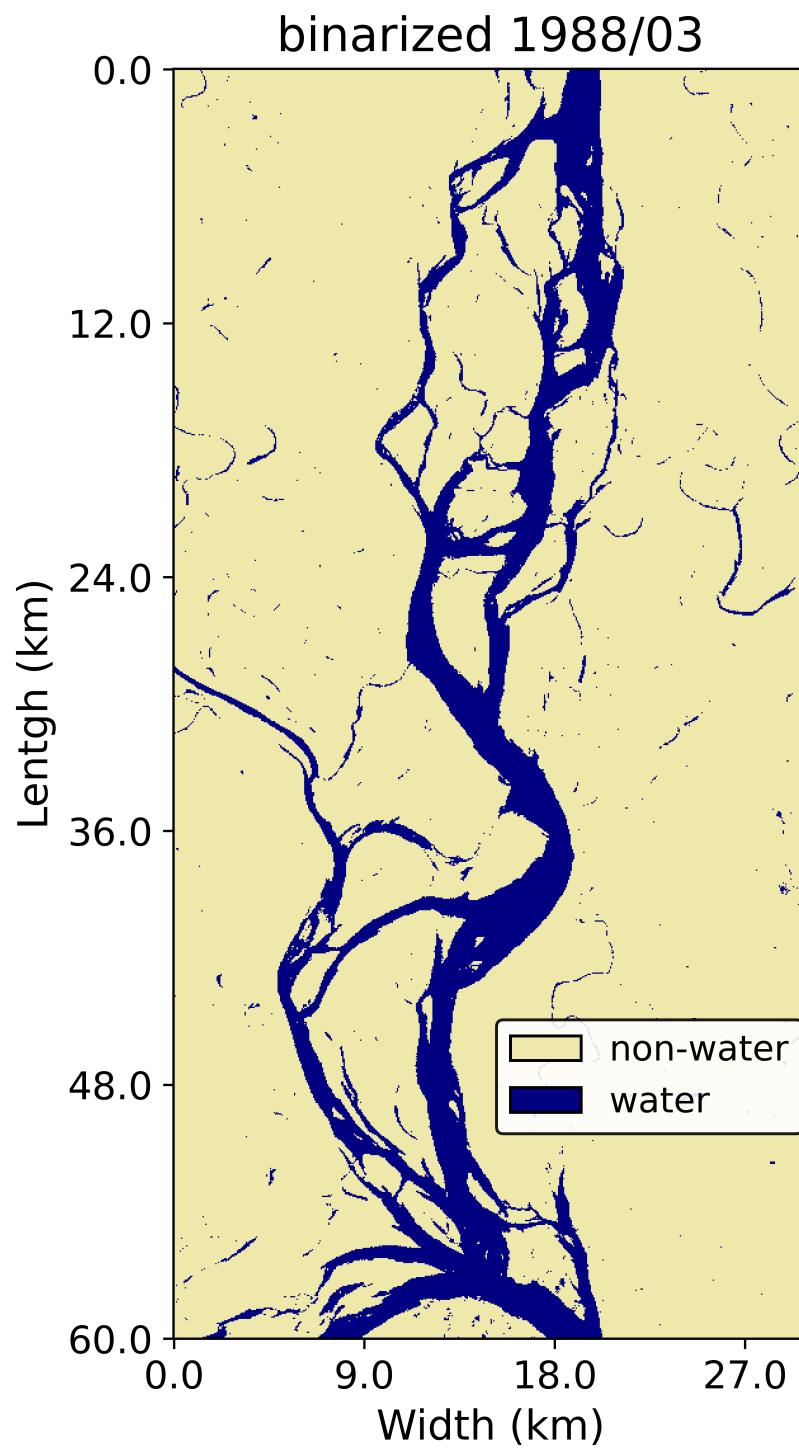


Figure B.5: Image representing the March 1988 conditions of the testing reach after "No data" replacement.

C

Dataset generation

This appendix contains details on the filtering procedure applied for generating all datasets and a mapping of the number of testing samples with the input and target years.

C.1. Quality filtering

All the images from 1994, 1998 and 1999 were completely full of “No data” pixels. This misclassification was related to the [GSWD](#) collection and not to any other preprocessing steps implemented so far.

For these images the “No data” replacement strategy described in Appendix B.2 returned fully “Non-water” images. To prevent these images from becoming part of any dataset, a further quality filtering step was applied. If any image within the input-target combination contained more than 480 000 “Non-water” pixels, the full combination was discarded. This value was adopted to make sure that only non-representative images were tracked and their combinations discarded. Figure C.1 shows an example of a discarded combination, where the image representative of March 1994 contains only “Non-water” pixels.

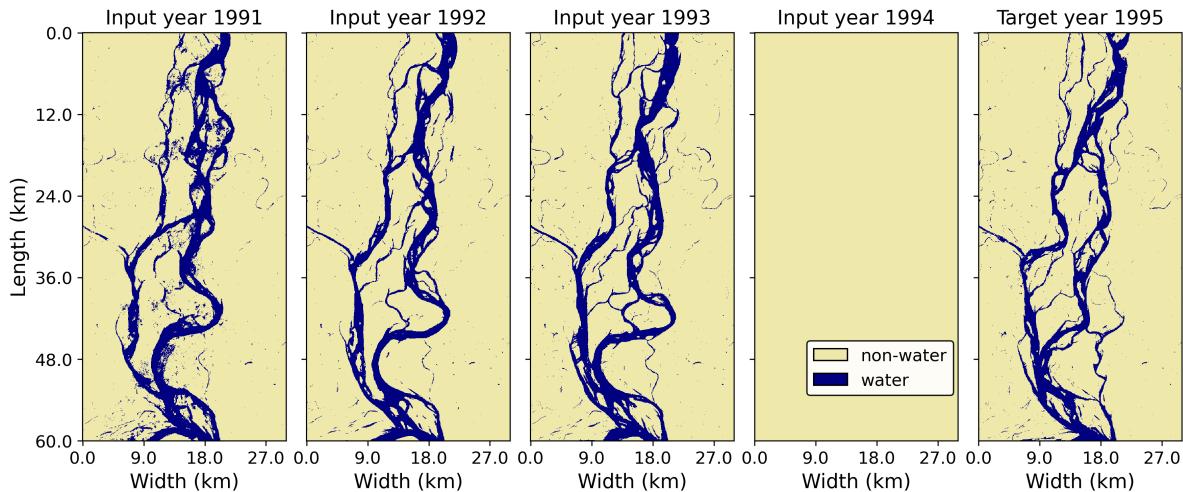


Figure C.1: Example of a discarded input-target combination containing at least one fully “Non-water” image.

C.2. Final datasets

The previous filtering operation reduced the original number of training, validation and testing samples from 812, 29, and 29, to 518, 19, and 19, respectively.

Figure C.2 shows a sketch of an input-target combination for the testing reach. The four images on the left are the inputs. Each input image was stacked as a single channel. The image on the right is the target.

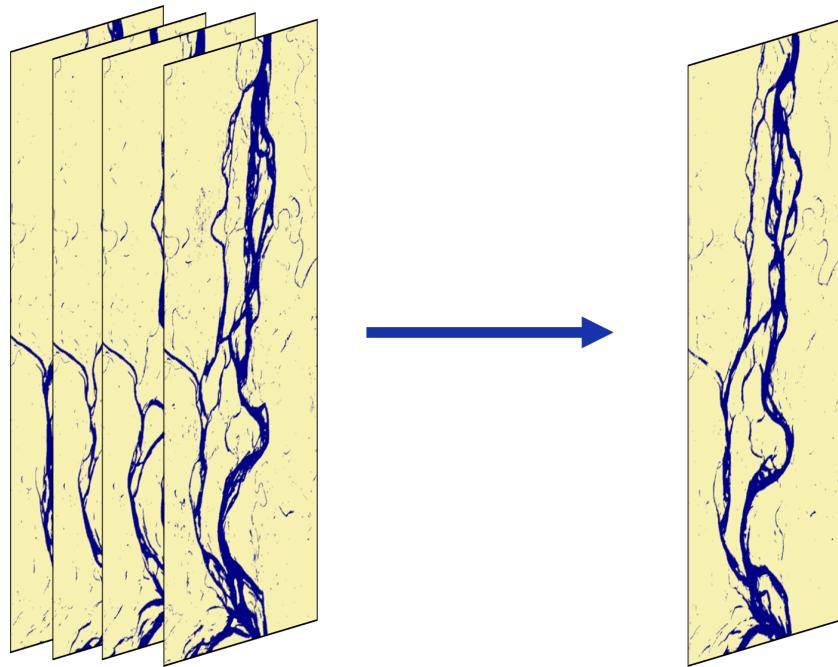


Figure C.2: Sketch of an input-target combination.

Table C.1 links the testing sample number with the input and target years.

Testing sample ID	Input years	Target year
1	1988, 1989, 1990, 1991	1992
2	1989, 1990, 1991, 1992	1993
3	2000, 2001, 2002, 2003	2004
4	2001, 2002, 2003, 2004	2005
5	2002, 2003, 2004, 2005	2006
6	2003, 2004, 2005, 2006	2007
7	2004, 2005, 2006, 2007	2008
8	2005, 2006, 2007, 2008	2009
9	2006, 2007, 2008, 2009	2010
10	2007, 2008, 2009, 2010	2011
11	2008, 2009, 2010, 2011	2012
12	2009, 2010, 2011, 2012	2013
13	2010, 2011, 2012, 2013	2014
14	2011, 2012, 2013, 2014	2015
15	2012, 2013, 2014, 2015	2016
16	2013, 2014, 2015, 2016	2017
17	2014, 2015, 2016, 2017	2018
18	2015, 2016, 2017, 2018	2019
19	2016, 2017, 2018, 2019	2020

Table C.1: Link between testing sample ID, input years, and target year.

D

JamUNet operations

This appendix contains more details on the operations performed by JamUNet model.

Figure D.1 is a sketch of a 2D convolution operating on multiple channels. The convolution slides through the image height and width and across channels. It computes the dot product of the filter with the local regions of the input across the channels. It then sums the output of each product and generates one feature map. The kernel has size 3×3 . This size represents the height and width of the filter.

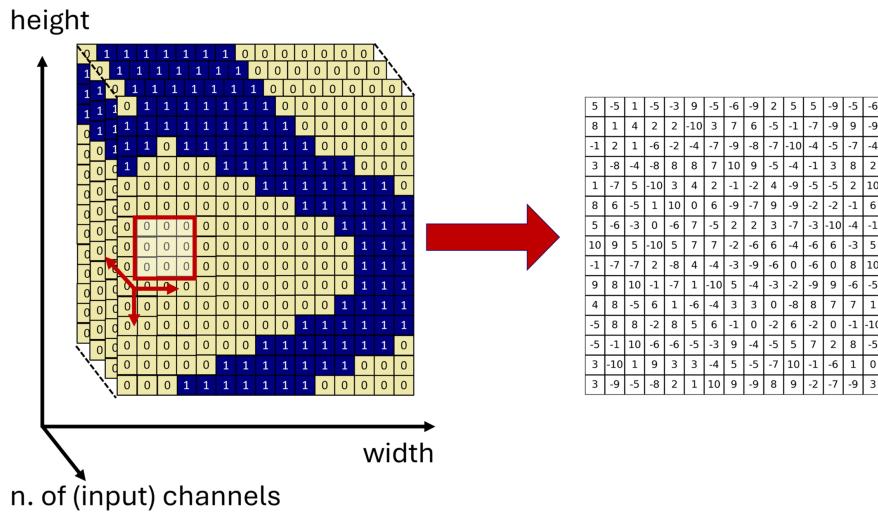


Figure D.1: Sketch of a 2D convolution operating on multiple input channels. The red arrows indicate the direction across which the convolution operates. The output pixels contain random values.

Figure D.2 is a sketch of a 3D convolution operating on multiple channels with a certain depth. The depth represents the temporal dimension. The 3D convolution operates in the same way of the 2D convolution but slides through the depth dimension as well. Given the dataset dimensions, the kernel has size $1 \times 3 \times 3$. Hence, the depth is equal to 1. As a result, the 3D convolution effectively operates as a 2D convolution.

Figure D.3 represents a sketch of the max pooling method. It uses a kernel size of 2×2 and stride 2×2 , which is the number of pixels that the filter shifts at each step. The pooling operation halves the image dimensions, height and width, by retaining only the maximum value of each window. The combination of convolutions and max pooling generates a downsampling layer. Within a downsampling layer the image dimensions are halved and the number of feature maps is doubled.

Figure D.4 represents a sketch of a transposed convolution used for upsampling the feature maps in the decoder path. Given a kernel size of 2×2 and stride 2×2 , the transposed convolution operates in the same way

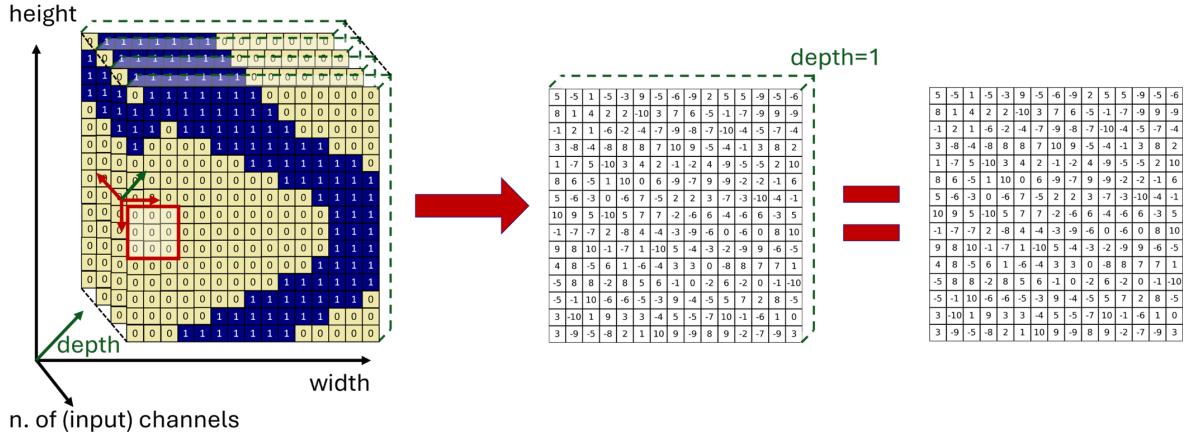


Figure D.2: Sketch of a 3D convolution operating on multiple input channels. The red and green arrows indicate the direction across which the convolution operates. As the input dataset has depth equal to 1, the final output of the 3D convolution resembles that of a 2D convolution. The output pixels contain random values.

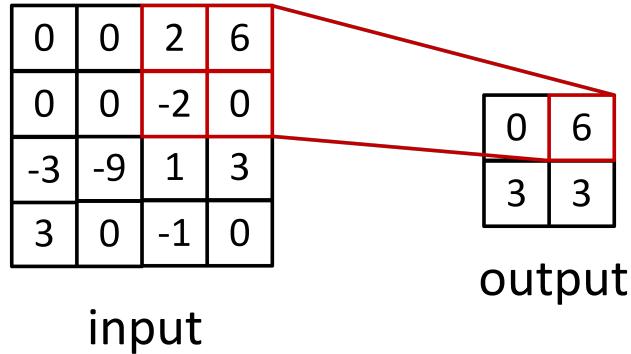


Figure D.3: Sketch of a max pooling operation. The input pixels contain random values.

of a convolution. Hence, it involves the presence of a kernel. The weights of this kernel are learned during the training process. The dot product of the weights with the input image pixel values generates the output. This operation doubles the feature map dimensions, height and width. The combination of a transposed convolution, a normal convolution, and the concatenation with skip connection (see Section 3.2.1) generates an upsampling layer. Within an upsampling layer, the feature map dimensions are doubled and the number of feature maps are halved.

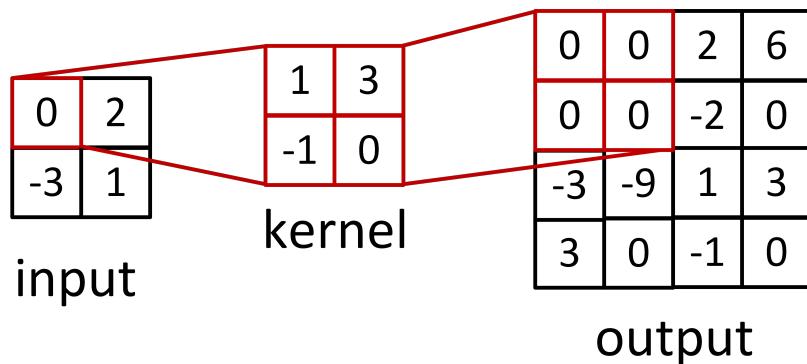


Figure D.4: Sketch of a transposed convolution. The input pixels and the kernel weights contain random values.

ReLU, Equation 3.1, and Sigmoid, Equation 3.2, are the two activation functions used within JamUNet. Figure D.5 illustrates the output of these functions. Note the difference in the scale of their outputs. ReLU generates positive unbounded values. Sigmoid generates only positive values bounded within the range 0 – 1. As a consequence, Sigmoid outputs can be interpreted as probabilities.

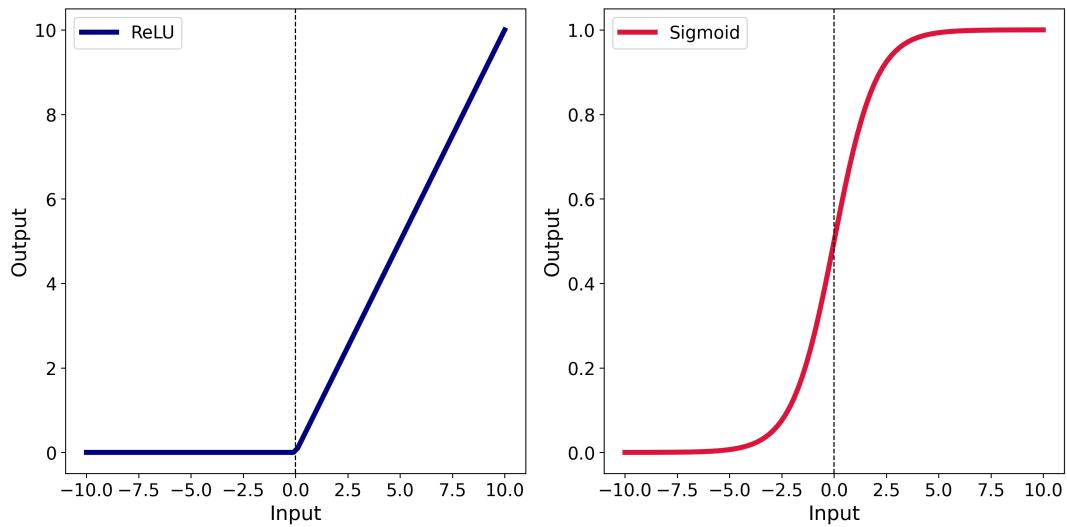


Figure D.5: Sketch of ReLU and Sigmoid activation functions.

E

Additional physics-induced loss

This appendix contains more information on the additional loss terms introduced in Section 3.2.3.

An attempt was made to include two physics-induced loss terms in addition to the BCE loss. The experiment aimed at checking whether introducing physics-inductive bias could help the training process. We expected that some physical information, retrieved from the available data, could improve the model learning and enhance the quality of predictions. The two additional terms represent the total area of erosion and the total area of deposition of the target year compared to the previous year.

The total area of erosion was computed by summing the number of “Water” pixel at the target year i that were “Non-water” at the last input year $i - 1$. This number was then multiplied by $60 \times 60 \text{ m}^2 = 3600 \text{ m}^2$, representing the image resolution area. A similar procedure was applied to the total area of deposition. All “Non-water” pixels in year i that were “Water” in year $i - 1$ were summed and multiplied by the same factor.

For these terms we no longer deal with a binary classification task but rather with a regression task. The Huber loss function (Equation E.3) was adopted. This loss balances two more common regression losses, namely MSE, Equation E.1, and MAE, Equation E.2. These are distinct measures of the difference between the observed and the predicted values. As for the performance metrics, the predictions had to be binarized before computing the Huber losses. The same water threshold of 0.5 was applied.

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (\text{E.1})$$

$$L_{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (\text{E.2})$$

$$L_{Huber} = \frac{1}{N} \sum_{i=1}^N \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{for } |y_i - \hat{y}_i| \leq \delta \\ \delta \cdot (|y_i - \hat{y}_i| - \frac{1}{2}\delta) & \text{otherwise} \end{cases} \quad (\text{E.3})$$

where y_i represents the target value, \hat{y}_i the binarized predicted value, and N represents the total number of samples. $\delta = 1$ is the threshold parameter that controls the transition between MSE and MAE behaviour of the loss function, and consequently the sensitivity of the loss function to outliers¹.

Equation E.4 represents the total training loss when physics-induced terms were included.

$$L_{total} = 1 \cdot L_{BCE} + \alpha_1 \cdot L_{Huber, er} + \alpha_2 \cdot L_{Huber, dep} \quad (\text{E.4})$$

where L_{BCE} is the classification loss, $L_{Huber, er}$ and $L_{Huber, dep}$ are the erosion and deposition areas regression losses, and α_1, α_2 the weights to scale erosion and deposition losses contributions to the total loss.

¹MSE loss is more sensitive to small outliers than MAE loss.

Simulations were run using the dataset introduced in Section 3.1.1. However, the evolution of the training and validation loss and metrics did not seem to bring improvements to the model training and learning. When the loss weights α_1, α_2 were too high, in the order of $1 \cdot 10^{-2}$ (-), the learning process worsened compared to the model trained without these additional terms. The training loss stagnated and the model predictions were not accurate. For lower values, in the order of $1 \cdot 10^{-4}$ (-), the influence of these additional terms was not visible and the training proceeded as for the case where no physics loss was included. Tuning these weights was done in order to scale the contributions of the individual terms to the total loss. Despite some attempts with different weights, proper values were not found and the experiment failed.

One possible reason for this failure is the complex combination of different losses, such as BCE for classification and Huber for regression. Additionally, as mentioned in Section 5.1, correctly predicting the total areas of erosion and deposition might not be enough informative for the model. These terms represent more abstract quantities that are not directly related to specific aspects of the inputs or outputs. The model might accurately predict the single values of the total areas of erosion and deposition but fail to capture their precise locations. In this case, the Huber losses would decrease, whereas the BCE loss would increase. Hence, the total areas of erosion and deposition seem to be weak and non-informative inputs for the model training. Using erosion and deposition locations across the input years as variables may be more suitable. This would allow for a purely classification-based loss, with no regression terms. Also, these variables are more closely related to the model inputs and outputs compared to the total areas of erosion and deposition.

F

Additional figures

This appendix contains additional Figures of the predictions generated by the min loss model trained with the dataset introduced in Section 3.1.3.

Figure F.1 shows the prediction of sample 2 using the water thresholds $w_{thr} = 0.402$. This threshold was found by optimising the F1-score in the PR curve (see Section 4.1.2). The image should be compared to Figure 4.6 to observe the influence of the last input year in the model output. It also shows that JamUNet tends to assign a larger weight to the “Water” areas occurrence rather than understanding the temporal evolution.

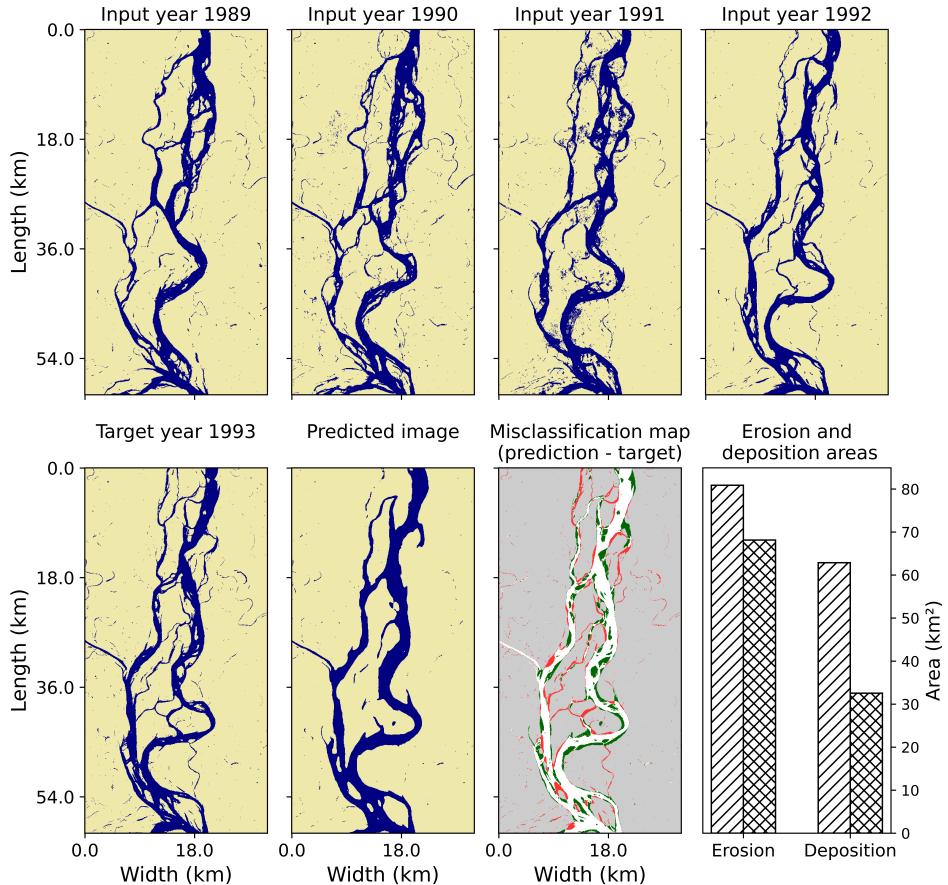


Figure F.1: Inputs, target, prediction, misclassification map and total areas of erosion and deposition of sample 2. The prediction is generated using the $w_{thr} = 0.402$.

Figure F.2 shows the prediction of testing sample 19. This is the only sample that can be visually compared to the outputs of the model trained with the temporal dataset (Figure H.4 in Appendix H).

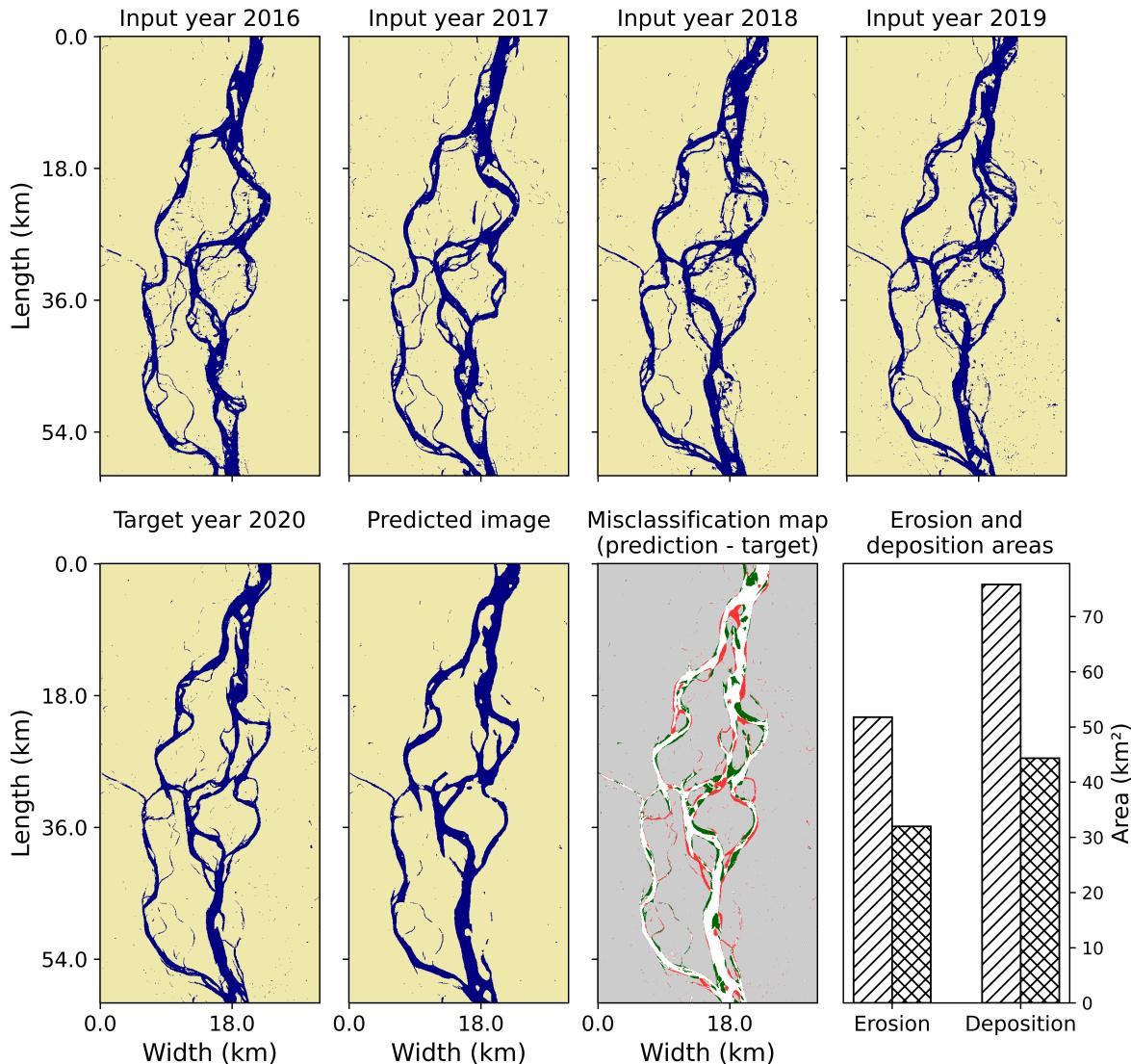


Figure F.2: Inputs, target, prediction, misclassification map and total areas of erosion and deposition of sample 19.

Figure F.3. In the first year after the bank stabilization in Chauhali, the model predicted a major erosion spot (point ① in Figure F4a). In the second year following the intervention, minor erosion was predicted as well (point ② in Figure F4b). This might be caused by the relevant meandering migration that the channel experienced within those input years (Figure F.5). However, the extent of the predicted erosion was significantly limited compared to the previous year prediction. This fact remarks again the significant impact of the last input year influence on the overall model output.

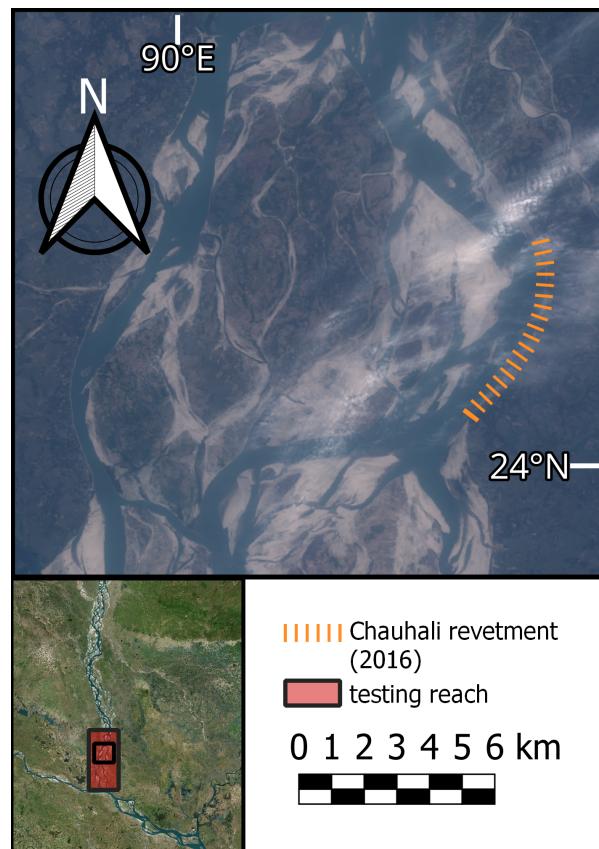


Figure E3: Location of the Chauhali revetment.

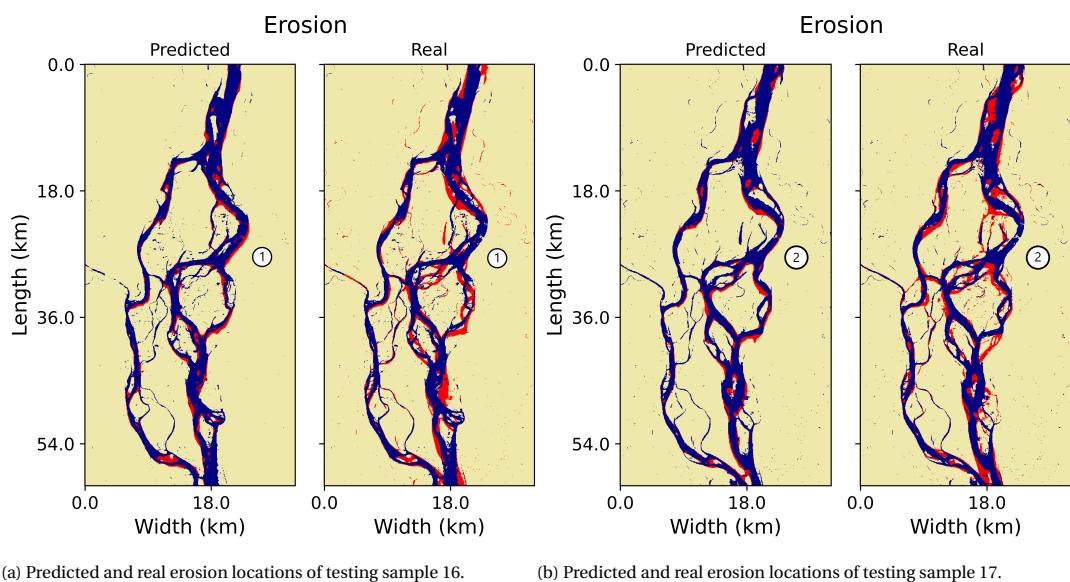


Figure F4: Predicted and real erosion locations of testing samples 16 and 17. Point ① and ② indicate the location of the Chauhali revetment and the predicted and real erosion.

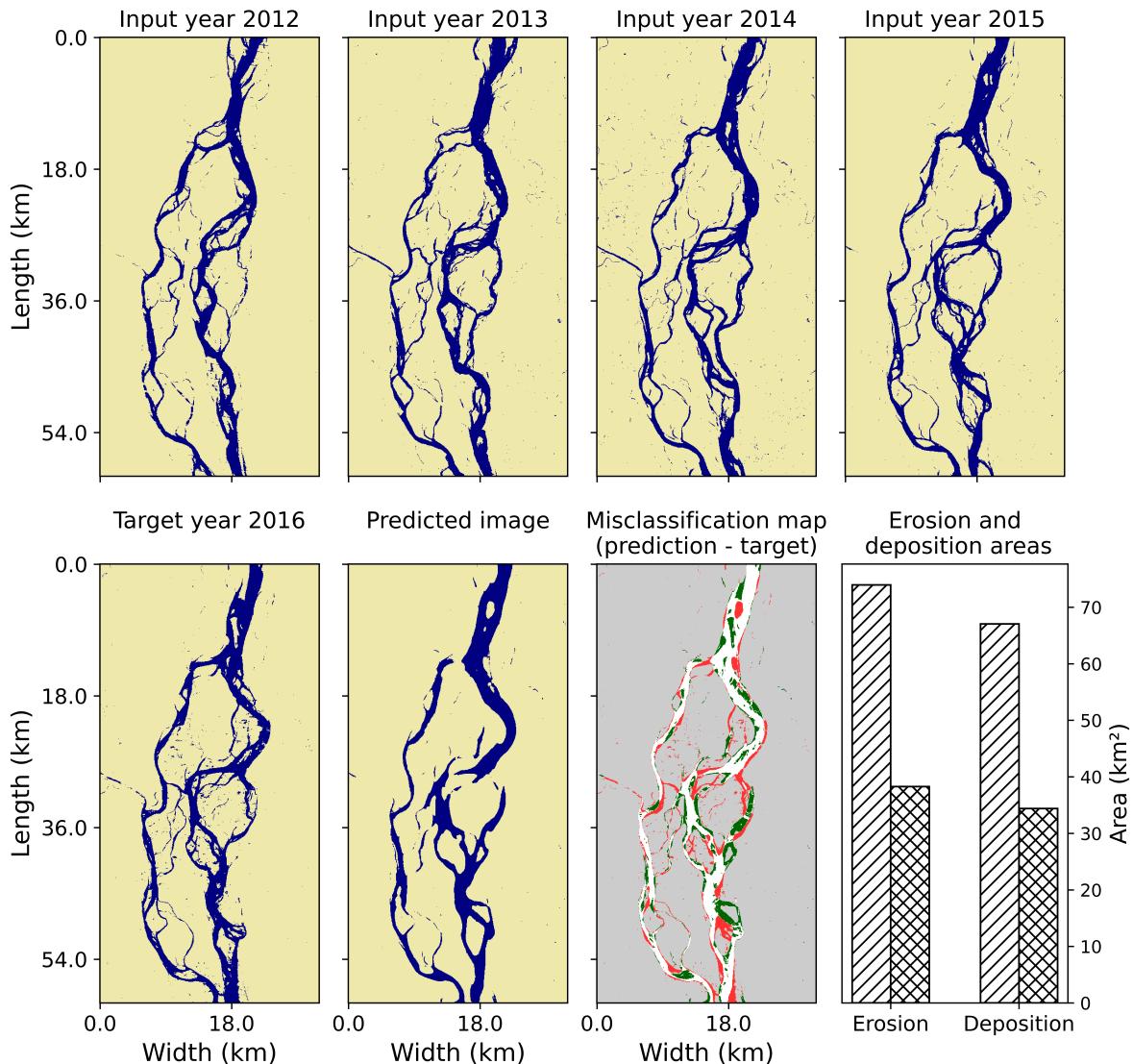


Figure F5: Inputs, target, prediction, misclassification map and erosion and deposition areas of sample 15. Note the evolution of the Chauhali channel during the input years (first row of images) before the implementation of the revetment.

G

Max recall model

This appendix contains information, results and figures of the max recall model.

This model was not selected as the main one because it achieves lower metric scores than those of the min loss model. Additionally, from a visual inspection, the results of the max recall model look similar compared to those of the min loss model. The difference lies in the amount of predicted “Water” areas, that is larger for the max recall one. As explained in Section 4.1.1, although this could be deemed as a positive feature due to the class imbalance, too conservative predictions are not desirable. Predicting “Water” areas indiscriminately within the braidplain is pointless, as it undermines the model ability to offer valuable insights into planform evolution. The only potential suggested application of the max recall model regards the prediction of meandering channel migration and bank erosion of the outermost areas of the braidplain. These are often underpredicted by the min loss model, while the max recall one produces more reliable forecast. As for the other processes such as channel shifting, channel formation and so on, the same issues exist for the max recall model. Furthermore, as this model predicts more “Water” areas, it underpredicts deposition process like channel abandonment.

Table G.1 contains the loss and metrics achieved on the test dataset by the max recall model. Figure G.1 shows the ROC and PR curves of the max recall model. The AUC values are 96.5% and 76.3%, respectively. Both values are lower than those obtained with the min loss model (see Section 4.1.2).

BCE loss	Accuracy	Precision	Recall	F1-score	CSI
$1.552 \cdot 10^{-1}$	0.930	0.670	0.777	0.718	0.561

Table G.1: Testing loss and metrics achieved by the max recall model. All quantities are adimensional.

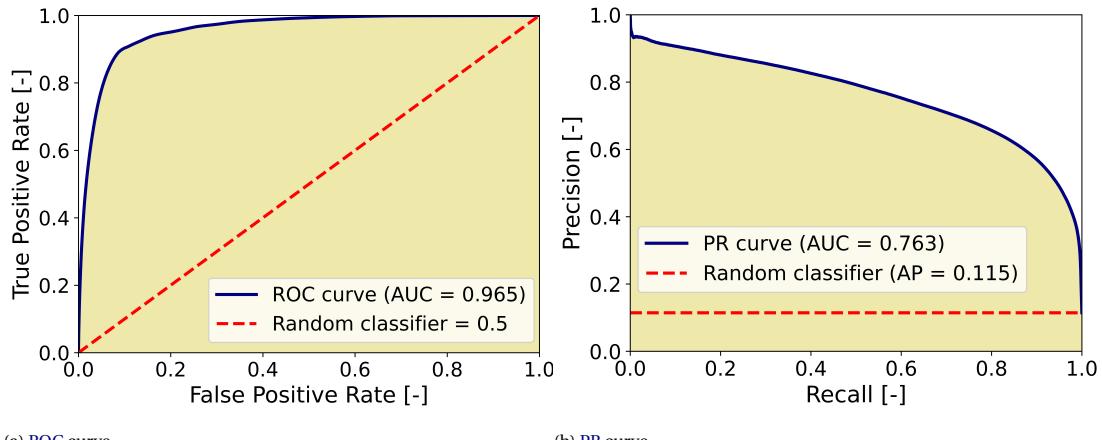


Figure G.1: ROC and PR curves for the max recall model.

Figure G.2 shows the boxplots of the loss and metrics achieved by the max recall model on the test dataset.

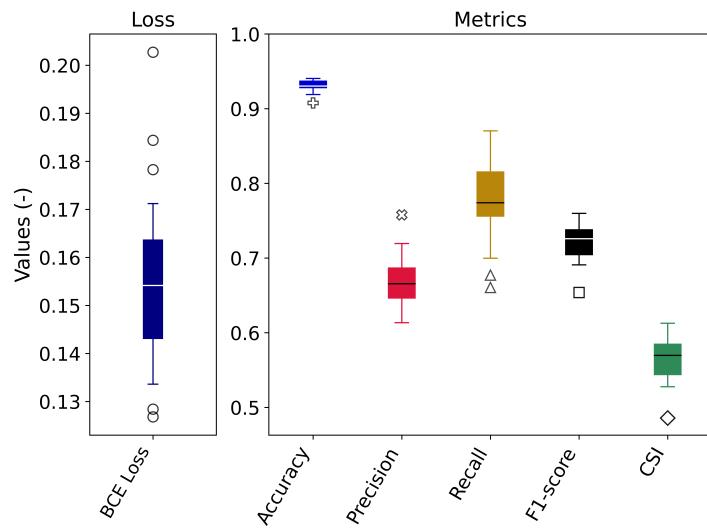


Figure G.2: Loss and metrics boxplots of the test dataset using the max recall model.

Figure G.3 shows the prediction of testing sample 2 by the max recall model. As compared to Figure 4.6, the same sample generated by the min loss model, it is possible to observe how more “Water” areas appear.

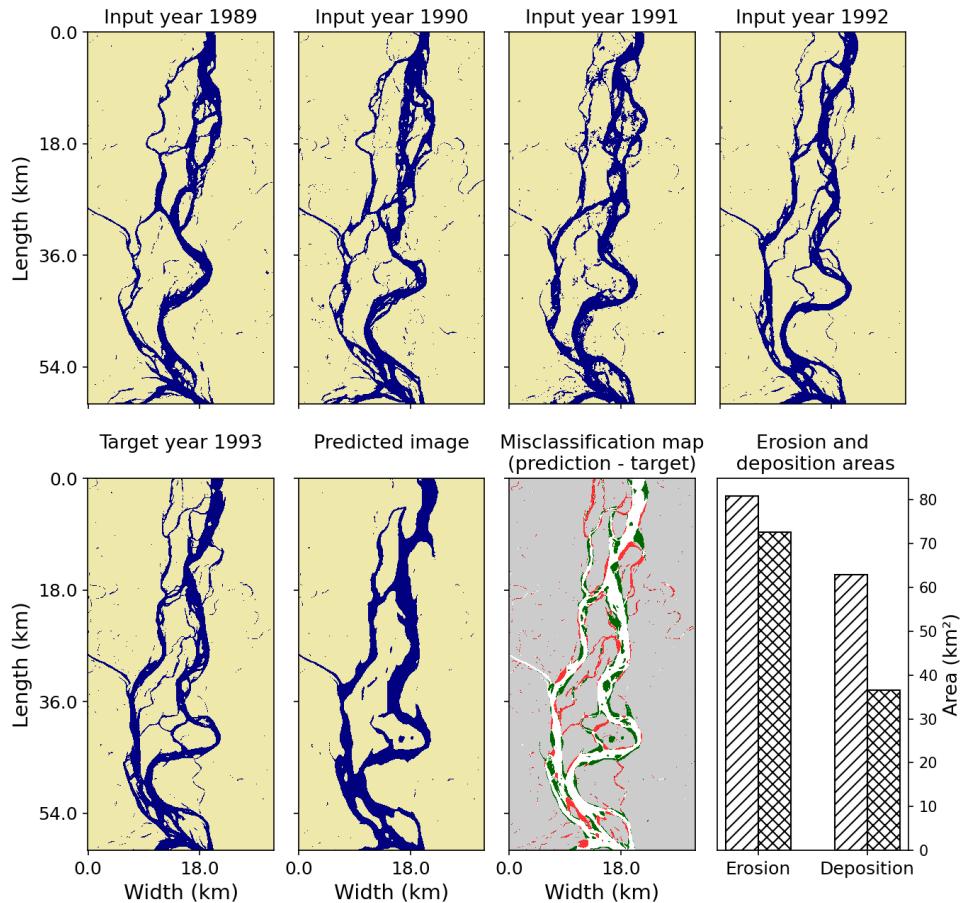


Figure G.3: Model prediction of testing sample 2 with max recall model. The image should be compared to Figure 4.6 to observe the differences between the min loss and the max recall model.

H

Temporal dataset

This appendix contains more information on the second splitting procedure adopted for generating the training, validation, and testing datasets. The results achieved with these datasets are reported as well.

H.1. Dataset generation procedure

The same methodology presented in Section 3 was applied. However, the training, validation, and testing datasets were generated by temporally splitting all available reaches. All input-target combinations from 1988 to 2009 were employed for the model training. Combinations from 2010 to 2015 were employed for the model validation. The remaining ones from 2016 to 2021 were used for testing the model. In this way we ensured that no image included in the validation and testing datasets was part of any training input-target combination, to avoid biased predictions and misleading results. The temporal splitting was introduced to check whether the temporal variation of the morphological features was more relevant and informative than the spatial variation. If that was the case, the model predictions were expected to improve. Also for this case only images from March were used. From here on, this dataset is referred to as the *temporal dataset*. This dataset is composed of 196, 30, and 30 samples for training, validation, and testing, respectively. It contains less training samples than the dataset introduced in Section 3.1.1¹ because no overlap between training and validation images was ensured. Moreover, all discarded combinations trace back to the initial years of record (see Appendix C.1). At the same time, the number of validation and testing samples is larger because one combination is created for each reach using the last years of record, which include all suitable images.

Testing sample number 1 represents the most downstream reach, just upstream of the confluence with the Ganges River at Aricha. This reach also represents the testing reach of the spatial dataset. The testing sample number identifier increases for each reach from 1 to 30 in the upstream direction.

H.2. Model results

The same model presented in Section 3.2, with the same hyperparameters, was trained with the temporal dataset. The evolution of the training and validation losses and validation metrics is presented in Figure H.1. Also in this case both the min loss and max recall models were stored.

Table H.1 contains the metrics of the min loss and max recall models achieved by training the model with the temporal dataset.

Model	BCE loss	Accuracy	Precision	Recall	F1-score	CSI
Min loss	$1.550 \cdot 10^{-1}$	0.926	0.646	0.621	0.635	0.466
Max recall	$1.930 \cdot 10^{-1}$	0.910	0.547	0.769	0.639	0.470

Table H.1: Testing loss and metrics achieved by the min loss and max recall models trained with the temporal dataset. All quantities are adimensional.

¹This dataset will be here referred to as the *spatial dataset*, to highlight the different conceptual splitting procedure.

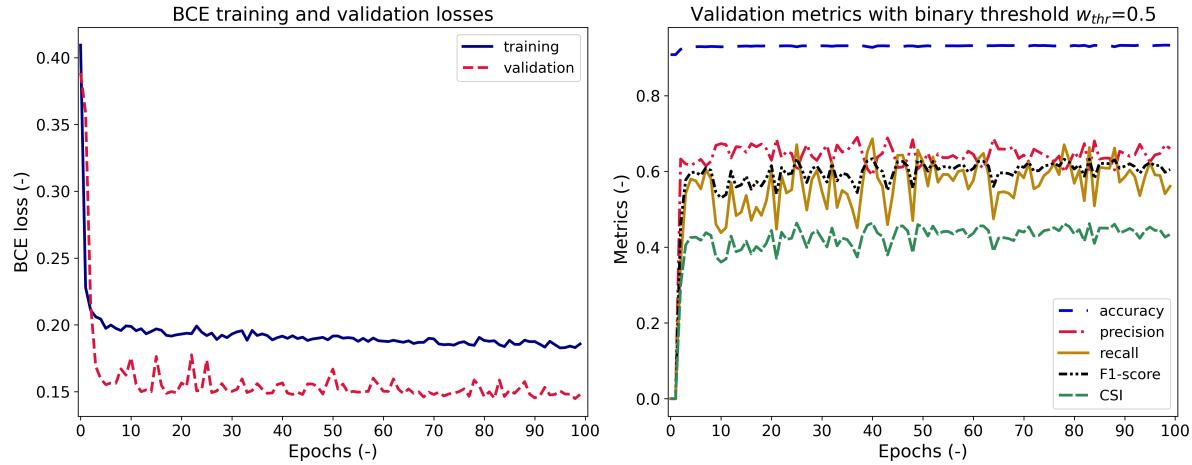


Figure H.1: Evolution throughout epochs of the training and validation losses (left plot) and validation metrics for the model trained with the temporal dataset.

Figure H.2 shows the ROC and PR curves of the min loss model trained with the temporal dataset. The achieved AUC are 95.6% and 68.1%, respectively. Both values are lower than those obtained with the spatial dataset and presented in Section 4.1.2.

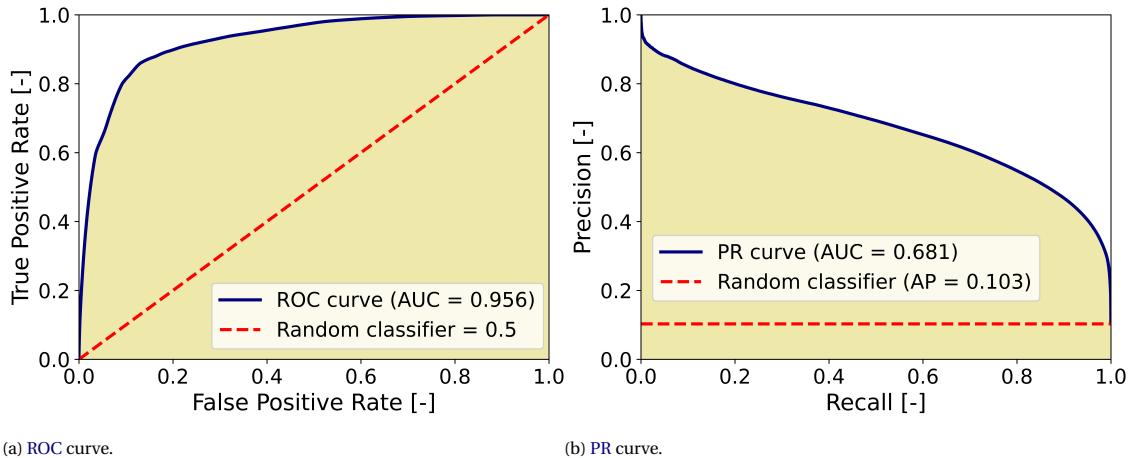


Figure H.2: ROC and PR curves for the min loss model trained with the temporal dataset.

Figure H.2 shows the ROC and PR curves of the min loss model trained with the temporal dataset. The AUC values are 94.8% and 62.3%, respectively. Both values are lower than those obtained with the spatial dataset and presented in Section G.

Figure H.4 shows the prediction of testing sample 1 using the min loss model trained with the temporal dataset. It should be compared to Figure F.2, which represents the same prediction achieved by the min loss model trained with the spatial dataset. The two predicted images look similar. The model trained with the temporal dataset predicts even less areas of erosion and deposition compared to the model trained with the spatial dataset. Both models generate disconnected channels, although these are found in different locations. It is still visible the lack of temporal evolution from the migration of the meandering channel that flows on the right side of the central bifurcation.

H.3. Motivation

Given the results achieved by the model trained with the spatial dataset presented in Section 4, we tested whether this different dataset would improve the performance. If that was the case, we could have concluded that, for predicting braided sand-bed rivers planform changes, the type of dataset, spatial or temporal, plays

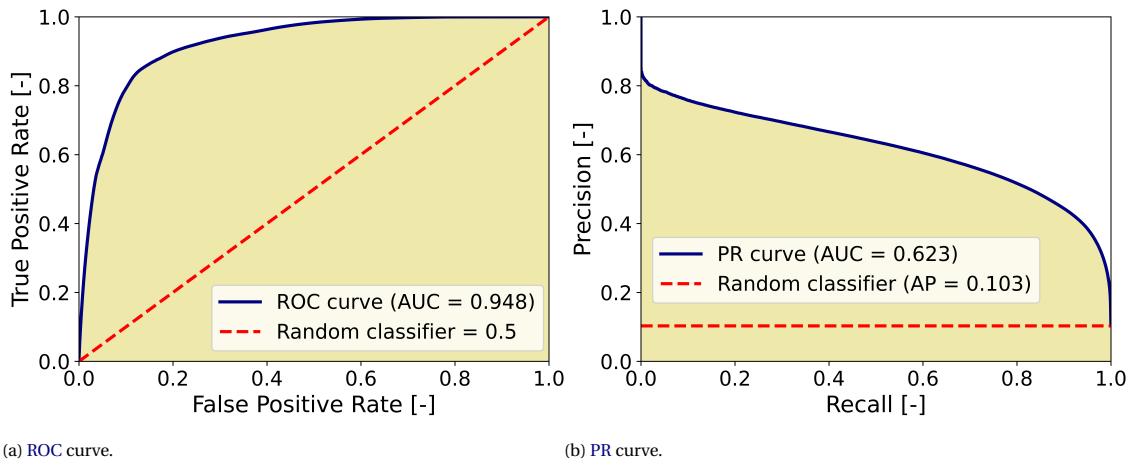


Figure H.3: ROC and PR curves for the max recall model trained with the temporal dataset.

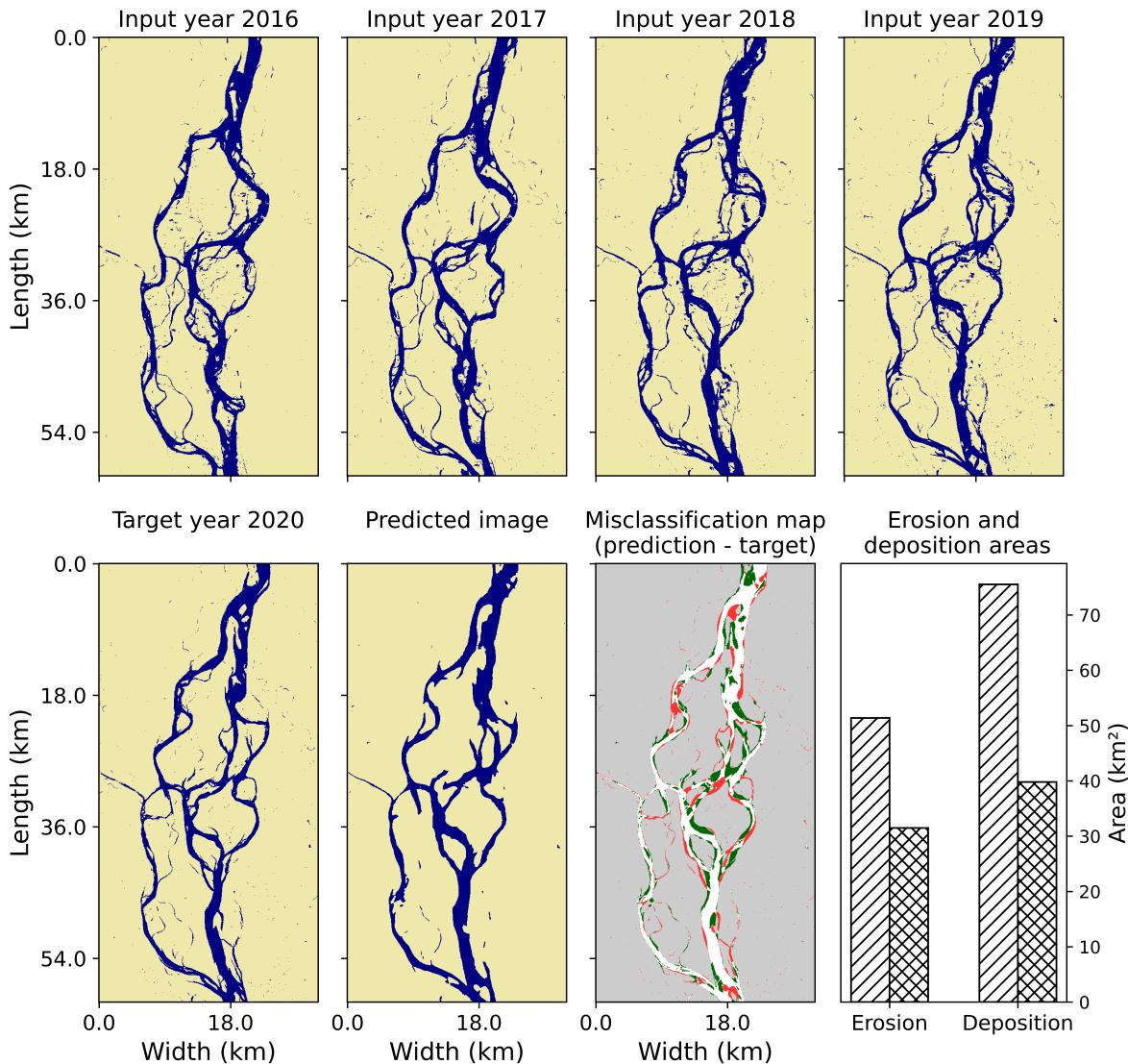


Figure H.4: Model prediction of testing sample 1 with min loss model trained with the temporal dataset.

a crucial role. In other words, we could have concluded that the spatial variability of the planform changes is less relevant and informative than the temporal variability. This was not the case eventually, as the same model trained with the temporal dataset achieved worse results. It is true that the temporal dataset had a smaller training dataset. This reflects in the learning process, which stops quite soon after the beginning of the simulation (see Figure H.1).

Eventually, the temporal dataset was not chosen for two reasons. The model clearly achieves lower metrics. Additionally, the desired generalizability of the model is reduced. With this splitting procedure, in every dataset the same reaches across different years are used to train, validate, and then test the model. For this, the performance on the validation and test datasets could be partly biased. The model might have learnt specific features of the training years and reproduced these features during the inference steps. Hence, for future applications on the validation or testing reaches, or for testing the model with other braided sand-bed rivers, the model is likely to generalise worse compared to the one trained with the spatial dataset.

Using the spatial dataset the model performance could have slightly degraded for the given case study, as the validation and testing steps are performed with completely unseen data. This is actually not the case, likely because the spatial dataset has a larger training dataset size compared to the temporal dataset. Regardless of this, a spatial split of the dataset is recommended to achieve a better generalizability of the model to other braided (sand-bed) rivers. Contrary, if the goal of the research is achieving a model for a specific case study then the temporal split could be desirable. In this case the model is not expected to generalise to other case studies as well as a model trained with a spatial dataset could do.