engineering proceedings

MDPI

Proceeding Paper

Towards an Automatic Tool for Resilient Waterway Transport: The Case of the Italian Po River †

Maria Luisa Villani [1,*], Ebrahim Ehsanfar [2], Sohith Dhavaleswarapu [2], Alberto Agnetti [3], Luca Crose [3], Giancarlo Focherini [3] and Sonia Giovinazzi [1]

1 ENEA-Italian National Agency for New Technologies, Energy and Sustainable Economic Development, 00123 Rome, Italy; sonia.giovinazzi@enea.it
2 Fraunhofer Institute for Material Flow and Logistics, 44227 Dortmund, Germany; ebrahim.ehsanfar@iml.fraunhofer.de (E.E.); sohith.dhavaleswarapu@iml.fraunhofer.de (S.D.)
3 Agenzia Interregionale per il Fiume Po, AIPO, 43121 Parma, Italy; alberto.agnetti@agenziapo.it (A.A.); luca.crose@agenziapo.it (L.C.); giancarlo.focherini@agenziapo.it (G.F.)
* Correspondence: marialuisa.villani@enea.it
† Presented at the 10th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 15–17 July 2024.

Abstract: Improved navigability can enhance inland waterway transportation efficiency, contributing to synchro-modal logistics and promoting sustainable development in regions that can benefit from the presence of considerable waterways. Modern technological solutions, such as digital twins in corridor management systems, must integrate functions of navigability forecasts that provide timely and reliable information for safe trip planning. This information needs to account for the type of vessel and for the environmental and geomorphological characteristics of each navigation trait. This paper presents a case study, within the EU project CRISTAL, focusing on the Italian Po River, of which the navigability forecast requirements of a digital twin are illustrated. Preliminary results to deliver navigability risk information were obtained. In particular, the statistical correlation of water discharge and water depth, computed from historical data, suggested that efficient forecast models for navigability risk, given some water discharge forecasts, could be built. To this aim, the LSTM (long-short-term-memory) technique was used on the same data to provide models linking water discharge and water depth predictions. Future work involves further testing these models with updated real data and integrating outcomes with climatic and infrastructure management information to enhance the accuracy of the risk information.

Keywords: resilience; navigability forecast; deep learning; waterways; time series analysis

Citation: Villani, M.L.; Ehsanfar, E.; Dhavaleswarapu, S.; Agnetti, A.; Crose, L.; Focherini, G.; Giovinazzi, S. Towards an Automatic Tool for Resilient Waterway Transport: The Case of the Italian Po River. *Eng. Proc.* **2024**, *68*, 64. https://doi.org/10.3390/engproc2024068064

Academic Editors: Olga Valenzuela, Fernando Rojas, Luis Javier Herrera, Hector Pomares and Ignacio Rojas

Published: 4 September 2024

1. Introduction

Inland waterways are essential for efficient, low-cost transportation; hence, they significantly contribute to economic growth and environmentally sustainable logistics. However, the navigability of waterways highly impacts the efficiency of the operations. Technological solutions, such as digital twins, can effectively support the monitoring and management of waterway operations as well as their resilience; digital twins need to address requirements of environmental, infrastructure, and social dimensions.

As far as the environmental perspective is concerned, rivers usually do not have a regular water depth, and shallow areas can impede the passage of large vessels. The draught (the depth of a vessel's hull below the waterline) of the various vessels must be considered when planning the trips, as insufficient water depth along the whole river path could lead to emergencies. Seasonal changes in water depth due to factors like rainfall or drought also affect river navigability. Low water depth during dry seasons may lead to restrictions or even the complete closure of certain river sections for navigation.

From the infrastructure management perspective, poorly maintained channels, locks, and dams can hinder smooth navigation, and the absence of proper dredging to maintain navigable depths can result in sedimentation and in the formation of sand dunes, reducing water depth. More generally, natural obstacles like sandbars, rocks, and fallen trees, which create hazards for vessels, need to be taken care of in infrastructure maintenance and resilience planning as these are hard to prevent, especially in case of extreme events. Thus, for a given case study, data analysis for the purpose of navigability forecasting requires expert knowledge that combines hydrology, climate, and infrastructure maintenance influencing factors.

Finally, from the social perspective, as different jurisdictions may have varying rules and restrictions, this creates challenges for the generalizability of the technical solution and/or adaptation to different contexts.

The European Commission funded the research project CRISTAL, "*Climate resilient and environmentally sustainable transport infrastructure with a focus on inland waterways*", focuses on the development of inland waterway transport (IWT) and its infrastructure with the vision of increasing the operability, sustainability, and resilience of IWT.

The crucial aim of the CRISTAL project is to create innovative technology for the monitoring and digitization of river transport as part of multimodal chains. The resulting solutions will be combined into a collaborative open data platform for synchro-modal resource planning and operation preparation and will include a comprehensive real-time monitoring and prediction system for water depth and hydrological conditions, including and integrating the current best practice requirements and standards, for instance, Digital Inland Waterway Area, DINA [1], and by the European Federated Network of Information exchange in Logistic, FENIX [2]. The project involves partners from nine countries, namely, Poland, Germany, Italy, Belgium, the Czech Republic, Hungary, Greece, France and UK. The case studies will cover Poland (Vistula and Odra), Italy (Po), and France (Mosele, Seine).

This paper presents the case study related to the Italian Po River, of which the navigability forecast requirements of a digital twin are illustrated in Section 2. Preliminary results to deliver navigability risk information were obtained by using Long-Short-Term-Memory (LSTM) forecast models. The source datasets, the methods, and the results are presented in Section 3. Plans for future work are given in Section 4.

2. The Free-Flowing Waterway Navigability Problem for Po River

Inland navigation in Northern Italy develops within and around the natural course of the Po River that crosses different Italian regions, having a primary role within the "Padano–Veneto Waterway System". The inland waterway mode of transport provided it has enhanced reliability (i.e., being able to continuously monitor and predict navigability issues) and resilience to extreme events, has a great potential to enhance the logistics in this area of high economic dynamism by reducing the frequent congestion of motorways and railways. Also, the inland waterway mode of transport, provided it has environmental sustainability, can make a significant difference in reducing harmful gas emissions from the logistics sector.

As far as reliability and resilience are concerned, in the Po River basin a cutting-edge technology is currently in operation to provide water forecast information via a multi-model approach. Such technology, referred to as the *Flood and Drought Early Warning System of the Po River*, EWA system, is fed by real-time data from monitoring networks. The sensor data is coupled with radar data and meteorological scenarios, and it is processed via different numerical models to provide forecasts during drought and flood events (Figure 1). The system is shared and maintained with all regional and national agencies involved in the flood and drought risk management in the Po River.

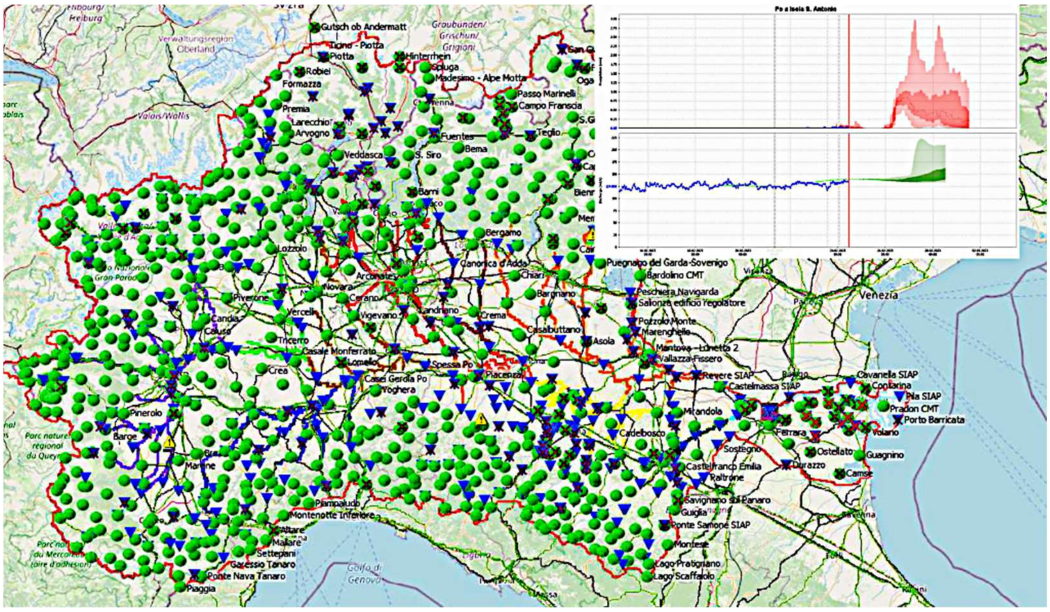


Figure 1. Flood and Drought Early Warning System of the Po River basin. The figure highlights the monitoring points and a type of chart provided by the tool by combining the source data.

Critical sections of the Po River, i.e., the locations where the accumulation of natural material in the riverbed can affect navigability, are known and constantly monitored, mainly during low regimes. The critical sections as shown in Figure 2.

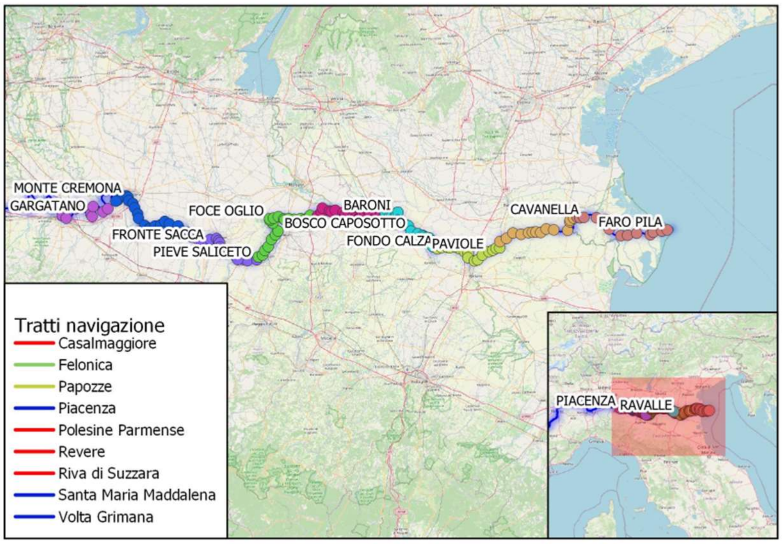


Figure 2. Monitored critical sections in the main Po River where different colors represent the different stretches.

The aim of a digital twin supporting the Po River logistic system is to elaborate the following: (i) the specific navigability level in each one of the several critical sections within each stretch (i.e., from one docking to the following one) and (ii) the overall navigability level in each stretch (Figure 2 shows the stretches, and Figure 3 provides a representation of the result as a navigability risk matrix). Navigability and alert information resulting from a digital twin will be related to different classes of vessels and different specific uses (e.g., transport, recreational uses, etc.) and loads.

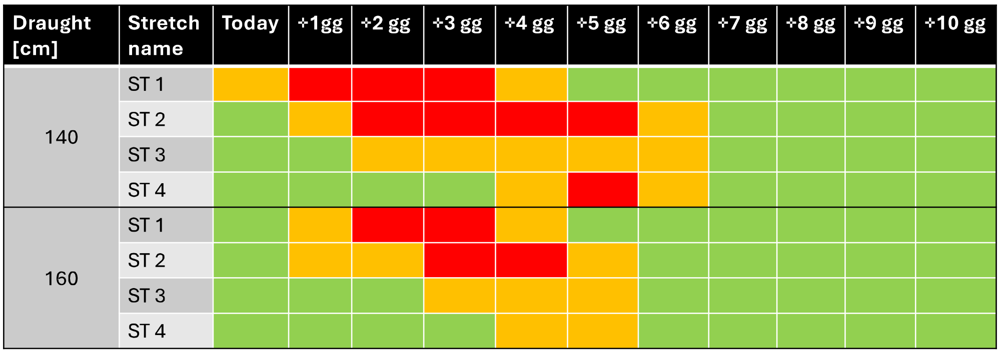


Figure 3. Draft of a navigability risk forecast matrix. A navigability risk level (red = high risk, yellow = medium risk, and green = low risk), based on the navigability forecast, will be provided for each stretch and for different vessel classes (according to the draught) for the next 10 days (gg = day).

3. Datasets and Methods

This section presents the datasets and a concise analysis, followed by an explanation of the methodology. Conducting data analysis, we focused on two pivotal time series datasets—water depth and water discharge—gathered from strategically positioned sensor stations spanning multiple years. The ensuing discussion outlines the intricate analytical procedures and frameworks applied to extract meaningful insights into the hydrological dynamics under investigation. The water depth time series delineates variations in depth, enabling an assessment of riverbed topography changes at specified sample stations. Concurrently, the water discharge time series, derived from additional river stations, provides insights into flow dynamics influenced by precipitation, snowmelt, and other hydrological factors. Employing descriptive statistics and visualization techniques, the objective is to uncover patterns and correlations within these datasets, contributing to a nuanced forecast of river navigability. The analysis is rooted in meticulously collected daily measurements from two distinct critical section stations, each representing unique segments. The dataset spans from 1 January 1988 to 12 May 2022, comprising 12,294 and 12,255 records for the respective stations. Recorded at daily intervals, the time series data underwent careful handling of missing values using the forward fill method to ensure data consistency. The time series of water discharge and water depth are represented Figure 4 and the main indicators of a descriptive analysis are reported in Table 1.

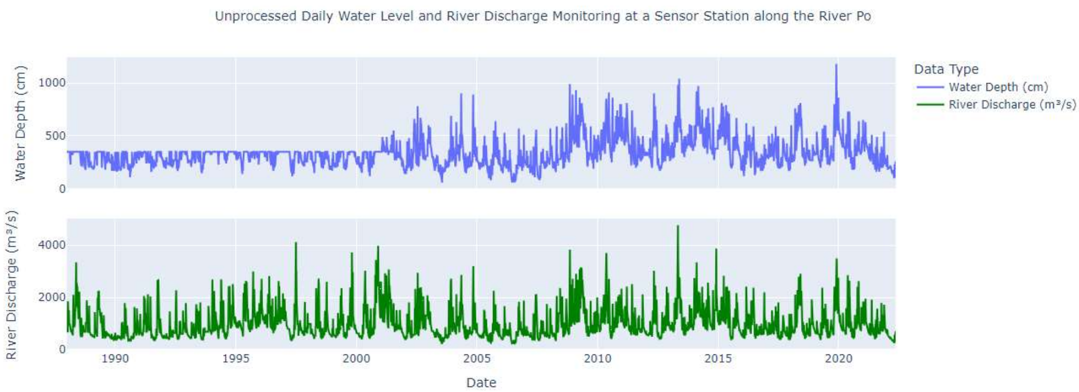


Figure 4. Surveyed daily water depth collected at each critical section and river discharge data recorded at each monitoring section of the river Po River derived from the monitoring network. Data source: www.agenziapo.it.

Table 1. Descriptive analysis of water depth and water discharge rates at a station.

Index	Depth (cm)	Discharge (m^3/s)
Records Count	12551	12551
Mean	333.32	950.67
Std	125.41	519.44
Minimum Value	60.00	221.50
25th Percentile	250.00	587.80
Median	330.00	796.61
75th Percentile	360.00	1157.08
Maximum Value	1180.00	4770.00

Observations revealed that the average depth and discharge spanning the years were 333 cm and 950 m^3/s, respectively, at a station. The correlation between depth and discharge attributes was notably high at 0.73, indicating a strong interdependence in hydrological relations (see Figures 5 and 6).

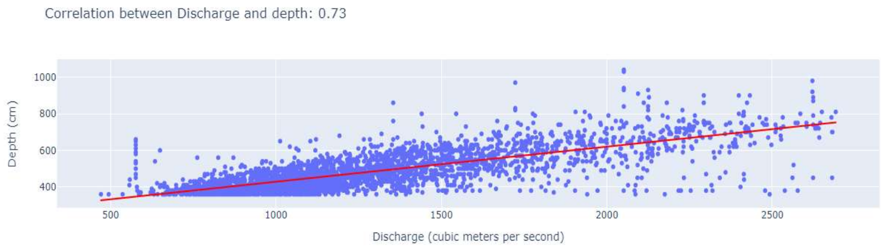


Figure 5. The correlation between daily water depth and river discharge in the Po River. The red trendline highlights the relationship with a computed Pearson correlation coefficient.

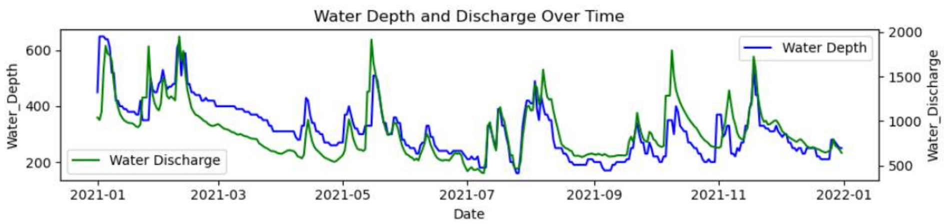


Figure 6. Year-long trend of water discharge and water depth data at a station.

The stationarity of the time series data was rigorously assessed through the Augmented Dickey–Fuller (ADF) test.

The test results, yielding p-values of 1.32×10^{-21} for depth and 9.45×10^{-22} for discharge rate, unequivocally substantiate the stationarity behavior within the recorded values, indicative of a stable and consistent hydrological pattern over time (see Figure 7).

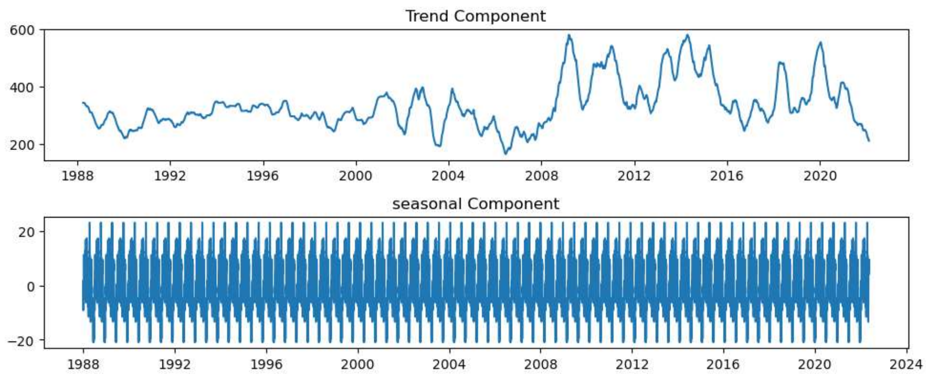
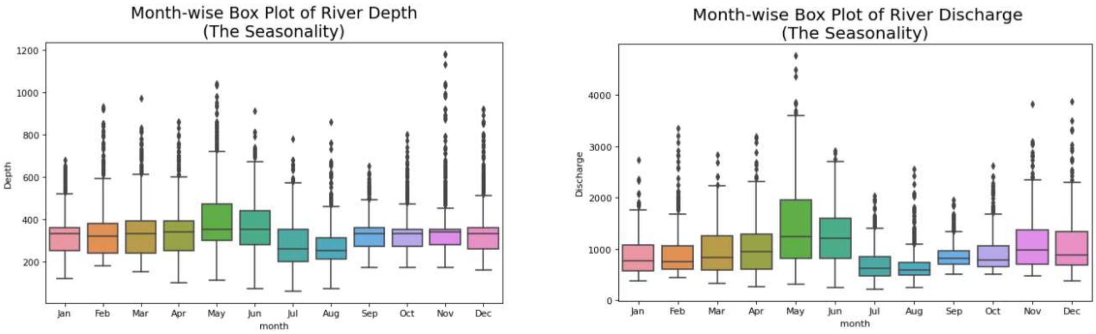


Figure 7. Water depth time series decomposition. Trend and seasonal components revealed through additive seasonal decomposition offer insights into temporal dynamics.

A meticulous seasonal decomposition analysis revealed a distinct 6-month pattern in river depths (see Figure 8). The trend indicated stable depths with a marginal increase. Notably, elevations in river depth and discharge were observed in May, June, November, and December, emphasizing a discernible seasonal trend during these months.



(**a**) (**b**)

Figure 8. Monthly Box Plots (2021) illustrating the distribution and variability of river depth (**a**) and discharge (**b**) at the station.

Exploring the dynamics of water depth and discharge aimed to uncover patterns and trends. This foundational analysis lays the groundwork for designing machine learning models intended to predict navigability, as described in the subsequent subsections.

3.1. Basic Probabilistic Method for Navigability Risks

A first approach used to evaluate the navigability of the Po River was provided using a statistical method. The time series of historical daily survey of water depth, collected for more than 40 yrs, coupled with discharge data available at each monitoring station, can be used to define a good estimation of probability along several critical sections on the Po River. That probability, joined with the predicted discharge data provided by the EWA system, based on hydrological and hydraulic models (cfr. Section 2), provides a basic method to compute the probability of navigability for each critical section for the next 10–15 days.

Figure 9 shows examples of navigability risks for the vessel classes (based on draught) given the discharge classes. The probability was computed based on the percentage of occurrences of the event *water depth* > *minimum draught* in the historical data. The navigability risk of a stretch is provided as the worst navigability risk value among those obtained for each critical section of the stretch. On the one side, this approach uses a huge amount of data and can define a good statistical correlation between discharges and critical sections for water depth in terms of overall performance.

Draught [cm]	Discharge [mc/s]																						
	50	100	150	200	250	300	350	400	450	500	600	700	800	900	1000	1500	2000	3000	4000	5000	8000	9000	10,000
140	0%	0%	3%	15%	36%	56%	77%	83%	87%	91%	95%	97%	98%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%
160	0%	0%	0%	13%	35%	54%	76%	82%	86%	90%	95%	97%	98%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%
180	0%	0%	0%	10%	33%	53%	76%	82%	86%	90%	95%	97%	98%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%
200	0%	0%	0%	5%	29%	50%	74%	81%	85%	90%	94%	97%	98%	99%	100%	100%	100%	100%	100%	100%	100%	100%	100%
220	0%	0%	0%	0%	22%	45%	72%	79%	84%	89%	94%	96%	98%	99%	99%	100%	100%	100%	100%	100%	100%	100%	100%
240	0%	0%	0%	0%	13%	39%	68%	76%	82%	87%	93%	96%	98%	99%	99%	100%	100%	100%	100%	100%	100%	100%	100%
260	0%	0%	0%	0%	0%	30%	64%	73%	79%	85%	92%	95%	97%	99%	99%	100%	100%	100%	100%	100%	100%	100%	100%

Figure 9. Discharge forecast from the probabilistic processing of DEWS/FEWS Po early warning systems data.

On the other side, this statistic does not consider morphological variations at each critical point due to artificial (dredging) or natural (floods) actions. To reduce the uncertainty related to those factors, methods that evaluate the daily probability based on the values of the previous days should be elaborated.

3.2. Deep Learning Method for Water Depth Predictions

Deep learning methods are being experimented with to compare or integrate the risk results obtained with the previous method. In particular, the first objective of the workplan was to generate a predictor for each critical point of each stretch of the river. Thus, for each vessel class, given a forecast of the water discharge of the next 10–15 days at each point, the navigability of a stretch would be the worst prediction of the water depth among the critical section results of that stretch.

The Long-Short-Term (LSTM) neural network [3] has been selected to generate the models. Indeed, it is well known that the LSTM is effective for modeling and predicting sequences, given its ability to retain information over extended periods. Other more recent methods, like Transformer models [4], are also employed for numerical time series analysis, but as they are especially effective in handling sequential data with complex dependencies (e.g., many variables or global dependencies), they were not the first choice for these experiments. Furthermore, recent works such as [5,6] have demonstrated that for river water level prediction, LSTM models trained on water depth measurements and time series alone may outperform other artificial neural network architectures that correlate weather data (e.g., temperature and humidity) with water level observations. This may be because the river's water discharge prediction at some locations is the result of a much more complex analysis that also considers non-local hydrological and climatic aspects. Figure 10 illustrates the problem statement of the LSTM method.

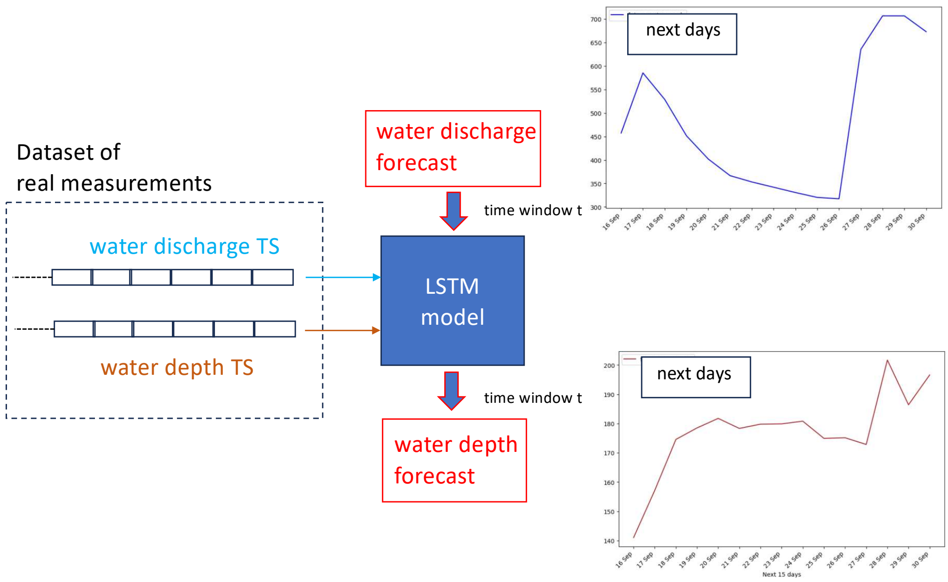


Figure 10. Problem statement representation of the deep learning method.

From the water depth forecast series, one can easily obtain the prediction of navigability for each type of vessel, described in Section 2. However, the confidence level of such prediction should be carefully analyzed to avoid both false positive risks (e.g., a vessel is not allowed to move, but then the navigability conditions result to be safe) and false negative risks (e.g., a vessel starts navigation in unsafe conditions).

The research questions posed to the ML-based forecast experiments are as follows:

[RQ1] How does the LSTM navigability predictor perform on the available real data?

[RQ2] What is the accuracy of the navigability results of the best predictor built?

[RQ3] How generalizable is a predictor to other critical sections or stretches of the river?

The TensorFlow Python implementation of LSTM has been used to build the code for training and testing the models. As it is common practice for this type of software development, the best LSTM configuration has been decided empirically. The development process of the models is an iteration over the quality of the predictions on the validation data by using various metrics, including accuracy (ACC score function) and f1. The general process followed is illustrated in Figure 11.

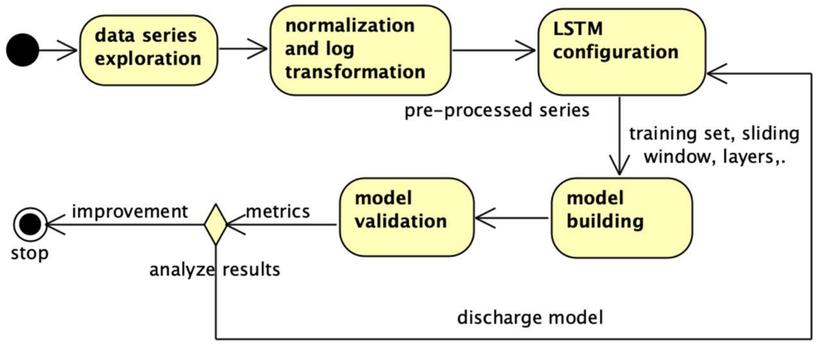


Figure 11. Development process of a prediction model.

To answer [RQ1], this process was applied for several critical sections of the same stretch and of different stretches, selected based on the quality of the input data (e.g., completeness). Among the experiments, the setting in Table 2 led to the most accurate model for more than one critical section.

Table 2. Model configuration.

Parameter	Value
Input dataset	series water discharge series water depth timeline: days from 1 January 2018 to 31 May 2022
window size	15
normalization	z-score
LSTM sequential model settings	bidirectional layers of 1024, 512, 256, 128, and 64 units adam optimizer, learning rate 0.002 20 epochs, early stopping
training-validation split	KFold cross-validation, 20 splits

The model is trained on the two-input series (to learn patterns of correspondence of their values), and the purpose of the resulting model is to take an input sequence of some length for one of the two series (series water discharge in our experiments) and learn to predict the corresponding values of the other series. The trained LSTM model could be used to predict future values of both water discharge and water depth series. In this study, the prediction of discharge is provided by the EWA system, which, as explained in Section 2, is based on a numerical model.

A preliminary answer to [RQ2] is provided with the following results. An example of the result of a critical section in the last 450 days of the dataset is shown in Figure 12. Please note that, as specified in Table 2, the KFold cross-validation technique was used so that the dataset was divided into "k = 20" subsets, and the model was trained and evaluated 20 times, using a different subset as the validation set in each iteration. The best model was then chosen based on accuracy scores. For this critical section, the subseries of real water depth of the last period were interesting as they featured the lowest values of the whole historical series, hence the choice to use the KFold cross-validation technique.

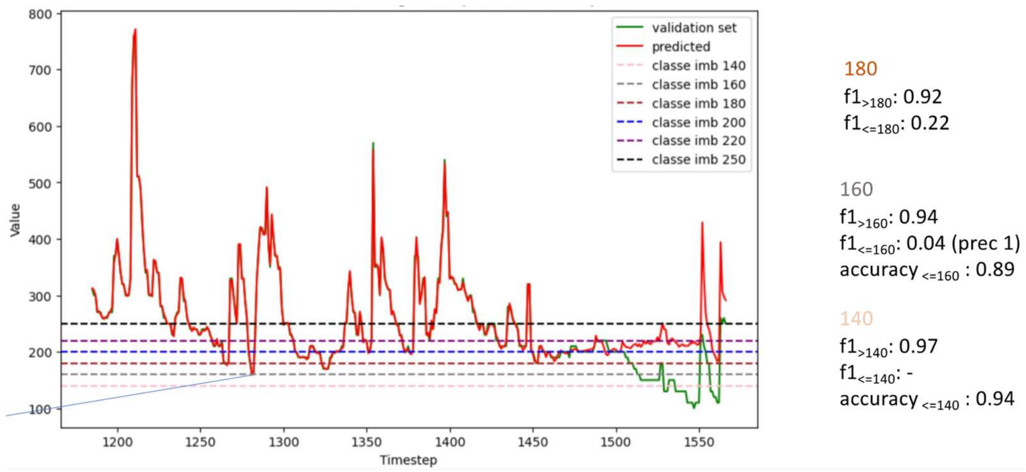


Figure 12. Water depth forecast vs. water depth observations for critical section 1 of stretch X. The blue line indicates the lowest water depth before the last 100 days.

Figure 12 shows the vessel drafts and the f1 and accuracy values for the 140, 160, and 180 cm classes. These lowest classes are the most critical for the confidence level of the predictions for non-navigability, as these predictions overestimate much of the real values and provide wrong negative estimates for the last 100 days (from February to May 2022). Indeed, from the analysis of the considered dataset, in this period, a similar decrease pattern in water depth was not present beforehand. Expert knowledge of the management of the infrastructure and/or a deeper investigation into weather and environmental aspects may help to better interpret these results.

From the data analysis only, better estimates for navigability risks of the same classes of vessels resulted in a critical section of another stretch of the river, as represented in Figure 13. It can be noticed that there is a higher frequency of low values in the validation dataset.

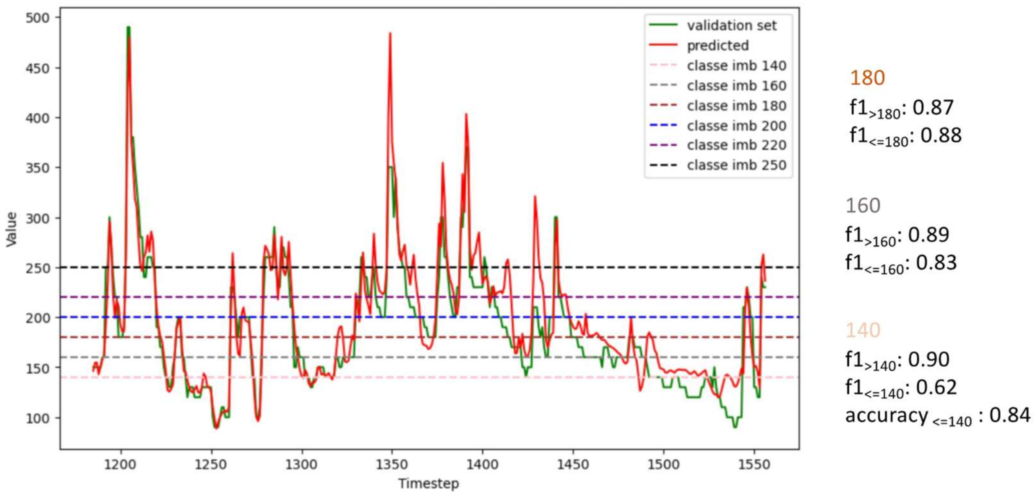
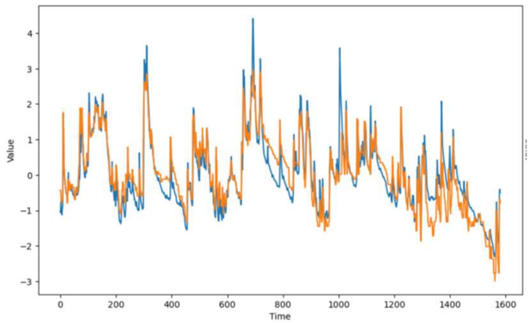
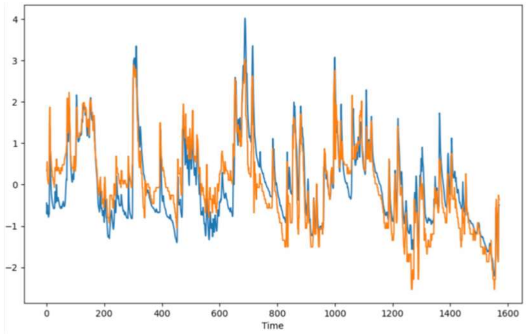


Figure 13. Water depth forecast vs. water depth observations for critical section 2 of stretch Y.

To address [RQ3], further tests should be made to search for models that can adapt to different critical sections and/or stretches. A first investigation of the datasets leads to the hypothesis that critical sections of the same stretch generally feature similar shapes (values and variability). However, critical sections of different stretches may feature different patterns to those used in Figures 12 and 13. The full datasets of the water discharge (blue line) and water depth (brown line) normalized series are displayed in Figure 14.



(a)



(b)

Figure 14. Water discharge (blue) and water depth (brown) time series: (**a**) critical section 1 of stretch X; (**b**) critical section 2 of stretch Y.

4. Conclusion and Future Work

In this paper, a real case study that requires the application of time series analysis and forecast methods to build a digital twin for the navigability of free-flowing waterways has been described. The problem is complex, so the challenge is to achieve reliable risk values. The data exploration and the preliminary results of the LSTM-based method presented in this paper look promising for that aim. However, further validation tests with real data are required, and the navigability forecast from the water data should always be complemented with further information, such as climate conditions and infrastructure management operations.

Author Contributions: Conceptualization, A.A., L.C., G.F., M.L.V., S.G. and E.E.; methodology, M.L.V., E.E. and S.D.; software, M.L.V., E.E. and S.D.; validation, M.L.V., E.E., S.D., A.A., L.C., G.F. and S.G.; data curation, A.A., L.C. and G.F.; writing—original draft preparation, M.L.V., E.E. and S.G.; writing—review and editing, M.L.V., E.E., A.A. and S.G. All authors have read and agreed to the published version of the manuscript.

Funding: The research activities have been funded by the project "CRISTAL, the Climate resilient and environmentally sustainable transport infrastructure, with a focus on inland waterways" funded by Horizon Europe—European Union Research and Innovation Program under grant agreement No. 101069838, HE10 CINEA. The sole responsibility for the content of this publication lies with the authors. It does not necessarily represent the opinion of the European Union. Neither the EASME, nor REA nor the European Commission is responsible for any use that may be made of the information contained therein.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study is partially publicly available and can be requested from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. European Commission; Directorate-General for Mobility and Transport; Punter, L.; Hofman, W. Digital Inland Waterway Areas–Towards a Digital Inland Waterway Area and Digital Multimodal Nodes–Final Report. 2017. Available online: https://data.europa.eu/doi/10.2832/26584 (accessed on 3 September 2024).
2. Fanti, M.P.; Ukovich, W.; Giansante, C.; Berardi, A.; Di Pierro, B. Common Requirements for the Federated Architecture of Platforms. 2019. Available online: https://fenix-network.eu/wp-content/uploads/2020/07/FENIX-Deliverable-D2.2.2_v2.0_FINAL.pdf (accessed on 3 September 2024).
3. Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
4. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.

5. Vizi, Z.; Batki, B.; Rátki, L.; Szalánczi, S.; Fehérváry, I.; Kozák, P.; Kiss, T. Water level prediction using long short-term memory neural network model for a lowland river: A case study on the Tisza River, Central Europe. *Environ. Sci. Eur.* **2023**, *35*, 92. [CrossRef]
6. Le, X.H.; Ho, H.V.; Lee, G.; Jung, S. Application of long short-term memory (LSTM) neural network for flood forecasting. *Water* **2019**, *11*, 1387. [CrossRef]