Two data-driven approaches for predicting bed levels in 3D for the Waal River

MSc Thesis

L.I. (Luuk) van Laar

April 24, 2024

Figure 1: Waal River near Ochten (Siebe Swart, 2019)

Graduation Committee:	
Prof. Dr. S.J.M.H. Hulscher	(University of Twente)
Dr. F. Huthoff	(University of Twente & HKV)
Ir. T. Stolp	(HKV)

UNIVERSITY OF TWENTE.

HKV LIJN IN WATER

Preface

Before you lies the MSc thesis for my Master's degree in Civil Engineering and Management with a specialisation in River and Coastal Engineering. The research was done in collaboration with the University of Twente and HKV lijn in water. I have learned a lot during this thesis project and would like to thank everyone involved.

Firstly, I would like to thank my supervisors Thomas Stolp, Freek Huthoff and Suzanne Hulscher for their help during the thesis. Thomas Stolp, you were very helpful and always made time for me. Your expertise in machine learning helped to gain a lot of insights into this new field for me. Also, you involved me in many of HKV's social activities, where I got to know fellow HKV colleagues, which was really nice. Freek Huthoff, you helped me maintain the academic quality of the study together with Suzanne Hulscher. You were critical at times of need, but supportive when things went well. Suzanne Hulscher, you were very insightful with the feedback you provided, concise and very good to the point. Thank you for that.

Next to my supervisors, I would like to thank my family and friends. They were always there for me and supported me in times of need. I would like to thank especially my parents, as they were much involved in the time during the thesis. For example, the evening walks 'just to get the 10,000 steps on the smart watches' were nice moments to relax.

Lastly, I would like to thank Andries Paarlberg, Pepijn van Denderen and Lieke Lokin at HKV. You were all very open to discuss my research with and provided interesting insights for my research. Lastly, I would like to thank everyone else who helped me during the thesis project.

I hope you enjoy reading this thesis.

Luuk van Laar

Enschede, April 2024

Abstract

The Dutch waterways, like the Waal River, are among Europe's busiest navigation routes. In these waterways, navigation faces challenges during periods of low flow due to shallow water depths. Dredging is conducted to keep the waterways navigable. As future projections indicate more extreme low-flow events, the need for efficient dredging strategies increases. Currently, short-term dredging decisions rely on real-time bed level data. Providing forecasts on bed levels can support this decision-making process by identifying potential future bottlenecks, deeper locations for sediment disposal and areas where subsequent measurements may not be necessary.

This study aimed to obtain a two-weekly bed level prediction in 3D for the Waal River, by developing and comparing two data-driven approaches. This included the development of a data-driven dune migration model, which only considers the horizontal displacement of bed features, alongside improving the bed level prediction accuracy of the machine learning model TrajGRU (Shi et al., 2017). The initial machine learning model TrajGRU produced highly blurred predictions with the bed level predictions converging to the mean bed level. The difference in the observed and predicted maximum bed levels was large (0.92m) and little of the bedform patterns were captured accurately. To improve these predictions, five data preprocessing experiments were set up: a sample selection, removal of outliers, coordinate transformation, wavelet reconstruction and consistent time intervals. Herein, each successive experiment built upon the previous experiments. Furthermore, three different loss functions were tested in the machine learning model: the root mean square error, a combination of the mean square error and structural similarity index and the Wasserstein loss.

The dune migration model showed to outperform the improved machine learning model. The machine learning model with the wavelet reconstructed data and the root mean square error loss function improved model performance most of all machine learning experiments. The maximum bed level error reduced from 0.92m to 0.77m and bedform patterns were more accurately captured compared to the other experiments. However, the blurring effect was still largely present. Compared to the dune migration model, the dune migration model had a substantially smaller maximum bed level error of 0.30m, the locations of the maximum bed levels were captured more accurately with a locational error of 14.5m instead of 25m for the machine learning model, and the overall bedform patterns were captured more accurately as well. Even when manually increasing the bed level spread of the machine learning prediction, the dune migration model outperformed the machine learning model.

Thus, the potential of a complicated machine learning model like TrajGRU for predicting bed levels in the Waal River shows to be small and the dune migration model proves more promising. Future research could focus more on the dune migration model and combine this method with an approach suitable for predicting vertical bed level changes as well.

Contents

1 Introduction

1.1 Context

The Dutch rivers belong to the busiest navigation routes in Europe (Sys et al., 2020; Vinke et al., 2022). Within the Netherlands, the Waal River is considered the busiest navigation route (Rijkswaterstaat, 2023). During periods of low flow, navigation encounters problems with low water depths at the Waal River, allowing less loading capacity and therefore increasing costs (Hiemstra et al., 2020; Ministry of Infrastructure and Water Management, 2022). To keep the Dutch waterways navigable, they are maintained via dredging (Rijkswaterstaat, 2023). Among the Dutch rivers, the Waal River undergoes the most frequent dredging (Rijkswaterstaat, 2023), with an annual dredging volume reaching 700,000m^3 (Rijkswaterstaat, 2020). Due to climate change, low flows in the Waal River are expected to become more extreme in the future (Buitink et al., 2023), increasing the need for dredging and therefore efficient dredging strategies.

Dredging strategies contain both long- and short-term perspectives. The long-term perspective focuses on the required budgeting for the upcoming year and determines whether large-scale dredging is necessary. On the other hand, the short-term perspective primarily focuses on local bottleneck dredging, requiring a short response time of a few days to two weeks (Klaassen & Sloff, 2000). The focus of this study is on the short-term perspective.

Current decisions on short-term dredging rely on real-time bed level data. Dredging is initiated when the bed level measurement exceeds the dredging reference level set by Rijkswaterstaat (Kruis et al., 2023). Providing a short-term prediction on the bed levels can contribute to effective channel regulation (Y. Liu et al., 2022). Such predictions can offer insights into potential future bottlenecks, identify deeper locations for sediment disposal and pinpoint areas where subsequent measurements may not be necessary due to low bed level predictions.

Bedforms in the Waal River consist of rhythmic features that propagate downstream. These features include primary river dunes, secondary bedforms and large-scale features such as river bars. The large-scale features such as river bars extend over kilometres long, are stable in location and have heights of ∼1m (de Ruijsscher et al., 2020). They change in shape with a dynamic timescale in the order of months to years (van Dijk et al., 2012). Secondary bedforms are smaller bed features superimposed on primary dunes. These bedforms have been observed in the Waal River with heights below 0.5m, lengths below 10m and downstream migration rates of around 1m/h (Zomer et al., 2021). Primary river dunes in the Waal River generally are observed from 10cm to 1.5 meters high, extend from 10 to 150m long, and migrate downstream several meters per day (Lokin et al., 2022; ten Brinke et al., 1999; Wilbers & ten Brinke, 2003).

Given the focus on short-term 14-day bed level predictions, capturing primary river dune development is most relevant for short-term prediction of bed levels. Including secondary bedform dynamics would be highly complex and would require hourly measurements to capture its dynamics. Modelling large-scale feature dynamics is also less relevant as their dynamic timescale extends beyond the 14-day prediction window.

Numerous numerical models have been developed to predict short-term river dune developments in 2DV (Giri & Shimizu, 2006; Nelson et al., 2011; Paarlberg et al., 2009) and 3D (Goll & Kopmann, 2012; Nabi et al., 2012, 2013a, 2013b). The 2DV models generally fail to capture the 3D dynamics of river dunes, featuring variability across the field in crest curvature, discontinuity, and height variations (Lefebvre, 2019). Still, Lokin et al. (2024) showed the capability of the 2DV model by Paarlberg et al. (2009) to accurately predict the growth rate of the highest dunes. However, migration rates were largely underestimated and growth rates of small dune dynamics were overestimated. The 3D models can capture the three-dimensional dynamics of river dunes better, but require excessive computational resources to implement for river-scale applications (Lokin, 2020). As there is a relatively large bed level dataset available for the Waal River, data-driven approaches are becoming attractive as well. Yet, very few studies have applied data-driven approaches for predicting bed levels (B. Liu et al., 2023).

In this study, one of the data-driven approaches used for predicting bed levels in the Waal River involves the application of the machine learning model TrajGRU (Shi et al., 2017). Previous to this study, HKV revisited the TrajGRU machine learning model structure to predict bed levels in the Waal River (Appendix A). The revisited model can provide a two-weekly prediction on the bed levels, using two sequential bed level measurements as input (Figure 2). Current model results show a highly blurred output with the bed levels converging to the mean bed level. Bed level differences between the model prediction and observation are in the order of decimeters to meters. Furthermore, the model tends to reconstruct the input frames and fails to capture newly developed bed features. Still, there are opportunities to improve model predictions by exploring data preprocessing and adapting loss functions.

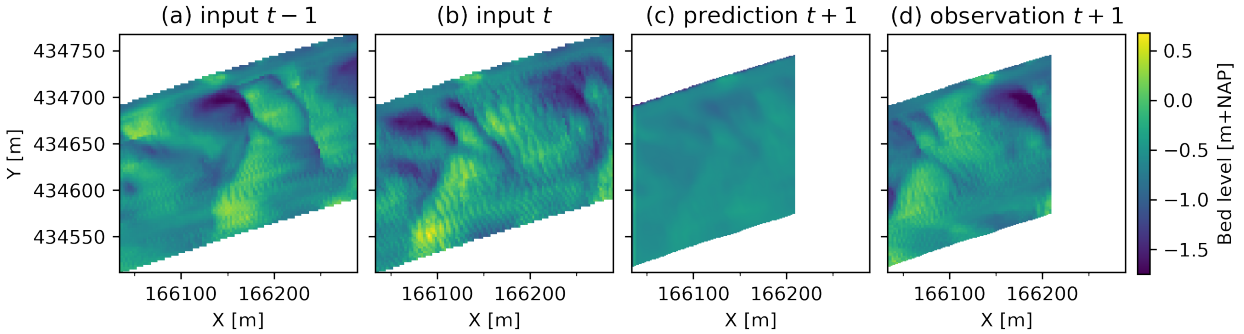


Figure 2: Example prediction of the TrajGRU model for a section of the Waal River near Ochten. t+1 denotes the predicted timestep, t the real-time situation and t-1 the previous timestep.

1.2 Research Objective

The objective of this research is to obtain a two-weekly bed level prediction in 3D for the Waal River, by developing and comparing two data-driven approaches. This includes the development of a dune migration model, alongside improving the bed level prediction accuracy of the machine learning model TrajGRU. The bed level accuracy of the machine learning model will be improved via data preprocessing and the adaptation of loss functions. Both data-driven approaches focus on a 3D prediction, such that the three-dimensional variability of the river bed can be captured. The lead time of the prediction is taken as two weeks, which provides a sufficient response time for dredging (Klaassen & Sloff, 2000).

1.3 Research Questions

To achieve the research objective, four research questions were formulated. The first question focuses on the bed level prediction accuracy of the dune migration model for predicting bed levels in the Waal River. The second and third research questions focus on improving the bed level prediction accuracy of the existing TrajGRU machine learning model via the implementation of data preprocessing and the adaptation of loss functions. The fourth research question focuses on a comparison between the machine learning model and the dune migration model for predicting bed levels in the Waal River. Specific research questions are:

RQ1: What is the bed level prediction accuracy of the data-driven dune migration model?

RQ2: How can data preprocessing be implemented to improve the bed level prediction accuracy of the TrajGRU machine learning model for the Waal River?

RQ3: How can the loss function be adapted to improve the bed level prediction accuracy of the TrajGRU machine learning model for the Waal River?

RQ4: How well does the improved TrajGRU machine learning model perform compared to the dune migration model for predicting bed levels for the Waal River?

1.4 Outline

The outline of this thesis is illustrated in Figure 3. In the next chapter, the research methodology is described. Herein, the study area and available data are described, together with the dune migration model, the applied machine learning approach, the data preprocessing and loss function experiments, and the performance evaluation method. The results are presented in chapter three. Example predictions of the dune migration model and the machine learning experiments are visually provided, together with supporting performance metric scores. Also, a comparison is given between the bed level predictions by the dune migration model and the machine learning model. Chapter four discusses the research results and chapter five gives the conclusions on the research. Finally, chapter six provides recommendations for future studies. The appendices contain a detailed overview of the machine learning model structure, a visualisation of the outlier locations in the bed level data, example bed level predictions of the two data-driven approaches, a brief analysis on the discharge implementation in the machine learning model and the poster of this research presented on the NCR days 2024.

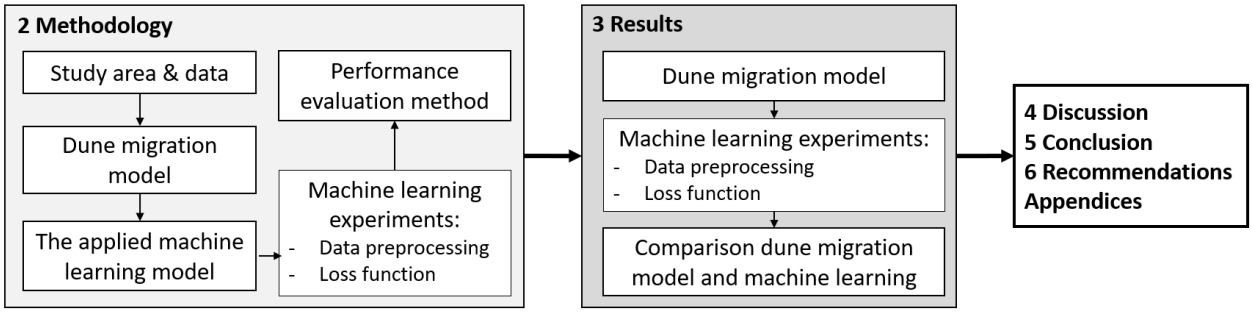


Figure 3: The outline of this thesis

2 Methodology

The methodology of this study is summarised in Figure 4. Two data-driven approaches are focused upon: a dune migration model and machine learning. The dune migration model applies a raster shifting approach, for which, a bed level measurement is shifted along the channel-wise axis with an empirically derived dune celerity to obtain the bed level prediction. For the machine learning model, five data preprocessing and three loss function experiments were set up to improve the bed level prediction accuracy. For the data preprocessing experiments, modifications were solely applied to the input data. Each successive experiment builds upon the modifications introduced in the preceding one. Such a systematic approach was thought to achieve a cumulative improvement of the model performance. The data preprocessing step showing the most improvement in the model performance was further used in the loss function experiments. Here, three different loss functions were tested. The best performing machine learning model was derived, which was tested against the dune migration model. The comparison between the two data-driven approaches indicates which one holds the most potential for predicting bed levels in the Waal River.

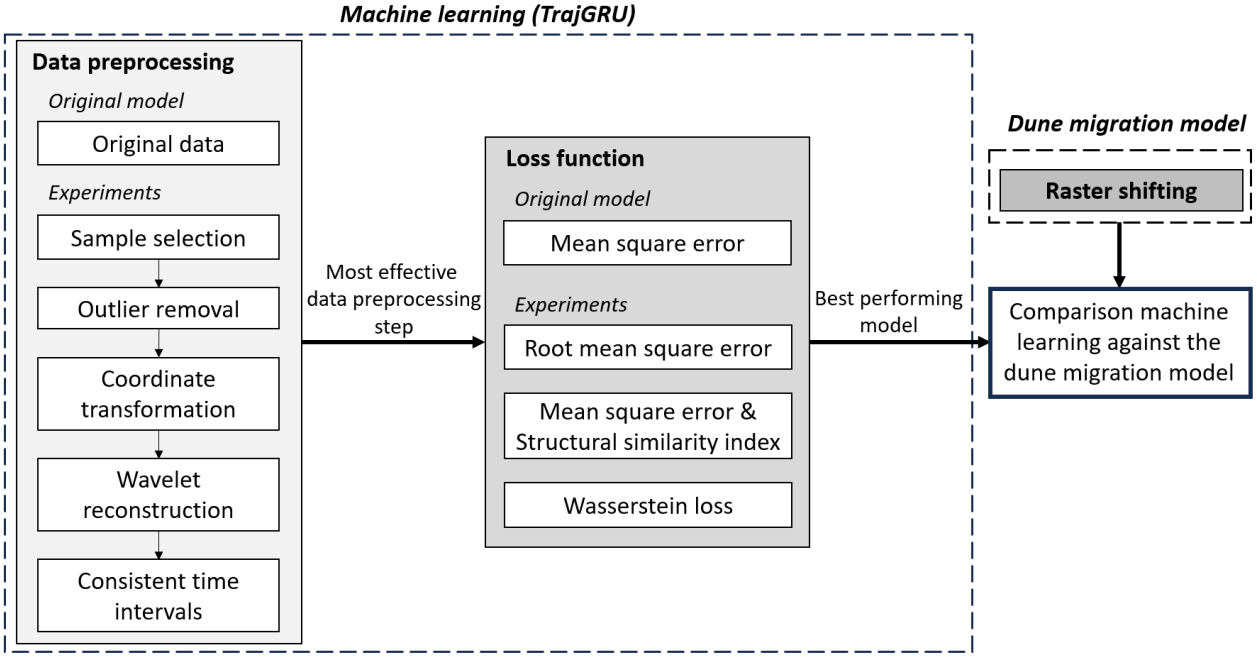


Figure 4: Flowchart of the research methodology of this thesis

This section starts with a description of the study area and available data. Then, the dune migration model is described. After this, the applied machine learning model is explained with a focus on data preprocessing and loss functions. Then, the data preprocessing and loss function experiments are elaborated upon. Finally, the method for evaluating the model performance is discussed.

2.1 Study area and data

This study focuses on a 270m long river section in the Waal River near Ochten (Figure 5). The study area was selected such that it represents a simple situation to test the suitability

of the data-driven approaches. The river section has not experienced ploughing or dredging, based on dredging and ploughing data between 2021 to 2023. Therefore, natural bedform developments are better captured in the data than for heavily dredged or ploughed areas (van Dijk et al., 2012). Also, the section is relatively straight, excluding complex flow processes like spiral flow effects in river bends (Beygipoor et al., 2013). The total study area covers an area of 256x256m.

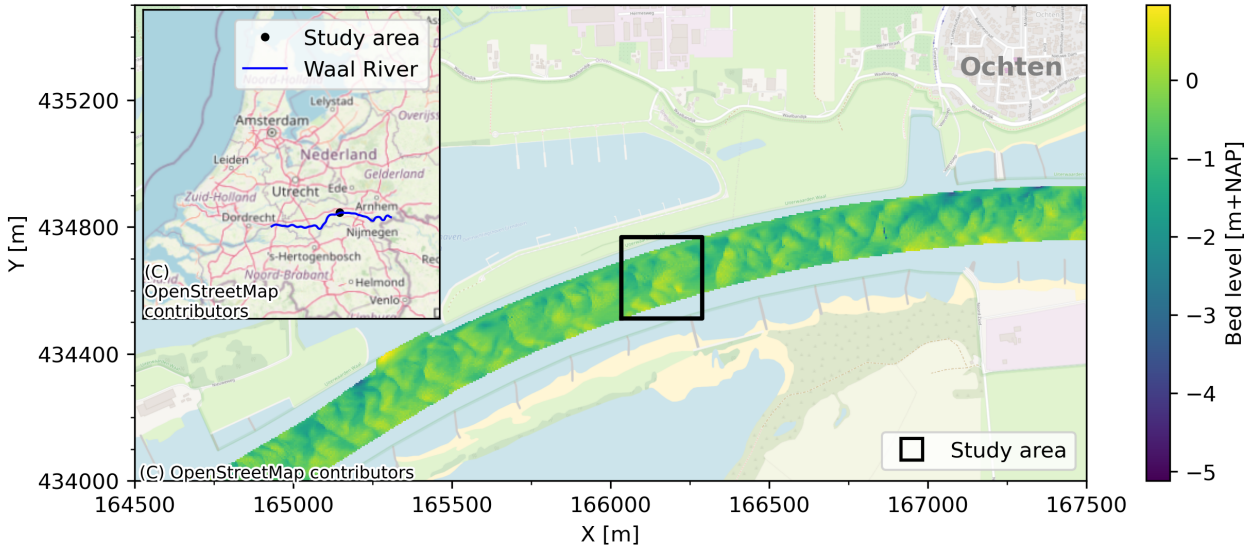


Figure 5: Study area of the Waal River section near Ochten in the Netherlands, plotted in the 'RD new' coordinate system. The multibeam echosounder measurement of 2005-07-19 is plotted here as example.

The bed level data consists of 394 MultiBeam EchoSounder (MBES) measurements aggregated to a 1x1m raster grid, covering the ∼200m wide navigation channel for the period 2005 to 2021. Time intervals between measurements are 14 days on median, but vary quite substantially from 5 to 55 days (Figure 6). The time interval histogram follows a normal distribution with a standard deviation of 2.9 days for time intervals smaller than 27 days. The time intervals of 27 days or more are due to missing MBES measurements between readings.

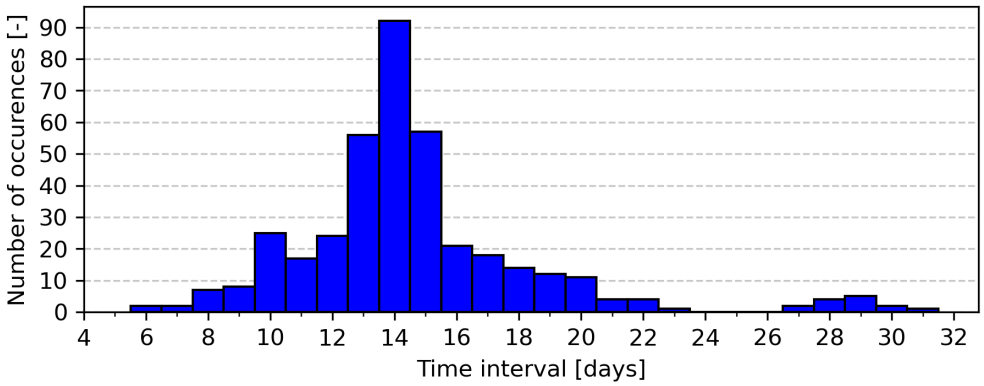


Figure 6: Histogram of the time intervals between MBES measurements for the period 2005 to 2021. Two outliers of 41 and 55 days, both occurring once, are excluded from the figure.

The riverbed in the Waal River consists of primary river dunes, secondary bedforms and large-scale features such as river bars. As mentioned in Section 1.1, the understanding of river dune dynamics is most relevant, given the focus on short-term 14-day bed level predictions. River dunes are rhythmic features in sand and gravel beds that propagate in a streamwise direction due to erosion at the stoss side and sedimentation on the lee side (Bradley & Venditti, 2021; Mohrig & Smith, 1996). River dunes in the Waal River are observed from 10cm to 1.5 meters high, extend from 10 to 150m long, and migrate downstream several meters per day (Lokin et al., 2022; ten Brinke et al., 1999; Wilbers & ten Brinke, 2003). A schematisation of a river dune is shown in Figure 7. Their development is driven by current-induced drag forces. They are therefore highly sensitive to river discharge. During low discharges, river dunes in the Waal River tend to diffuse and reduce their migration rate, as the relative strength of gravitation on sediment particles is higher than the current-induced drag force (Lokin et al., 2022). During high flow conditions with dominant bed load transport, river dunes will grow in height and develop until equilibrium is reached (van Duin et al., 2021). Dune lengths decrease (Cisneros et al., 2020; Lokin et al., 2022) and migration rates increase (Lokin et al., 2022; ten Brinke et al., 1999; Wilbers & ten Brinke, 2003). For extremely high flows (i.e., during dike design conditions), the dominant mode of transport shifts from bed load towards suspended transport (van Rijn, 1984), leading river dunes to decay in height (Naqshband et al., 2014). This continues until the dunes have been washed out and a flat riverbed remains (Hulscher et al., 2017; van Duin et al., 2021).

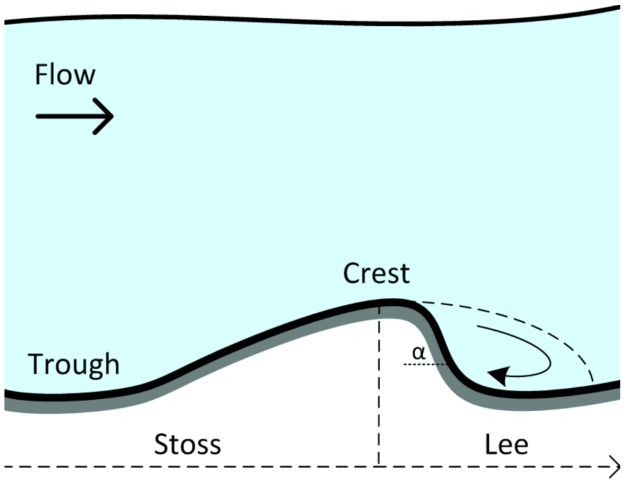


Figure 7: River dune schematisation. α represents the lee side angle (Naqshband et al., 2011)

2.2 Data-driven dune migration model

A data-driven dune migration model was developed to predict the bed levels in 3D for the Waal River. When trying to predict short-term bed evolution, key processes to be captured are migration speed and growth/decay of the river dunes (length and height). Under modest changes in flow conditions dunes still migrate meters per day but may not change significantly in shape (Lokin et al., 2022; ten Brinke et al., 1999). Therefore, the dune migration model only considers the horizontal displacement of the river bed. The model applies a method further referred to as raster shifting, for which, a bed level measurement is shifted downstream along the channel axis by a distance L_{sh}, giving the bed level prediction (Figure 8). The shifting distance (L_{sh} $[m]$) is by summing the daily dune celerities (c_i $[m/d]$) over the 14 day prediction period (equation 1).

$$L_{sh} = \sum_{i=1}^{14} c_i \tag{1}$$

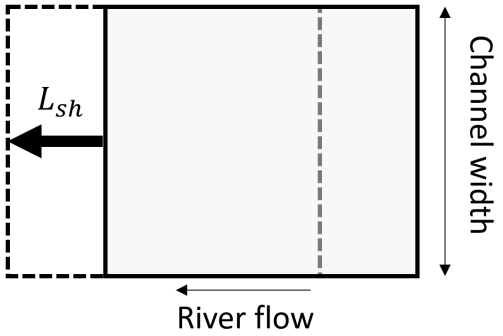


Figure 8: Schematisation of the raster shifting prediction approach. The grey box indicates the bathymetry measurement. L_{sh} is the distance in meters with which the bed level measurement is shifted.

Data on daily dune celerities is highly limited. Therefore, a relation was established between the dune celerity and river discharge, for which daily measurements are available for the period 2005 to 2021. Equation 2 depicts an empirically derived third-order relation between the river discharge (Q $[m^3/s]$) and dune celerity (c $[m/d]$) for the Waal River, based on data by Lokin et al. (2022). This relation provides a reasonable estimate of the dune celerity (Figure 9a), despite the relatively high uncertainty ($R^2 = 0.61$). Dune celerities increase with increasing discharges and stabilize near discharges of $3000 m^3/s$ and higher. Figure 9b shows a discharge time series and the derived dune celerities by equation 2. The dune celerity generally follows the discharge curve and shows high variability over a short period. This high variability indicates the importance of accumulating daily dune celerity as applied in equation 1.

$$c = 0.748 + 4.26 \cdot 10^{-3} Q - 1.20 \cdot 10^{-6} Q^2 + 1.11 \cdot 10^{-10} Q^3 \tag{2}$$

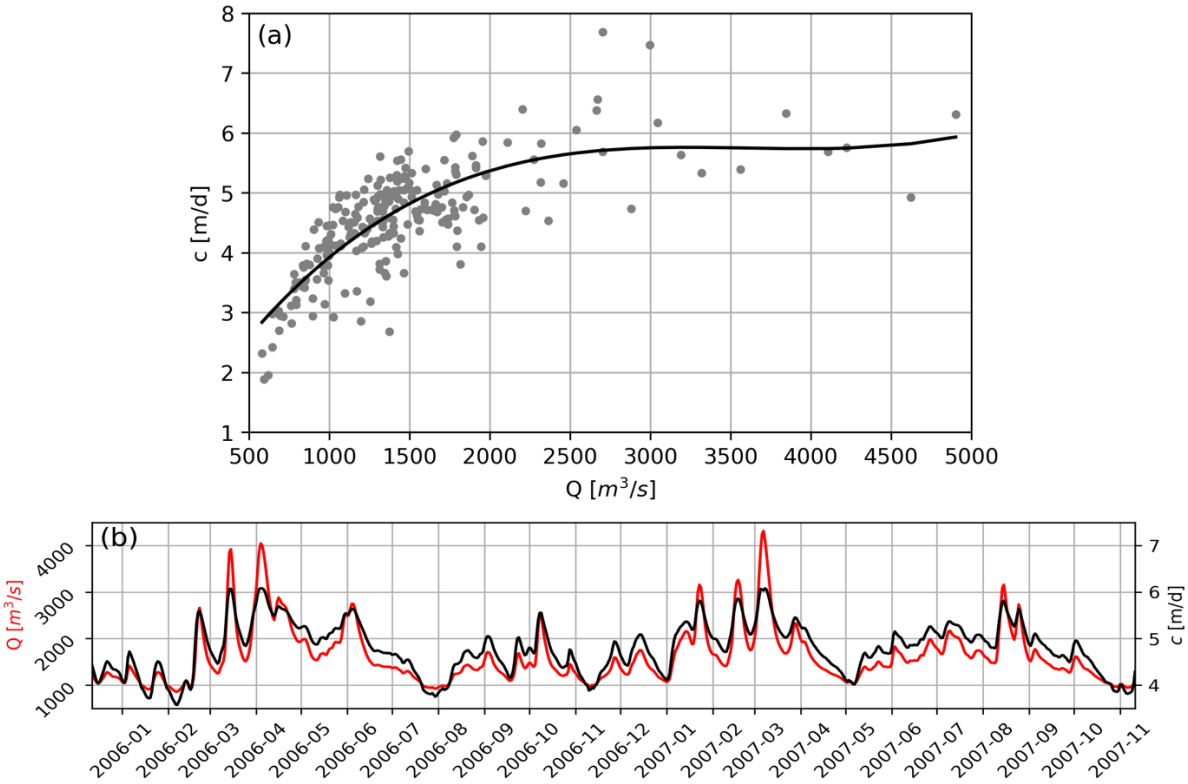


Figure 9: (a) Dune celerity against the discharge. The grey dots indicate the dune celerity data by Lokin et al. (2022). The black line indicates equation 2. (b) The measured discharge time series plotted in red and derived dune celerities from equation 2 plotted in black.

2.3 The applied machine learning approach

Next to the dune migration model, machine learning is applied as a data-driven approach to predict bed levels in the Waal River. This study applies the machine learning model TrajGRU (Shi et al., 2017). This model was originally developed for short-term forecasting of precipitation, but was revisited by HKV for predicting bed levels in the Waal River (Appendix A). The revisited model accepts two sequential bathymetry measurements as input and provides a single prediction for the next time step (Figure 10). The machine learning model is based on a video prediction model, meaning that time steps between frames are assumed constant (Srivastava et al., 2015). Model time steps are denoted as t-1 (input), t (input) and t+1 (prediction or target). This section will give an understanding of the fundamental principles in this machine learning algorithm with a focus on data preprocessing and loss functions.

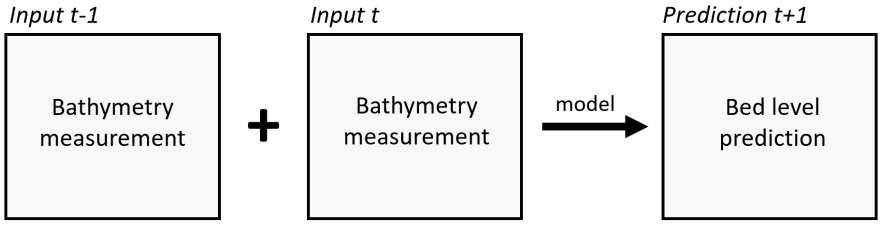


Figure 10: Schematic overview of the machine learning model input and output.

2.3.1 Steps in machine learning

Within machine learning, there are generally five steps followed to obtain a good performing model (Figure 11). Firstly, all relevant data is obtained. Then, the data is presented in such a way suitable for the model to accept. Consequently, the model is trained on a set of training samples, where the model tries to improve its predictions. A trained model follows from this, which is then tested to assess its performance. Finally, the model can be adapted to improve predictions further. Data preprocessing is applied in the preparation of the data and loss functions are applied in the training of the model.

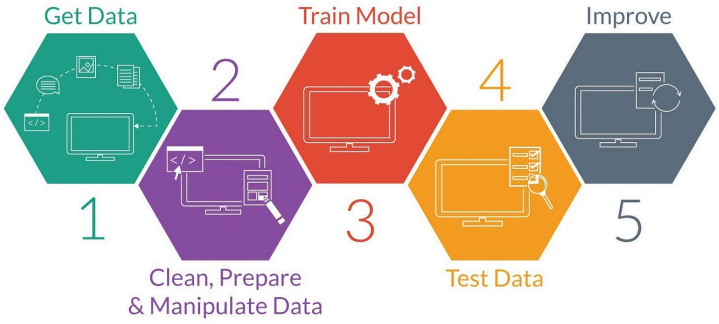


Figure 11: Five core machine learning steps (Maloo, 2018)

2.3.2 Data preprocessing

Data preprocessing is done in the preparation of the model data. The data is transformed before feeding it to the model to train. It is done to improve the data quality and model performance (Maharana et al., 2022). Steps in the data preprocessing generally include the removal of outliers in the data and a selection of the model domain (Dekker et al., 2022; Shi et al., 2017; van der Kooij et al., 2021).

The data is provided to the model in the form of samples. Each sample contains three sequential bathymetry frames. The first two frames are the input of the model and the last frame is used for evaluation against the model prediction (Figure 12).

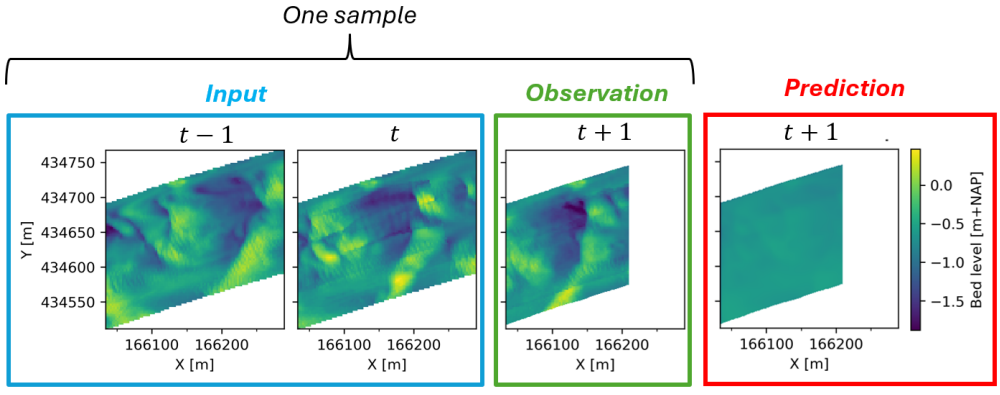


Figure 12: The model input together with observed and predicted bed levels. One sample includes the input and observation.

The data is split into training, validation and test samples.

- **Training samples**: These are the samples used for the training of the model. This sample set contains around 80% of all data (e.g., Shi et al., 2017)
- **Validation samples:** These are the samples which are used to validate the model against overfitting during training. This sample set contains around 5% of all data.
- **Test samples:** The test samples are used to test the model. This set is completely independent of the training of the model. This sample set generally contains around 15% of all data.

In this study, only samples containing time intervals between frames within the range of 10 to 18 days were included for the model training, validating and testing. This selection was done to keep some consistencies in time intervals, as the machine learning model is video-based and therefore assumes constant time steps between frames. This selection resulted in a total of 275 samples, which were split into 223 training samples, 13 validation samples, and 39 testing samples.

The model accepts bathymetry raster frames of size 256x256, containing cells values between 0 and 1 (Dekker et al., 2022). Therefore, bed levels are scaled between values of 0 to 1, based on the maximum and minimum bed levels in the dataset, before entering the model.

2.3.3 Model training and loss functions

To understand the purpose of a loss function, a deeper understanding of the model training is required. A flowchart of the training scheme for the applied machine learning approach is shown in Figure 13. This scheme uses a classification in the data as presented in Figure 12. First, the model starts with a random set of model weights (Figure 13), which are the internal connections in the neural network. Training samples are provided to the machine learning model in batches (two samples at a time). The model provides a prediction for each sample in the batch, which is evaluated via a loss function. The loss function quantifies a loss between the prediction and observation. The larger the loss, the worse the model performs. During training, this loss is minimised by adjusting the model weights after each batch. After all training samples have passed through, the final set of model weights result, which define the newly trained model. These model weights can be loaded into the model structure, and the model is ready to provide a prediction. The model training uses a predefined set of hyperparameter values, which guide the training process and remain constant during training. More information on the hyperparameters and model structure can be found in Appendix A.

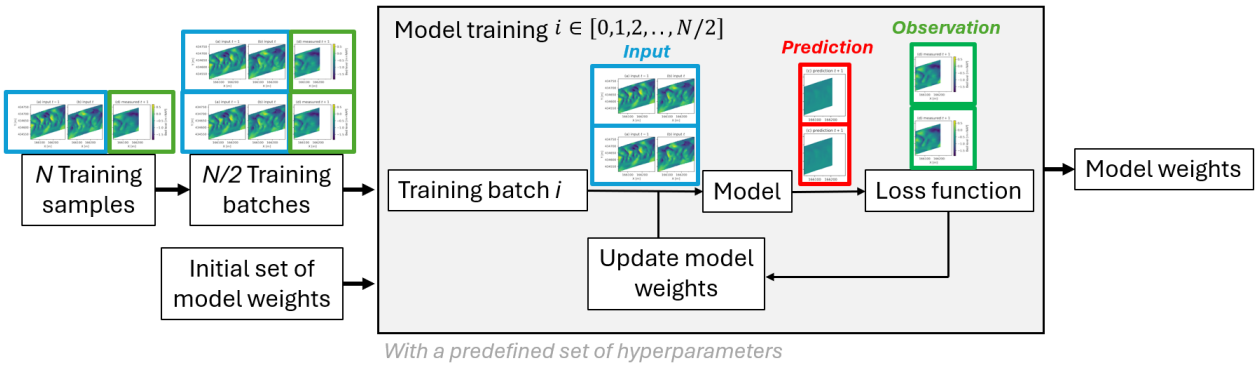


Figure 13: Training scheme for the applied machine learning approach. N indicates the total number of training samples.

The flowchart indicated in Figure 13 shows one entire passing of training data through the machine learning algorithm, also referred to as one epoch. By increasing the number of epochs, the model can iterate its weights more often to achieve better predictions. However, increasing the number of epochs increases the risk of overfitting. Overfitting in machine learning occurs when a model learns the training data too well, capturing noise or random fluctuations instead of general patterns, leading to poor performance on unseen data. To prevent overfitting, a model validation step is included in the training scheme (Figure 14). The model will continue training for another epoch until the model validation loss has not decreased for four consecutive epochs, at which point it outputs the final model weights.

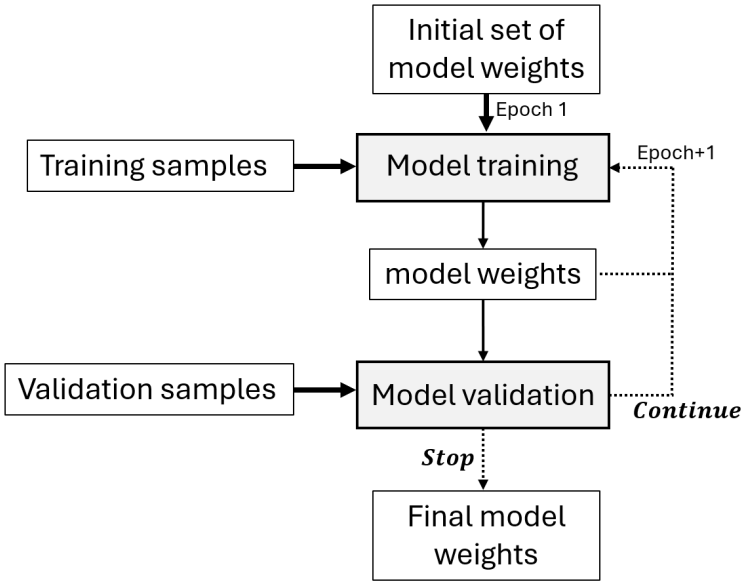


Figure 14: Model training including the model validation step

2.3.4 The applied loss function

The machine learning model applies the Mean Square Error (MSE) loss function, which is one of the most popular loss functions applied in image forecasting applications (Tran & Song, 2019). The MSE loss function can make a decent estimation of the global similarity between two images, but not about the local structure (Jonnalagadda & Hashemi, 2023; Tran & Song, 2019). A small misplacement in features may result in a higher loss than for a highly blurred prediction. Therefore, many studies argue that the blurriness in these types of machine learning models arises due to the use of the mean square error (Jing et al., 2019; Ma et al., 2022; Tian et al., 2019; Tran & Song, 2019).

The MSE loss function is depicted in equation 3. Here, $P_{i,j}$ denotes the predicted bed level and $O_{i,j}$ the observed bed level for i^{th} row and j^{th} column. k is a masking parameter, which excludes the most upstream part of the model domain for evaluation (Figure 15a). As the model domain is stationary and dune features move downstream, the model cannot predict what features enter from the upstream model boundary. This is different compared to other forecasting applications, such as precipitation forecasting, where features can enter from all model boundaries (Figure 15b).

$$\mathcal{L}_{\text{MSE}}(P, O) = \frac{1}{256 \cdot (256 - k)} \sum_{i=1}^{256} \sum_{j=1}^{(256-k)} (P_{i,j} - O_{i,j})^2 \tag{3}$$

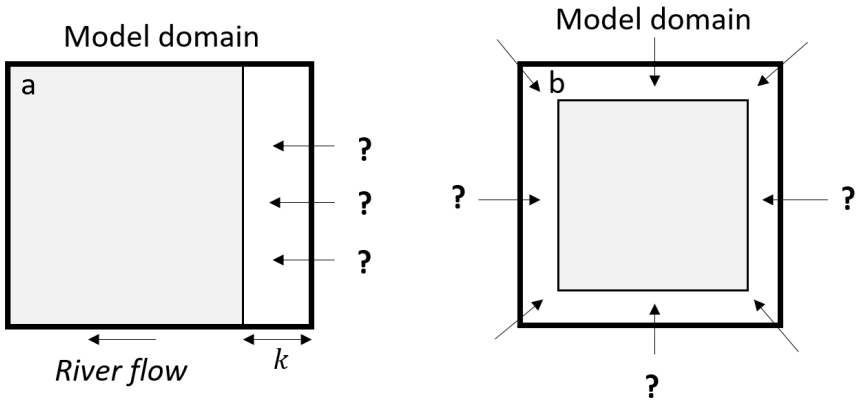


Figure 15: Schematisation of the applied mask for (a) the bed level prediction and (b) precipitation forecasting applications. The white box indicates the excluded section for evaluation of the model performance, with k being the masking parameter.

2.4 Data preprocessing experiments

Five data preprocessing experiments were set up to improve the bed level prediction accuracy for the machine learning model. The five data preprocessing experiments include: sample selection, removal of outliers, coordinate transformation, wavelet reconstruction, and consistent time intervals. Each experiment built upon the modifications introduced in the preceding one, as this was thought to achieve a cumulative improvement of the model performance. The data preprocessing experiment showing the best model performance was further used in the loss function experiments. Below, each of the data preprocessing experiments is further discussed.

2.4.1 Sample selection

In this experiment, samples with high flow conditions were excluded from the training sample set. Modelling of these periods is considered irrelevant, as navigation mainly experiences problems during low flow periods (Lokin et al., 2023). The inclusion of high flow periods could complicate the learning process, as river dune dynamics are highly sensitive to flow conditions. Also, during high flow periods, river dunes encompass their lengths between measurements (Figure 16a), further complicating the recognition of dune development for the machine learning model. Therefore, samples are removed from the training set for which the 14-day dune celerity exceeds the dune length. Figure 16b shows the scaled dune celerity over the discharge. If this factor equals 1 or larger, the 14-day dune celerity exceeds the dune length. There is quite some deviation in the data (R^2=0.62), but the scaled dune celerity remains generally below 1 for discharges below $2000 m^3/s$. Discharges of $2000 m^3/s$ or higher are therefore considered as a high flow condition for the Waal River and are removed from the training samples. In total, 162 training samples remain for training the machine learning model.

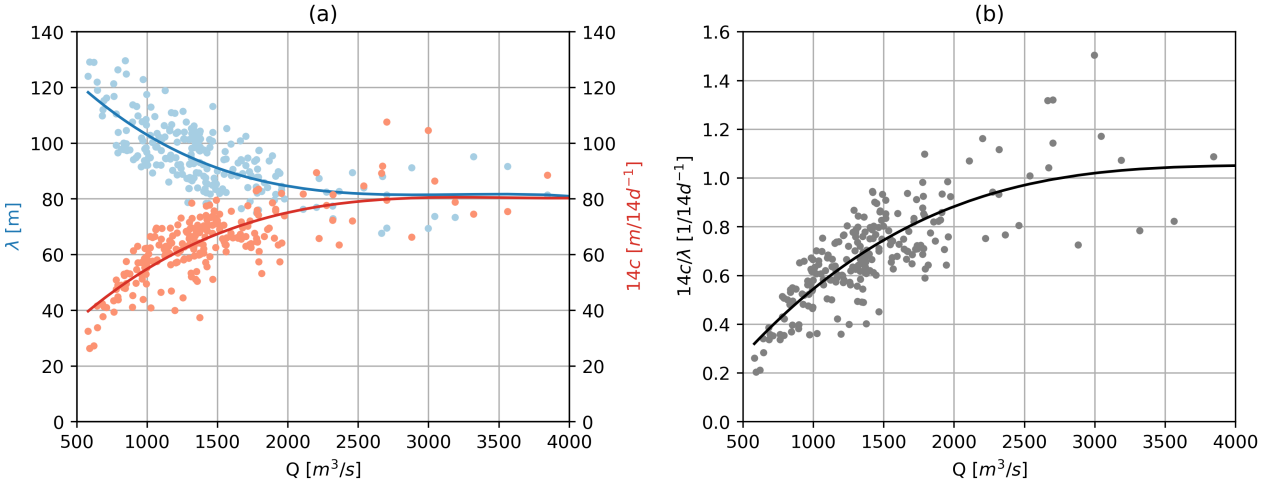


Figure 16: (a) The dune length, λ, and 14-day dune celerity, $14c$, as a function of the discharge for the Waal River, (b) Scaled dune celerity against the discharge for the Waal River. The plots are based on data by Lokin et al. (2022)

2.4.2 Removal of outliers

Far outliers sometimes occur in MultiBeam EchoSounder (MBES) datasets, even though they are more than often filtered out as part of the acquisition process (Le Deunf et al., 2020). These outliers misinform the machine learning model in the training process, potentially leading to less accurate predictions. Therefore, in this second experiment, outliers were identified and removed from the MBES dataset. This experiment built upon the data from the sample selection experiment.

Outliers in the MBES data were identified by applying a median-based filter. The median bathymetry was extracted over all measurements and the difference to the measured bed levels was plotted (Figure 17). Bed level deviations from the median generally range up to 2m, which can be attributed to the constantly evolving river dunes and ripples. Deviations larger than 2m are mostly observed in low bed levels, which could be due to the presence of scour holes. As this study focuses on predicting the shallow areas more correctly, the threshold deviation is set to 3m, which excludes deep scour holes and actual outliers. Also, the spatial surrounding 3 meters of the outliers are removed, limiting the sloping effects. The outliers are replaced by interpolation via inverse distance weighting, which assigns values to the outlier locations based on the weighted average of surrounding cells, see also Lu and Wong (2008) and Maleika (2020).

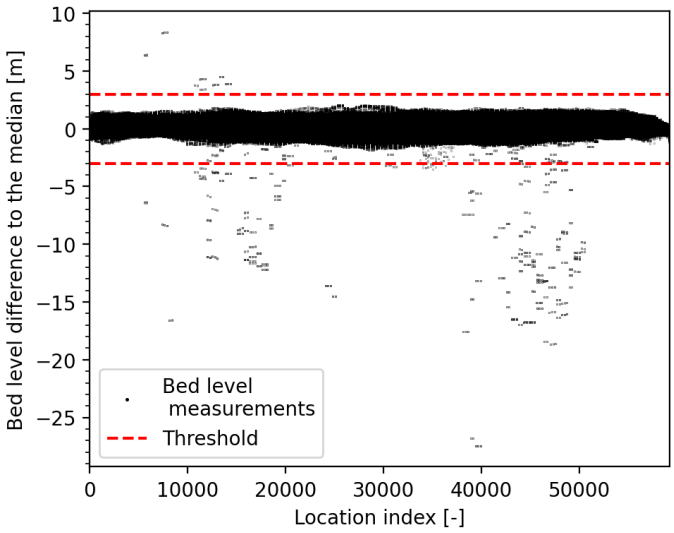


Figure 17: The differences in bed level measurements to the median bed for the study area, together with the upper and lower threshold for outlier detection.

The bathymetry measurements also contain missing data within the navigation channel. In the original data preprocessing, these missing values were replaced by zeros during normalisation of the data, which is equal to the minimum observed bed level. Thus, biases towards low bed levels were introduced, leading to inaccuracies in the representation of the navigation channel bed. To mitigate this effect, these missing values were also replaced using inverse distance weighting. For a visualisation of the locations of the missing data and the identified outliers within the navigation channel, please refer to Appendix B.

2.4.3 Coordinate transformation

The raw input data is given as an orthogonal grid in RD new coordinates (X,Y). To simplify the learning of bedform developments, the grid is transformed to a channel-fitted coordinate system (s, n) proposed by Dietrich and Dungan Smith (1984). Here, s is the channel centerline, which can be translated to the river kilometre (rkm), and an 'n' axis perpendicular to that centerline (Figure 18). This coordinate transformation aligns the main direction of flow, and thus dune propagation, along a single horizontal axis. This simplifies the representation of dune propagation in contrast to the RD coordinate system, as the dunes now propagate horizontally over the frame instead of diagonally. Due to the stream coordinate transformation, the study area was slightly adjusted to keep consistency in the frame size of 256x256m (Figure 18). Again, this experiment built upon the data provided by previous experiments: the sample selection and the removal of outliers.

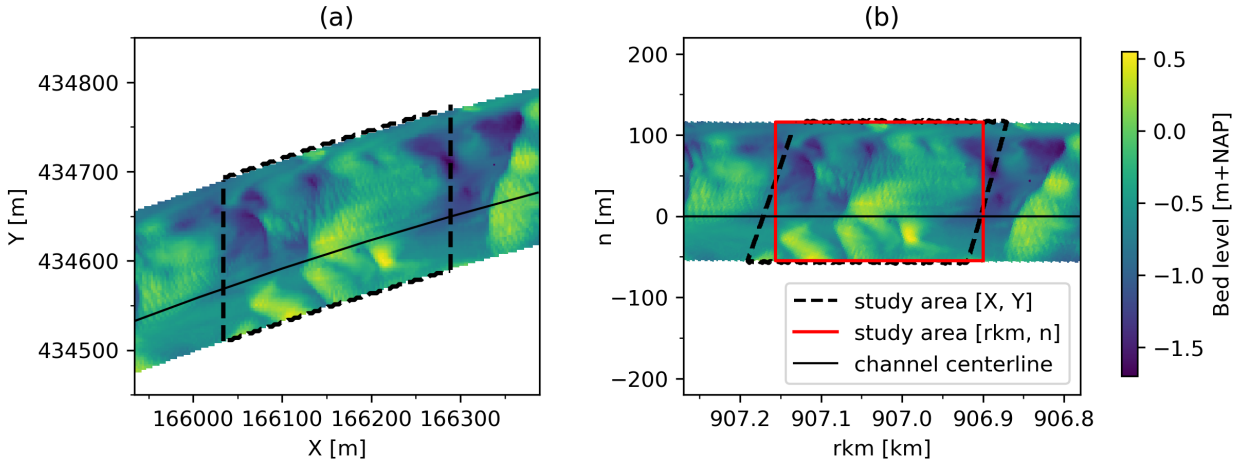


Figure 18: Stream coordinate transformation. (a) Raw bed level data in RD new coordinates [X, Y], with the study area indicated in the black dotted line. (b) Bed level data transformed to the channel-fitted coordinate system [rkm,n], where n represents the perpendicular distance from the channel centerline and rkm denotes the river kilometre. The bed level measurement of 2005-07-19 is plotted here as example.

2.4.4 Wavelet reconstruction

In this experiment, secondary bedforms are removed from the bed level data, as their bedform dynamics can hardly be captured within measurements taken at 14-day intervals. By removing these features from the data, the training data is simplified and the model can shift its focus more towards relevant bed feature dynamics such as primary dune development.

The removal of secondary bedforms is achieved by applying a 2D wavelet transform (van Denderen et al., 2022). Kruis et al. (2023) demonstrated the capability of such analysis to accurately reconstruct the river bathymetry, whilst removing secondary bedforms. There exist other dune analysis methods capable of removing secondary bedforms as well (Gutierrez et al., 2018; van der Mark et al., 2008). However, these require additional data preprocessing steps (removal of large-scale bedforms) to analyse and reconstruct dune features.

Secondary bedforms in the Waal River typically have wavelengths up to 10m (Zomer et al., 2021). Therefore, the riverbed is reconstructed with wavelengths between 10 to 500m. An example reconstruction of the riverbed is depicted in Figure 19b. Figure 19d shows that smaller bedform deviations are filtered, whilst capturing the larger bedform shapes well. However, as shown in Figure 19c, the wavelet reconstruction does result in an overestimation of higher bed levels and underestimates lower bed levels by up to three decimeters.

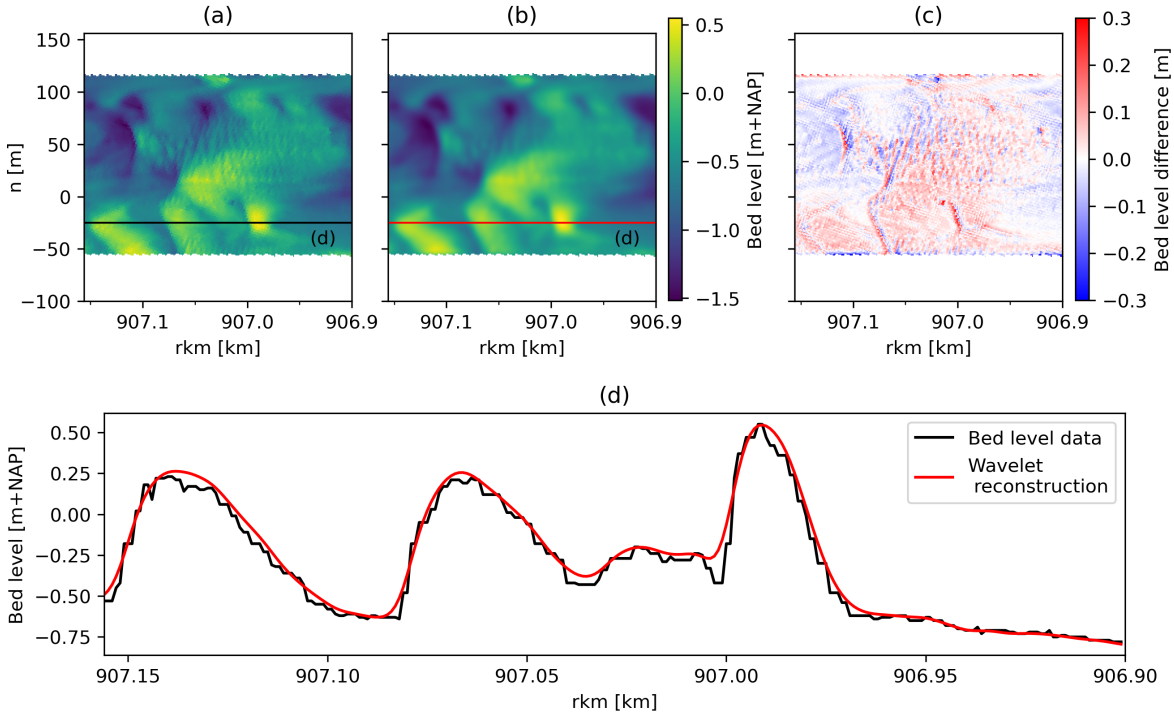


Figure 19: The bed level data and wavelet reconstruction of the study area riverbed, using wavelengths of 10-500m. (a) The bed level data in channel-fitted coordinates. (b) Reconstructed bed levels with the wavelet transform (c) Absolute differences in bed level between the bed level data and Wavelet reconstruction. A negative value indicates a lower bed level by the reconstruction compared to the data and a positive value indicates a higher bed level. (d) 1D bed level profile along the river axis, showing the bed level data and wavelet reconstruction. The bed level measurement of 2005-07-19 is plotted here as example.

2.4.5 Consistent time intervals

The machine learning model TrajGRU assumes constant time intervals between consecutive frames, since it is based on a video prediction model (Srivastava et al., 2015). As the time intervals in the training samples deviate between 10 to 18 days (Section 2.3.2), it is proposed to adjust the dataset such that time intervals become consistently 14 days. Time interval inconsistencies are removed via the raster shifting approach as described in Section 2.2. If the time interval between consecutive measurements deviates from the 14 day median, the bathymetry raster is shifted along the channel axis, such that the time intervals become consistently 14 days. The input frames are shifted with respect to the prediction frame, preventing data augmentation of the evaluation frames. The shifting distance is determined with equation 4.

$$L_{sh} = \sum_{i=1}^{\Delta t} c_i \tag{4}$$

The deviation in time interval from the 14 day median (Δt) is determined by subtracting the time interval between measurements (t_{int}) with 14 days ($\Delta t = |t_{int} - 14|$). For time intervals smaller than 14 days, the raster is shifted downstream and for intervals larger than 14 days, the raster is shifted upstream.

2.5 Loss function experiments

Many studies argue that blurriness in image machine learning predictions arises due to the use of the Mean Square Error (MSE) (Jing et al., 2019; Ma et al., 2022; Tian et al., 2019; Tran & Song, 2019). Dekker et al. (2022) has shown that the adaptation of loss functions in TrajGRU can overcome the blurring effect to some extent. Therefore, three adaptations in the loss functions were tested for: the Root Mean Square Error (RMSE), a combination of the Mean Square Error (MSE) and Structural Similarity index (SSIM), and the Wasserstein loss. Each of these loss functions is further discussed below. For the loss evaluation, 80m of the most upstream part of the model domain is excluded (Section 2.3.4, $k = 80m$), as this is near to the maximum distance river dunes migrate over 14 days (Figure 16).

2.5.1 Root mean square error

The Root Mean Square Error (RMSE) is equal to the square root of the MSE. It was observed during the training of TrajGRU that the MSE losses within the model were in the order of 10^{-3}. A change in the loss was hardly recognisable at this scale. By taking the RMSE as loss function, differences between losses become more accentuated. The study by van der Kooij et al. (2021) tested the RMSE loss function in the TrajGRU model for precipitation nowcasting and showed substantial improvement in the mean absolute error and RMSE metrics. However, they found that peak rainfall intensities were further underestimated compared to the benchmark model. Despite this, the RMSE loss function is tested due to its ability to accentuate differences in the loss. The RMSE loss function is denoted in equation 5, where $P_{i,j}$ denotes the predicted bed level and $O_{i,j}$ the observed bed level for i^{th} row and j^{th} column. k is the masking parameter as indicated in Section 2.3.4.

$$\mathcal{L}_{\text{RMSE}}(P, O) = \sqrt{\frac{1}{256 \cdot (256 - k)} \sum_{i=1}^{256} \sum_{j=1}^{256-k} (P_{i,j} - O_{i,j})^2} \tag{5}$$

2.5.2 Mean square error & structural similarity index measure

In this experiment, a combination of the Mean Square Error (MSE) and Structural Similarity Index Measure (SSIM) is tested as loss function. The Structural Similarity Index Measure (SSIM) is a metric that measures the differences in local structure, luminosity and contrast between two images (Dekker et al., 2022; Jonnalagadda & Hashemi, 2023). Tran and Song

(2019) and Dekker et al. (2022) compared several loss functions and found that the implementation of the SSIM improved the model performance with respect to applying MSE only, as the blurring effect in the model predictions was reduced.

The equation for SSIM is depicted in equation 6, where P represents the predicted bed levels and O the observed bed levels. The luminance of the image is depicted by l(P,O), the contrast by c(P,O) and the structural similarity by s(P,O). The value of SSIM ranges between 1 to -1, with 1 indicating perfect similarity, 0 no similarity and -1 perfect anti-correlation. C_1, C_2 and C_3 ($C_3 = C_2/2$) are small positive constants for numerical stability. These are set to $C_1 = 0.01$ and $C_2 = 0.03$, as proposed by Z. Wang et al. (2003) for normalised pixel values.

$$\begin{aligned}\mathrm{SSIM}(P_{i,j}, O_{i,j}) =& l(P,O) \cdot c(P,O) \cdot s(P,O) = \left(\frac{2\mu_P\mu_O + C_1}{\mu_P^2 + \mu_O^2 + C_1}\right) \cdot \left(\frac{2\sigma_P\sigma_O + C_2}{\sigma_P^2 + \sigma_O^2 + C_2}\right) \cdot \left(\frac{\sigma_{PO} + C_3}{\sigma_P\sigma_O + C_3}\right) = \\ & \frac{(2\mu_P\mu_O + C_1)\,(2\sigma_{PO} + C_2)}{(\mu_P^2 + \mu_O^2 + C_1)\,(\sigma_P^2 + \sigma_O^2 + C_2)}\end{aligned} \tag{6}$$

The SSIM is computed for each cell in the frame by considering its surrounding neighbourhood, where cells more nearby are accounted more heavily towards the SSIM score. For this, a Gaussian filter is applied with a window size of 21 and a standard deviation of 5. The window size and standard deviations are set higher than generally observed in the literature (Dekker et al., 2022; Z. Wang et al., 2003), as the grid is highly detailed (1x1m) whilst the interest lies into the larger scale river dune slopes. Finally, the SSIM loss can be computed by averaging the SSIM scores over all cells (equation 7). A loss of 0 indicates perfect similarity, 1 no similarity and 2 perfect anti-correlation. The SSIM loss is combined with the MSE loss in equation 8.

$$\mathcal{L}_{\mathrm{SSIM}}(P,O) = 1 - \frac{1}{256 \cdot (256-k)} \sum_{i=1}^{256} \sum_{j=1}^{(256-k)} \mathrm{SSIM}(P_{i,j}, O_{i,j}) \tag{7}$$

$$\mathcal{L}_{\mathrm{MSE+SSIM}}(P,O) = \mathcal{L}_{\mathrm{MSE}}(P,O) + \beta\mathcal{L}_{\mathrm{SSIM}}(P,O) \tag{8}$$

β is a factor in equation 8 indicating the importance of the SSIM loss. Different values for β are suggested in the literature, ranging from 0.02 to 0.5 (Dekker et al., 2022; Le et al., 2022; Tran & Song, 2019; Zhang et al., 2023). From experimental analysis, it followed that a scaling factor of $\beta = 0.5$ resulted in best model performance. Further increasing of the scaling factor did not improve results.

2.5.3 Wasserstein loss

Another method to capture the deviations in bed levels more accurately is via the use of the Wasserstein loss. The idea of this loss function is to treat each cell value as a mass located at a specific point, calculating the cost required to move these masses to their correct positions to match the observation. In the context of this study, the bed levels represent the masses in the system. If a patch of higher bed levels would be predicted on locations with low observed bed levels, the Wasserstein loss would calculate the lowest transport costs to allocate these

masses to the correct location (Figure 20). The Wasserstein loss is implemented as depicted in equation 9. Herein, $\gamma_{i,j}$ denote the transport plans, which describe how much mass should be moved from position x_i to y_j. The distance between the two positions is computed with $\|x_i - y_j\|$. All possible transport plans ($\gamma \in \Pi$) are tried and the transport plan with the lowest transport costs gives the Wasserstein loss.

$$\mathcal{L}_{wass}(P, O) = \min_{\gamma \in \Pi} \frac{1}{256 \cdot 256} \sum_{i=1}^{256} \sum_{j=1}^{256} \gamma_{i,j} \|x_i - y_j\| \tag{9}$$

An advantage of this loss function is that it is less sensitive to displacement errors compared to the MSE (Dekker et al., 2022). If a prediction is slightly shifted, the Wasserstein loss would penalize the model less than the MSE loss function, as the transportation costs would be small whilst the point-wise error would be large. Also, blurred prediction would be higher penalised by the Wasserstein loss during training, as relatively large masses should be converted to the right locations, giving a high Wasserstein loss.

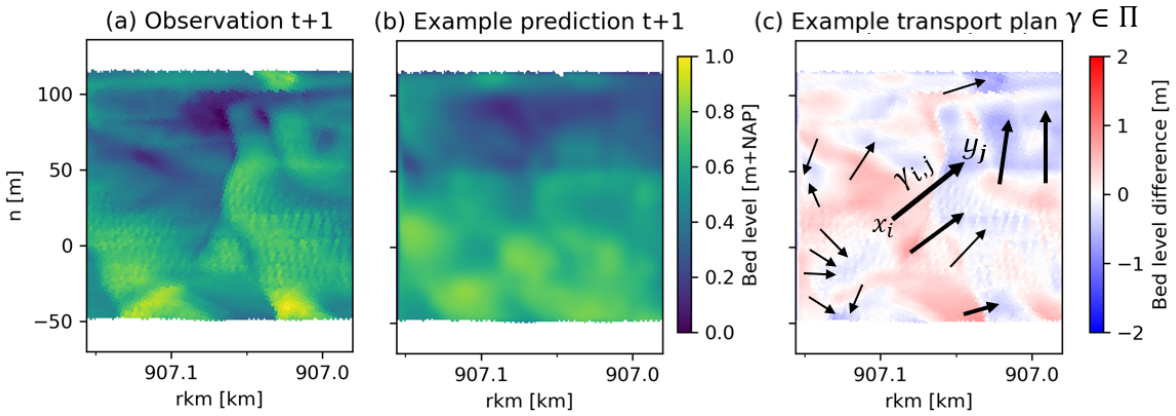


Figure 20: Visualisation of an example transport plan of the Wasserstein loss. (a) shows observed bed levels, (b) shows a random example prediction and (c) gives an example transport plan, see also equation 9. The arrows indicate where masses are relocated. The thicker the arrow, the higher the transport cost. Also, the bed level differences are indicated between the prediction and observation, where a negative value indicates a lower bed level by the prediction compared to the observation and a positive value indicates a higher bed level.

2.6 Performance evaluation

In this study, the evaluation of model performance focuses on two key points: the maximum bed level and the overall bed level patterns. For the dredging companies and Rijkswaterstaat, it is highly relevant to obtain an accurate prediction of the maximum bed level and its location, as this indicates where future dredging might be necessary. From a scientific perspective, the ability of the model to effectively capture bedform patterns is also relevant, as current studies have been struggling to develop predictive models capable of accurately predicting bedform patterns (e.g. Paarlberg et al., 2009). For the performance evaluation, 80m of the most upstream part of the model domain is excluded, similar to the loss functions.

The maximum bed level error serves as an indicator of how well the model predicts the highest points of bed elevation. The maximum bed level error was determined by extracting the maximum bed levels over 32-meter-long stretches of the river and computing the Mean Absolute Error between the observed and predicted maximum bed levels (MAE_b). Also, the error in the locations of the maximum bed levels was determined. This was done via visual inspection and quantification of the locational error. Two 80x80m regions were manually identified in the observation for a total of 10 testing samples (Appendix C), capturing the highest river dunes. Figure 21 depicts an example observation with corresponding 80x80m regions. For each region, the location of the maximum bed level is determined for both the prediction and observation and the distance between the predicted and observed locations gives the locational error.

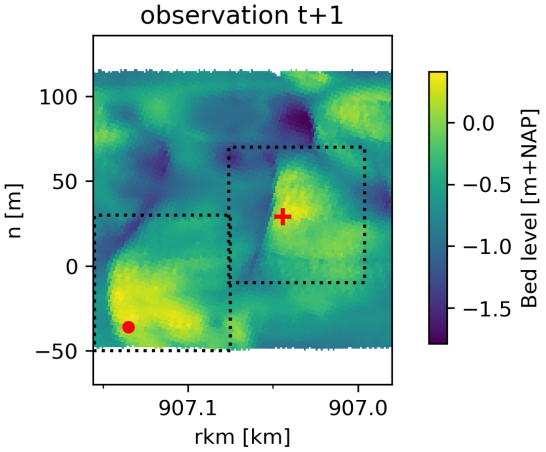


Figure 21: Example observation with two 80x80m regions, indicated by the black dotted lines, capturing the highest river dunes. The red dot and plus indicate the location of the maximum bed level within these regions.

The error in the overall bed level patterns was assessed via visual inspection and two supporting point-wise performance metrics: the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE). Dekker et al. (2022) and van der Kooij et al. (2021) both showed the importance of visual inspection in TrajGRU forecasting applications and indicated that an improvement in point-wise performance metrics does not necessarily lead to visual improvement as well. Bosboom (2019) mentioned that point-wise metrics for morphodynamic model evaluation generally penalize, rather than reward, the features of interest (dune peaks/troughs). The prediction of a morphological feature that is correct in terms of timing and size, but is misplaced in space, may not outperform even a flat bed, which is inconsistent with the common judgement of morphologists. However, applying only visual inspection could include a subjective bias. Therefore, combining visual inspection alongside the MAE and RMSE metrics ensures a comprehensive evaluation of the overall bed level patterns.

3 Results

The results of the dune migration model and the machine learning model experiments are presented in this chapter. First, the dune migration model results are presented. This is followed by the data preprocessing and loss function experiment results for the machine learning model. After this, a comparison is done between the best-performing machine learning model and the dune migration model for predicting bed levels in 3D for the Waal River.

3.1 Dune migration model

The dune migration model predicts the horizontal displacement of the river bed, applying the raster shifting method described in Section 2.2. Figure 22 shows an example prediction of this model. The model visually provides a good prediction of the bed levels. The locations of maximum bed levels are effectively captured as well as the locations of the overall higher and lower bed levels. Furthermore, bedform shapes remain similar in the prediction compared to the previous time step. In Appendix C, extra example predictions are provided, also showing the potential of the dune migration model. The maximum bed level error (MAE_b) and the overall bed level errors (MAE & $RMSE$) are relatively similar, around 30cm (Table 1).

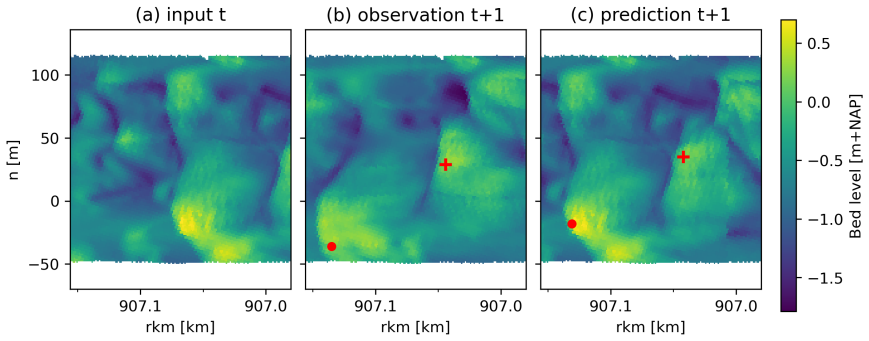


Figure 22: Example prediction by the dune migration model, note that the 80m upstream of the input frame is excluded for visual purposes. The red plus and red dot indicate the location of two maximum bed levels.

	Dune migration model
MAE_b [m]	0.30
MAE [m]	0.26
RMSE [m]	0.34

Table 1: Performance metric scores for the dune migration model on the complete test sample set.

3.2 Data preprocessing experiments

Five data preprocessing experiments were set up and tested against the machine learning model trained on the original data. Each successive experiment was built upon the modifications introduced in the preceding one, as this was thought to achieve a cumulative improvement of the model performance. An example prediction of each experiment is provided in Figure 23. The resulting performance metric scores are summarised in Table 2. The main observation is that all model experiments provide a highly blurred output (Figure 23). The blurring effect is not reduced between experiments. The maximum bed level prediction therefore remains largely underestimated with the lowest MAE_b score of 0.92m for the original data. Also, the location of the maximum bed levels is little accurately predicted.

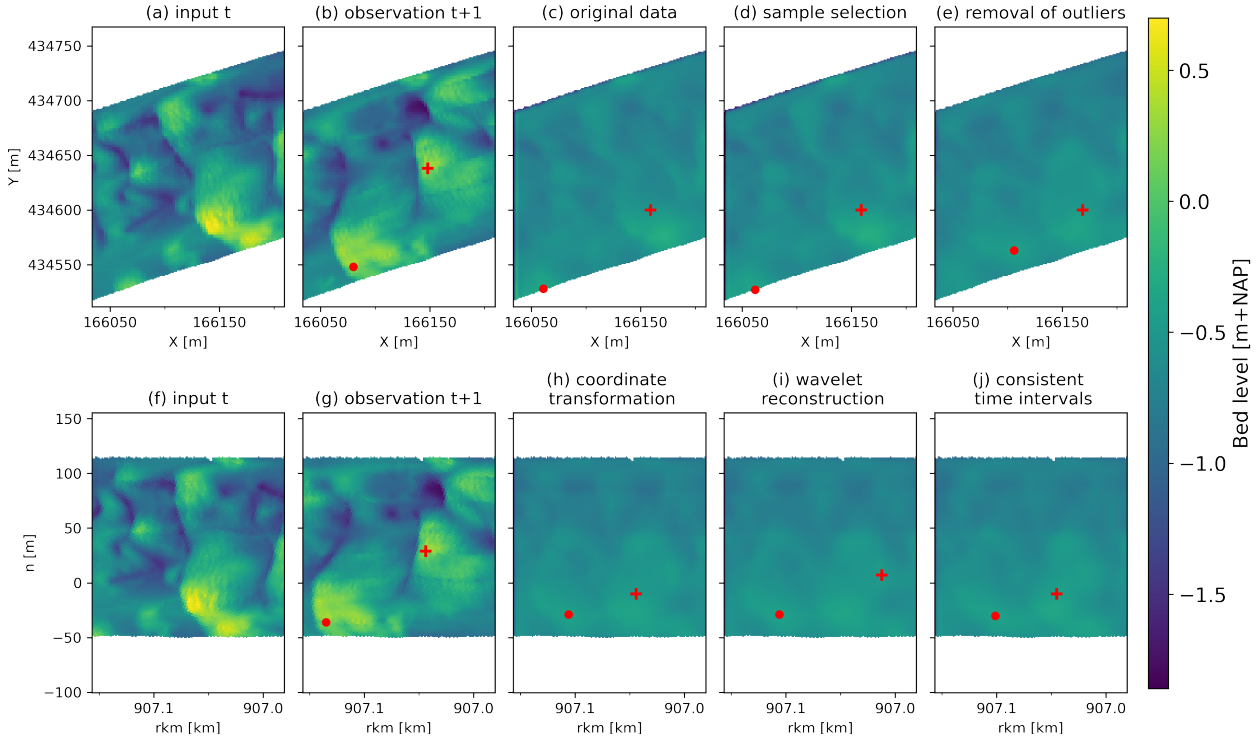


Figure 23: Example bed level prediction for the data preprocessing experiments. (a) and (f) represent the real-time bathymetry measurement, used as the second input frame for TrajGRU. 80m upstream of the input frame is excluded for visual inspection purposes. (b) and (g) represent the observed bathymetry for the next time step. (c) shows the model results of the model trained on the original data and (d-e) and (h-j) represent the predictions by the five data preprocessing experiments. The red plus and red dot indicate the location of two maximum bed levels.

Table 2: Performance metric scores for the data preprocessing experiments and the original dataset, tested on the test sample set. Lowest metric scores are indicated in bold.

	Original data	Sample selection	Removal of outliers	Coordinate transformation	Wavelet reconstruction	Consistent time intervals
MAE_b [m]	**0.92**	0.96	1.03	1.01	0.95	0.95
MAE [m]	0.30	0.31	0.27	**0.26**	0.28	0.28
RMSE [m]	0.41	0.44	**0.34**	**0.34**	0.35	0.35

All model experiments appear to largely reconstruct the input frame. However, the removal of outliers (Figure 23e) resulted in more deviation of the input frame (Figure 23e) compared to previous experiments (Figure 23c-d), visually resulting in a more accurate prediction of the river bed patterns. This is also reflected in the MAE and RMSE metric scores, which are reduced compared to the original data (Table 2). This is likely caused by the narrower bed level range in the data normalisation for the removal of outliers. With outliers removed, bed levels are scaled to values between 0 and 1 based on a range of -3.64 to 3.64m+NAP, as opposed to the wider range in the original data of -10.00 to 3.64m+NAP. Differences between the predicted and observed bed levels are therefore higher penalised in the loss function. Noticeably, the maximum bed level error did increase from 0.92m to 1.03m with the removal of outliers, whilst visually the location of the maximum bed level got closer to the observation with the removal of outliers (Figure 23b-c and 23e).

The coordinate transformation, wavelet reconstruction and consistent time interval experiments (Figure 23h-j) show negligible visual changes compared to the removal of outliers. The MAE and RMSE metric scores support this observation, as these remain relatively similar between these experiments. However, the maximum bed level error does show quite a reduction from 1.03m to 0.95m for the wavelet reconstruction location. The predicted location of the maximum bed level (Figure 23h-e) neither improved nor deteriorated.

Overall, the model trained on the wavelet reconstructed data is considered to perform the best of all experiments. The model scores single best on the maximum bed level error, whilst predicting the riverbed patterns relatively well. The wavelet reconstructed data was therefore further used for the evaluation of the loss functions. The wavelet reconstructed data includes the successive modifications by the sample selection, removal of outliers, coordinate transformation and the wavelet reconstruction.

3.3 Loss function experiments

The original machine learning model by HKV applied the Mean Square Error loss function ($\mathcal{L}_{\text{MSE}}$) for training. The three tested loss functions are the Root Mean Square Error ($\mathcal{L}_{\text{RMSE}}$), a combination of the MSE and Structural Similarity Index Measure ($\mathcal{L}_{\text{MSE+SSIM}}$) and the Wasserstein loss ($\mathcal{L}_{\text{wass}}$). The machine learning model was trained on each loss function separately. An example prediction of each trained model is provided in Figure 24. The performance metric scores are summarised in Table 3.

The spread in the bed levels increased considerably with the implementation of the RMSE loss and Wasserstein loss (Figure 24d and 24h). As the spread in bed levels increased, the maximum bed level error reduced from 0.95m to 0.77 and 0.78m (Table 3). Between the two loss functions, the RMSE loss resulted in a more accurate prediction of the bedform patterns. The implementation of the RMSE loss showed a similar bedform pattern prediction as the MSE loss, but with a larger spreading in the bed levels and a more accurate prediction on the locations of the maximum bed levels (Figure 24d). The MAE and RMSE scores between the MSE and RMSE loss remain similar. The Wasserstein loss resulted in a generic reconstruction of the second input frame, giving the worst RMSE and MAE metric scores among all loss function experiments. Also, the location of the maximum bed level was visually captured worst of all losses (Figure 24h).

The implementation of the SSIM loss visually shows little changes to the MSE loss. Contrasts in the prediction are slightly more accentuated with the SSIM loss, and the maximum bed level error has reduced from 0.95 to 0.90m.

Concluding, the implementation of the RMSE loss function resulted in the best model performance. The bed level spread increased considerably, resulting in the lowest maximum bed level error (0.77m) of all loss function experiments. Furthermore, the location of the maximum bed levels was visually captured best of all experiments, whilst bedform patterns were also predicted relatively well. Therefore, the model including the wavelet reconstructed data (Section 3.2) and the RMSE loss was used for comparison with the dune migration model.

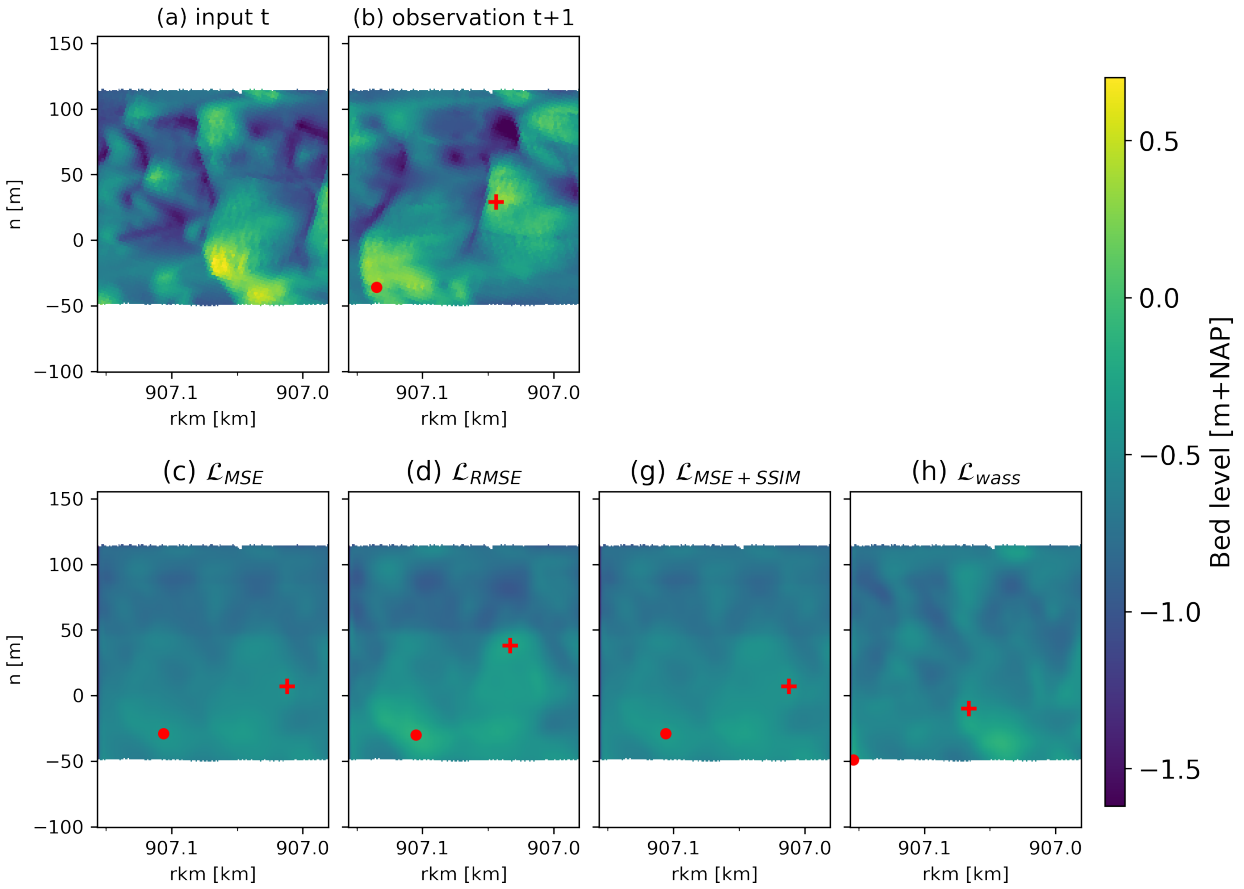


Figure 24: Example bed level prediction for the TrajGRU wavelet model trained on four loss functions: (c) MSE, (d) RMSE, (e) MSE+SSIM and (h) the Wasserstein loss. (a) represents the real-time bathymetry measurement, used as the second input frame for TrajGRU. Note that 80m upstream of the input frame is excluded for visual inspection purposes. (b) represents the observed bathymetry for the next time step. The red plus and red dot indicate the location of two maximum bed levels.

Table 3: Performance metric scores of the TrajGRU machine learning model for the four loss functions on the test sample set. The lowest metric scores are indicated in bold.

	$\mathcal{L}_{MSE}$	$\mathcal{L}_{RMSE}$	$\mathcal{L}_{MSE+SSIM}$	$\mathcal{L}_{wass}$
MAE_b [m]	0.95	**0.77**	0.90	0.78
MAE [m]	**0.28**	**0.28**	0.29	0.32
RMSE [m]	**0.35**	**0.35**	0.36	0.39

3.4 Comparison dune migration model and machine learning

Here, the best performing machine learning model, using the wavelet reconstructed data and the RMSE loss, is compared with the dune migration model for predicting bed levels in the Waal River. As the machine learning model still provided highly blurred predictions, a blown-up model prediction was added for comparison. This blown-up prediction is similar to the machine learning prediction, but with the bed level standard deviation manually increased in the prediction to match the standard deviation in the bed level measurement (input t).

Figure 25a and 25b illustrate an example prediction of both the machine learning model prediction and the blown-up model prediction. The blown-up model prediction visually shows a substantial improvement in the bed level prediction. Contrasts in the bed levels are more accentuated and the maximum bed level error reduces from 0.77m to 0.47m (Table 4). Based on Figures 25a-d, the locations of the maximum bed levels in the observations are relatively well captured in the machine learning prediction, with a locational error of 25m.

Comparing the machine learning dune migration model predictions, the dune migration model outperforms the machine learning model both visually and on the metric scores. Visually, the dune migration model predicts the locations of the higher and lower bed levels more accurately than the machine learning model (Figure 25c-d). Figure 25g does indicate relatively large differences in predicted and observed bed levels, but these are more on a local scale compared to the machine learning predictions (Figure 25e-f). The larger patches of over-prediction in bed levels, observable in Figures 25e-f, are less recognised. The distribution in bed level spread (Figure 25j) also follows the observed distribution better. The performance metric scores support all the above observations, as all performance metrics score lowest for the dune migration model (Table 4). Especially the maximum bed level error is substantially smaller by the dune migration model predictions (0.30m) compared to the machine learning predictions. Also, the location of the maximum bed levels was more accurately predicted by the dune migration model, as the error in the location was 15m instead of 25m for the machine learning predictions.

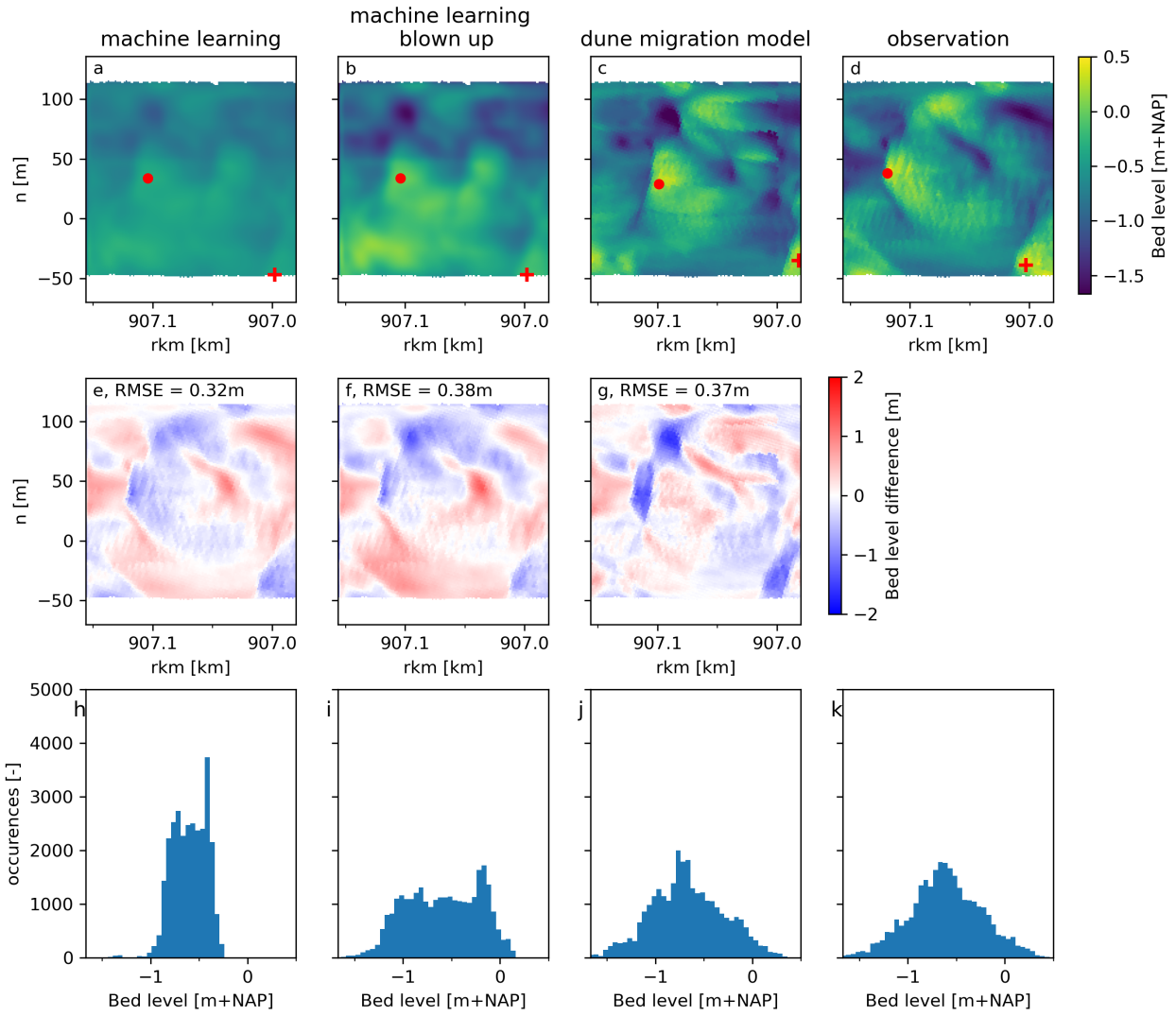


Figure 25: Example bed level predictions by (a) the machine learning model, (b) the machine learning model, with a manual increase in bed level spread, and (c) the dune migration model. (d) shows the observed bed levels for the predicted time step. In (a-d), the red dot and red plus indicate two maximum bed levels. (e-g) depict the bed level difference between the predictions and the observation. (h-k) are the histograms showing the spread in bed levels.

Table 4: Performance metric scores for the machine learning model and the dune migration model on the test sample set. The lowest metric scores are indicated in bold.

	Machine learning	Machine learning blown up	Dune migration model
MAE_b [m]	0.77	0.47	**0.30**
MAE [m]	0.28	0.33	**0.26**
RMSE [m]	0.35	0.41	**0.34**

4 Discussion

This chapter discusses the dune migration model, the machine learning model, the practical implications and research limitations. First, the advantages and disadvantages of the dune migration model and the machine learning model are discussed upon. For the machine learning model, also a comparison is done between the loss function experiment results of this study and studies that similarly evaluated the effect of these loss functions. The practical implications and research limitations follow upon this.

4.1 Dune migration model

A data-driven dune migration model was developed, which predicts the horizontal displacement of the river bed via the raster shifting method described in Section 2.2. The model can provide a good prediction of the bedform patterns and the locations of the maximum bed levels. One of the main advantages of this model is that it can easily be adapted for applications with different prediction lead times and can easily be applied for larger sections of the Waal River than currently considered.

A limitation regarding this model, is that this model only predicts the horizontal displacement of bed features. The vertical changes in the maximum bed level are not included, which are of primary interest to dredging companies and Rijkswaterstaat. Also, the empirical relation for the dune celerities (equation 2) is based on a 16.5km river section of the Waal River upstream of Tiel. The study area of this study is located within this section. However, for river sections further downstream or upstream, the empirical relation might not be as accurate anymore. Lastly, raster shifting currently shifts stationary bed features, whilst these remain stationary in place. For example, a stationary bed feature was found in the study area at coordinates [X=166160m, Y=434720m] or [rkm=907.25km, n=110m]. This bed feature would be shifted in place with raster shifting, whilst the observation indicates no horizontal displacement (e.g. Figure 22).

4.2 Machine learning model

The second data-driven approach applied in this study, was the machine learning model TrajGRU. An advantage to this model compared to the dune migration model is that the machine learning model predicts both the horizontal and vertical displacement of the bed levels. However, the maximum bed level error remained relatively large (0.47m) and model predictions showed little accuracy in capturing the locations of the overall higher and lower bed level patches. Regarding the model structure itself, the machine learning model is relatively little flexible. The prediction window of the machine learning model is tied to the time intervals in the data and the model requires two sequential bed level measurements. If either of these measurements is missing as input, the model cannot provide a prediction.

4.2.1 Loss function experiments

The adaptation of loss functions showed to have a larger effect upon the model results than the data preprocessing steps. Primarily the spreading in the bed level prediction was in-

creased more in the loss function experiments than in the data preprocessing experiments. Other studies have evaluated the effect of these loss functions in the TrajGRU machine learning model as well, but then for the precipitation forecasting application (Dekker et al., 2022; van der Kooij et al., 2021). Below, the findings of this research regarding loss functions are compared with those of the aforementioned studies.

The implementation of the Root Mean Square Error (RMSE) loss function increased the bed level prediction accuracy most of all adopted loss functions. The RMSE loss resulted in a similar bedform pattern prediction as for the Mean Square Error (MSE) loss, but with a substantially larger spread in bed levels. Consequently, the maximum bed level error decreased from 0.95m to 0.77m. These improvements upon the maximum bed level error contradict the findings by van der Kooij et al. (2021), who observed that while the RMSE loss function improved the estimation of low rainfall intensities, it led to further underestimation of peak rainfall intensities. However, in this study, both the overall bed level errors and maximum bed level errors were reduced.

The implementation of the Structural Similarity Index Measure (SSIM) improved model results slightly, as the maximum bed level error decreased from 0.95m to 0.90m and contrasts in the visual inspection were slightly more accentuated. This is in accordance with the observations by Dekker et al. (2022) and Tran and Song (2019). Additionally, Ma et al. (2022) mentioned that the SSIM implementation is designed for the overall clarity of the image, which just plays an auxiliary role. The MSE loss generally plays a leading role in the losses, which is also seen back in the results of this study as model results change very slightly with the implementation of the SSIM.

The Wasserstein loss implementation resulted in a general reconstruction of the second input frame for TrajGRU, resulting in worse model performance compared to the use of the original mean square error loss. This is in accordance with the study by Dekker et al. (2022), which also showed the worse performance of the Wasserstein loss implementation in TrajGRU for precipitation forecasting.

4.3 Practical implications

This study demonstrates that the machine learning model TrajGRU is far from applicable for implementation in short-term decision-making processes for dredging. The model predictions remain highly blurred, resulting in a high maximum bed level error of 0.77m. Even with a manual increase in the spread of the bed level prediction, the machine learning predictions kept a relatively large maximum bed level error of 0.47m. Also, the model structure is little flexible to iterate upon. Therefore, further improvements on the TrajGRU machine learning model are unlikely to yield a functional model for dredging purposes.

The dune migration model provides a more promising approach for predicting bed levels in the Waal River. The bedform patterns were better captured and the maximum bed level error was substantially smaller ($MAE_B = 0.30m$). Also, this model can more easily be adapted for applications with differentiating prediction lead times and can easily be applied

for larger sections of the Waal River. However, this model only predicts the horizontal displacement of bed features. The vertical changes in the maximum bed level are not included, which is essential for identifying future bottlenecks for navigation. To include the vertical displacement as well, a combination of the developed dune migration model and a machine learning or numerical modelling approach could be sought. The dune migration model can give a preliminary prediction on the location of the bedforms, whilst the other approach predicts the change in vertical displacements. For example, the 2DV numerical model of Paarlberg et al. (2009) captures the growth and decay of the highest dunes well (Lokin et al., 2024). The model generally tends to underestimate the horizontal displacement of river dunes, but this can be overcome by implementing the raster shifting method of the dune migration model. Thus, both model limitations of the developed dune migration model and the model by Paarlberg et al. (2009) could be overcome by seeking a relation between the two. However, the model by Paarlberg et al. (2009) provides a 2DV prediction of the bed levels, whilst raster shifting can provide a 3D prediction. By combining the two, the three-dimensionality in the raster shifting prediction would therefore be lost. Other than a combination with a numerical model, a machine learning model could be combined with raster shifting. A machine learning model could be trained on predicting the vertical displacements based on relevant features such as discharge, river curvature and grain size. Note that this machine learning model could either be a simple regression analysis, coupling the discharge and dune height variations, or a complicated machine learning algorithm like TrajGRU. The combination with a machine learning algorithm could ensure a 3D prediction, with no loss of data. In conclusion, combining the raster shifting method with either a numerical model or machine learning model could result in a functional model for dredging purposes, capable of both predicting the maximum bed level and the spatial configuration of the bedforms accurately.

4.4 Research limitations

Four research limitations have been identified and are discussed here. Firstly, the raster shifting method of the dune migration model was applied on a measured discharge time series to derive the shifting distance. In reality, the predicted daily discharges will be provided, introducing relatively high uncertainties, thereby impacting the accuracy of the raster shifting distance. Consequently, the dune migration model predictions might perform less well in practice than currently shown in this research.

Regarding the machine learning model application, the main limitation is the limited availability of data. For example, Singh et al. (2017) and Tian et al. (2019) applied 7000 or more training samples for training the TrajGRU machine learning model, whilst this study applied only 162 to 226 training samples. By increasing the dataset, machine learning performance generally improves (e.g. B. Liu et al., 2023).

Also, interventions in the data preprocessing experiments were accumulated under the assumption that each intervention would improve the prediction accuracy of the TrajGRU machine learning model. However, the sample selection intervention slightly reduced the model performance compared to the use of the original data. This is likely due to the de-

creased number of training samples in the sample selection. Skipping the sample selection intervention for the resulting wavelet model could have resulted in a slightly better performing machine learning model.

Furthermore, the current study area was chosen such that it represents a simple situation to test the suitability of the data-driven approaches. It is located in a section where no dredging and ploughing activities occur. However, in practice, the data-driven approaches would mainly be relevant to areas requiring frequent dredging and ploughing. In such locations, the bed level data can contain disturbances in the bed due to dredging. For the machine learning model, this can complicate the learning of bed developments, possibly decreasing the model performance further. The dune migration model is relatively little sensitive to dredging activities. The prediction is only influenced if the original measurement was iterated by dredging or ploughing.

Lastly, the locational error in the maximum bed level indicated a relatively low error of 25m for the machine learning model predictions. However, the locational error considered the 80x80m regions as shown in Appendix C, whilst the machine learning predictions predicted extra patches of high bed levels outside the range of these 80x80m regions. These patches were therefore not included in the evaluation of the locational error, meaning that the locational error would be larger for the machine learning model than currently stated in this thesis. For the dune migration model, most high bed levels in the predictions are captured within the 80x80m regions. The locational error of 15m would therefore remain the same.

5 Conclusion

This research aimed to obtain a two-weekly bed level prediction in 3D for the Waal River by developing and comparing two data-driven approaches: a dune migration model and machine learning. Below, the findings of this research are presented for each research question, with the main conclusion of this research summarised at the end.

RQ1: What is the bed level prediction accuracy of the dune migration model?
The dune migration model predicts the bed levels by applying raster shifting, for which, a bed level measurement is shifted downstream along the channel axis based on empirically derived dune celerities. The bed level predictions visually showed relatively accurate predictions on the bedform plan. The location of maximum bed levels and overall higher and lower bed level patches were accurately captured. The point-wise metrics showed errors of MAE=0.26m and RMSE=0.34m, and the maximum bed level error (MAE_b) was 0.30m.

RQ2: How can data preprocessing be implemented to improve the bed level prediction accuracy of the TrajGRU machine learning model for the Waal River?
Five data preprocessing experiments were conducted to improve the machine learning model performance: sample selection, removal of outliers, coordinate transformation, wavelet reconstruction and consistency in time intervals. Each experiment builds upon the interventions introduced in the preceding one. The model trained on the wavelet reconstructed data, and thus also sample selection, outlier removal and coordinate transformation, performed the best of all experiments. The bedform patterns were better captured compared to the model trained on the original data. However, model results remained highly blurred with a slight increase in the maximum bed level error of 0.92 to 0.95m.

RQ3: How can the loss function be adapted to improve the bed level prediction accuracy of the TrajGRU machine learning model for the Waal River?
Three loss function adaptations were tested to improve the machine learning model performance. The tested loss functions were the Root Mean Square Error (RMSE), a combination of the Mean Square Error (MSE) and the structural similarity, and the Wasserstein loss. The original model applied the MSE loss function only. Of all experiments, the RMSE loss improved model predictions the most. The bed level spread increased considerably, reducing the maximum bed level from 0.95m to 0.77m, and the location of the maximum bed levels was captured best of all experiments. Still, the model produced highly blurred results.

RQ4: How well does the improved TrajGRU machine learning model perform compared to the dune migration model for predicting bed levels for the Waal River?
The best performing machine learning model, using the wavelet constructed data and the RMSE loss function, was tested against the dune migration model. It resulted that the dune migration model performed better on both predicting the bedform patterns and the maximum bed levels. The maximum bed level error was substantially smaller for the dune migration model (0.30m) than for the machine learning model (0.77m), with a locational error of only 15m instead of 25m. Even with extra spreading manually induced in the machine learning bed level prediction, the dune migration model outperformed the machine learning model.

Summarising, the potential of a complicated machine learning algorithm like TrajGRU is small for predicting bed levels in the Waal River. The dune migration model proves a more promising approach for predicting bed levels in the Waal River. Future research could focus more on the dune migration model and could combine its method with an approach suitable for predicting vertical bed level changes as well.

6 Recommendations

In this section, recommendations are given on the outcomes of this research. The two main recommendations of this research are to delve deeper into the dune migration model and to have a consultation with dredging companies and Rijkswaterstaat on their requested model requirements for bed level predictions. Also, recommendations are given upon future machine learning studies focussing on bed level predictions.

6.1 Dune migration model

Four recommendations for future research on the dune migration model were identified. Firstly, it is recommended to search for a combination between the raster shifting method of the dune migration model and a numerical or machine learning approach for predicting bed levels in rivers. The vertical changes in the maximum bed level are not included in the current raster shifting methodology, which is essential for identifying future bottlenecks for dredging. Therefore, raster shifting could be used to predict the horizontal displacement of bed features, whilst the numerical or machine learning approach could focus on predicting the vertical displacement. As discussed in Section 4.3, the 2DV numerical model of Paarlberg et al. (2009) could be suitable to combine with raster shifting, taking into account the limitation of the prediction becoming in 2DV rather than 3D. Regarding a machine learning approach for vertical bed level developments, a simple regression method could be used, coupling the discharge and dune height variations, or a complicated machine learning algorithm like TrajGRU. An advantage of the machine learning approach to the numerical approach is that the three-dimensionality in the bed level prediction can be maintained.

Secondly, future research could focus on the effect of the uncertainty in discharge predictions for the dune migration model predictions. As mentioned in Section 4.4, this study applied the measured discharge time series to obtain the shifting distance. However, in practice, the predicted discharge will be provided, introducing more uncertainties in the shifting distance, impacting the model prediction.

Thirdly, it would be recommended to include a method for filtering stationary bed features before shifting the raster. These stationary bed features are currently shifted in location, whilst these bed features remain stationary in place.

Lastly, the empirical relation for deriving the dune celerities from the river discharge (equation 2) is based on measurements of a 16.5km river section of the Waal River upstream of Tiel (Lokin et al., 2022). For river sections outside of this range, it would be recommended to test the accuracy of the dune migration model with the current coefficients of equation 2 and recalibrate these coefficients if the model predictions do not accurately capture the horizontal displacement of the river bed.

6.2 Consultation with key users

It is recommended to have a consultation with the key users of the bed level prediction model. The understanding of the model requirements by the dredging companies and Rijkswaterstaat was limited in this study. By conducting a consultation with the key users, a better understanding can be obtained of e.g. the required prediction lead time and desired prediction accuracy.

6.3 Future machine learning approaches for predicting bed levels

During this study, five main recommendations were derived for future studies applying similar machine learning models for bed level predictions. Firstly, it is recommended to increase the dataset by finding similar river sections and combining these datasets to create extra training samples. By increasing the dataset, machine learning performance generally improves (e.g. B. Liu et al., 2023). Before combining datasets of separate river sections, careful considerations should be taken on similarities based on e.g. river curvature, groyne locations and shipping intensities. An other method that could be considered to obtain more training data is by considering a 2DV machine learning model for predicting bed levels in the Waal River. More training data can be generated by providing several longitudinal profiles of the Waal River as input to the model. However, this removes the three-dimensional dynamics of river dunes and can result in a loss of data in the prediction as the model does not capture the entire riverbed. Lastly, CoVadem data could be used, which contains a large bathymetry dataset of daily measurements of the Waal River bathymetry in 3D for the past several years. However, the uncertainty in the CoVadem bed level measurements has shown to be substantially high (van Middendorp, 2020). This uncertainty should be reconsidered before applying such data for machine learning applications.

Secondly, it is recommended to decrease the grid resolution from 1x1m to 10x10m or larger. A lower grid resolution of 10x10m or larger could also suffice to capture both the bed feature dynamics like river dunes and maximum bed levels. This reduces the number of model weights in the model, as the number of grid cells in the model domain reduces. Therefore, model predictions will likely improve for the same amount of training data.

Thirdly, it is recommended to focus more towards a correct choice in the loss function, rather than to change data preprocessing steps. This study has shown that the choice of loss function improved model results more than data preprocessing. Two loss functions that have not been tested in this research, but also show potential for improving image predictions are the balanced loss function by Shi et al. (2017) and the use of adversarial losses. Especially, the use of adversarial losses has shown much potential for providing accurate predictions (Jing et al., 2019; Singh et al., 2017; Tian et al., 2019). However, these types of loss functions require a thousand or more training samples before they provide good model results (Karras et al., 2020; B. Liu et al., 2023). Thus, the bed level dataset should increase before adversarial losses become attractive to apply.

Furthermore, it would be recommended to provide extra data on relevant features/variables to the machine learning model. This could enhance model performance, as more information about the system is given. For example, the discharge could be added to the model as this highly influences river dune dynamics (Lokin et al., 2022). However, it should be noted that the machine learning model applied in this study is not suitable to include such extra features. The model accepts raster inputs only and generally looks at spatial structures and relations, whilst features such as discharge are spatially invariant (Shi et al., 2017). This was also tested by implementing the discharge as extra feature, which resulted in worse model performance (Appendix D). Nevertheless, alternative machine learning approaches, which are more suitable for implementing features such as discharge, may improve in their results.

References

Beygipoor, G., Shafai Bajestan, M., & Nazari, S. (2013). The effects of submerged vane angle on sediment entry to an intake from a 90 degree converged bend. *Advances in Environmental Biology*, *7*(9), 2283–2292. https://www.researchgate.net/publication/288109917

Bosboom, J. (2019). *Quantifying the quality of coastal morphological predictions* [Doctoral dissertation, Delft University of Technology]. https://doi.org/10.4233/uuid:e4dc2dfc-6c9c-4849-8aa9-befa3001e2a3

Bradley, R. W., & Venditti, J. G. (2021). Mechanisms of Dune Growth and Decay in Rivers. *Geophysical Research Letters*, *48*(20). https://doi.org/10.1029/2021GL094572

Buitink, J., Tsiokanos, A., Geertsema, T., ten Velden, C., Bouaziz, L., & Sperna Weiland, F. (2023). *Implications of the KNMI'23 climate scenarios for the discharge of the Rhine and Meuse* (tech. rep.). Deltares.

Cisneros, J., Best, J., van Dijk, T., Almeida, R. P. d., Amsler, M., Boldt, J., Freitas, B., Galeazzi, C., Huizinga, R., Ianniruberto, M., Ma, H., Nittrouer, J. A., Oberg, K., Orfeo, O., Parsons, D., Szupiany, R., Wang, P., & Zhang, Y. (2020). Dunes in the world's big rivers are characterized by low-angle lee-side slopes and a complex shape. *Nature Geoscience*, *13*(2), 156–162. https://doi.org/10.1038/s41561-019-0511-7

Dekker, D. D., Schleiss, M. A., Fioranelli, F., Taormina, R., Basu, S., & van Hoek, M. (2022, July). *Perceptual losses in precipitation nowcasting (MSc Thesis)*. Delft University of Technology. http://resolver.tudelft.nl/uuid:5e61e2a7-bf04-41cd-a66a-68d1ceeb3c99

de Ruijsscher, T. V., Naqshband, S., & Hoitink, A. J. F. (2020). Effect of non-migrating bars on dune dynamics in a lowland river. *Earth Surface Processes and Landforms*, *45*(6), 1361–1375. https://doi.org/10.1002/esp.4807

Dietrich, W. E., & Dungan Smith, J. (1984). Bed Load Transport in a River Meander. *Water Resources Research*, *20*(10), 1355–1380.

Giri, S., & Shimizu, Y. (2006). Numerical computation of sand dune migration with free surface flow. *Water Resources Research*, *42*(10). https://doi.org/10.1029/2005WR004588

Goll, A., & Kopmann, R. (2012). *Dune simulation with TELEMAC-3D and SISYPHE: A parameter study.* (tech. rep.). Federal Waterways Engineering and Research Institute.

Gutierrez, R. R., Mallma, J. A., Núñez-González, F., Link, O., & Abad, J. D. (2018). Bedforms-ATM, an open source software to analyze the scale-based hierarchies and dimensionality of natural bed forms. *SoftwareX*, *7*, 184–189. https://doi.org/10.1016/j.softx.2018.06.001

Hiemstra, K. S., van Vuren, S., Vinke, F. S. R., Jorissen, R. E., & Kok, M. (2020). Assessment of the functional performance of lowland river systems subjected to climate change and large-scale morphological trends. *International Journal of River Basin Management*, 1–22. https://doi.org/10.1080/15715124.2020.1790580

Hulscher, S. J. M. H., Daggenvoorde, R. J., Warmink, J. J., Vermeer, K., & van Duin, O. (2017). River dune dynamics in regulated rivers, 26–28. https://doi.org/10.1061/(ASCE)HY.1943

Jing, J., Li, Q., Ding, X., Sun, N., Tang, R., & Cai, Y. (2019). AENN: A Generative Adversarial Neural Network for Weather Radar Echo Extrapolation. *International Archives*

of the Photogrammetry, Remote Sensing and Spatial Information Sciences, *42*(3/W9), 89–94. https://doi.org/10.5194/isprs-archives-XLII-3-W9-89-2019

Jonnalagadda, J., & Hashemi, M. (2023). Quality-Aware Conditional Generative Adversarial Networks for Precipitation Nowcasting. *Engineering Proceedings*, *39*(1). https://doi.org/10.3390/engproc2023039011

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., & Aila, T. (2020). Training Generative Adversarial Networks with Limited Data. http://arxiv.org/abs/2006.06676

Klaassen, G. J., & Sloff, C. J. (2000, September). *Voorspelling bodomligging en herstelrelaties voor baggeren op de Waal* (tech. rep.). Deltares.

Kruis, W. S., Lokin, L. R., van Denderen, R. P., & Hulscher, S. J. M. H. (2023). *River Dune Behaviour in Dredged Areas (MSc Thesis)*. University of Twente.

Le, Q., Chan, F., & D'Amours, C. (2022). *Novel Deep Learning Models for Spatiotemporal Predictive Tasks (MSc Thesis)*. University of Ottawa.

Le Deunf, J., Debese, N., Schmitt, T., & Billot, R. (2020). A review of data cleaning approaches in a hydrographic framework with a focus on bathymetric multibeam echosounder datasets. *Geosciences*, *10*(7), 254. https://doi.org/10.3390/geosciences10070254

Lefebvre, A. (2019). Three-Dimensional Flow Above River Bedforms: Insights From Numerical Modeling of a Natural Dune Field (Río Paraná, Argentina). *Journal of Geophysical Research: Earth Surface*, *124*(8), 2241–2264. https://doi.org/10.1029/2018JF004928

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A*, *379*(2194). https://doi.org/10.1098/rsta.2020.0209

Lim, B., Zohren, S., & Roberts, S. (2020, January). *Recurrent Neural Filters: Learning Independent Bayesian Filtering Steps for Time Series Prediction*. University of Oxford. http://arxiv.org/abs/1901.08096

Liu, B., Zhang, T., Yu, Y., & Miao, L. (2023). A data generation method with dual discriminators and regularization for surface defect detection under limited data. *Computers in Industry*, *151*. https://doi.org/10.1016/j.compind.2023.103963

Liu, Y., Yang, Z., Li, W., & Li, M. (2022). Applying machine learning methods to river topography prediction under different data abundances. https://doi.org/10.22541/au.164873886.65750516/v1

Lokin, L. R. (2020). *Dune dynamics under high and low flows: Literature report*. (CE&M Research report; No. 2020R-003/WEM-003). University of Twente.

Lokin, L. R., Bomers, A., Warmink, J. J., & Hulscher, S. J. M. H. (2024). *Forecasting the development of observed river dunes with a process based model*. (Submitted). University of Twente.

Lokin, L. R., Warmink, J. J., Bomers, A., & Hulscher, S. J. M. H. (2022). River Dune Dynamics During Low Flows. *Geophysical Research Letters*, *49*(8). https://doi.org/10.1029/2021GL097127

Lokin, L. R., Warmink, J. J., & Hulscher, S. J. M. H. (2023). The Effect of Sediment Transport Models on Simulating River Dune Dynamics. *Water Resources Research*, *59*(12). https://doi.org/10.1029/2023WR034607

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers and Geosciences*, *34*(9), 1044–1055. https://doi.org/10.1016/j.cageo.2007.07.010

Ma, Z., Zhang, H., & Liu, J. (2022). Focal Frame Loss: A Simple but Effective Loss for Precipitation Nowcasting. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 6781–6788. https://doi.org/10.1109/JSTARS.2022.3194522

Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), 91–99. https://doi.org/10.1016/j.gltp.2022.04.020

Maleika, W. (2020). Inverse distance weighting method optimization in the process of digital terrain model creation based on data collected from a multibeam echosounder. *Applied Geomatics*, *12*, 397–407. https://doi.org/10.1007/s12518-020-00307-6

Maloo, J. (2018). Machine Learning For Beginners From Zero Level. https://medium.com/@maloojinesh/machine-learning-for-beginners-from-zero-level-8be5b89bf77c

Mehta, P., Bukov, M., Wang, C.-H., Day, A. G. R., Richardson, C., Fisher, C. K., & Schwab, D. J. (2019). A high-bias, low-variance introduction to Machine Learning for physicists. *Physics Reports*, *810*, 1–124. https://doi.org/10.1016/j.physrep.2019.03.001

Ministry of Infrastructure and Water Management. (2022). *Uitvoeringsagenda Klimaatbestendige Netwerken Hoofdvaarwegennet Hoofdwatersysteem Hoofdwegennet* (tech. rep.). https://www.rijksoverheid.nl/documenten/rapporten/2022/03/29/202242087-1-bijlage-bij-kamerbrief-rws-uitvoeringsagenda-klimaatbestendige-netwerken

Mohrig, D., & Smith, J. D. (1996). Predicting the migration rates of subaqueous dunes. *Water Resources Research*, *32*(10), 3207–3217. https://doi.org/10.1029/96WR01129

Nabi, M., De Vriend, H. J., Mosselman, E., Sloff, C. J., & Shimizu, Y. (2012). Detailed simulation of morphodynamics: 1. Hydrodynamic model. *Water Resources Research*, *48*(12). https://doi.org/10.1029/2012WR011911

Nabi, M., De Vriend, H. J., Mosselman, E., Sloff, C. J., & Shimizu, Y. (2013a). Detailed simulation of morphodynamics: 2. Sediment pickup, transport, and deposition. *Water Resources Research*, *49*(8), 4775–4791. https://doi.org/10.1002/wrcr.20303

Nabi, M., De Vriend, H. J., Mosselman, E., Sloff, C. J., & Shimizu, Y. (2013b). Detailed simulation of morphodynamics: 3. Ripples and dunes. *Water Resources Research*, *49*(9), 5930–5943. https://doi.org/10.1002/wrcr.20457

Naqshband, S., Ribberink, J. S., Hurther, D., & Hulscher, S. J. M. H. (2014). Bed load and suspended load contributions to migrating sand dunes in equilibrium. *Journal of Geophysical Research: Earth Surface*, *119*(5), 1043–1063. https://doi.org/10.1002/2013JF003043

Naqshband, S., Sterlini, F., Dohmen-Janssen, M., & Hulscher, S. J. M. H. (2011). A model study on the influence of suspended sediment transport on river dune morphology and evolution. *7th IAHR Symposium on River, Coastal and Estuarine Morphodynamics (RCEM), Bejing*. https://www.researchgate.net/publication/280612121

Nelson, J. M., Logan, B. L., Kinzel, P. J., Shimizu, Y., Giri, S., Shreve, R. L., & Mclean, S. R. (2011). Bedform response to flow variability. *Earth Surface Processes and Landforms*, *36*(14), 1938–1947. https://doi.org/10.1002/esp.2212

Paarlberg, A. J., Dohmen-Janssen, C. M., Hulscher, S. J. M. H., & Termes, P. (2009). Modeling river dune evolution using a parameterization of flow separation. *Journal of Geophysical Research: Earth Surface*, *114*(1). https://doi.org/10.1029/2007JF000910

Rijkswaterstaat. (2020, September). Baggeren om te kunnen varen. https://www.bodemplus.nl/actueel/nieuwsberichten/2020/baggeren-varen/

Rijkswaterstaat. (2023). Baggeren en terugstorten houdt de rivier stabiel. https://www.rijkswaterstaat.nl/nieuws/archief/2023/05/baggeren-en-terugstorten-houdt-de-rivier-stabiel

Salinas, D., Flunkert, V., & Gasthaus, J. (2019). DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks. http://arxiv.org/abs/1704.04110

Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.-Y., Wong, W.-k., & Woo, W.-c. (2017). Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model. http://arxiv.org/abs/1706.03458

Siebe Swart. (2019). Luchtfoto Overnachtingshaven IJzendoorn. https://www.siebeswart.nl/image/I0000fq_onLj1yks

Singh, S., Sarkar, S., & Mitra, P. (2017). A deep learning based approach with adversarial regularization for Doppler weather radar ECHO prediction. *2017 IEEE International Geoscience and Remote Sensing Symposium*, 5205–5208. https://doi.org/10.1109/IGARSS.2017.8128174

Srivastava, N., Mansimov, E., & Salakhutdinov, R. (2015). Unsupervised Learning of Video Representations using LSTMs. http://arxiv.org/abs/1502.04681

Sys, C., Van de Voorde, E., Vanelslander, T., & van Hassel, E. (2020). Pathways for a sustainable future inland water transport: A case study for the European inland navigation sector. *Case Studies on Transport Policy*, *8*(3), 686–699. https://doi.org/10.1016/j.cstp.2020.07.013

ten Brinke, W. B. M., Wilbers, A. W. E., & Wesseling, C. (1999). Dune growth, decay and migration rates during a large-magnitude flood at a sand and mixed sand–gravel bed in the Dutch Rhine river system. *Special publications of the international association of sedimentologists*, *28*, 15–32. https://doi.org/10.1002/9781444304213.ch2

Tian, L., Li, X., Ye, Y., Xie, P., & Li, Y. (2019). A Generative Adversarial Gated Recurrent Unit Model for Precipitation Nowcasting. *IEEE Geoscience and Remote Sensing Letters*, *17*(4), 601–605. https://doi.org/10.1109/LGRS.2019.2926776

Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2021). Deep Learning for Time Series Forecasting: A Survey. *Big Data*, *9*(1), 3–21. https://doi.org/10.1089/big.2020.0159

Tran, Q. K., & Song, S. K. (2019). Computer vision in precipitation nowcasting: Applying image quality assessment metrics for training deep neural networks. *Atmosphere*, *10*(5). https://doi.org/10.3390/atmos10050244

van der Kooij, E., Schleiss, M. A., Fioranelli, F., Taormina, R., Lugt, D., & van Hoek, M. (2021, July). *Nowcasting heavy precipitation in the Netherlands: a deep learning approach (MSc Thesis)*. Delft University of Technology. http://resolver.tudelft.nl/uuid:536b1a77-625c-4476-9354-4d5b259a1080

van der Mark, C. F., Blom, A., & Hulscher, S. J. M. H. (2008). Quantification of variability in bedform geometry. *Journal of Geophysical Research: Earth Surface*, *113*(3). https://doi.org/10.1029/2007JF000940

van Denderen, R. P., Kater, E., Jans, L. H., & Schielen, R. M. (2022). Disentangling changes in the river bed profile: The morphological impact of river interventions in a managed river. *Geomorphology*, *408*. https://doi.org/10.1016/j.geomorph.2022.108244

van Dijk, T. A. G. P., van der Mark, C. F., Doornenbal, P. J., Menninga, P. J., Keppel, J. F., Rodriguez Aguilera, D., Hopman, V., & Erkens, G. (2012). *Onderzoek Meetstrategie en Bodemdynamiek* (tech. rep.). Deltares.

van Duin, O. J. M., Hulscher, S. J. M. H., & Ribberink, J. S. (2021). Modelling Regime Changes of Dunes to Upper-Stage Plane Bed in Flumes and in Rivers. *Applied Sciences*, *11*(23). https://doi.org/10.3390/app112311212

van Middendorp, D. C. (2020). *Using CoVadem to increase the value of Rijkswaterstaat's least sounded depth in the river Waal (MSc Thesis)*. University of Twente.

van Rijn, L. C. (1984). Sediment Transport, Part III: Bed forms and Alluvial Roughness. *Journal of Hydraulic Engineering*, *110*(12), 1733–1754.

Vinke, F., van Koningsveld, M., van Dorsser, C., Baart, F., van Gelder, P., & Vellinga, T. (2022). Cascading effects of sustained low water on inland shipping. *Climate Risk Management*, *35*. https://doi.org/10.1016/j.crm.2022.100400

Wang, Y., Smola, A., Maddix, D. C., Gasthaus, J., Foster, D., & Januschowski, T. (2019). Deep Factors for Forecasting. In C. Kamalika & R. Salakhutdinov (Eds.), *Proceedings of the 36th international conference on machine learning* (pp. 6607–6617). PMLR.

Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multi-Scale Structural Similarity for Image Quality Assessment. *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, 1398–1402. https://doi.org/10.1109/ACSSC.2003.1292216

Wilbers, A. W. E., & ten Brinke, W. B. M. (2003). The response of subaqueous dunes to floods in sand and gravel bed reaches of the Dutch Rhine. *Sedimentology*, *50*(6), 1013–1034. https://doi.org/10.1046/j.1365-3091.2003.00585.x

Zeng, Q., Li, H., Zhang, T., He, J., Zhang, F., Wang, H., Qing, Z., Yu, Q., & Shen, B. (2022). Prediction of Radar Echo Space-Time Sequence Based on Improving TrajGRU Deep-Learning Model. *Remote Sensing*, *14*(19). https://doi.org/10.3390/rs14195042

Zhang, T., Liew, S. Y., Ng, H. F., Qin, D., Lee, H. C., Zhao, H., & Wang, D. (2023). GraphAT Net: A Deep Learning Approach Combining TrajGRU and Graph Attention for Accurate Cumulonimbus Distribution Prediction. *Atmosphere*, *14*(10). https://doi.org/10.3390/atmos14101506

Zomer, J. Y., Naqshband, S., Vermeulen, B., & Hoitink, A. J. F. (2021). Rapidly Migrating Secondary Bedforms Can Persist on the Lee of Slowly Migrating Primary River Dunes. *Journal of Geophysical Research: Earth Surface*, *126*(3). https://doi.org/10.1029/2020JF005918

Appendices

A TrajGRU model structure

The TrajGRU model structure, for predicting bed levels, combines two deep learning techniques, a convolutional neural network, and a recurrent neural network. Convolutional Neural Networks (CNN) are designed to extract local features from image datasets, which are invariant over spatial dimensions (features can be recognised regardless of their position in the input) (Mehta et al., 2019; Torres et al., 2021). Predefined filters are moved across the image and the recognition of the filters is quantified and stored in a convolution matrix. This matrix is reduced in dimension via 'max pooling', which simplifies and summarizes the information extracted by the convolutional layers. With max pooling, the convolution matrix is split into smaller regions and for each region, the maximum value is outputted. Features such as river dunes can be extracted by the CNN, which are fed to the recurrent neural network to provide predictions for. A schematisation of a CNN is depicted in Figure 26.

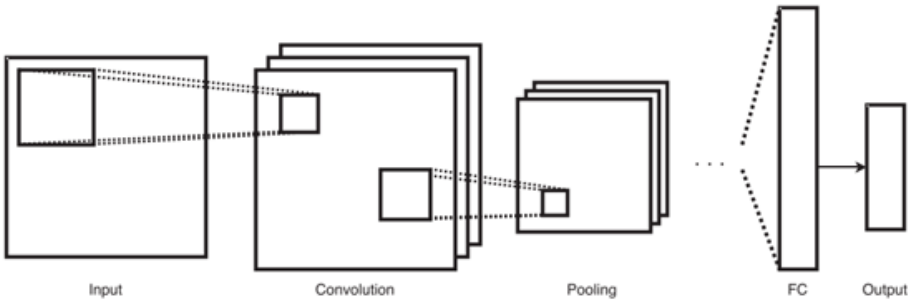


Figure 26: Schematisation of a convolutional neural network (Torres et al., 2021)

Recurrent Neural Networks (RNN) have specifically been designed to deal with sequential data such as time series for forecasting (Torres et al., 2021), and have been widely used in forecasting applications (Lim et al., 2020; Salinas et al., 2019; Y. Wang et al., 2019). At their core, RNN cells have an internal memory state, containing a compact summary of past information. This memory state is continually updated with new observations at each time step (Lim & Zohren, 2021). Recurrent neural networks therefore show the ability to capture temporal dependencies well (Torres et al., 2021). A schematisation of the RNN structure, as applied in the TrajGRU model by HKV, is shown in Figure 27.

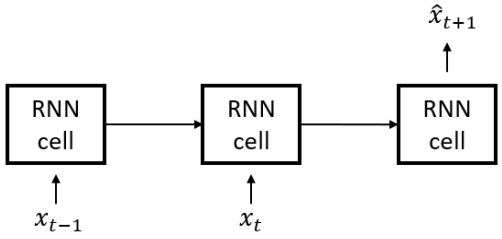


Figure 27: Schematisation of the recurrent neural network in the TrajGRU by HKV. x depicts the input frame and $\hat{x}$ the predicted frame.

Recurrent neural networks generally have problems regarding their infinite lookback window, causing a vanishing gradient to find their optimal solution. Therefore, a Gated Recurrent Unit (GRU) is implemented. A GRU controls the information flow between input and output via a gating mechanism. It consists of a reset gate, which controls what information from the previous inputs should be kept or disregarded, and an update gate which controls what information from the new input should be used. The reset gate controls the short-term memory of the model, and the update gate the long-term memory (van der Kooij et al., 2021).

The model uses an encoder-forecaster structure, where downsampling of the input data is performed in the encoder, and upsampling in the forecaster. A schematisation of the model structure is depicted in Figure 28 and Table 5.

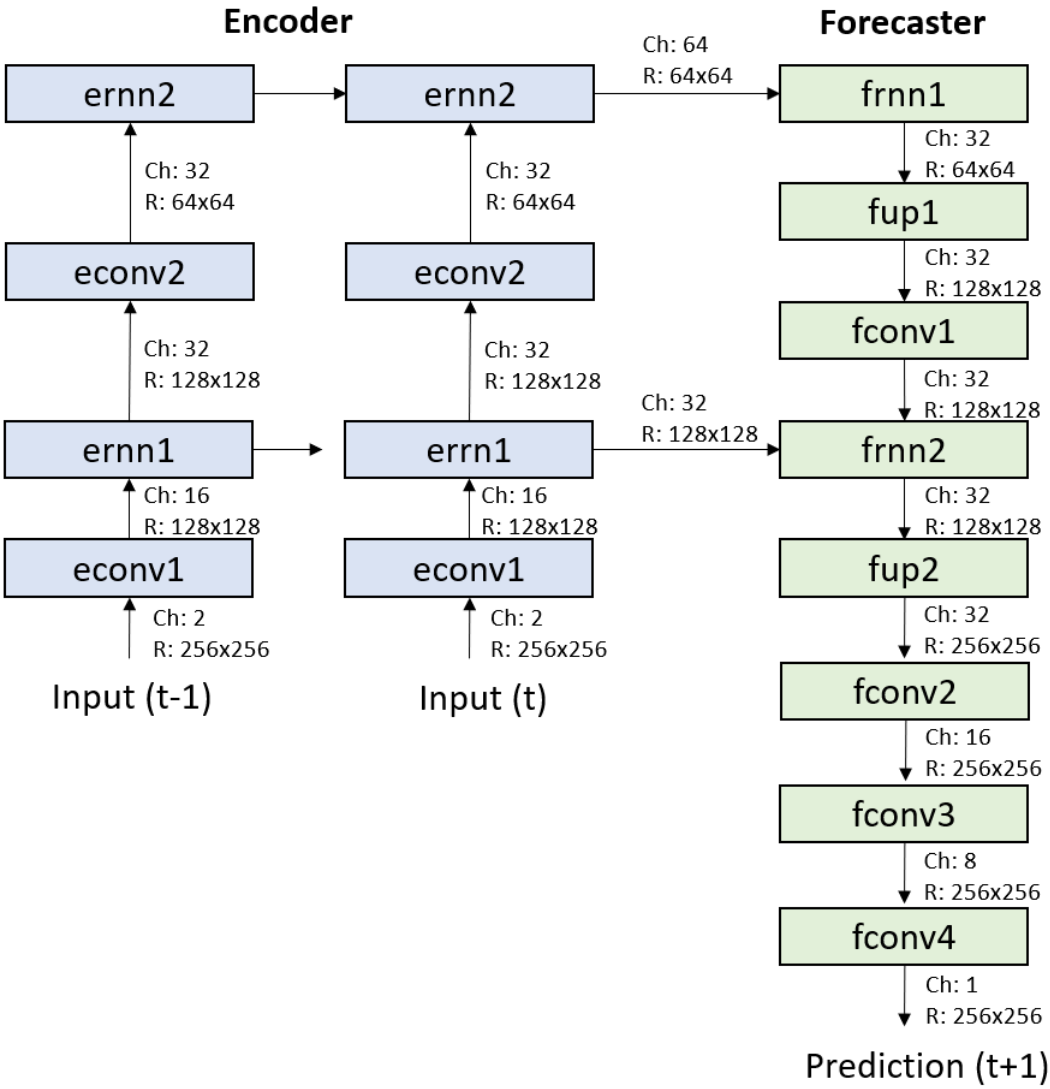


Figure 28: Figure of overview TrajGRU model structure. R: represents the resolution of the matrices, Ch: represent the number of matrices (channels). e denotes the encoder and f the forecaster. conv is a convolutional layer including pooling, and rnn a recurrent neural network cell. Please refer to Table 5 for more information on the figure.

Table 5: Encoder forecaster structure of the employed TrajGRU model

	Kernel	Stride	Padding	Channels In/Out	Res In	Res Out	Type
econv1	4x4	2x2	1x1	2/16	256x256	128x128	Convolution
ernn1	3x3	1x1	1x1	16/32	128x128	128x128	RNN
econv2	4x4	2x2	1x1	32/32	128x128	64x64	Convolution
ernn2	3x3	1x1	1x1	32/64	64x64	64x64	RNN
frnn1	3x3	1x1	1x1	64/32	64x64	64x64	RNN
fup1	-	-	-	32/32	64x64	128x128	Upsample
fconv1	3x3	1x1	1x1	32/32	128x128	128x128	Convolution
frnn2	3x3	1x1	1x1	32/32	128x128	128x128	RNN
fup2	-	-	-	32/32	128x128	256x256	Upsample
fconv2	3x3	1x1	1x1	32/16	256x256	256x256	Convolution
fconv3	3x3	1x1	1x1	16/8	256x256	256x256	Convolution
fconv4	1x1	1x1	0x0	8/1	256x256	256x256	Convolution

A.1 Model hyperparameters

Hyperparameters are parameters that dictate the learning process. They are set before training the model and remain constant during the training of the model. In Table 6, hyperparameter values of the TrajGRU model are summarised. The batch size of 2 indicates that each batch contains two samples. An epoch refers to the one entire passing of training data through the machine learning algorithm. Generally, model training of TrajGRU ends after 2-8 epochs, due to stopping the validation criterion (see Section 2.3.3). The learning rate controls the rate of learning or speed at which the model learns. It regulates the amount of allocated error with which the model's weights are updated each time they are updated. A learning rate of 0.0001 is commonly used for TrajGRU applications (Shi et al., 2017; Tran & Song, 2019; Zeng et al., 2022). Lastly, a constant random seed was set for all experiments. The seed determines the initial set of model weights in the neural network. By keeping this constant between experiments, the difference between experiments is visible without the interference of randomness. The model is implemented using PyTorch and is trained on an 8 GB Nvidia GeForce RT2080 GPU.

Table 6: The hyperparameter values used in this study

Parameter	value
Batch size	2
Maximum epochs	15
Learning rate	0.0001
Seed	3547236857

B Outlier locations in the bed level data

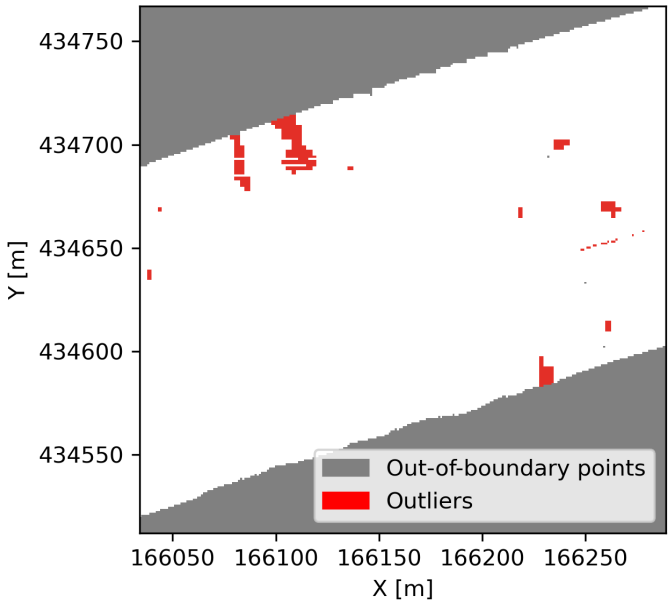


Figure 29: The locations of the identified outliers in the multibeam echosounder bed level measurements, based on the threshold deviation of 3m from the median bed (see Section 2.4.2). Plotted in the RD new coordinate system.

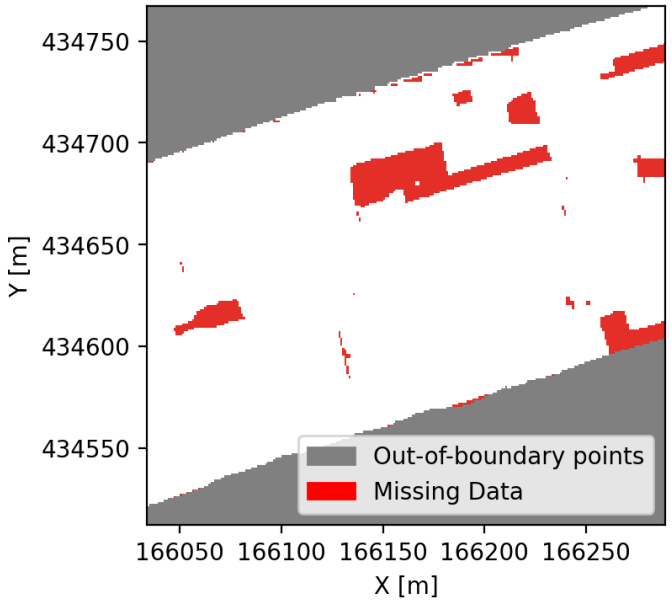


Figure 30: The locations of missing data in the multibeam echosounder bed level measurements. Plotted in the RD new coordinate system.

C Example predictions of the two data-driven approaches

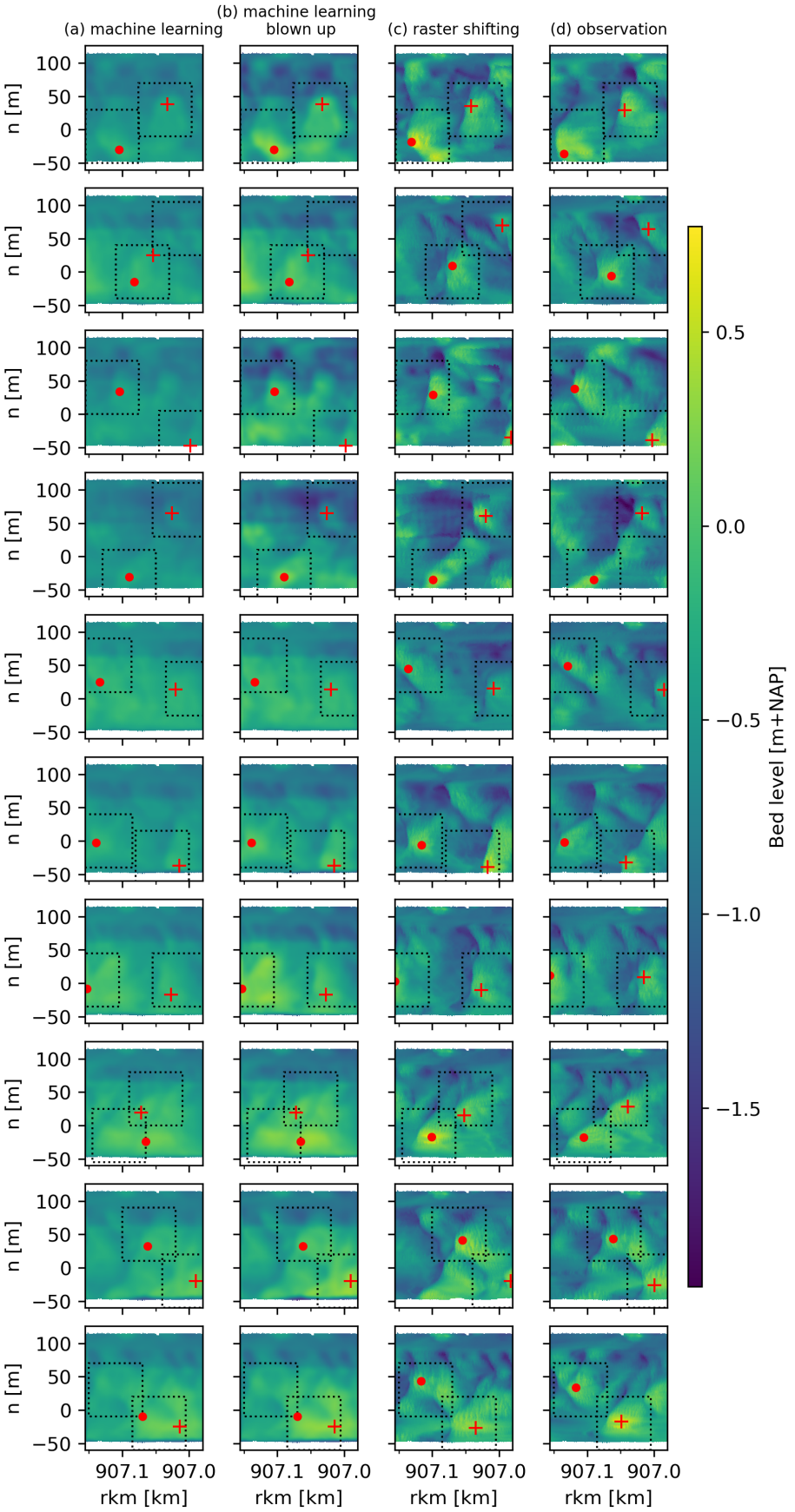


Figure 31: Example predictions by (a) the machine learning model using the wavelet reconstructed data and RMSE loss function, (b) the machine learning model with an extrapolation in the bed levels and (c) the dune migration model. (d) depicts the observed bed levels. The red dot and red plus indicate the location of the maximum bed levels in the black dotted regions.

D Discharge feature implementation

To see if the TrajGRU machine learning model picks up the relation between relevant features and bathymetry developments, a two-weekly averaged discharge was added as feature into the model. This discharge represents the mean discharge expected for the next two weeks and was obtained from actual discharge measurements. In practice, the predicted discharge should be provided, which would introduce more uncertainties in the bed level prediction. The two-weekly averaged discharge was implemented as an extra 256x256 raster frame containing a spatially constant discharge value.

Figure 32 depicts a prediction by the TrajGRU model with and without discharge implementation. The implementation of discharge shows little visual changes compared to the TrajGRU model without discharge implementation. Bedforms do not change shapes and model predictions are further blurred, with the maximum bed level error increasing from 0.77m to 0.86m (Table 7). Thus, including extra features into the TrajGRU model does not show potential to further improve the bed level prediction accuracy.

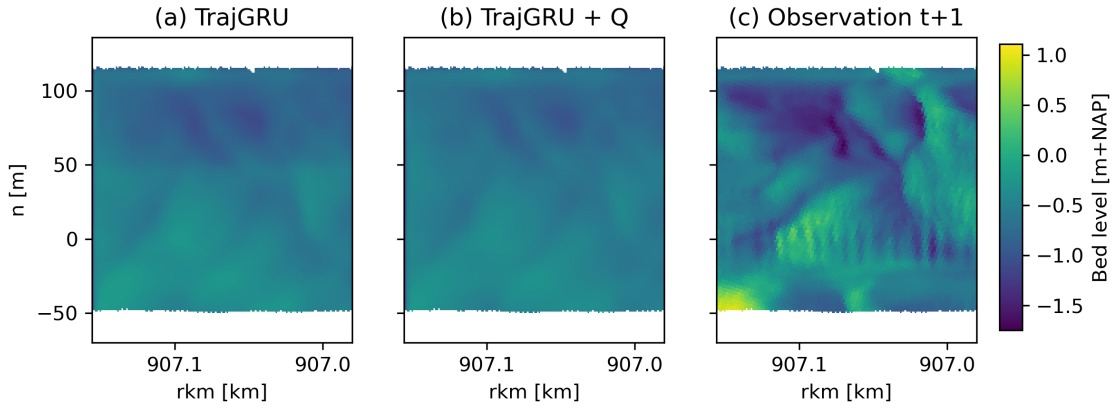
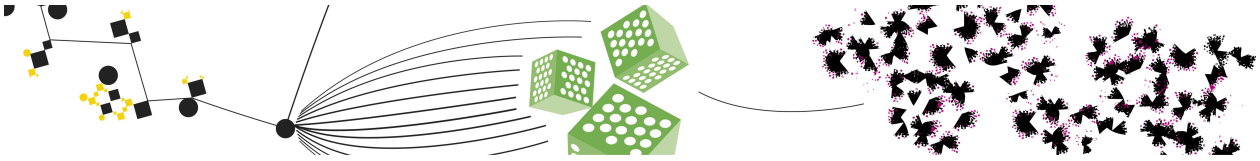


Figure 32: Example bed level prediction of (a) the TrajGRU machine learning model, using the wavelet reconstructed data and RMSE loss function and (b) the same model, but with the discharge, Q, implemented as extra feature. (c) depicts the observed bed level.

Table 7: Performance metric scores for the TrajGRU machine learning model with and without discharge implementation

	TrajGRU	TrajGRU (+ discharge)
MAE [m]	0.77	0.86
MAE [m]	0.28	0.29
RMSE [m]	0.35	0.35

E NCR days 2024 poster



Short-term bed level predictions for the Waal River: a machine learning approach

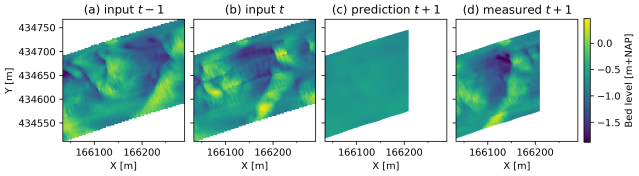
Luuk I. van Laar[a]*, Freek Huthoff[a,b], Thomas Stolp[b], Suzanne J.M.H. Hulscher[a]
[a]University of Twente, Department of Marine and Fluvial Systems, Faculty of Engineering Technology
[b]HKV
*l.i.vanlaar@student.utwente.nl

Problem context

- Extensive dredging is taking place in the Waal River
  - More extreme low-water periods are expected
  - Requires more frequent dredging → demand for efficient dredging strategies.
- Current method:
  - Real-time data on bed levels
  - If bed levels exceed the dredging reference level → dredging required
- Improvement:
  - Short-term bed level prediction
  - Machine learning - TrajGRU

Machine learning model (TrajGRU)

- 2 input bathymetry frames
  - 256x256m
  - Δt = 14days
- Predicts bathymetry 14 days ahead



Objective

"The objective is to improve the accuracy of the TrajGRU bed level predictions for the Waal River"

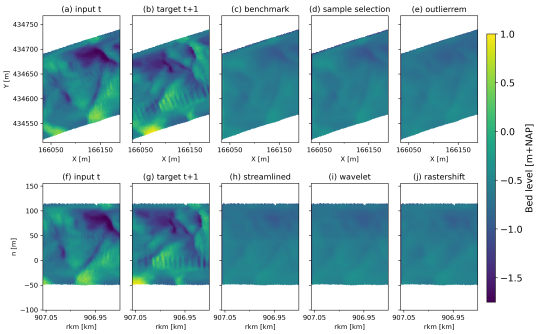
Method

Data preprocessing:
(1.1) Selection data based on discharge (Q<2000m³/s)
(1.2) Outlier removal
(1.3) Coordinate transformation
(1.4) Removal of bedform noise (wavelet)
(1.5) Raster shifting

Loss function adaptation (to be continued):
(2.1) RMSE
(2.2) MSE + SSIM
(2.3) Wasserstein loss

Results

- Remains highly smoothed
- Few observable changes between experiments
- Largely reconstructs input frame
- Few bedform characteristics are captured



Discussion

- Raster shifting VS machine learning
- Machine learning technology is showing some potential in predicting bed levels and progress should be made for practical application. Is more data needed? Or a different modelling approach?
- Possible further iterations TrajGRU:
  - Increase raster size (currently 1x1m)
  - Further down- and upsampling within model structure
  - Include relevant features (discharge, ship movements etc.)