Damova, Mariana; Stankov, Stanko

River Dynamics Forecasting using TELEMAC, Earth Observation and AI

River Dynamics Forecasting using TELEMAC, Earth Observation and AI

Mariana Damova, Stanko Stankov
mariana.damova@mozajka.co, Sofia, Bulgaria
stanko.stankov@mozajka.co, Sofia, Bulgaria
Mozaika, Ltd.
http://www.mozajka.co

Abstract – **The changes in the runoff and alluvial outflow lead to changes in the slope, the depth, meandering, the width of the riverbed and the vegetation. Being able to predict the changed of the river dynamics is of crucial importance for the sustainable monitoring, maintenance and operation of rivers.**

We adopt deep learning architectures pipeline consisting of GAN, CNN and LSTM to actually generate forecasts for river discharge, water level and sediment deposition by using historic satellite data of the meteorological features listed above, and in-situ measurements for water level, discharge and turbidity. To leverage the applicability of the forecasts on the river morphology in integrated models, we calibrate hydrodynamic models using TELEMAC-2D, and we demonstrate how the fusion of a complex EO4AI method and geometry mapping produces a solution for a real user need of being aware of upcoming changes in the navigable channel of the downstream of the Danube. The satellite data are provided by ADAM via the NoR service of ESA.

We provide comparison between the generated hydrodynamic models with real and with forecasted river data, and analyse them. Finally, we demonstrate a visualisation of the forecasted pathway on a GIS component using OpenLayers.

Keywords: hydrodynamic modelling, TELEMAC, deep learning, earth observation, AI, forecast, GIS visualisation, navigable channel.

I. INTRODUCTION

The changes in the runoff and alluvial outflow lead to changes in the slope, the depth, meandering, the width of the riverbed and the vegetation. The bed load and the suspended load can change the morphology of the riverbed as a result of high runoff. This has a direct impact on the determination of the channel in navigable rivers. That is why it is of great importance for assisting the maintenance of the navigable rivers to provide with instruments to predict the modifications in the river morphology that will potentially impact the channel. To address this problem, it is necessary to forecast the sediment deposition amounts and the river runoff and to determine how they will change the river morphology. Predicting sediment deposition potential depends on a variety of meteorological and environmental factors like turbidity, surface reflectance, precipitations, snow cover, soil moisture, vegetation index. Satellite data offer rich variety of datasets, supplying this information.

We adopt deep learning architectures pipeline consisting of Generative adversarial networks (GAN), Convolutional Neural Networks (CNN) and Long Short Term memory (LSTM), further described below, to generate forecasts for river discharge, water level and sediment deposition by using historic satellite data of the meteorological features listed above, and in-situ measurements for water level, discharge and turbidity. To leverage the applicability of the forecasts on the river morphology in integrated models, we calibrate hydrodynamic models using TELEMAC-2D, and we demonstrate how the fusion of a complex EO4AI method, a method that makes use of earth observation data to train AI models, conversely to the usually employed in earth observation methods using AI models in order to better detect objects or phenomena on earth through earth observation – AI4EO[10] - and geometry mapping produces a solution for a real user need, such as being alerted of upcoming changes in the navigable channel of the downstream of the Danube. The structure of the paper is as follows: first we present the adopted method, then we describe the data that have been used for the experiments to demonstrate the method, third we outline the experiments and the experiment results, including GIS (Geographical Information System) visualization of the forecasted navigable channel, finally we discuss related work and conclude.

II. METHOD

We address the problem of river dynamics prediction by forecasting the river discharge and the sediment deposition amounts and then proceed to hydrodynamic modelling with TELEMAC-2D by using the forecast data, the river bathymetry including sediment sizes. (see Figure 1)

[10] Satellite data provide a very rich source of information about the earth. However, they are not perfect. That is why a big segment of the research dedicated to earth observation is concerned with developing AI methods that will improve and maximize the Earth observation using satellite data. This field of research is referred to with AI4EO. In our approach we are using satellite data of meteorological features to train AI methods and obtain predictions for their behavior. This field of research is referred to with EO4AI, and is in its early stages of development.

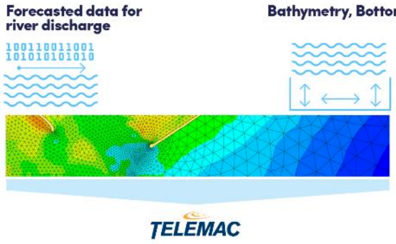


Figure 1. Figure 13 Forecast of river dynamics

For the forecasts we adopt a pipeline of deep learning architectures consisting of GAN, CNN and LSTM by using historic satellite data of meteorological features, that have impact on the river dynamics, such as precipitations and soil moisture, and in-situ measurements for the hydraulic input feature to be forecasted e.g. water level, discharge and turbidity (see Figure 2).

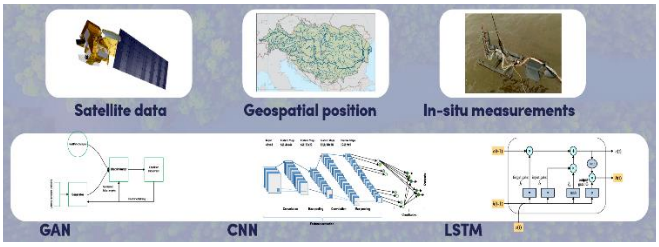


Figure 2. Figure 14 Forecast pipeline of NN

A. Generative adversarial network (GAN)

A generative adversarial network (GAN) [1] is a type of construct in neural network technology that offers a lot of potential in the world of artificial intelligence. It is composed of two neural networks: a generative network and a discriminative network (see Figure 2). The discriminator function compares real data with generated sample data optimizing the model towards reaching a state of no discrepancy. Thus, the one network generates data, e.g. models a transform function that takes a variable and produces another variable following the target distribution and the other network is a discrepancy evaluator that models a discrepancy function that returns the probability of a generated data to be true. The benefits of adopting GANs are that they generate data that looks similar to original data. If a GAN is given an image, then it will generate a new version of the image which looks similar to the original image. Similarly, it can generate different versions of the text, video, audio. However, they are much harder to train since it is needed to provide different types of data continuously to check if it works accurately or not.

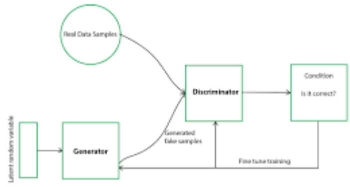


Figure 3. Figure 15 Generative adversarial network architecture

We use GAN to address one specificity of satellite data to be inconsistent, e.g. they do not produce harmonized timeseries, for example daily data, as no data as being provided for some days. But to run neural network architecture to generate forecasts using satellite data, we need consistent timeseries. So, GAN is the first step of our pipeline. It harmonizes the timeseries of the satellite data by generating the missing values in the timeseries.

B. Convolutional neural networks (CNN)

CNN [2] is a type of artificial neural network used for descriptive and generative tasks very often adopted in computer vision. The term “convolutional” means mathematical function derived by integration from two distinct functions. It includes rolling different elements together into a coherent whole by multiplying them. CNNs are made from neurons with trainable weights where each neuron receives input data and takes a weighted sum over them and passes it through an activation function returning the output (see Figure 4.

The convolution applies a kernel [19] over the input data, performing elementwise multiplication with the part of the input that is currently on, after that it sums up the result into a single value (see Figure 4).

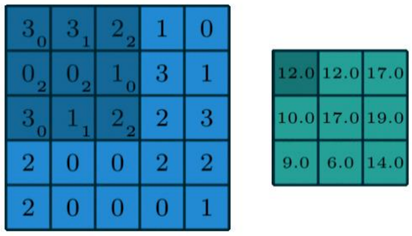


Figure 4. Figure 16: Convolution

After the convolutional layer in CNNs a pooling layer is applied [4]. This layer is responsible for reducing the size of the convoluted feature. It is used to decrease the computation power. There are two types of pooling layers – max pooling and average pooling. They are used to extract the average values or the max values from a kernel (see Figure 5).

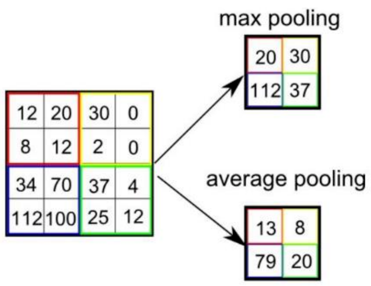


Figure 5. Figure 17: Pooling layer

CNN architectures are used for generating forecasts (see Figure 6)

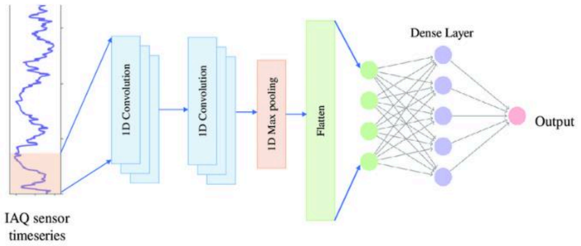


Figure 6. Figure 18: Convolutional neural network used for forecasting

The benefits of adopting CNNs are mainly the high accuracy in tasks that require image recognition and the weight sharing. The disadvantages of adopting CNNs are the need of large datasets to obtain good performance, and long training time, that typically requires a specialized hardware (GPU) to speed up the training process.

CNNs are used as a first step in order to optimally consume the geopositioned input data, as they are effective with geospatial data.

C. *Long Short Term Memory (LSTM)*

LSTM [5] is a type of Recurrent Neural Networks (RNN) model [6] to remember each information throughout time, which is very helpful in any time series predictor. Recurrent Neural Networks (RNN) are designed to recognize sequence patterns and stock markets [7] (see Figure 7).

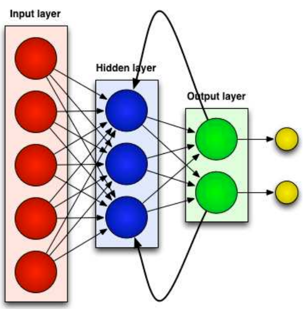


Figure 7. Figure 19: RNN architecture

Instead of neurons, LSTM networks have memory blocks that are connected through layers (see Figure 7). A block has components that make it smarter than a classical neuron and a memory for recent sequences. A block contains gates that manage the block's state and output. A block operates upon an input sequence and each gate within a block uses the sigmoid activation units to control whether they are triggered or not, making the change of state and addition of information flowing through the block conditional. There are three types of gates within a unit: 1) Forget Gate: conditionally decides what information to throw away from the block. 2) Input Gate: conditionally decides which values from the input to update the memory state. 3) Output Gate: conditionally decides what to output based on input and the memory of the block.

The advantages of using LSTMs can be summarized in the following points:

- They can remember each information throughout time, which is very useful for time series predictions
- Native Support for Sequences. LSTMs are a type of recurrent network, and as such are designed to take sequence data as input, unlike other models where lag observations must be presented as input features.
- Multivariate Inputs. LSTMs directly support multiple parallel input sequences for multivariate inputs, unlike other models where multivariate inputs are presented in a flat structure.
- Vector Output. Like other neural networks, LSTMs are able to map input data directly to an output vector that may represent multiple output time steps.

D. *ConvLSTM*

In the adopted architecture, the forecast model of our solution includes both CNNs and LSTMs [8]. Firstly, convolutional layers are used for feature extraction on the input data and are combined with a LSTM layer allowing the architecture to support sequence prediction (see Figure 7).

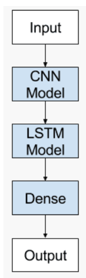


Figure 8. Figure 20: ConvLSTM architecture

ConvLSTMs are used when the input is 2D structure such as image, or 1D as word, sentence or some other sequential input data with the task for classification or forecast. The advantages of using ConvLSTM architectures are that they achieve high accuracy, they operate with input data in 1D, 2D or 3D and since they contain both CNN and LSTM layers they can look back and forward on 1D.

E. *The open TELEMAC system*

TELEMAC-2D is a well-known and established hydrodynamic model solving the shallow water equations. It is used to simulate free-surface flows in two dimensions of horizontal space. At each point of the mesh, the program calculates the depth of water and the two velocity components. It can perform simulations in transient and permanent conditions. The algorithms used by TELEMAC-2D are extremely efficient and are constantly being improved in order to incorporate the latest developments made in the field of numerical computations. TELEMAC-2D meet the specific requirements of each model: specification of initial conditions or complex boundary conditions, links up with other modelling systems, allow the introduction of new functions.

As shown on Figure 9 we combine the physical model of the river bathymetry and sediments with the forecasted water data for river discharge and water level, employing the Composite Modelling (CM) method that is defined as the integrated and balanced use of physical and numerical models [9]. It is important to emphasize the novelty of our approach to make use of forecast hydraulic input data instead of using hydrological formulas in TELEMAC to obtain the river dynamics prediction.

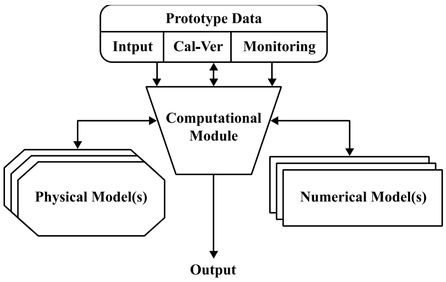


Figure 9. Figure 21 Model integration leading to an integrated model11

III. DATA

Predicting discharge, water level and sediment deposition depends on a variety of meteorological and environmental factors such as precipitations, snow cover, soil moisture, vegetation index, turbidity, surface reflectance. Satellite data offer rich variety of datasets, supplying this information. That is why we make use of satellite data to feed our forecast models with this kind of information. In addition, we make use of in-situ measurements for discharge, water level and turbidity for the forecast models, as well as gain size and bathymetry for the hydrodynamic modelling.

The data used for the experiments and the forecasting prototype as presented below.

A. Satellite data

The satellite data are provided by ADAM (http://adamplatform.eu). ADAM allows accessing a large variety of multi-year global geospatial collections enabling data discovery, visualisation, combination, processing and download. It permits to exploit data from global to local scale. Table I shows the satellite datasets that are provided and made use of in our experiments. The data collections have been selected in order to guarantee:

- - full coverage of the project spatial domain
- - the best available spatial resolution per time
- - data availability for at least 10 years.

Table I Satellite data collections

Meteorological Feature	Satellite Data Collection
Temperature	MODIS land surface temperature day
Soil Moisture	SMOS
	CCI Soil Moisture
precipitation	imerg liquid precipitation daily
	imerg liquid precipitation 30 min
	imerg solid precipitation daily
	imerg solid precipitation 30 min
Snow cover	MODIS Snow cover
Solar irradiance	MSG Downwelling Shortwave Surface Flux
vegetation index	MODIS NDVI
	Sentinel 2 NDVI
Copernicus water turbidity	MSI based 300m (CLMS)
	OLCI based 300m (CLMS)
MODIS surface reflectance	MODIS based 250m

The core of ADAM is a Data Access System (DAS), a software module that manages a large variety of geospatial information that feature different data format, geographic / geometric and time resolution. It allows accessing, visualising, sub-setting, combining, processing, downloading all data sources simultaneously. The DAS exposes OGC Open Search (CSW-compliant) and Web Coverage Service (WCS 2.x) interfaces that allow discovering available collections and subset them in any dimension with a single query.

B. In-situ measurements

As the experimental setting of our work was the Danube River, the in-situ measurements were taken from measurement points on the Bulgarian segment of the Danube. They are outlined below.

Historic daily measurements for water discharge in cm^3/s, water levels in m and water temperature in ^oC are provided for the period 2015-2020 by the National Agency for Exploration and Maintenance of the Danube River from the hydrometric stations Lom (km 743,300), Svistov (km 554,300), and Silistra (km 375,500). The current daily measurements from 2021 of the same categories are being collected from the site of the Agency on a daily basis.

[11] Gerritsen, H., Sutherland, J., Deigaard, R., Sumer, B.M., Fortes, J, Sierra, J-P and Prepernau, U, (2009). Guidelines for Composite Modelling of the Interactions Between Beaches and Structures. Final Report, September, 66 pages, (HYDRALAB-III Deliverable JRA1.4).

The turbidity and grain size historic data daily sampling from the same hydrometric stations are provided for the period 2015-2020 inclusively by The National Institute for Meteorology and Hydrology. The turbidity [g/l] is being determined via laboratory analysis in The National Institute for Meteorology and Hydrology.

For the purposes of the geometry modelling of the Danube River, to address the use case for predicting of the navigable channel on the Danube River, the bathymetry of the Bulgarian segment of Danube River has been provided by the National Agency for Exploration and Maintenance of the Danube river.

IV. EXPERIMENTS AND RESULTS

Experiments have been carried out with different ConvLSTM architectures. The best performing one turned out to be the one using two convolutional layers. We show the results of the generation of missing data with GAN, and the ConvLSTM architecture with two convolutional layers. Subsequently, we show the results of the river dynamics prediction with the adopted method using TELEMAC 2D.

A. Missing data generation with GAN

The missing data generation with GAN was carried out with daily satellite data timeseries from 2014-2019 and the tests were the daily timeseries from 2020. The results are shown on Table II. It is evident that the precision of the generated missing data is very satisfactory with MinMae score close to 0. It is important to point out that the spatial resolution for the satellite data in the reported experiments was 1.1 km.

Table II Missing data generation results

Meteolologial feature	Satellite dataset name	MinMae
Soil moisture	ESACCI-SOILMOISTURE-L3S-SSMV-COMBINED_4326_025	0
Liquid precipitations	IMERG_DAY_LIQUID_SCALED_4326_01	0,92
Snow cover	MODIS_SNOW_4326_001	0,73
NDVI	MOD13_NDVI_4326_005	0,05
Solid precipitations	IMERG_DAY_ICE_SCALED_4326_01	0,07
Reflectance	MODIS based ISMoSeDe_reflectance	0,04

B. Forecast with ConvLSTM

The forecasting experiments were carried out with the harmonized satellite data, as described in the previous subsection, and with in-situ measurements for discharge, water level and turbidity as per the description in section III. Table III below shows the annual average deviation in the results for discharge, water level and turbidity for the three hydrometric posts on the Danube with a model calibrated to generate forecast for 7 days ahead. The experiments are carried out with daily historic data from 2014-2019 and the test is performed on year 2020.

Table III Forecast performance

Feature	Deviation		
	Lom	Svishtov	Silistra
Turbidity (NTU)	0,01	0,02	0,04
Discharge (cm3/sec)	224,70	181,70	86,90
Water level (cm)	26,74	20,40	18,83

Figure 10 shows the performance of the discharge forecast compared to real measurements for 30 days ahead for the three hydrometric stations on the Danube.

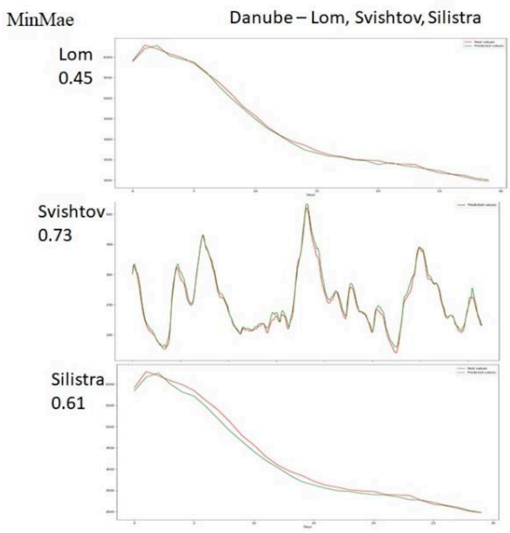


Figure 10. Figure 10 Forecast performance for 30 days ahead

Here we also notice that the performance is satisfactory.

C. River dynamics prediction with TELEMAC D

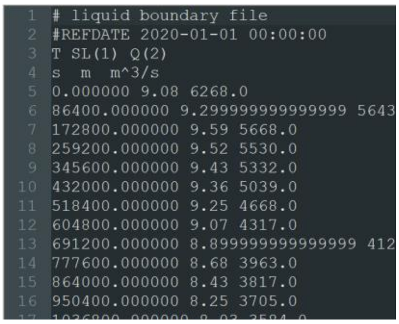
TELEMAC simulation is performed for the critical areas around the hydrometric points Lom, Svishtov and Silistra. We have used the bathymetry of the critical areas to draw the mesh, in-situ measurements for substrate and the forecasted data for discharge and water level, as turbidity was not relevant for the type of river segments of interest.

For the simulation a time step value of 10 and Nikuradse for the law of bottom friction. The liquid boundary is represented with three parameters – time, free surface and discharge measured in s, m and m^3/s. The temporal resolution of the simulation is one day, represented in seconds (0-86400 seconds). We introduce the forecasted value for discharge and calculate the free surface from the forecasted water levels using the following formula:

$$Fs = wl / 100 + K$$

where wl is the forecasted water level and K is the station elevation.

The created output file is shown on Figure 11



```
# liquid boundary file
#REFDATE 2020-01-01 00:00:00
T SL(1) Q(2)
s  m  m^3/s
0.000000 9.08 6268.0
86400.000000 9.299999999999999 5643
172800.000000 9.59 5668.0
259200.000000 9.52 5530.0
345600.000000 9.43 5332.0
432000.000000 9.36 5039.0
518400.000000 9.25 4668.0
604800.000000 9.07 4317.0
691200.000000 8.899999999999999 412
777600.000000 8.68 3963.0
864000.000000 8.43 3817.0
950400.000000 8.25 3705.0
```

Figure 11. Figure 11 Liquid boundary file with forecasted data

The generated 6 features: riverbed, velocity U, velocity V, friction velocity, surface elevation, and water depth are produced. Figure 12 shows the predicted riverbed for the critical area around Svishtov, and Figure 13 shows the comparison between the real and predicted depth for January 2020 for the critical area around Svishtov for 1, 10, 20 and 30 days ahead.

The precision in the prediction of the depths is evident from the similarity of the images. It is important to point out that we do not observe deterioration of the performance results as the prediction period increases to 30 days.

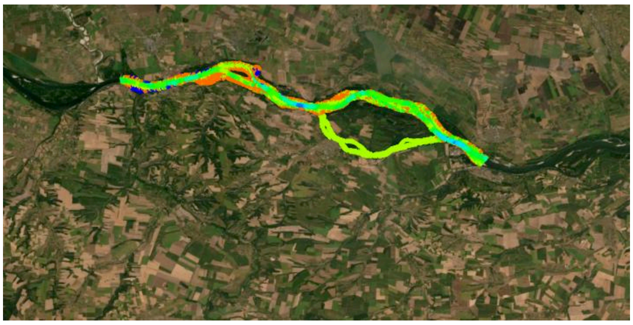


Figure 12. Figure 12 Predicted riverbed of a critical area around Svishtov

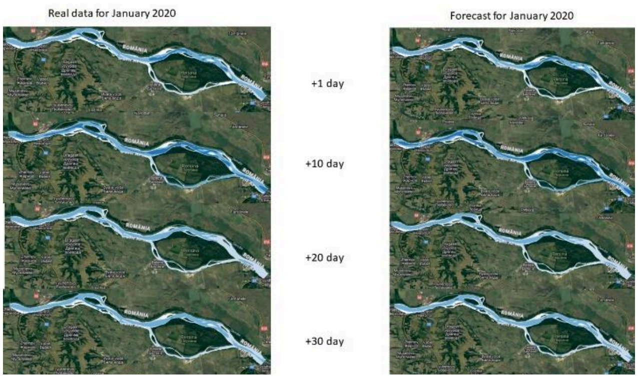


Figure 13. Figure 13 Comparison of real vs predicted depth of a critical area around Svishtov

Based on the forecasted data as TELEMAC output we proceeded to the generation of the forecasted navigable channel by identifying the middle path on the XY axis as shown on Figure 14.

Figure 14. Figure 14 Visualization of the forecasted fairway on the critical area of Svishtov

V. Related work

Hydrodynamic modelling [9] forms the basis for many modelling studies, whether sediment transport, morphology, waves, water quality and / or ecological changes are being investigated. Research is being carried out to improve the representation of tides, waves, currents, and surge in coastal waters. A variety of coastal models are available, and the modelling techniques have become sufficiently mature [10]. Composite Modelling (CM) is defined as the integrated and balanced use of physical and numerical models [11]. Our approach blends the CM method and demonstrates how it can be applied to hydrodynamic modelling.

River dynamics has been observed and forecasted with various mathematical models. They are gaining popularity for solving a wide range of natural fluid mechanical problems. When it comes to free-flow currents and sediment transport processes in open channels, single-dimensional (1D) and two-dimensional (2D) digital models are widespread. To the best of our knowledge our approach is the first to attempt to forecast river dynamics by inserting forecasts of hydraulic input into TELEMAC. Moreover, the forecasts of the hydraulic input are produced by means of using historic satellite data of meteorological features and in-situ measurements from given hydrometric stations. This constitutes a cutting-edge novel method in the field of hydrodynamic modelling.

VI. Conclusion and future work

We presented a method for predicting river dynamics using TELEMAC, AI and Earth observation. It consists of producing TELEMAC simulation by using forecasted measurements for discharge and water level. For the forecast we adopt neural networks pipelines and historic satellite data with meteorological information and in-situ measurements.

The satellite data input to the neural network is harmonized in a first processing step also using neural networks - GAN. We succeed to obtain very good forecasting results that are consequently blended in the TELEMAC simulations, based on which we are capable to derive hands-on determination and visualization of forecasted fairway. We have demonstrated how the fusion of a complex EO4AI method and geometry mapping produces a solution for a real user need of being aware of upcoming changes in the river fairway of the downstream of the Danube. To the best of our knowledge our approach is the first to make use of forecast hydraulic input data to produce a hydrodynamic simulation with TELEMAC. Our results demonstrate the viability and robustness of our method. Our plans are to integrate the method in our e-Infrastructure for monitoring dams and rivers for sustainable development – ISME-HYDRO (http://isme-hydro.com).

ACKNOWLEDGEMENT

This work has been carried out within ESA Contract No 4000133836/21/NL/SC. For the hydrodynamic modelling we received training and consulting from IMDC for which we thank Alexander Breugem, Theofano Koutrouveli, Boudewijn Decrop.

REFERENCES

[1] https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/
[2] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
[3] https://towardsdatascience.com/intuitively-understanding-convolutions-for-deep-learning-1f6f42faee1
[4] https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/
[5] https://wiki.pathmind.com/lstm
[6] https://stackabuse.com/time-series-analysis-with-lstm-using-pythons-keras-library/
[7] https://towardsdatascience.com/understanding-recurrent-neural-networks-the-prefered-neural-network-for-time-series-data-7d856c21b759
[8] https://machinelearningmastery.com/cnn-long-short-term-memory-networks/
[9] http://www.coastalwiki.org/wiki/Modelling_coastal_hydrodynamics
[10] Chau, KW Modelling for coastal hydraulics and engineering, Taylor and Francis Ltd., 231 pp., 2010.
[11] Gerritsen, H., Sutherland, J., Deigaard, R., Sumer, B.M., Fortes, J, Sierra, J-P and Prepernau, U, (2009). Guidelines for Composite Modelling of the Interactions Between Beaches and Structures. Final Report, September, 66 pages, (HYDRALAB-III Deliverable JRA1.4).
[12] Barthel, V and E.R. Funke. Hybrid modelling as applied to hydrodynamic research and testing. In: Recent Advances in Hydraulic Physical Modelling, R. Martins (Ed), NATO ASI Series E: Applied Sciences – vol. 165, Kluwer Academic Publishers, 1989.
[13] Oumeraci, H. Strengths and Limitations of Physical Modelling in Coastal Engineering - Synergy Effects with Numerical Modelling and Field Measurements. Proc. of the Hydralab -Workshop on Experimental Research and Synergy Effects with Mathematical Models, Eds.: K.-U. Evers, J. Grune, A. van Os, Coastal Research Centre, Hannover, 1999.
[14] https://www.mikepoweredbydhi.com/download/mike-2016/mike-11?ref=%7B181C63FF-2342-4C41-9F84-F93884595EF3%7D
[15] https://webspace.clarkson.edu/~wwu/CCHE1D.html
[16] https://www.usbr.gov/tsc/techreferences/computer%20software/models/srh1d/index.html