technische universität
dortmund

Fraunhofer
IML

MASTER'S THESIS

Advancing Inland Waterways Logistics: Monitoring Navigability Through Water Level Forecasting, Predictive Analysis and Feature Engineering of Hydrological Data

Author
Sohith Dhavaleswarapu
Master's Data Science

Supervision
Prof. Dr. Christine Müller
Ebrahim Ehsanfar (M.Sc.)

Faculty of Statistics,
Technische Universität Dortmund

In cooperation with
Department of Transport Logistics,
Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund

Dortmund, May 14, 2024

Contents

List of Figures

List of Tables

List of Abbreviations

ACF	Auto Correlation Function
ADF	Augmented Dickey Fuller
AIPo	Agenzia Interregionale del fiume Po Authority
ANN	Artificial Neural Network
API	Application Programming Interface
ARIMA	Auto Regressive Integrated Moving Average
CNN	Convolutional Neural Network
CRISTAL	The Climate Resilient and Environmentally Sustainable Transport Infrastructure, with a Focus on Inland Waterways
CSV	Comma Separated Values
ECMWF	European Centre for Medium-Range Weather Forecasts
EDA	Exploratory Data Analysis
EFAS	European Flood Awareness System
ENEA	Italian National Agency for New Technologies, Energy and Sustainable Economic Development
ERA	European Centre for Medium-Range Weather Forecasts Reanalysis
EU	European Union
FLAML	Fast and Lightweight AutoML Library
Fraunhofer IML	Fraunhofer Institute for Material flow and Logistics
Geo-Package	Geospatial Package
GeoJSON	Geospatial JavaScript Object Notation
GRDC	Global Runoff Data Base
GRU	Gated Recurrent Unit
ITISE	International Conference on Time Series and Forecasting
KB	Kilo Byte
LAT	Latitude

LAT	Longitude
LOCF	Last Observation Carried Forward
LSTM	Long Short Term Memory
LSTNet	Long and Short term Time series Network
MB	Mega Byte
MDPI	Multidisciplinary Digital Publishing Institute
MLP	Multi Layer Perceptron
NetCDF	Network Common Data Form
PACF	Partial Auto Correlation Function
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RQ	Research Question
UK	United Kingdom
USA	United States of America
UTC	Universal Time Coordinated
VAR	Vector Auto Regression

1. Introduction

The European Commission has set ambitious targets for the transportation sector. By the year 2030, it aims to achieve a 30% modal shift away from road freight transport towards more environmentally friendly and socially responsible modes of transportation such as rail and inland waterways (Ambra et al., 2019). The commission has set an even more substantial goal of achieving a 50% shift by the year 2050 (Ambra et al., 2019). In the complex web of the European supply chain and logistics industry, inland waterways emerged as a pivotal player due to many factors, including cost-effectiveness, ecological benefits and a reputation for reliability (Zwicklhuber and Kaufmann, 2023).

The CRISTAL project, officially known as The Climate Resilient and Environmentally Sustainable Transport Infrastructure, with a Focus on Inland Waterways (CRISTAL project team (2022)), is a research initiative funded by the European Commission and developed in collaboration with European union partners. The project aims to encourage freight transport on inland waterways. The Fraunhofer Institute for Material flow and Logistics (Fraunhofer IML) and other EU partners actively shape the project's trajectory as part of the consortium. This thesis is the data driven study of inland water navigation within the CRISTAL project. It focuses on river depth forecasting using machine learning algorithms through feature engineering of hydrological data, followed by predictive analysis of navigational risk.

Within the CRISTAL project, the River Po in northern Italy offers significant opportunities to support the supply chain and inland logistics corridor (CRISTAL project team, 2022). The potential navigable stretch of the Po River comprises shallow areas that, under specific hydro-metric conditions, pose navigation risks. This navigable segment is divided into branches with potential shallow points, also called critical points, that can disrupt overall navigability. To facilitate effective navigation planning whilst considering logistical challenges, having a prediction model that provides timely and forecasted information on river depth levels in each branch is necessary. Such a model ensures accurate planning and holds promising potential for future advancements in the logistics sector.

Over the previous three decades, the Agenzia Interregionale del fiume Po Authority (AIPo) (Interregional Agency of the Po River, 2011), also a partner in the CRISTAL project, has identified critical shallow points along each branch of the navigable section of the Po River. Continuous monitoring of hydrological information at these locations involves measuring upstream water levels, river depths and river discharge flow rates.

This thesis focuses on three shallow points, namely Piacenza, Monte P.Te Revere and Cavanella, each from a different branch by examining historical data from the last 10 years. Additionally, daily records of climate data from Copernicus climate services (Hersbach et al., 2023) are sourced to identify the impact of climate attributes like total precipitation, temperature and snow depth on shallow points.

The main objective of this thesis is to provide a proof of concept to establish the use of inland navigation in River Po, a task highly dependent on monitoring water depth levels at crucial river points. Various ship classes, determined by their draft lengths, rely on specific water level thresholds (Kirilenko and Epifantsev, 2023). Utilising forecasted water depth levels allows for calculating the likelihood of future navigation, optimising ship scheduling and refining routing strategies. This process minimises delays and improves the overall efficiency of corridor management.

This thesis focuses on advanced feature engineering, incorporating cross-correlation of hydrological and climate attributes to assess their impact on shallow points. Various machine learning models, including Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Long and Short term Time series Network (LSTNet) and Vector Auto Regression (VAR) methods, are considered to forecast daily river depth levels over 14 days and hourly water levels for the next 24 hours. The study explicitly addresses navigability concerns for different types of ships by generating future estimates of water depth levels at critical points. Furthermore, the probability of navigation risk for each ship class is calculated using the best-performing model based on forecasted water depth levels. This study presents a probabilistic approach to developing a proactive warning system that can provide timely alerts. Such a system can significantly improve navigational planning and safety in water transportation, benefiting logistics and related stakeholders.

1.1. Problem Statement

According to the CRISTAL project's analysis, the Po River in north Italy is a vital waterway with a crucial navigable section for efficient transport that connects the Mediterranean Sea with the inland of Italy. The AIPo authority has divided the navigable section into nine branches, namely Piacenza, Polesine Parmense, Casalmaggiore, Riva di Suzzara, Revere, Felonica, Santa Maria Maddalena, Papozze and Volta Grimana (Interregional Agency of the Po River, 2011). There are 157 shallow points (see figure 1) spread across the navigable branches. In 2022, Po River faced the heaviest drought situation in the past two centuries (see figure 2) that is part of a long-term trend of more

frequent and severe drought in the area (Montanari et al., 2023). Because of climate change and difficult hydrological conditions, each critical point poses unique challenges to navigability due to potential record low water depth levels. According to the AIPo, it is nearly impossible to implement the navigation of ships with heavy cargo and long draft length at these crucial points.

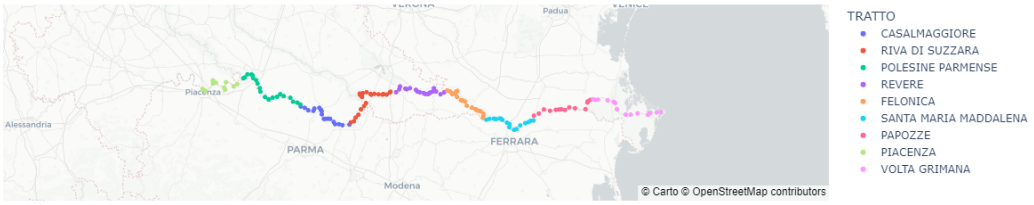


Figure 1: Shallow points identified by the AIPo authority across navigable branches of River Po in northern Italy ('TRATTO' in figure translates to navigational branch).

(a) 2020 (b) 2021 (c) 2022

Figure 2: Copernicus Sentinel-2 satellite images of the River Po Valley near Piacenza, revealing significant river shrinkage between 2020 and 2022 (Drusch et al., 2012).

The AIPo authority recognises the importance of addressing navigational uncertainties along waterways by proactively anticipating issues occurring from low water depth levels. To achieve this, feature engineering on hydrological components that affect water depth levels is necessary. Additionally, a predictive model needs to be developed to forecast future water depth levels at specific points. The historical data collected by the AIPo on river depth, river discharge rates and upstream water levels can be used to train the model. This model can then be integrated with future river discharge rates from the European Flood Awareness System (EFAS) (Copernicus Climate Change Service, 2019)

and the AIPo's existing water level estimation system (Interregional Agency of the Po River, 2011) to predict future river depth accurately.

The water depth level is influenced by various atmospheric variables, including air temperature, precipitation volume and snow depth (Atashi et al., 2022). Recognising the significance of these atmospheric factors, the feature engineering process will examine hydrological components and map their influence based on a geometrical radius of impact within the river catchment area. This approach aims to provide a subtle understanding of how each component contributes to fluctuations in water depth level at shallow points. Furthermore, the hydrological dynamics of river discharge and the accumulation of sand sediment on riverbeds also notably impact the river depth at these shallow points (Vezzoli et al., 2015). Understanding and incorporating these factors into the predictive model is crucial for a comprehensive and effective solution to the navigability challenges faced along the Po River.

The work in this thesis contributes to the practical use of data science knowledge by studying advanced machine learning algorithms to solve challenges in inland navigation. Comparing the forecasting capabilities of machine learning techniques like LSTM, GRU, LSTNet and VAR for predicting water depth levels enables the identification of optimal approaches for operational forecasting. By analysing the influence of key hydrological components in river depth at specific locations along the River Po, the study establishes a foundation for predictive modelling. Furthermore, by assessing navigational risk probabilities for different ship classes based on forecasted water depths, the thesis demonstrates a practical application of data-driven decision-making in hydrology using statistical knowledge. This interdisciplinary research advances state-of-the-art hydrological modelling and provides actionable insights that can inform policy about enhancing safety protocols and optimising navigation strategies.

1.2. Research Questions

This thesis aims to investigate and address the following research questions comprehensively:

RQ1 : How much do the different hydrological factors, such as river discharge rates, upstream water levels, precipitation, temperature and snow depth in the catchment area, affect the daily variations in river depth at the three important shallow points (Piacenza, Monte P.Te Revere, Cavanella) in the River Po?

RQ2 : How well do machine learning algorithms like LSTM, GRU, LSTNet and VAR predict daily and hourly water depth levels at three shallow points based on the influential hydrological components identified in RQ1?

RQ3 : What are the probabilities of navigational risk for different ship classes at the three shallow points in the River Po, based on forecasted water depth levels using the best-performing machine learning algorithm identified in RQ2?

1.3. Organisation of Thesis

The thesis is divided into ten sections, each with a specific purpose in presenting the research and findings. The "Introduction" chapter 1 provides a brief overview of the significance of this thesis work, followed by the problem statement, which explains the existing challenges, objectives and the systematic approach of the study by framing research questions. The "Background" section 2 outlines the author's role and the importance of the work. The "Related Work" section 3 critically examines existing literature on the research topic to identify gaps and establish the theoretical framework.

The section labelled "Dataset" 4 provides an overview of the data utilised in this thesis. It includes information on data collection methods, necessary processing and machine learning model development preparation. The "Methodology" section 5 describes the statistical techniques used in the study, including time series analysis, machine learning approaches like RNN and its types such as LSTM and GRU, the hybrid method called LSTNet and a statistical approach named VAR. It also explains how the models are tuned for optimal performance and the evaluation metrics used in the study.

The "Feature Engineering of Hydrological Components" section 6 addresses RQ1, detailing the process of identifying influential components at shallow points through cross-correlation analysis through graphs and explaining findings by setting inputs for RQ2. The "Exploratory Data Analysis (EDA)" section 7 checks the assumptions necessary for time series analysis and extracts knowledge of spatial and temporal patterns required for machine learning processes. The "Machine Learning Approach for River Depth Forecasting" section 8 addresses RQ2, providing a workflow overview of the forecasting approach and delving into the design and training using LSTM, GRU, LSTNet and VAR models.

In the subsequent "Findings and Discussion" section 9, the forecasting results of the models are presented through data visualisations and statistical analysis to determine the most effective model based on evaluation metrics and cross-validation outcomes. This chapter also addresses RQ3 by explaining the probability approach that is used

to calculate the navigation risk of ships based on forecasted river depth by the best model identified in RQ2. Furthermore, the findings are discussed in detail, incorporating comparative analysis, considerations of limitations and validity and future scope of importance. Additionally, this section outlines a recent publication resulting from this research, highlighting the significance and potential contributions to the scientific field.

Finally, the "Summary" section 10 concludes the key findings, discusses their broader implications and suggests avenues for future research. Additionally, the thesis includes appendices containing supplementary material, such as additional tables and figures used in the analysis, further enhancing the comprehensiveness of the study presented.

2. Background

The CRISTAL project involves the development of various technological advancements and digital solutions for transport infrastructure. The project is co-created by 15 partners from 9 European countries, including Poland, Germany, Italy, Belgium, the Czech Republic, Hungary, Greece, France and the UK. Notably, the Po River in Italy is one of the project's pivotal sites in planning the navigation route and the governance procedures. The project aims to digitalise logistics and transport services, focusing on increasing the effective utilisation of River Po's water resources by planning navigability. The outcomes of the project make substantial contributions to enhancing logistics planning for sustainability and infrastructure resilience (CRISTAL project team, 2022).

As a partner of the CRISTAL project, Fraunhofer IML represents Germany and is responsible for developing Digital Twin Technology for transport infrastructure. To study River Po and its navigational feasibility, Fraunhofer IML is closely associated with Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) and the AIPo authority (Interregional Agency of the Po River, 2011). The author's entitlement at the Fraunhofer IML is as a student research assistant for machine learning solutions and actively engaged in the CRISTAL project team (2022).

The team at Fraunhofer IML focuses on developing geospatial solutions and visualising analytical aspects of river dynamics for navigation across European waterways. In addition to that, this thesis introduces aspects of time to predict future water depth levels using machine learning techniques based on hydrological aspects, intending to enhance navigability. This process contributes to the development of a warning system for shallow points, serving as a dashboard tailored for various types of ships to facilitate advanced logistics planning and prepare for extreme weather events.

3. Related Work

Water level prediction in rivers is pivotal in various domains, ranging from flood management to ensuring navigability (Ghimire, 2017). The complexity of this task becomes particularly evident when considering rivers with unique geographical characteristics, such as the Po River (Ravazzani et al., 2015). In pursuing compelling water level predictions, researchers have explored diverse methodologies and technologies to address the multiple challenges rivers pose in different regions.

The Po River basin expands into two distinct regions: the Upper Po, encompassing 75% of the area, sourcing the water drainage from mountainous tributaries that originate from the Alps and the rest 25% of the area, distinguished by expansive, flat plains (Arttna et al., 1990). Geographical properties of the River Po exhibit distinctive features that significantly influence its hydrological dynamics (Montanari, 2012). "River Research and Applications" by Castellarin et al. (2011) explains the effects of different floodplain management in the Po River. Flood management techniques also consider hydrological aspects in defining their objectives, which draws similarities to the analysis for water depth level prediction regarding navigability (Castellarin et al., 2011). While flood management studies contribute valuable insights, there exists a noticeable gap in the literature concerning the explicit consideration of water depth level in the shallow points of the Po River for navigability purposes.

In recent decades, adopting forecasting methods in hydrology and water resource management has seen significant attention, mainly by using machine learning techniques (Atashi et al., 2022). Statistical models such as ARIMA have been applied to predict river discharge at two stations along the Schuylkill River, USA (Ghimire, 2017). Additionally, multivariate time series models, such as VAR, have been utilised to forecast rainfall flow discharge in various locations, including Sojomerto, Juwero and Glapan in Central Java Province, Indonesia (Hartini et al., 2015). ANN and RNN have shown effectiveness in hydrological tasks, including flood forecasting in the Red River, USA (Atashi et al., 2022) and estimation of water levels in Japanese rivers (Borwarnginn et al., 2022). In addition, deep learning models such as the GRU have been used to forecast groundwater levels in the Qoşaçay plain, Iran (Lin et al., 2022). Similarly, the hybrid approach, namely the LSTNet model, showed prominent results when predicting groundwater Levels in the Middle and Lower Reaches of the Heihe River in China (Yang and Zhang, 2022). However, applying such techniques to the Po River, with its unique geographical context, remains unexplored.

The review outlines the importance of water level predictions, particularly in the context of the Po River and underscores the necessity of understanding its geographical attributes. This work seeks to fill the research gap by combining insights from hydrological studies and machine learning methodologies with a specific focus on the characteristics of the River Po.

Following an extensive review of the literature and a comprehensive assessment, this study integrates an ANN approach, including RNN such as the LSTM architecture and its optimized variant, the GRU. Additionally, a hybrid model called LSTNet is considered to address existing gaps in research and explore new avenues for water level forecasting. In contrast to that, a traditional statistical approach, VAR, is included for comparative analysis with deep learning methods.

4. Dataset

This study investigates the hydrological dynamics at shallow points along the River Po. The primary objective is to explore the relationships between river depth at critical shallow points and key hydrological components, including river discharge rates and water levels at measuring station points. Additionally, the influence of climate data, such as precipitation patterns, temperature variations and snow depth across the radius of influence in the catchment area are measured. In the perspective of this thesis, strategically, three individual shallow points, namely Piacenza, Monte P.Te Revere and Cavanella (see figure 3) that belong to the navigable branches of Piacenza, Revere and Volta Grimana respectively, are considered. These specific locations serve as crucial transit points between major inland ports and the connection to the Mediterranean Sea.

Figure 3: Geographical locations of shallow points along the Po River considered within the scope of this thesis.

The following sections will examine each dataset in detail, comprehensively analysing the relationship between water depths and various hydrological and climatic factors. This investigation aims to enhance the understanding of the factors controlling water dynamics at selected shallow points in the River Po.

4.1. Data Components

Table 2 presents an overview of the data components utilised in this thesis, followed by detailed descriptions of individual attributes that highlight their purpose and features.

Category	Data Component	Source	File format	Summary
Geographical	Shallow water points	AIPo	Geo-Package	Geographic information about shallow water points.
	Navigable section	AIPo	Geo-Package	Geographic information about the navigable branches along the Po River.
	Po river basin	GRDC	GeoJSON	Geographic information of Po River basin.
Hydrological	River depth	AIPo	CSV	River depth recorded at a shallow water point.
	River discharge rates	AIPo	CSV	River discharge rates at a shallow water point.
	Water level	AIPo	API	Upstream water levels at monitoring point nearest to shallow water point.
Climate	Temperature	ECMWF	NetCDF	Temperature recorded 2m above of land, sea or inland waters in a geographical grid.
	Total precipitation	ECMWF	NetCDF	Total precipitation recorded in a geographical grid.
	Snow depth	ECMWF	NetCDF	Depth of snow accumulated in a geographical grid.
Logistics	Ship classes	AIPo	CSV	Category of ships classified based on their draft length.

Table 2: Summary of data components sourced for the scope of this thesis.

4.1.1. Geographic Information of River Po

The AIPo Authority provided crucial geographic information on navigable branches along the Po River (Interregional Agency of the Po River, 2011). Two Geo-Package

files contain details about shallow water points and navigable sections (see table 3 and 4). These files are processed using the `geopandas` package (Jordahl et al., 2020) in a Python 3.11.8 environment (Python Software Foundation, 2023).

Geo-Package File for Shallow Water Points:

The file is 124 KB and consists of 157 records and five variables.

Variable Name	Data Type	Description
`"NAME_LOCAL"`	String	Names of each shallow point.
`"CODE"`	Numeric	Unique encoded values for identifying each shallow point.
`"BRANCH"`	String	Names of navigable branches where specific shallow points are located.
`"Program_km"`	Numeric	Distance from the origin of the river to shallow point, measured in kilometres.
`"geometry"`	Geometry Array	Geographical coordinates in `POINT` format (LAT, LAT) object representing the location of each shallow point.

Table 3: Description of variables stored in the Geo-Package file for shallow water points.

Geo-Package File for Navigable Sections:

The file is 128 KB and consists of 9 records and three variables.

Variable Name	Data Type	Description
`"BRANCH"`	String	Names of navigable branches.
`"Length"`	Numeric	Length of branches in meters.
`"geometry"`	Geometry Array	Geometrical coordinates in `MULTILINESTRING` format (LAT, LAT), representing the branches along the Po River.

Table 4: Description of variables stored in the Geo-Package file for navigable branches.

Both files contain complete information and the data quality has been verified and validated by the AIPo authority. The original variable names, initially in Italian, have been translated into English for easier understanding. The data is intended for academic and research purposes.

GeoJSON File for the Po River Basin:

Additionally, for a comprehensive understanding of the Po River's geography, a geographical information file in GeoJSON format for the entire Po River drainage basin from the GRDC (Federal Institute of Hydrology (BfG), 2020) is sourced. This file is publicly open to download and information can be used for academic and research perspectives. This data enhances the comprehension of the river's water catchment area and the geometric influence on water collection into the river. Figure 4 plotted using the `Folium` library (Qiusheng Wu, 2021) in Python 3.11.8 environment (Python Software Foundation, 2023), explains the River Po basin along with the water catchment areas and drainage system into the river.

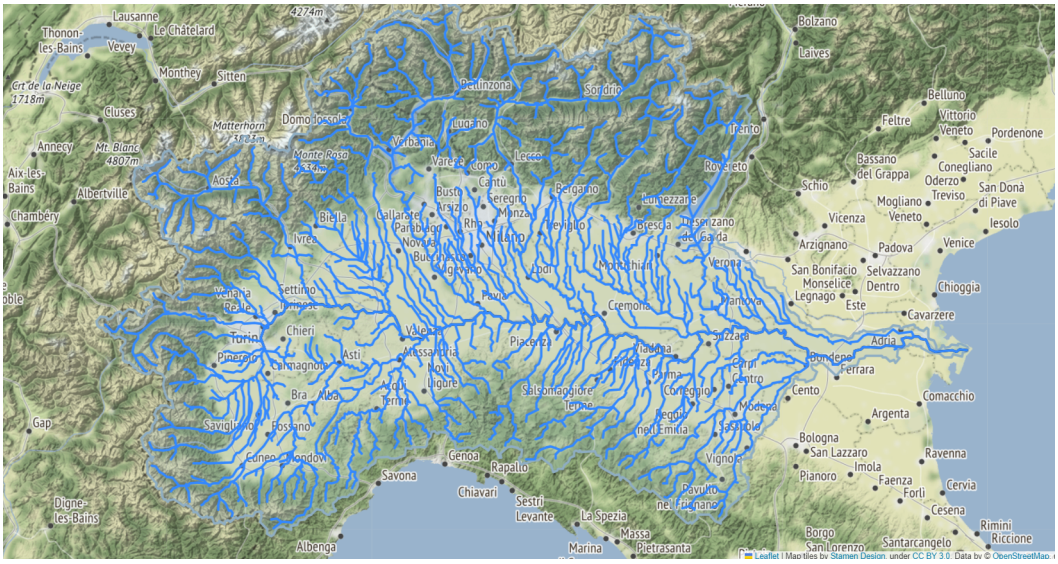


Figure 4: Plot of the entire River Po basin in northern Italy and its water catchment areas.

The mentioned data sets offer valuable insights into the Po River's geography, providing visual information on regions influencing shallow points and navigable branches.

4.1.2. Hydrological Components

River dynamics at shallow points are critical components in understanding the hydrological characteristics of the Po River (Montanari, 2012). High discharge rates often lead to increased river depth, indicating elevated water levels. Similarly, low discharge rates may result in shallower river depths (Vezzoli et al., 2015). These attributes are crucial

determinants in influencing the formation of shallow points within the river system, a focal point of the study. By comprehensively understanding how discharge rates and water levels influence water depth levels at critical shallow points in Po River, the aim is to forecast these river depths from the navigability perspective.

River Depth

River depth at a water point refers to the distance from the water surface to the riverbed at a specific location (Arttna et al., 1990). Shallow water points are typically areas where the river is relatively low and the riverbed is closer to the surface (Chow, 2017). The formation of sandbars and dunes due to sediment deposition can significantly alter the morphology of the river channel, influencing patterns of depth. The accumulation of sand and other sediments on the riverbed may lead to aggradation, which decreases water depth (Mugade and Sapkale, 2015). This phenomenon can affect water flow and navigation (Whitmeyer and FitzGerald, 2006). Understanding river depth dynamics at these points is essential, as fluctuations can significantly impact navigation channel depth and accessibility, posing risks to maritime activities (Whitmeyer and FitzGerald, 2006). In this thesis, river depth at critical shallow points acts as a target variable, resulting in outcomes that allow ships to transit safely within permissible thresholds.

River Discharge Rate

River discharge rate refers to the volume of water flowing through a river per unit of time, often measured in cubic meters per second (m^3/s) (Mazzetti et al., 2023). The discharge rates at a shallow water point provide insights into the overall flow dynamics, helping to assess the water's force and potential impact on the navigable branches (Kazimierski et al., 2012). Monitoring discharge rates is crucial for understanding variations in shallow water points, especially during different seasons or due to external influences.

Since 1988, the AIPo authority has recorded and maintained crucial information related to river discharge rates and river depth readings at shallow points in the Po River. For this analysis, the authority has shared a CSV file format (see table 5) (Interregional Agency of the Po River, 2011). This file contains daily time series values encompassing river discharge rates and river depth at each shallow point, as described in the accompanying Geo-Package file for shallow water points.

The dataset provided by the river authority is substantial, with a file size of 58.3 MB. The dataset is organised into two sheets: "1 set" and "2 set", each containing distinct

sets of records. The first set encompasses Casalmaggiore, Felonica, Papozze, Piacenza, Polesine Parmense and Revere branches. The second set includes shallow points within branches like Revere, Riva di Suzzara, Santa Maria Maddalena and Volta Grimana.

Sheet "1 set" comprises 10,28,378 records and six columns. Sheet "2 set" contains 7,96,709 records organised into six columns. The columns in both sheets convey consistent information, ensuring uniformity and facilitating ease of analysis.

Variable Name	**Data Type**	**Description**
`"BRANCH"`	String	Names of navigable branches.
`"NAME"`	String	Names of shallow points.
`"HYDROMETER"`	String	Nearby Hydrometer station responsible for measuring water levels in perspective of navigability.
`"DATE"`	Datetime	Date and time (`DD-MM-YY HH:MM:SS` format) for which the value is recorded.
`"DEPTH"`	Numeric	River depth recorded in centimetres (cm) for a specific shallow point.
`"DISCHARGE"`	Numeric	River discharge rate recorded in cubic meters per second (m^3/s) for a specific shallow point.

Table 5: Description of variables stored in the dataset shared by the AIPo authority related to river depth and river discharge rates.

This comprehensive dataset spans from 01 January 1988 to 12 May 2022 in daily frequency, covering all recorded shallow points within the specified branches. The file's original variable names are initially in Italian and have been translated to English for easier understanding.

Water Level

Accurate water level readings are crucial for ensuring navigability in river systems, particularly at critical points (Cuppini et al., 2015). The emphasis on water levels, rather than depth alone, is motivated by the nature of vital points, where variations in sand sedimentation play a central role in ensuring safe and effective navigation. Based on strategic locations, water levels are measured using hydraulic gauges or scales to monitor flood-related activities in the river (Castellarin et al., 2011). In this thesis, the focus on water level data is driven by its significance in unravelling complexities at critical points along the River Po.

Water level data in this thesis is collected through the AIPo authority's dedicated API, accessed via their web servers (Interregional Agency of the Po River, 2011). This API is a comprehensive source, providing real-time and historical water level measurements since 1988 for several known critical shallow points from existing sensors. By making specific requests using the station identifier attribute of API, the process results in historical water level data measured in meters for the defined period (see table 6 and 7). These station identifiers are located near the described shallow water points within the navigable branch, ensuring the collected data is linked to the relevant geographic locations of interest.

Variable Name	**Data Type**	**Description**
`"elementName"`	String	Type of measurement (as per API request 'Water level')
`"elementId"`	Numeric	Unique identifier number for each station per the corresponding element.
`"stationName"`	String	Unique name for each station.
`"stationId"`	Datetime	Unique identifier number for each station.
`"lat"`	Numeric(decimal)	Geographical latitude location.
`"lon"`	Numeric(decimal)	Geographical longitude location.
`"decimals"`	Numeric	Number of decimal value reading is rounded.
`"measUnit"`	String	Measurement unit in which data is measured in meters.
`"time"`	Datetime	Date and time `(DD-MM-YY HH:MM:SS UTC +1` format) for which the value is recorded.
`"value"`	Numeric(decimal)	Water level at a station for a specific time (only latest data)
`"trend"`	Numeric(decimal)	Variation in current value compared to previous hour recording.

Table 6: Description of variables obtained through an API request for station identifiers.

The station identifiers required for water level reading are attained from the same API web server (see table 6) dedicated to all the river basins in Italy. For this analysis, only details of the Po River basin are permitted. Upon request, a response with 14 records describing station identifiers of water level corresponding to 11 variables is retrieved.

Another API collects the historical water levels at a particular station (see table 7) which takes the corresponding "elementId" from the previous API (see table 6) of the unique station and results in necessary information of variables.

Variable Name	Data Type	Description
"0"	Datetime	Date and time (DD-MM-YY HH:MM:SS UTC +1 format) for which the value is recorded.
"1"	Numeric(decimal)	Recorded Water level at a station for a specific time.

Table 7: Description of variables obtained through an API request for historical water level data.

The dataset encompasses a comprehensive record of measurements taken at the requested station, with data points recorded every 30 minutes from 01 January 2001 to the present day.

4.1.3. Climate Components

Climate changes can strongly impact rivers, leading to notable shifts in the availability of water resources (Fiseha et al., 2014). Navigating the complexities of the River Po requires an in-depth exploration of the relationship between climate dynamics and water depth levels, particularly at shallow points. According to studies, projections suggest that summer river discharge rates in the Po River will decrease from 2040 to 2050 compared to the baseline period of 2000 to 2010 (Ravazzani et al., 2015). This decline is linked to a significant reduction in seasonal precipitation and an accelerated snow melt.

The focus extends beyond prediction to practical implications, offering navigators insights into how climatic shifts influence river depth at shallow water points in the Po River. The objective of this study is to investigate how critical climate factors such as temperature, precipitation and snow depth affect the hydrological conditions of the River Po. By explicitly analysing the features and calculating the impact period through cross-correlation, this research aims to identify the influence of climate attributes at crucial points.

The essential climate data utilised in this study is retrieved from the Copernicus Climate Data Storage (Hersbach et al., 2023), derived from ERA5, the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) (2022) tool, covering global climate and weather patterns over the past eight decades. Access to the data is facilitated through a dedicated API request to the Copernicus Climate Data Storage (Hersbach et al., 2023). This dataset is generated under the framework of the Copernicus Climate Change Service (C3S) (2023). The data is available in a gridded format with

a geographical resolution of 0.25° latitude by 0.25° longitude and a temporal resolution of one hour. This thesis focuses on daily climate attributes, specifically temperature, total precipitation and snow depth. The geographical subset of interest encompasses the region within North 50°, West 0°, East 40° and South 20° latitudes and longitudes, covering the Po River basin from 2013 to 2022. Due to resource limitations, three different API requests have been made for temperature (see table 8), total precipitation (see table 9) and snow depth (see table 10) attributes outputs are generated in NetCDF format and accessed through `xarray` library (Hoyer and Hamman, 2017) in Python 3.11.8 environment (Python Software Foundation, 2023).

Temperature

In the Alpine region, the increasing temperatures have reduced over half of the glacier's volume since 1900 (Beniston, 2012). If global temperatures rise by 2–4°C, it is projected that between 50% and 90% of the ice mass from mountain glaciers could vanish by the close of this century (Beniston, 2012). In the context of the Po River, the simulation results from the "Hydrological Model for Assessing Climate Change" by Ravazzani et al. (2015) exhibit an overestimation of monthly temperatures during winter and an underestimation from late spring to the end of summer. This discrepancy in the actual and simulated meteorological forecast leads to overestimating average monthly discharge in March and April, while underestimations are followed in September, November and December. The profound impact of temperature fluctuations on the Po River defines the importance of incorporating temperature data in the context of river depth level forecasting.

According to Copernicus Climate Data Storage (Copernicus Climate Change Service (C3S), 2023), the "2m temperature" variable is the air temperature at 2m above the surface of land, sea or inland waters. It is calculated by interpolating between the lowest model level and the Earth's surface, taking account of the atmospheric conditions.

When an API request is made, the file named `2m_temperature.nc` (Hersbach et al., 2023) is downloaded in NetCDF format. The file is 555.5 MB and contains 291,079,008 rows for four columns.

Variable Name	Data Type	Description
`"longitude"`	Numeric(decimal)	Geographical longitude location.
`"latitude"`	Numeric(decimal)	Geographical latitude location.
`"time"`	Datetime	Date and time (`DD-MM-YY HH:MM:SS` format) for which the value is recorded.
`"t2m"`	Numeric(decimal)	Recorded temperature measured in units of kelvin (K) for a specific time.

Table 8: Description of variables obtained through an API request for temperature data.

Total Precipitation

Intense precipitation over the Alps in the Mediterranean region has drawn significant attention concerning the Po River (Isotta et al., 2014). This focus is proved by the recurrent incidence of destructive floods, which profoundly impact the river's water levels (Isotta et al., 2014). The annual average precipitation volume is recorded at 78 cm^3, with 60% of this volume being converted into outflow at the closure section of the Po River (Montanari, 2012). The hydrological characteristics of the Po River, particularly concerning the flood regime, have been extensively studied (Montanari, 2012). However, despite these efforts, several significant questions persist regarding the river's hydrology. Analysing precipitation patterns is crucial for comprehensively understanding water depth levels at critical shallow points along the river and estimating navigation risk.

Copernicus Climate Data Storage (Copernicus Climate Change Service (C3S), 2023), has stated the "Total precipitation" variable in their data source as the accumulated liquid and frozen water, comprising rain, snow, sleet, hail, drizzle and any other forms of water, that falls to the Earth's surface and reach the ground over a specific period. The units of this parameter are depth in metres of water equivalent.

The API request resulted in file name `2m_Total_PRECIPITATION.nc` in NetCDF format (Hersbach et al., 2023). The file is 555.5 MB and contains 291,079,008 rows for four columns.

Variable Name	Data Type	Description
`"longitude"`	Numeric(decimal)	Geographical longitude location.
`"latitude"`	Numeric(decimal)	Geographical latitude location.
`"time"`	Datetime	Date and time (`DD-MM-YY HH:MM:SS` format) for which the value is recorded.
`"tp"`	Numeric(decimal)	Recorded total precipitation measured as depth in metres of water equivalent for a specific time.

Table 9: Description of variables obtained through an API request for total precipitation.

Snow Depth

Snow constitutes a predominant contributor to seasonal runoff in hydrological basins (Dettinger and Cayan, 1995), such as the rivers originating from the Alpine region. It is particularly evident when the snow-pack releases water during the spring and summer melt (Dettinger and Cayan, 1995). The Po River at Piacenza and Pontelagoscuro exhibits changes in water level over seasonal patterns, featuring a minor peak in spring, typically occurring towards the end of March. This phenomenon is likely attributed to snow melting from mid-altitude mountains in the surrounding region (Montanari, 2012). The timing and rate of snow melt can influence the volume and flow of water in rivers, contributing to fluctuations in water levels. Therefore, understanding the patterns and characteristics of snow melt is essential for accurate and reliable water level forecasting,

In Copernicus Climate Data Storage (Copernicus Climate Change Service (C3S), 2023), the variable "snow depth" is defined as the amount of snow covering a particular area in a grid. It is measured in meters of water equivalent and represents the depth that the water would reach if the snow melted and is evenly distributed over the entire grid.

The API request resulted in file name `snow_depth.nc` in NetCDF format (Hersbach et al., 2023). The file is 555.5 MB and contains 291,079,008 rows for four columns.

Variable Name	Data Type	Description
`"longitude"`	Numeric(decimal)	Geographical longitude location.
`"latitude"`	Numeric(decimal)	Geographical latitude location.
`"time"`	Datetime	Date and time (`DD-MM-YY HH:MM:SS` format) for which the value is recorded.
`"sd"`	Numeric(decimal)	Recorded snow depth measured as depth in metres of water equivalent for a specific time.

Table 10: Description of variables obtained through an API request for snow depth accumulation.

The climate data for all attributes originates from the Copernicus Climate Data Storage (Copernicus Climate Change Service (C3S), 2023). This data is publicly accessible and explicitly intended for academic and climate research purposes. The data quality is closely monitored and validated regularly to ensure its accuracy and reliability, with feedback loops established back to the data providers for continual quality assurance.

4.1.4. Ship Classes

This study uses a comprehensive dataset on various ship classes to investigate their probability of navigation risk at particular shallow points, considering water depth level readings and ship draft values. The ship class data is crucial for understanding the composition and characteristics of the navigation in inland waterways. According to the AIPo authority, ship classes are differentiated based on their draft values, representing the vertical distance between the waterline and the deepest part of a ship's hull, usually measured at the midpoint of the vessel's length (Kirilenko and Epifantsev, 2023). This draft reading is crucial for maintaining the ship's balance and ensuring safe navigation (Kirilenko and Epifantsev, 2023).

In this thesis, ship draft data is pivotal in understanding the required water depth levels at shallow points to navigate safely. By integrating forecasted water depth levels with the necessary draft values, the aim is to calculate the probability of navigational risk for specific ship classes. This approach enhances the understanding of the maritime challenges at shallow points and contributes valuable insights to improve the overall safety and efficiency of inland waterways navigation.

As part of the CRISTAL documentation (CRISTAL project team, 2022), ship classes are categorised according to their draft measured in centimetres and cargo capacity

measured in tons. The dataset, sourced from the project database, is stored in a CSV file format (see table 11). This file is a valuable resource for the study, providing essential information about the characteristics of different ship classes, which is vital for the investigation into navigability at shallow points.

The file is 10 KB and contains the classification of classes based on draft measured in centimetres and cargo weight measured in tons.

Draft [cm]	Class IV		Class V	
	From [ton]	To [ton]	From [ton]	To [ton]
140	370	620	790	880
160	700	750	960	1060
180	820	870	1130	1230
200	950	1000	1290	1410
220	980	1130	1460	1600
250	1280	1320	1720	1860

Table 11: Description of ship classes categorised by a draft length in centimetres and cargo capacity in tons.

The data related to sand sedimentation, which studies such as Whitmeyer and FitzGerald (2006) and Mugade and Sapkale (2015) suggest has a significant impact on river depth at shallow points, is unavailable and cannot be provided by the AIPo authority for this analysis due to technical difficulties. Alternative strategies for addressing this limitation are discussed in detail in section 8.1.

4.2. Data Description

This thesis focuses on three strategically selected shallow points along the Po River: Piacenza, Monte P.Te Revere and Cavanella. These points reside within the Piacenza, Revere and Volta Grimana navigable segments as described in table 12. The selection aimes to obtain representative data for each navigable segment. Piacenza and Cavanella lie at the ends of the navigable section, capturing potential variations in water depth due to their starting and finishing positions. Monte P.Te Revere occupies a central location, providing insight into conditions within the segment.

Shallow point	Segment	Program km	Latitude	Longitude
Piacenza	Piacenza	342.259	45.057632	9.709441
Monte P.Te Revere	Revere	523.660	45.055091	11.134256
Cavanella	Volta Grimana	630.987	45.019353	12.135474

Table 12: Critical shallow points examined in this thesis, along with their distances from the origin river and corresponding geographical coordinates.

To determine the optimal upstream water level station for monitoring each shallow point, the geographical coordinates of potential stations are acquired and the distances are calculated using the Haversine formula. The findings are detailed in table 13.

Shallow Point	Water Level Station	Distance (km)
Piacenza	Piacenza	0.428
Cavanella	Cavanella SIAP	2.489
Monte P.Te Revere	Revere SIAP	0.021

Table 13: Proximity distances in kilometres from each shallow point to the respective upstream water level monitoring stations.

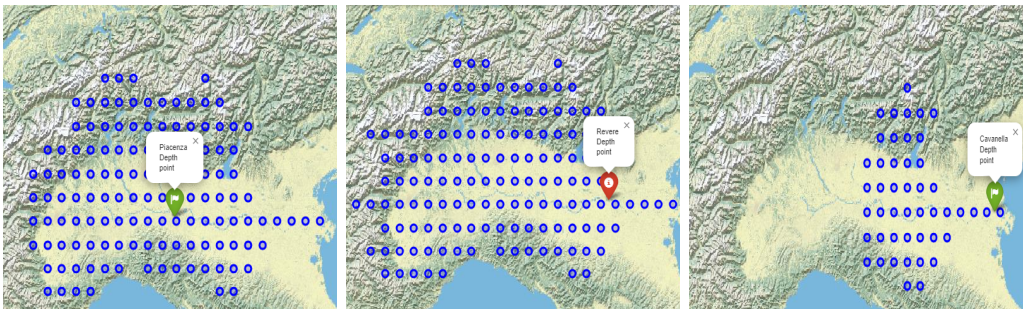
As outlined in section 4.1, the required river depth and discharge data for these points are extracted from the file provided by the AIPo authority and upstream water level data is sourced from API request. Upon consolidating this data, the descriptive statistics of the unprocessed sourced data for each shallow point are presented in the table 14.

	Piacenza			Monte P.Te Revere			Cavanella		
	Depth (cm)	Discharge rates (m^3/s)	Water level (cm)	Depth (cm)	Discharge rates (m^3/s)	Water level (cm)	Depth (cm)	Discharge rates (m^3/s)	Water level (cm)
Count	6170	6170	1165433	8455	8455	153648	11369	11369	314878
Mean	257.24	852.64	4232.46	448.18	1293.20	905.13	339.04	1448.78	574.57
Std	103.56	677.62	106.09	155.33	960.30	160.45	68.99	1037.50	74.03
Minimum	100.00	178.92	4058.00	140.00	218.07	571.00	170.00	0.00	266.00
25%	200.00	484.85	4169.00	350.00	721.95	802.00	300.00	834.8 7	529.00
Median	230.00	648.20	4203.00	410.00	990.24	869.00	340.00	1126.6 8	554.00
75%	290.00	968.85	4260.00	520.00	1541.49	976.00	350.00	1707.38	597.00
Maximum	980.00	7728.19	5846.00	1280.00	11752.20	1785.00	990.00	9516.62	976.00

Table 14: Descriptive statistics of raw data collected from all three shallow points.

The climate dataset, sourced from `2m_temperature.nc`, `2m_Total_PRECIPITATION.nc` and `snow_depth.nc` files, is consolidated into a single dataset based on geographical coordinates. The climate data is merged using the Haversine formula with a radius of influence set at 200 km from each shallow point. This consolidated dataset comprises information from 131, 80 and 46 and weather locations from Piacenza, Monte P.Te

Revere and Cavanella, shallow points, respectively. Each of these stations represents geographical locations within the radius of the influence of the river catchment area. These data frames encompass daily observations of temperature, precipitation and snow depth spanning from 01 January 2013 to 31 December 2022 in daily frequency.



(a) Piacenza (b) Monte P.Te Revere (c) Cavanella

Figure 5: Geographical map displaying considered weather data points (blue dots) within the river catchment area, located within a 200 km radius from the shallow points (represented by markers).

4.3. Data Refinement

The data refinement phase encompasses essential processes to ensure the quality and reliability of the dataset used in this study. It includes better data-cleaning procedures to address inconsistencies, missing values, or outliers within the collected data. Additionally, efforts are made to manage data variations using imputation techniques across all attributes. By thoroughly refining the dataset through these procedures, it establishes a robust foundation for accurate analysis and quality of results.

4.3.1. Missing Data

Although the raw data sourced appeared appropriate, initial analysis of the unprocessed information revealed significant missing values across all attributes associated with each shallow point. Notably, the Monte P.Te Revere point lacked complete river depth and discharge data from 1988 to 1999, representing a substantial 48.5% gap. Similar patterns are observed with Piacenza and Cavanella, where crucial river depth and discharge data are absent for portions of 2004 and 2005 in Piacenza and 2019 to 2022 in Cavanella. Further analysis revealed missing values related to all three critical shallow points in the upstream water level data. In particular, Cavanella exhibits the highest proportion of

missing entries at 24.33%, compared to 9.69% and 1.48% for Monte P.Te Revere and Piacenza, respectively. Table 15 describes the complete information on missing data.

Shallow Point	Variable	Available Data	Freq	Total	Missing	Percentage
Piacenza	River Depth	2004-01-03 to 2022-05-11	Daily	6170	535	8.67%
	River Discharge	2004-01-03 to 2022-05-11	Daily	6170	535	8.67%
	Water Level	2006-01-01 to 2024-01-01	30 Min	1183032	17600	1.48%
Monte P.Te Revere	River Depth	1988-01-01 to 2022-05-12	Daily	8455	4097	48.45%
	River Discharge	1988-01-01 to 2022-05-12	Daily	8455	4097	48.45%
	Water Level	2009-04-15 to 2024-01-01	30 Min	280018	27136	9.69%
Cavanella	River Depth	1988-01-01 to 2022-05-05	Daily	11369	1176	10.34%
	River Discharge	1988-01-01 to 2022-05-05	Daily	11369	1176	10.34%
	Water Level	2002-04-10 to 2024-01-01	30 Min	415348	101090	24.33%

Table 15: Details on missing data in the raw dataset for river depth, river discharge rate and water level variables at the selected shallow points.

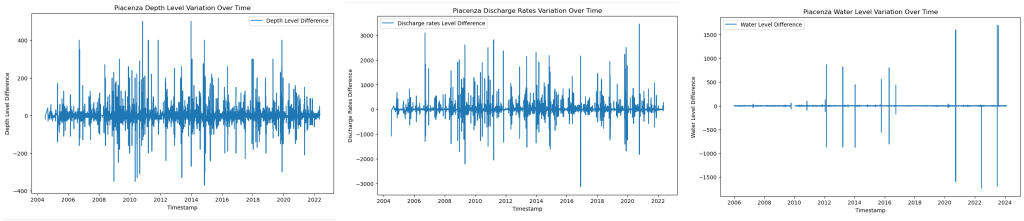
Due to quality concerns, information regarding continuous missing data for river depth and discharge at all three shallow points is excluded from the analysis. Instead, the forward-fill method addresses missing values in intermittent data based on the Last Observation Carried Forward (LOCF) approach. This method replaces missing values with the most recent valid data point, assuming no change during the missing period. Given its 30-minute frequency, the forward fill method is again employed for water level data. Here, missing values are substituted with the preceding valid data point, assuming short-term stability in water level changes.

The consolidated climate dataset within the River Po basin is of high quality, with complete and accurate information. There are no missing values or data inconsistencies.

4.3.2. Data Variation

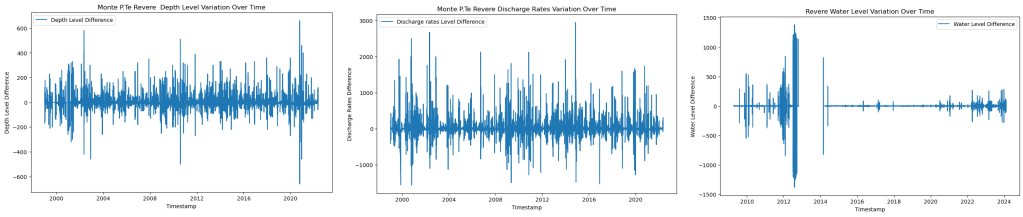
Descriptive statistics from the raw dataset (see table 14) reveal significant variability in water depth, discharge rates and water levels data collected from three locations. Notable differences are observed in the values between the present day and its preceding day at each location, as described in the accompanying graphs (see figures 6, 7 and 8). Both artificial and natural morphological factors influence these variations. For instance, Piacenza exhibits a shallower range of depth measurements (100 cm to 980 cm) compared to Monte P.Te Revere (140 cm to 1280 cm) and Cavanella (170 cm to 990 cm). Similarly, the standard deviation of discharge rates is notably higher for Cavanella (1037.50 m^3/s) compared to Piacenza (677.62 m^3/s) and Monte P.Te Revere (960.30 m^3/s), indicating

more significant variability in flow rates at shallow points. These findings underscore the importance of understanding data variability in hydrological analysis, as it influences feature engineering and the quality of forecasting results.



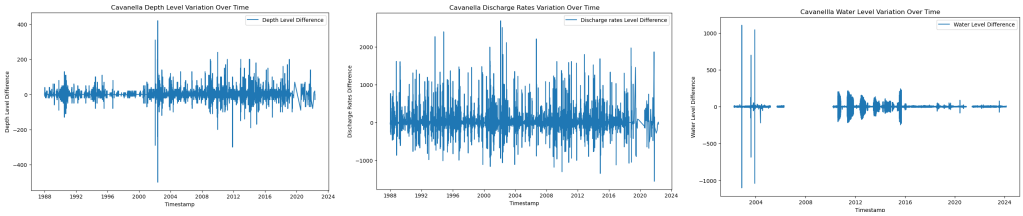
(a) River depth (b) River discharge rates (c) Water level

Figure 6: Plot explaining the data variation observed across different components in the Piacenza dataset.



(a) River depth (b) River discharge rates (c) Water level

Figure 7: Plot explaining the data variation observed across different components in the Monte P.Te Revere dataset.



(a) River depth (b) River discharge rates (c) Water level

Figure 8: Plot explaining the data variation observed across different components in the Cavanella dataset.

In the quality analysis of hydrological data, variability is standardised and managed through imputation techniques. Specifically, when the variation in river depth exceeds 200 cm with the preceding day, the data is imputed with values from the previous day. Similarly, for discharge rates, if the change exceeds 1000 m^3/s, imputation methods are

applied to address the variability. Additionally, in the case of water level fluctuations exceeding 100 cm, imputation techniques are employed to ensure data consistency and reliability. These standardised procedures to address variability ensure the integrity and precision of the hydrological dataset, facilitating robust analysis for forecasting.

4.4. Data Integration and Preparation

In preparation for daily and hourly time series forecasting, the initial step involves resampling the 30-minute water level data by averaging values into daily intervals and hourly intervals to match the temporal granularity of the depth and discharge data. Concurrently, the daily depth and discharge data are linearly interpolated to create hourly data points, ensuring a consistent temporal resolution across all variables. This process entailed aggregating the 30-minute measurements to calculate daily averages or totals. Subsequently, individual consolidated data frames are created for hourly and daily analysis, containing the interpolated hourly depth and discharge data and the resampled water level data. These consolidated data frames provide a comprehensive dataset suitable for time series forecasting at hourly and daily intervals, enabling the application of appropriate forecasting models and techniques to derive insights and make informed decisions.

In the context of climate data and considering the geographic characteristics of the Po River's catchment area detailed in section 4.1.1, the GeoJSON file outlines the geological boundaries of the Po River basin. This information is utilised to consolidate a unified dataset comprising temperature, total precipitation and snow depth. This dataset is filtered and restructured based on geographical coordinates to encompass all three shallow points within the basin's territory.

5. Methodology

In this section, the in-depth discussion of the methodology includes time series analysis approaches like stationarity testing, seasonal decomposition and auto correlation functions to validate assumptions for the time series data. Various functional machine learning algorithms like ANN, RNN and CNN are discussed, and differences among LSTM, GRU, LSTNet and VAR models are explained in detail from a multivariate time series forecasting point of view. Finally, the tuning approach and evaluation metric are discussed in detail and applied in the study. The motive behind selecting the statisti-

cal methods are thoroughly explained, emphasising their alignment with the research questions.

5.1. Time Series Analysis

Referring to the data described in the section 4, aside from the geographical information of the case study, the fundamental nature of the data of hydrological variables is time series data.

According to Brockwell and Davis (1991), a time series is typically defined as a family of random variables (RV's) that can be real-valued ($\mathbb{R}$), vector-valued in $\mathbb{R}^k$, or even complex-valued ($\mathbb{C}$), denoted as $X_t, t \in T$. The index set T can be the set of natural numbers ($\mathbb{N}$) or the set of integers ($\mathbb{Z}$). This collection is called a time series or a time series process. This section contributes to the necessary methods utilised to analyse the time series data.

5.1.1. Stationarity

Understanding the behaviour of time-series data is a primary procedure to gain insights into the nature of data. Stationarity is a fundamental concept in time series analysis, playing a crucial role in understanding temporal patterns in data. A time series is considered stationary when its key statistical characteristics, such as mean, variance or auto correlation, remain consistent and do not vary with time (Witt et al., 1998).

According to Brockwell and Davis (1991, def 1.3.2), stationarity is defined as time series $\{X_t, t \in \mathbb{Z}\}$, with the index set $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$, is said to be stationary if:

1. $E[X_t] = \mu$ for all $t \in \mathbb{Z}$,
2. $E|X_t|^2 < \infty$ for all $t \in \mathbb{Z}$, and
3. $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ for all $r, s, t \in \mathbb{Z}$.

Where $E[X_t] = \mu$ as expectation; $E|X_t|^2$ as variance and the auto covariance function $\gamma_X(r, s)$ of $\{X_t\}$ is defined by $\gamma_X(r, s) = \text{Cov}(X_r, X_s) = \mathbb{E}[(X_r - \mathbb{E}[X_r])(X_s - \mathbb{E}[X_s])], \quad r, s \in \mathbb{Z}$ for the series (Brockwell and Davis, 1991, p.12).

5.1.2. Augmented Dickey Fuller (ADF) Test

The stationarity of a time series plays a crucial role in shaping its properties and forecasting its behaviour. Failing to transform a time series into the appropriate form of stationarity can lead to misleading results (Greunen et al., 2014). Testing the stationarity of the data is necessary to identify underlying patterns and meaningful trends. The ADF is a statistical approach that helps determine whether the data is stationary. In this methodology, the ADF test is considered optimal due to its wide recognition as a valuable tool to check stationarity in time series data (Shumway and Stoffer, 2017).

According to Ajewole et al. (2020), the unit root test sets the foundation for the ADF test, which is performed based on a first order autoregressive AR(1) process. The goal is to eliminate dependence between the current value and its lagged value, which can indicate the series has a unit root.

Based on the equation defined by Ajewole et al. (2020), the relation between current value x_t at time t and its last lagged value x_{t-1} an be represented as:

$$x_t = \phi x_{t-1} + w_t \tag{1}$$

Here x_t is the observation of the current value at time t, x_{t-1} as the last value with time $t-1$, ϕ is the autoregression coefficient and w_t as white noise in the time series. In the perspective of the unit root test, the time series x_t converges to a stationary time series as $t \to \infty$, if $|\phi| < 1$. This condition ensures that the series stabilises over time and its statistical properties, such as mean and variance remain constant. Conversely, if $|\phi| > 1$, the series x_t is not stationary and its variance becomes time dependent. A value of $|\phi| = 1$ indicates the presence of a unit root, leading to a non-stationary (Ajewole et al., 2020).

Equation 1 reformed as differenced autoregressive AR(1) process with equation $\Delta x_t = \delta x_{t-1} + w_t$, where Δx_t is $x_t - x_{t-1}$ and δ is $\phi - 1$ which acts a basis for unit root test in ADF.

The ADF test involves evaluating the stationarity of a time series using ARIMA process, which typically includes additional terms such as a constant, trend and moving average components to comprehensively analyse the time series dynamics (Ajewole et al., 2020). The ADF test involves checking and testing based on the following equation:

$$\Delta x_t = \alpha + \delta t + \lambda^{-1} x_{t-1} + \sum_{j=1}^{p} \beta_j t_j \Delta x_{t-j} + w_t \tag{2}$$

Where α represents the constant term or intercept, δt represents a linear trend component with δ as the coefficient of time t, $\lambda^{-1} x_{t-1}$ is the autoregressive (AR) component with λ^{-1} as the coefficient of the lagged value x_{t-1}, $\sum_{j=1}^{p} \beta_j t_j \Delta x_{t-j}$ represents the moving average component of the model, where β_j as coefficients for the lagged differenced terms Δx_{t-j} up to order p (Ajewole et al., 2020).

According to statistical testing methodology from Mushtaq (2011) and Ajewole et al. (2020), the procedure involves formulating the null hypothesis H_0 when autoregressive coefficient $\phi = 1$ indicating a unit root and non-stationarity with the alternative hypothesis H_1 when $\phi < 1$ suggesting stationarity and is stated as:

$$H_0 : \phi = 1 \quad \text{vs} \quad H_1 : \phi < 1 \tag{3}$$

The ADF test generates a test statistic based on estimated values of constant term α, linear trend coefficient δ, autoregressive coefficient λ, and moving average coefficient β. The test statistic is compared to appropriate critical values in the Dickey Fuller table for decision making based on the significance level. If the test statistic is below the critical value, the null hypothesis of non-stationarity is rejected, indicating stationarity. Conversely, if the test statistic exceeds the critical value, the null hypothesis cannot be rejected, suggesting non-stationarity. This process provides a robust means to evaluate the stationarity of a time series based on ADF test results (Ajewole et al., 2020).

5.1.3. Seasonal Decompose

The initial step in time series analysis involves plotting the data for visual analysis. If discontinuities, such as changes in level, are observed in the series, it helps to improve the analysis by partitioning the series into homogeneous segments (Brockwell and Davis, 1991). This methodology aims to facilitate a clearer understanding of the underlying patterns and variations in the time series data.

As per Montgomery et al. (2011, p.42), the additive decompose method is the fundamental approach to breaking down a time series into seasonal, trend and residual

components. For a time series $X_t, t \in T$ with x_t observations at time period $t = 1, 2, \ldots$, the additive model is expressed as:

$$x_t = S_t + L_t + \epsilon_t \tag{4}$$

Where $S_t = S_{t+s} = S_{t+2s} = \ldots$ for $t = 1, \ldots, s-1$, $t, s \in T$ with s as a length of the period of cycles, $L_t = \beta_0 + \beta_1 t$ represent the linear trend component with β_0, β_1 as coefficients of time t and ϵ_t represents residual component which is uncorrelated with mean 0 and constant variance (Montgomery et al., 2011, p.210).

In this thesis, the additive model is chosen based on stationarity analysis, as it is suitable when the magnitude of the seasonal variation does not vary over time.

5.1.4. Sample Auto Correlation Function (ACF)

In terms of data knowledge exploration, understanding the presence of patterns or dependencies within the time series and any cyclic behaviour over the period to be addressed. Auto correlation is a valuable approach to exploring and analysing time series data. It helps to indicate the degree of similarity between a time series and a delayed version of itself (Montgomery et al., 2011).

According to Montgomery et al. (2011), If a time series is stationary, it implies that the joint probability distribution of any two observations, x_t and x_{t+h}, remains constant for any two time periods t and $t + h$ separated by the lag h. This condition allows for the satisfaction of the assumption underlying the sample autocorrelation function.

As per Shumway and Stoffer (2017, def 1.15), the sample auto correlation function has a sampling distribution that allows to assess whether the data comes from a completely random source. The correlation between a sample time series X_t and its lagged values can be calculated using the sample auto correlation function. The sample auto correlation function denoted as $\rho(h)$ of a stationary process can be expressed as:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}, \tag{5}$$

Here $\rho(h)$ is the correlation coefficient. Where $\gamma(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})$ is the auto-covariance for the observations x_t and x_{t+h}, at time $t = 1, \ldots, n$ and for lag $h = 0, 1, \ldots, n-1$. $\bar{x}$ is the mean of all observations and also $\gamma(-h) = \gamma(h)$, for lag h. Auto correlation coefficient $\rho(h)$ lies in range of $-1 < \rho(h) < 1$ and $\rho(h = 0) = 1$ by definition. When plotted, auto correlation is interpreted by observing decay or spikes at specific lags. These patterns indicate auto-regressive behaviours or seasonality within the time series (Shumway and Stoffer, 2017).

5.1.5. Partial Auto Correlation Function (PACF)

The partial auto correlation function is similar to the sample auto correlation function. For deeper understanding within the time series, the partial auto correlation function considers only the direct correlation at each lag after removing the correlations explained by the intermediate lags (Shumway and Stoffer, 2017, p.105).

As per Shumway and Stoffer (2017, def 3.9), the partial auto correlation function of a stationary time series X_t with observation x_t and x_{t+h} for time $t = 1, 2, \ldots, n$ and for lag h. The PACF is denoted as $\phi(hh)$ for $h = 1, 2, \ldots, n-1$ and is expressed as:

$$\phi(hh) = \text{corr}(x_{t+h} - \bar{x}_{t+h}, x_t - \bar{x}_t) \tag{6}$$

Where $\phi(hh)$ is the correlation coefficient between x_{t+h} and x_t, for $h \geq 1$ is the lag and $\bar{x}_{t+h}$, $\bar{x}_t$ being the mean of all observations of x_{t+h} and x_t respectively. Partial auto correlation measures the correlation (corr) between x_{t+h} and x_t with eliminating the linear connection of $x_{t+1}, \ldots, x_{t+h-1}$. This is achieved by $x_{t+h} - \bar{x}_{t+h}$ where the deviation of x_{t+h} from its mean $\bar{x}_{t+h}$, indicating how much x_{t+h} varies from its average value. Similarly, $x_t - \bar{x}_t$ term represents the deviation of x_t from its mean $\bar{x}_t$ (Shumway and Stoffer, 2017, p.105).

In terms of interpretation, PACF coefficient $\phi(hh)$ ranges from $-1 < \phi(hh) < 1$. When plotted, each significant spike in the PACF plot corresponds to the correlation between the series at the current time point and the series at that specific lag, with the influence of the intermediate lags removed. The lag values associated with these significant spikes indicate the potential order of the auto-regressive component in the time series model. The PACF values also exhibit an exponential decay to zero after a certain lag. The

point where these PACF values become negligible explains the order of the autoregressive component in the time series model (Shumway and Stoffer, 2017, p.160).

5.1.6. Cross Correlation Function

In the context of this thesis, the problem statement necessitates addressing multiple variables within the dataset, highlighting the need for multivariate time series analysis. It is crucial to look into the relationships between variables, highlighting the complex interplay among various factors. The cross correlation is a valuable method to quantify the metrics that identify the relationships between two variables. This approach enables the interpretation of the underlying patterns and influential attributes, particularly concerning model training and forecasting objectives (Shumway and Stoffer, 2017).

The approach is similar to auto correlation, focusing on the lag feature. However, in cross correlation, one attribute remains fixed and the lag of the other attribute is systematically shifted. This technique identifies and quantifies the effects based on the lag, offering insights into the relationships between the two attributes over time.

Based on the notation defined in section 5.1.4 and as per Shumway and Stoffer (2017, def 1.11), The cross correlation function denoted as $\rho_{xy}(h)$ of jointly stationary time series$(X_t)_{t\in\mathbb{N}}$ and $(Y_t)_{t\in\mathbb{N}}$ with observations x_t and y_t respectively for time $t = 1, 2, \ldots, n$ and for lag h is defined as:

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_x(0)\gamma_y(0)}} \tag{7}$$

Where $\gamma_{xy}(h)$ is the cross covariance function and as per Shumway and Stoffer (2017, def 1.10) it is defined as :

$$\text{cov}(x_{t+h}, y_t) = E\left[(x_{t+h} - \mu_x)(y_t - \mu_y)\right] \tag{8}$$

Here $\text{cov}(x_{t+h}, y_t)$ calculates expectation value (or mean) for product of $(x_{t+h} - \mu_x)$ and $(y_t - \mu_y)$. Where x_{t+h} is the observations at time $t = 1, 2, \ldots, n$ and lag $h = 1, 2, \ldots, n-1$. $\mu_x = E[x_{t+h}]$ is the mean of all x_{t+h} observations and y_t is the fixed observation at time $t = 1, 2, \ldots, n$ and with mean $\mu_y = E[y_t]$.

The cross correlation coefficient $\rho_{xy}(h)$ takes values between -1 and 1. It is not symmetric about zero, meaning that $\rho_{xy}(h) \neq \rho_{xy}(-h)$. This coefficient provides valuable insights into the relationship between two attributes, x_{t+h} and y_t, over a specific lag h. The correlation coefficient equals 1, which indicates a strong linear relationship between the attributes. Conversely, a coefficient of -1 indicates a negative relationship, while 0 indicates no relationship (Shumway and Stoffer, 2017, p.26).

5.2. Time Series Forecasting

As described in section 5.1, time series data contains information on real-time observations captured over different periods. This data, which is collected from past events (also referred to as historical data), allows to identify trends, seasonality and cyclic temporal patterns that occurred in the past and may also project similar behaviour in the future. Based on the past observations, it is possible to predict future observations (Brockwell and Davis, 1991).

5.2.1. Multivariate Time Series and Forecasting

In a multivariate time series forecasting context, two or more variables containing historical information are involved in forecasting future instances of a target variable.

According to Brockwell and Davis (1991, def 11.1), a multivariate time series $\mathbf{X}_t$ is expressed in a vector form as:

$$\mathbf{X}_t = (X_{1t}, \ldots, X_{mt})' \text{ for } t = 0, \pm 1, \pm 2, \ldots \tag{9}$$

Where $X_{1t}, \ldots, X_{mt}$ represents m individual components of the vector $\mathbf{X}_t$ for the time period t. In multivariate time series, the serial dependence of each component series X_{it}, for $i = 1, ..., m$ and also the interdependence between different component series X_{it} and X_i are considered. To forecast $\mathbf{X}_{t+h}$, where h represents the desired forecasting horizon ahead of the current timestamp t, it is assumed that $(X_{1t}, \ldots, X_{mt})'$ are available (Brockwell and Davis, 1991).

Artificial Neural Networks (ANN)

The utilisation of ANN is prevalent in the realm of time series forecasting (Atashi et al., 2022). These networks offer the advantage of linear and nonlinear modelling without requiring prior information or assumptions about the correlation between input and output variables. In multivariate time series forecasting, a Multi Layer Perceptron (MLP) is a type of ANN that is structured with input, hidden and output layers accompanied by an activation function connected in a finite acyclic graph. Through these layers, input variables can be trained to predict upcoming instances of target variable (Hamzaçebi et al., 2009).

Based on the explanation from (Hamzaçebi et al., 2009), the forecasting of future instances is achieved by:

$$F_{t+l} = \alpha_l + \sum_{j=1}^{m} v_{jl} f \left(\sum_{i=1}^{k} w_{ij} \mathbf{X}_{t-i} + \theta_j \right) \tag{10}$$

In this formula, the variable $\mathbf{X}_{t-i}$, for $(i = 1, 2, \ldots, k)$ represents previous observations of the multivariate time series for the past k periods. The predictions for the current and future n periods are denoted by F_{t+l}, where $l = 0, 1, 2, \ldots, n$. The term $\sum_{i=1}^{k} w_{ij} \mathbf{X}_{t-i}$ computes a weighted sum of values of the multivariate time series $\mathbf{X}_t$ where weights of connections from input layer neurons to hidden layer neurons are represented by w_{ij}, for $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, m$. The weights of connections from hidden-layer neurons to output layer neurons are represented by v_{jl}, for $j = 1, 2, \ldots, m$ and $l = 0, 2, \ldots, n$. The weights of bias connections are denoted by α_l and θ_j, where $l = 0, 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. Finally, f represents the activation function (Hamzaçebi et al., 2009). Based on the estimated weights and bias when trained on historical data, MLP network can forecast future instances.

However, ANNs such as MLPs can effectively address specific tasks. Nevertheless, they present some drawbacks when utilised for multivariate time series forecasting challenges. One particularly challenging aspect is capturing long-term dependencies in sequential time series, as ANNs treat each input as independent (Box et al., 2015). Complex designs with more neurons may also lead to overfitting and increased computational costs. Such disadvantages led to the development of CNN and RNN, which can handle sequential modelling and efficient training (Yao et al., 2017).

Convolutional Neural Network (CNN)

The possible alternative for traditional ANNs is Convolutional Neural Network (CNN)) (Yao et al., 2017). CNNs, a type of ANN, have gained popularity due to their success in classification problems such as image recognition (Krizhevsky et al., 2012) and time series classification (Wang et al., 2016). The CNN comprises a series of convolutional layers. These layers are designed to only connect to local regions within the input data. The connection is achieved by sliding a weight matrix and filter over the input data. A dot product is computed between the input and filter at each point, essentially a convolution. The structure enables the model to learn filters to identify specific patterns in the input data. The CNNs has an advantage over RNNs due to its convolutional structure, which results in fewer trainable weights, making it more efficient for training and predicting (Borovykh et al., 2018).

According to Lai et al. (2018) and in the perspective of multivariate time series input $\mathbf{X}_t$, the computation of a convolutional layer in a neural network is represented as:

$$h_k = \text{ReLU}(W_k * \mathbf{X}_t + b_k) \tag{11}$$

CNN use a convolution operation, denoted by $*$, where $W_k * \mathbf{X}_t$ represents the convolution operation between the filter W_k and the input $\mathbf{X}_t$ for the time $t \in T$ and b_k is a bias which adds to output. The output of this operation is a vector and is represented as h_k. To ensure that each h_k vector has a length of T, the input matrix $\mathbf{X}_t$ is zero-padded to the left (Lai et al., 2018).

The convolutional layer contains multiple filters, each with a width of w and a height of m (the same as the number of variables). Each filter moves across the input matrix $\mathbf{X}_t$ and the ReLU function is applied to each element of the resulting vector. The ReLU function is defined as $\text{ReLU}(x) = \max(0, x)$. The output matrix of the convolutional layer has a size of $d_c \times T$, where d_c is the number of filters used and T is the length of the time series (Lai et al., 2018).

The output from all convolutional layers is represented as output matrix H and global pooling is applied to summarise the learned features across all filters. These layers are followed by a dense layer that learns to map the extracted features from the convolutional layers to predict future values of the target variable (Lai et al., 2018).

Recurrent Neural Networks (RNN)

RNN are initially proposed by Elman (1990) as a type of ANN. Unlike traditional ANNs and CNNs, RNNs incorporate an internal state, called the hidden state, to capture temporal dependencies using recurrent connections in sequential data.

Based on the explanation from Hewamalage et al. (2021), basic RNN is formulated as:

$$h_t = \sigma(W_i \cdot h_{t-1} + V_i \cdot x_t + b_i) \tag{12}$$

$$z_t = \tanh(W_o \cdot h_t + b_o) \tag{13}$$

Where $h_t \in \mathbb{R}^d$ represents the hidden state at time t for d cell dimension; σ denotes sigmoid activation function, which generates the output between 0 and 1. $W_i \in \mathbb{R}^{d \times d}$ is the weight matrix for the recurrent connections for the previous hidden state h_{t-1}. The term $x_t \in \mathbb{R}^m$ (where $x_t \in \mathbf{X}_t$ for multivariate with m size of the input) is the input for the cell with $V_i \in \mathbb{R}^{d \times m}$ as weight matrix. The $b_i \in \mathbb{R}^d$ denotes the bias vector for the hidden state h_t. Likewise, $z_t \in \mathbb{R}^m$ represents output of the cell at time step t for the result from tanh as activation function with $W_o \in \mathbb{R}^{d \times d}$ signify the weight matrix for the h_t hidden state at time t for d cell dimension and $b_o \in \mathbb{R}^d$ signify the bias vector of the cell output. The current hidden state depends on the hidden state of the previous time step and the current input. After training the model, the last known input-output sequence from the training set iteratively predicts future outputs by feeding the predicted output z_t back into the model as the following input x_{t+1} (Hewamalage et al., 2021).

However, RNNs difficulty capturing long-term dependencies because of the vanishing gradient problem, where gradients diminish as they move back through time. This restricts their effectiveness in tasks that require memory over long sequences. Since the data used in this work depends on long term sequences, traditional RNNs do not support the current requirement. The LSTM and GRU are types of RNNs that overcome the limitations of RNNs in capturing long-term dependencies (Yao et al., 2017).

5.2.2. Long Short Term Memory (LSTM)

LSTM, proposed by Hochreiter and Schmidhuber (1997), is a type of RNN that has been widely utilised for time series forecasting purposes and its success has led to its adaptation in this thesis. Based on the tutorial for LSTM from Staudemeyer and Morris

(2019), it is summarised as an artificial neuron cell with internal memory and consists of three gates namely the input gate, the forget gate and the output gate. The previous time's output is considered input for the current time. The structure contains a loop that repeats the same task for all data across the sequence of input vectors and the output from the previous computation. This architecture enables LSTM to effectively capture and retain information over extended sequences, making them well-suited for tasks involving time series data. The mechanism allows to store or forget information selectively. In this work, the sequence of historical time series is trained on an LSTM architecture-based model to forecast future steps based on the previous sequence.

According to Borwarnginn et al. (2022), the following equations describe the flow of information through an LSTM cell at different time steps:

$$\begin{aligned}
\text{Input gate } (i_t) &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
\text{Forget gate } (f_t) &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
\text{Output gate } (o_t) &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
\text{Candidate cell state } (\tilde{C}_t) &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\
\text{Updated cell state } (C_t) &= f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \\
\text{Hidden state } (h_t) &= o_t \odot \tanh(C_t)
\end{aligned} \tag{14}$$

The provided equations describe the computations involved in the neural network's cell state and hidden state evolution. This process integrates the current input x_t, the previous hidden state h_{t-1} and the preceding cell state C_{t-1}. These variables are crucial in generating the new cell state C_t through the candidate cell state vector $\tilde{C}_t$. At the same time, the hidden state h_t is computed. The significance of weights W_i, W_f, W_o, W_c and biases b_i, b_f, b_o, b_c are essential throughout the calculations, particularly at the gates and cell states. The input gate i_t, forget gate f_t and output gate o_t are computed using the sigmoid function σ, which is crucial for regulating information flow (Staudemeyer and Morris, 2019).

The candidate cell state $\tilde{C}_t$ is determined through the hyperbolic tangent function tanh and the updated cell state C_t is a combination of the previous cell state and the candidate cell state, controlled by the input and forget gates. Finally, the hidden state h_t is generated by applying the output gate to the hyperbolic tangent of the updated cell state. Prediction or forecasting can be achieved by passing the final hidden state through additional layers or by applying a linear transformation, depending on the specific architecture and task of the network (Staudemeyer and Morris, 2019).

For multivariate time series forecasting, the historical dataset $x_t \in \mathbf{X}_t$ is divided into sequences based on identified patterns in the EDA section 7. The first input sequence initialises the LSTM cell's hidden state h_0 and cell state C_0. Each element of the input sequence is sequentially processed through the network. For each time step t, the updated cell state C_t and hidden state h_t are computed using the provided equations 14.

The last historical data sequence is fed to the trained LSTM model to forecast future values. The LSTM will use its learned parameters (weights W_i, W_f, W_o, W_c and biases b_i, b_f, b_o, b_c) to update its internal states and generate predictions for each subsequent time step. The predicted output h_t at time t is used to predict h_{t+1} at the next time step. This process is repeated to forecast multiple future time steps (Staudemeyer and Morris, 2019).

5.2.3. Gated Recurrent Unit (GRU)

The GRU, a type of RNN, introduced by Cho et al. (2014a) distinguishes itself from LSTM networks through its simplified structure, featuring only two gates namely reset and update gates. These gates combine aspects of the input and forget gates found in LSTM. GRU has demonstrated notable performance in handling long sequences, proving efficient and effective outcomes in time-series forecasting (Lin et al., 2022). In the context of this thesis, GRU emerges as the preferred method due to its faster training times and efficient results. Additionally, GRU is considered a potential alternative for time series forecasting, providing a basis for comparing forecasted results against LSTM outcomes. This choice allows for a research contribution regarding model comparison, specifically in hydrological components like river depth forecasting.

According to Cho et al. (2014b), the structure of GRU is described in the following equations:

$$\begin{aligned} \text{Update gate } (z_t) &= \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ \text{Reset gate } (r_t) &= \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ \text{Candidate hidden state } (\tilde{h}_t) &= \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \\ \text{Hidden state } \ (h_t) &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \end{aligned} \tag{15}$$

In the GRU, the update gate z_t, determined by applying the sigmoid function σ to a weighted sum of the input x_t, controls necessary information from the weighted sum of the previous hidden state h_{t-1}. This gate controls the amount of information from

the last hidden state h_{t-1} to retain and how much of the candidate hidden state $\tilde{h}_t$ to incorporate into the new hidden state h_t. The reset gate r_t helps in deciding which parts of the past hidden state h_{t-1} to forget when computing the candidate hidden state $\tilde{h}_t$. Additionally, the candidate hidden state $\tilde{h}_t$ is computed based on the current input x_t, the reset gate r_t and the previous hidden state h_{t-1} (Cho et al., 2014a).

The update and reset gates are essential components of the GRU architecture, allowing it to update and utilise information from the past selectively and the current input. The network weight matrices and bias vectors indicated by (U_z, U_r, U_h), (W_z, W_r, W_h) and (b_z, b_r) are indeed parameters that are learned during the training phase through back-propagation and optimisation. These parameters are crucial for the GRU to adapt and capture patterns in sequential data effectively (Cho et al., 2014a).

Similar to LSTM, the final hidden state obtained after processing the input sequence can be used to forecast the next observation in GRU. This process is iterated using the input variables of the last training sequence combined with the newly forecasted value to predict the target feature for future observations.

5.2.4. Long and Short Term Temporal Network (LSTNet)

LSTNet is a hybrid neural network architecture designed especially for time series forecasting. It is proficient at capturing both long term dependencies and short term patterns in temporal data. Initially introduced by Lai et al. (2018), this architecture (see figure 9) integrates a one-dimensional CNN layer as described in 5.2.1 for handling local patterns and RNN layers 5.2.1 for capturing global patterns. Skip-RNN connections facilitate communication between these components, enabling the model to share information between short-term and long-term representations seamlessly.

Recent empirical studies from Yin et al. (2019) underscore the superior forecasting accuracy of LSTNet compared to individual implementations of CNN and RNN. A linear component Auto-regressive (AR) is also incorporated into the LSTNet model to overcome the drawback of deep neural networks, i.e., the scale of outputs is not sensitive to the scale of inputs. LSTNet has been selected as an effective approach for forecasting water levels due to its ability to capture both short-term and long-term patterns from historical data. The capability of LSTNet to handle missing attributes in forecasting scenarios makes it a suitable choice for utilisation in this thesis.

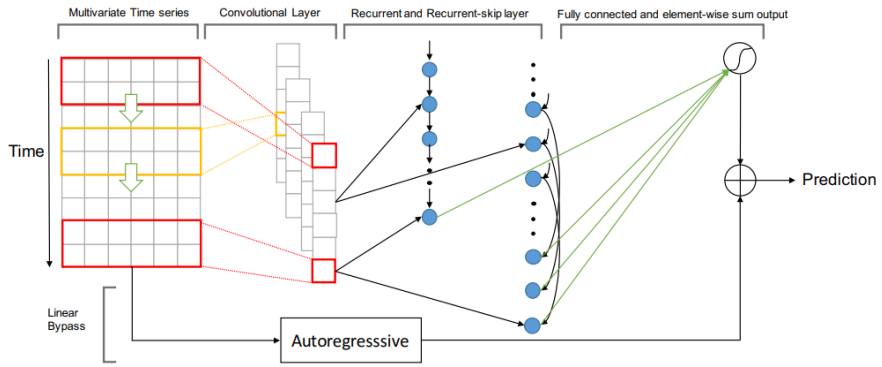


Figure 9: The architecture of the Long- and Short-term Time-series network (LSTNet), as described in Lai et al. (2018).

5.2.5. Vector Auto Regression (VAR)

VAR model is a statistical technique commonly employed to analyse multivariate time series data to tackle forecasting problems. The approach utilised in this work is based on the methodology proposed by Granger (1969) for analysing economic time series. The VAR model is a regression algorithm that studies the influence of two or more time series variables on each other. It considers both the moving average and autoregressive components of a time series, which can then be used to predict future observations in the variables. In this analysis, is perfectly suitable for understanding complex systems where variables interact based on the characteristics of the data.

According to the definition described by Lütkepohl (2005, def 2.1.1) and the notation used for multivariate time series in section 5.2.1, $\mathbf{X}_t$ for m different time series variables that are observed at discrete time points ($t = \pm 1, \pm 2, \ldots$). The Vector Auto Regression (VAR) model of order p, also known as VAR(p), can be expressed as follows:

$$\mathbf{X}_t = \boldsymbol{\nu} + A_1\mathbf{X}_{t-1} + A_2\mathbf{X}_{t-2} + \cdots + A_p\mathbf{X}_{t-p} + \boldsymbol{u}_t \tag{16}$$

Where $\mathbf{X}_t = (X_{1t}, \ldots, X_{mt})^\top$ is a $m \times 1$ multivariate time series vector for time t, variables m and also $(\mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \ldots, \mathbf{X}_{t-p})$ are the lagged vectors of $\mathbf{X}_t$. The $\mathbf{A}_i$ are fixed $m \times m$ coefficient matrices, $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_m)^\top$ is a fixed $m \times 1$ vector of intercept terms allowing for a possibly non zero mean $\mathbb{E}(\mathbf{X}_t)$. Additionally, $\mathbf{u}_t = (u_{1t}, \ldots, u_{mt})^\top$ represents a m-dimensional white noise characterized by $\mathbb{E}(\mathbf{u}_t) = \mathbf{0}$, $\mathbb{E}(\mathbf{u}_t\mathbf{u}_t^\top) = \Sigma_u$

and $\mathbb{E}(\mathbf{u}_t\mathbf{u}_s^\top) = \mathbf{0}$ for $s \neq t$. The covariance matrix $\Sigma_{\boldsymbol{u}}$ is assumed to be non singular (Lütkepohl, 2005, p.31).

Estimation of the coefficient $A_1, A_2, .., A_p$ is estimated using the maximum likelihood estimate (MLE) or ordinary least squares (OLS) regression based on input data. Once the model is estimated, it is used to forecast future values of the multivariate time series by recursively applying the estimated coefficients. The forecast of the new period h is described as:

$$\mathbf{X}_t(h) = \boldsymbol{\nu} + \mathbf{A}_1\mathbf{X}_t(h-1) + \cdots + \mathbf{A}_p\mathbf{X}_t(h-p) \tag{17}$$

The VAR(p) model follows recursive calculation to forecast future observations in horizon h. Prior known values of $(\mathbf{X}_t(0), \mathbf{X}_t(-1), \ldots, \mathbf{X}_t(-p+1))$ from historical data is used as the input for equation 17. Further, forecasted vector $\mathbf{X}_t(h)$ for the time t and horizon h is computed using lagged vector $\mathbf{X}_t(h-1), \ldots, \mathbf{X}_t(h-p)$. The process is recursively performed by updating lagged vectors $\mathbf{X}_t(h-1), \ldots, \mathbf{X}_t(h-p)$ with the newly forecasted values for the next iteration until the forecast horizon h (Lütkepohl, 2005, p.37).

5.3. Model Tuning

To optimise the performance and select optimal hyper-parameters for the Vanilla LSTM model, the `FLAML` library, which stands for Fast and Lightweight AutoML Library (FLAML) (Wang et al., 2021) is used in python 3.11.8 environment (Python Software Foundation, 2023). This library efficiently explores the hyper-parameter search space.

FLAML leverages the inherent structure of the hyper-parameter space, enabling it to intelligently determine an optimised search order that balances computational cost and prediction error. Given the constraints of the system's computational resources, FLAML streamlines the process of finding the best hyper-parameters for the model. This optimisation ensures better performance and significantly reduces the training time, allowing to predict results within a shorter time frame.

5.4. Evaluation Metrics

Evaluation metrics involve establishing criteria for measuring the performance or effectiveness of machine learning models. In this thesis, only one target value will be forecasted by a model. To measure the difference between the actual value (split of

original data) and the forecasted values (by models) or performance among the models, the Root Mean Square Error (RMSE) is considered. It measures the average magnitude of errors between predicted and actual values for each model. The RMSE is calculated as the square root of an average squared error difference (Borwarnginn et al., 2022).

As stated in Borwarnginn et al. (2022), the mathematical formulation of the RMSE is given by:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{h}(x_{it} - \hat{x}_{it})^2}{h}}$$

In a multivariate time series context, where x_{it} represents the actual value of the target variable X_{it} at time t, $\hat{x}_{it}$ is the forecasted value of X_{it} by a model trained on historical data from the time series $\mathbf{X}_t$ and h is the number of forecasted time steps. The RMSE serves as a critical performance metric based on calculated difference. A lower RMSE signifies higher model accuracy, indicating that the model's predictions closely align with the actual observed values (Borwarnginn et al., 2022).

When considering the data outlined in data description 4.1, RMSE values are calculated in centimetres as the metric for river depth forecast.

6. Feature Engineering of Hydrological Components

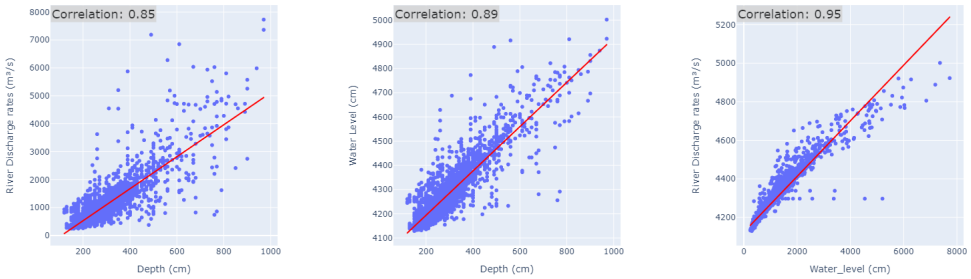
In river depth forecasting, which mainly focuses on shallow points, identifying influential factors, such as water level, river discharge rates and climate variables, holds critical importance. Based on studies described in the section 4.1.2, it is assumed that hydrological and climate attributes influence the river depth at critical points. To address the research question RQ1 of the thesis, the feature engineering process is structured to meet the specific requirements.

This section explains the significance of hydrological and climate attributes contributing to river depth data at shallow points such as Piacenza, Monte P.Te, Revere and Cavanella. This study aims to uncover valuable insights by comprehensively exploring the mutual relationship between these attributes, which involves calculating correlation coefficients. These insights will facilitate understanding and mitigation of the impacts of hydrological factors on river dynamics. For this study, scatter plots are generated using the `Plotly` library (Plotly Technologies Inc., 2015) of Python 3.11.8 (Python Software Foundation, 2023), enabling visual inspection of the relationships between the attributes.

6.1. Hydrological Components

To assess the effects of river discharge rates and upstream water levels on river depth levels, the cross correlation method described in section 5.1.6 is performed with a lag of 0. By aligning river depth with discharge and water level, the analysis captures the instantaneous response of river depth to discharge and water level changes.

The correlation coefficient values provide valuable insights into the relationship between river depth, river discharge rate and water level at the shallow points. From the scatter plot (see figure 10a), a correlation coefficient of 0.85 between river depth and river discharge rate at the Piacenza station suggests a strong positive linear relationship. The resultant correlation value indicates that changes in river discharge rates significantly impact river depth and higher discharge rates generally correspond to increased river depths. This association highlights the importance of considering river discharge dynamics when assessing variations in river depth.



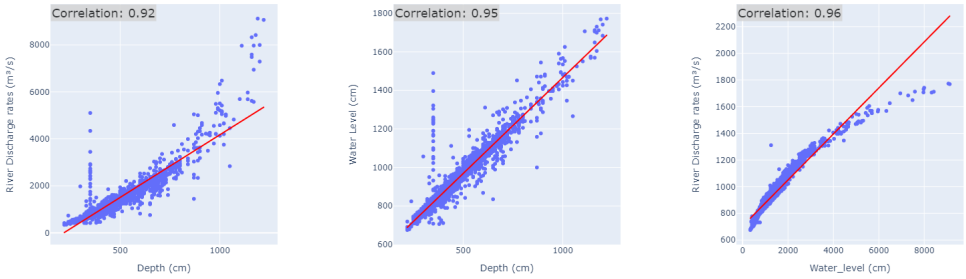
(a) River depth vs River discharge rate (b) River depth vs Water level (c) Water level vs River discharge rate

Figure 10: Scatter plots show the daily frequency correlation analysis between variables at the Piacenza shallow point.

Similarly, the correlation coefficient 0.89 between river depth levels and upstream water levels at Piacenza (see figure 10b) reflects a strong positive correlation, implying that water level fluctuations closely mirror river depth changes. This finding underscores the direct influence of upstream water level variations on river depth, highlighting the critical role of water level in predicting and managing river depth fluctuations.

Furthermore, the notably high correlation coefficient of 0.95 between water level and river discharge rate at Piacenza (see figure 10c) indicates a very strong positive relationship, suggesting a close linkage between these two variables. It implies that changes in river discharge rates have a pronounced effect on water levels, impacting river depth. Such

a tight coupling between water level and river discharge rate underlines the dynamic nature of river systems and the importance of understanding the complex interactions between hydrological variables.

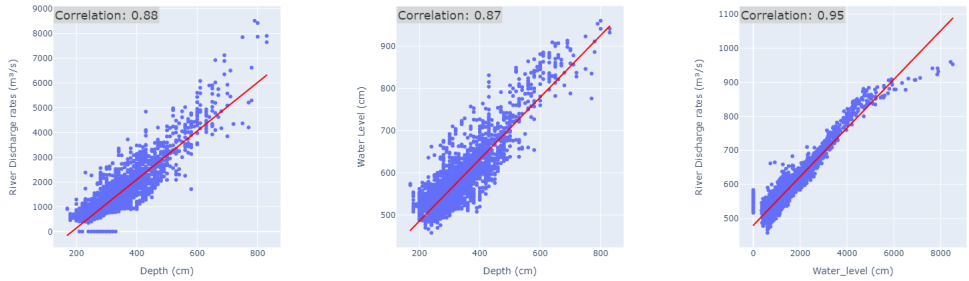


(a) River depth vs River discharge rate (b) River depth vs Water level (c) Water level vs River discharge rate

Figure 11: Scatter plots show the daily frequency correlation analysis between variables at the Monte P.Te Revere shallow point.

The consistency of these relationships across different stations, as evidenced by correlation coefficients of 0.92 (see figure 11a), 0.95 (see figure 11b) and 0.96 (see figure 11c) at Monte P.Te Revere. Similarly, correlation coefficients of 0.88 (see figure 12a), 0.87 (see figure 12b) and 0.95 (see figure 12c) at Cavanella, further emphasises the robustness of the findings. In visual representations of scatter plots, Ordinary Least Squares (OLS) regression lines (in red) capture the linear trend and show the direction, strength and linearity of the relationship between variables. The steepness of the OLS regression line reflects the strength of the relationship, with steeper slopes indicating stronger correlations. However, it is essential to note minor deviations from the OLS line across all three stations between depth and discharge. These deviations may arise from the imputation methods used to handle missing values and natural or artificial hydrological fluctuations.

The observed consistency implies that the relationships between river depth, river discharge rate and water level might not be confined to selected locations but assumes to be apparent at other points along the river. Overall, the results suggest that changes in river discharge rates and upstream water levels significantly influence river depth at all three stations, with potential implications for developing a robust river depth-level forecasting model.



(a) River depth vs River discharge rate (b) River depth vs Water level (c) Water level vs River discharge rate

Figure 12: Scatter plots show the daily frequency correlation analysis between variables at the Cavanella shallow point.

6.2. Climate Components

To investigate the correlation between climate factors and river depth, a systematic approach is developed as follows:

The cross correlation analysis method, detailed in section 5.1.6, is applied with a lag range from 0 to 365, representing the potential lag effects of up to one year. The analysis examines the relationship between river depth at shallow points and the three climate attributes within the radius of influence. For the cross correlation calculation, the climate attributes are fixed and river depth values are shifted to identify the effect of climate on river depth levels. The calculation is performed between three climate attributes of each geometric location and river depth at a specific critical point.

This analysis focuses on identifying the strongest positive correlations between precipitation and river depth and determining the specific time lags that reflect when the maximum impact occurs. Similarly, the strongest negative correlations between temperature and snow depth variables. By examining these negative correlations, insights can be gained into the effects of temperature and snow depth changes on river depth and the time delays associated with these impacts.

The map in figure 13 displays the regions (highlighted in red) where the total precipitation correlates with the river depth at Piacenza point (marked in blue). Additionally, the map also demonstrates the impact observed across the river basin. Similarly, in the perspective of Monte P.Te Revere (see figure 38 from the appendix), it is observed that the total precipitation from different geographical regions impacts the river depth. Cor-

responding correlation values and lags are identified to quantify this relationship. For Cavanella (see figure 39 in the appendix), a similar analysis is conducted to assess the impact of total precipitation on river depth.

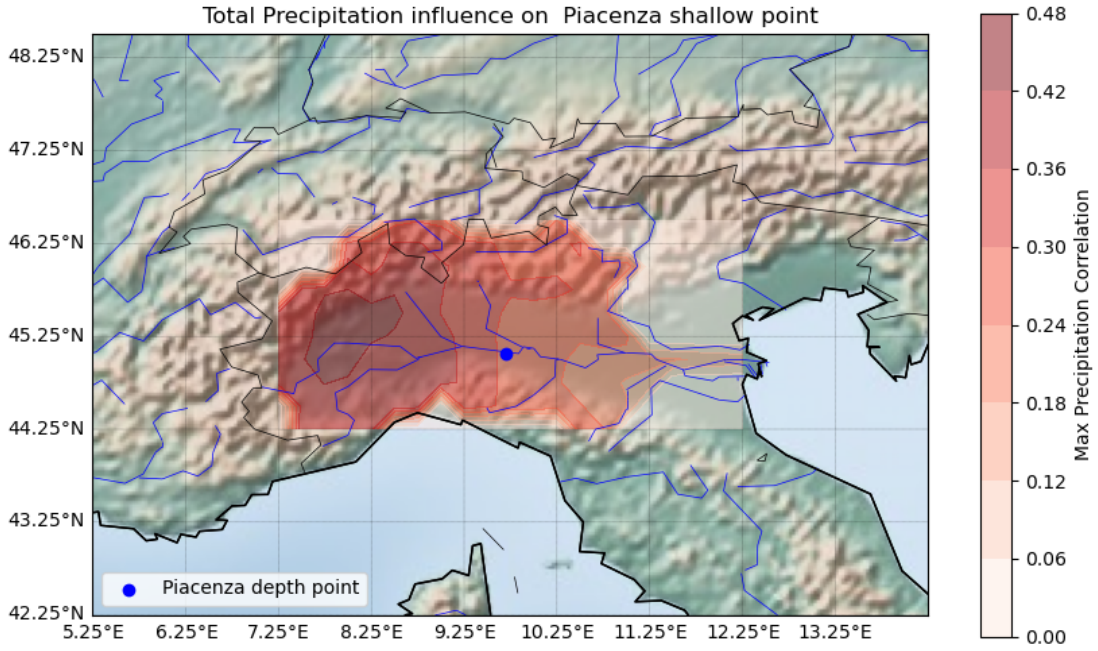


Figure 13: Map highlighting regions (in red) with maximum cross correlation between total precipitation and river depth with the Piacenza shallow point (in blue).

The results obtained through cross correlation analysis are sorted according to the highest to lowest correlation values for each geographical region. Specifically, the results revealed that the Piacenza point showed the highest correlation values with a 3-day lag (see table 27 in the appendix), while the Monte P.Te Revere (see table 30 in the appendix) and Cavanella points (see table 33 in the appendix) exhibited the highest correlation values with a 4-day lag for specific geographical regions. Notable correlations for Piacenza include 0.465 (see figure 14) from the river water catchment in the Graian Alps mountains, which are situated west of Piacenza, with coordinates 45.25° latitude and 8.0° longitude at 135.31 kilometres distant. Similarly, a correlation of 0.460 with coordinates 45.25° latitude and 7.75° longitude, around 154.70 kilometres distant (see table 27 in the appendix).

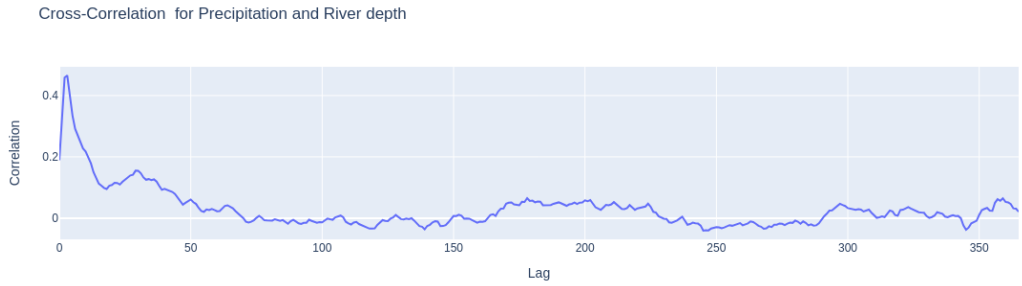


Figure 14: Line chart displaying the cross correlation value for related lag between total precipitation and river depth at Piacenza shallow point from most influencing region.

Approximately 188.20 kilometres from Monte P.Te Revere, water originating from precipitation at the Apennines Mountains (45.25° latitude, 8.75° longitude) has a maximum correlation value of 0.404 (see table 30 in the appendix) with the effect of 4 days. Similarly, at nearby coordinates such as 45.0° latitude and 8.75° longitude, the maximum correlation value is 0.391 with the same lag of 4 days (see table 30 in the appendix). Regarding Cavanella, noteworthy correlations include 0.219 from the southern Apennines Mountains at coordinates 44.75° latitude and 10.0° longitude, roughly 170.88 kilometres away and 0.217 from the northern Dolomite mountains at 44.75° latitude and 9.75° longitude, approximately 190.30 kilometres distant (see table 33 in the appendix). These correlation values, alongside their respective distances and lag periods, offer insights into the spatial and temporal dynamics of precipitation's influence on river depth at shallow points.

Looking at the results of temperature, the negative correlation of -0.127 (see figure 15) is observed between Piacenza and at coordinates 45.0° latitude and 8.0° longitude, approximately 134.08 kilometres away (see table 28 in the appendix). Similarly, at 44.5° latitude and 8.75° longitude, around 97.62 kilometres distant from Monte P.Te Revere, a correlation of -0.127 is observed with a lag of 359 days (see table 31 in the appendix). For Cavanella, correlations include -0.209 at coordinates 44.5° latitude and 8.75° longitude, approximately 198.05 kilometres away, with a lag of 361 days and -0.280 at 44.25° latitude and 10.75° longitude, approximately 139.05 kilometres distant, also with a lag of 361 days (see table 34 in the appendix).

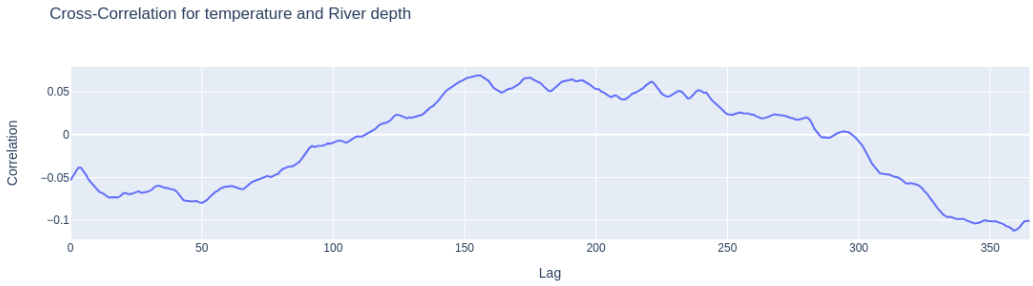


Figure 15: Line chart displaying the cross correlation value for related lag between temeprature and river depth at Piacenza shallow point from most influencing region.

Upon reviewing the temperature results, nearly all selected geographical regions across the Po River catchment area exhibit consistent outcomes (see tables 28, 31 and 34 in the appendix). The cross correlation between temperature and shallow point depth ranges from -0.127 to 0.2 across all regions. However, the results do not provide a clear impact and the lag between 350 and 360 indicates that drawing conclusions may not be suitable. Notably, there's a minimal seasonal pattern observed. For instance, at Piacenza, where the correlation value tends to spike for lags between 0 to 180 days and then declines from 180 to 360 days, indicating the subtle relation with temperature and river depth (see figure 15).

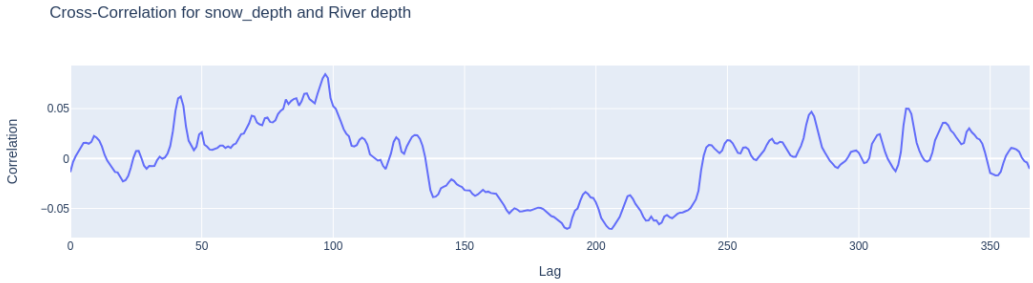


Figure 16: Line chart displaying the cross correlation value for related lag between snow accumulation and river depth at Piacenza shallow point from most influencing region.

Due to sparse measurement locations, the analysis found limited correlations between river depth at all three critical points and snow accumulations at geographical coordinates, especially in the Alpine mountain region. For instance, at coordinates 46.25° latitude and 10.0° longitude, approximately 134.46 kilometres from the Piacenza point, a correlation of -0.2607 with a lag of 360 days is observed (see table 29 in the appendix). At Monte P.Te Revere (see table 32 in the appendix), the minimum correlation coefficient of -0.141 with snow accumulation over a lag period of 179 days is found to be at a distance of 159.43 kilometres north at 46.25° latitude and 10.0° longitude. Around 120 km distant from the Cavanella point, a correlation of -0.1765 is observed at 45.5, 10.75 coordinates with a lag of 216 days (see table 35 in the appendix).

These findings explain the complex relationship between temperature, snow depth and river depth dynamics. However, it is crucial to exercise caution when interpreting these correlation values, as their relatively weak magnitudes may limit the conclusiveness of the effects.

Overall, the analysis of cross correlation results across the three shallow points reveals that precipitation alone makes only a slight difference, with the highest correlation around 0.4, indicating a relatively modest relationship. Additionally, the lag period helps define weather events and their temporal effects on shallow points, which is relevant for navigating ships in shallow waters. The results suggest that other factors beyond precipitation may influence the dynamics of the shallow points, revealing the system's complexity.

Furthermore, the examination of temperature and snow depth cross correlation results indicates that their relationship with the selected shallow points is not evident enough to consider them as influencing factors. However, climate attributes, especially temperature and snow depth, may have some effect in reality. Still, based on available data, their contribution to the dynamics of the shallow points appears to be relatively minor. From a positive perspective, the findings lead to the interpretation that climate attributes have some systematic behaviour on the shallow points.

Moreover, results from hydrological components like river discharge rates and upstream water levels show significant correlations in the depth of river level at selected shallow points. Further research considering additional factors and conducting a more comprehensive data analysis may provide a clearer understanding of the dynamics at play in shallow waters. In this thesis, only river discharge rates and upstream water levels are considered influential variables for further exploratory data analysis and river depth forecasting in the context of multivariate time series forecasting.

7. Exploratory Data Analysis (EDA)

EDA is an essential step in developing a machine learning model for river depth forecasting at shallow points, playing a crucial role in uncovering insights, patterns and relationships within datasets (Chatfield, 1986). This section explores temporal trends, seasonal fluctuations and spatial patterns in upstream water levels, river discharge rates and river depth data at shallow points. Exploring summary statistics, time series plots and spatial visualisations aims to identify recurring patterns and behaviour for river depth level forecasting. Insights derived from EDA contribute to selecting appropriate forecasting techniques and parameter optimisation strategies, thereby enhancing model performance and reliability (Chatfield, 1986).

Building upon the feature engineering results discussed in section 6, which highlighted the influence of hydrological variables like water level and river discharge rates on water depth at shallow points, this study aims to understand deeper relations and further validate the findings visually. Additionally, the aim is to verify the assumptions necessary for multivariate time series forecasting, including stationarity, normality, trends, seasonality, auto correlation and partial auto correlation across all three variables. This analysis is based on data described in the preprocessing section 4.4 and is crucial for developing robust forecasting models that contribute to improved river depth prediction and ship navigability perspective.

7.1. Stationarity

The ADF test described in methods section 5.1.2 is employed using `statsmodels` library (Seabold and Perktold, 2010) in Python 3.11.8 (Python Software Foundation, 2023) to assess the stationarity of water level, river discharge rates and river depth data at Piacenza, Monte P.Te, Revere and Cavanella shallow points. The null hypothesis of the ADF test is that the variable under consideration is non-stationary, meaning it possesses a unit root. Conversely, the alternative hypothesis suggests that the variable is stationary.

The results from table 16 explain p-values less than the conventional significance level of 0.05 or even 0.01 for all variables, providing strong evidence against the null hypothesis. Therefore, the null hypothesis is rejected in favour of the alternative, indicating that the statistical properties of the variables remain relatively constant over time, particularly in terms of their mean and trend components. These findings are crucial for multivari-

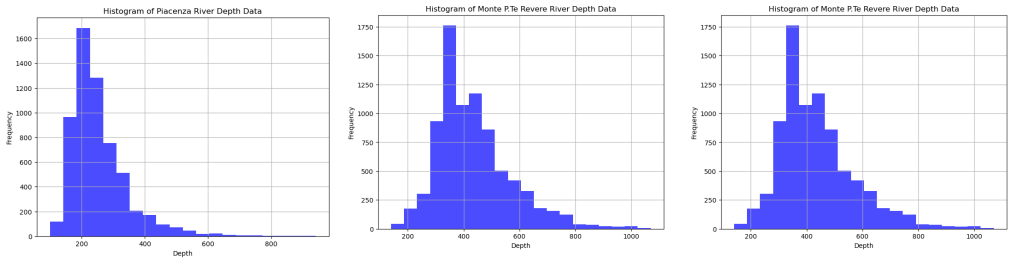
Attribute	Piacenza	Monte P.Te Revere	Cavanella
River Depth	4.64×10^{-13}	1.43×10^{-13}	9.16×10^{-10}
River Discharge rate	9.36×10^{-14}	3.63×10^{-16}	1.43×10^{-16}
Water level	2.58×10^{-24}	8.58×10^{-14}	2.71×10^{-21}

Table 16: Results of p-values from the ADF test for each attribute associated with all three shallow points.

ate time series analysis, as they establish a solid foundation for further modelling and forecasting efforts. However, the ADF test confirms stationarity and visual inspections of the time series graphs and auto correlation plots complement these results to ensure the absence of other underlying patterns or trends. Nonetheless, confirming stationarity in the water level, river discharge rates and river depth data at three shallow points underscores the reliability of subsequent analyses.

7.2. Normality

The data collected from all three shallow points exhibits a bell-shaped curve when plotted in histograms (see figure 17), indicative of a normal distribution. This characteristic bell curve suggests that the distribution of depth measurements at these points follows a pattern commonly observed in datasets conforming to a normal distribution. The symmetrical nature of the curves further supports that most depth values cluster around the mean, with fewer occurrences of values at the extremes.



(a) Piacenza (b) Monte P.Te Revere (c) Cavanella

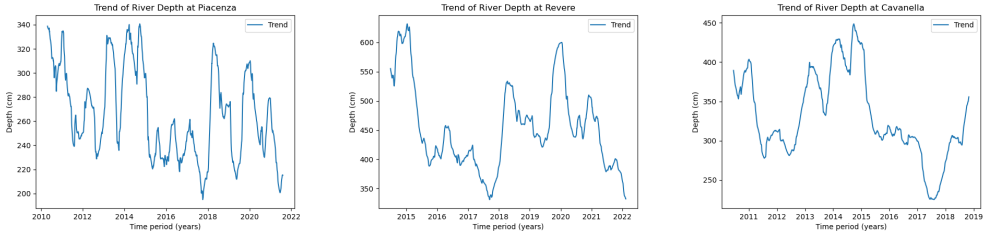
Figure 17: Histogram illustrating the normal distribution of river depth values across all three shallow points.

7.3. Seasonal Decompose

Seasonal decomposition as described 5.1.3 is applied to data using `statsmodels` library (Seabold and Perktold, 2010) in Python 3.11.8 (Python Software Foundation, 2023) to separate individual time series into its components, namely trend, seasonality for all three shallow points.

7.3.1. Trends

To visually explore noticeable trends, the seasonal decomposition method with the additive model is used to study the water level, depth and discharge with a seasonal period of 180 days (6 months). The graphs project the trend component at the Piacenza shallow point from 2010 to 2022 (see figure 18a), at Monte P.Te Revere from 2014 to 2022 (see figure 18b) and at Cavanella from 2010 to 2019 (see figure 18c). Observing the strong correlation among water level, depth and discharge from section 6, similar and identical trends are observed across all three attributes at each shallow point (see figures 35,36,37 in the appendix). The results showed a slightly decreasing and non-linear trend across all three river depth points (see figure 18).



(a) Piacenza (b) Monte P.Te Revere (c) Cavanella

Figure 18: Graphs show the trends of river depth levels across the years at three shallow points.

Based on visual interpretation, between 2015 and 2017, the river depth levels at Piacenza were relatively low, with the lowest point recorded in 2017. At Monte P.Te Revere, there was a significant decline in river depth levels in 2015 and later in the years between 2018 and 2021, depth levels remained stable with minor fluctuations. Cavanella, located at sea level, experiences fluctuations in river depth levels across the years, but a significant drop was observed in 2017. The pattern across shallow points (except Cavanella due to data unavailability) is downward from 2020 to 2022, with the lowest levels recorded in 2017 and 2022.

7.3.2. Seasonality

Analysing the results from the seasonality component of the seasonal decomposition process, a noticeable spike in the graph (see figure 19) indicates a clear seasonal trend for 180 days at the Piacenza shallow point. Similar observations are also noted at other shallow points (see figures 35,36,37 in the appendix). Box plots are generated by grouping monthly data over the years to understand seasonality across specific months. Upon visualising the box plots, for Piacenza (see figure 20), Monte P.Te Revere (see figure 21) and Cavanella (see figure 22), it is evident that shallow points experience higher river depth, discharge and water levels in May, June, November and December. Conversely, the lowest levels are observed in July and August, followed by January and February. These findings suggest a consistent seasonal pattern in river depth, discharge and water levels at the shallow points under consideration.

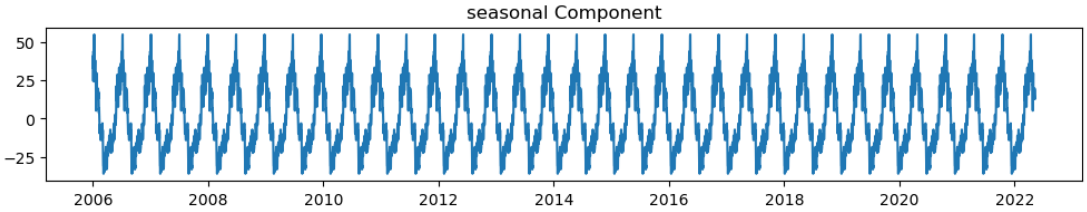
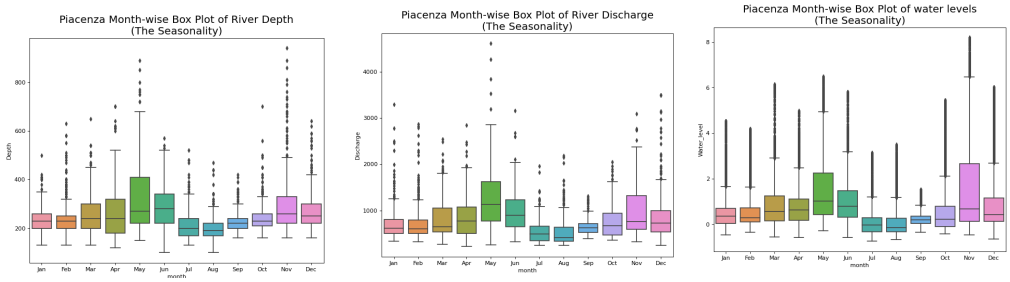
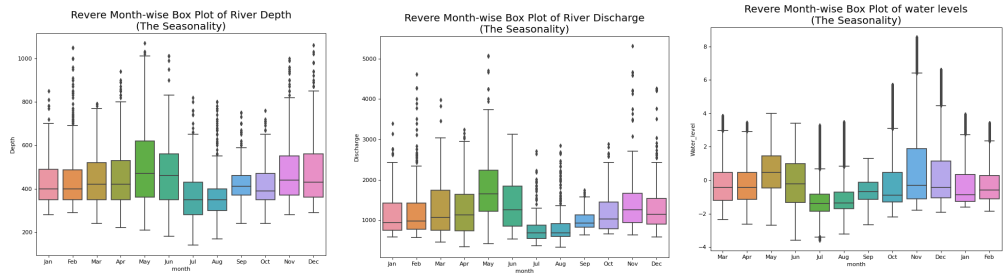


Figure 19: Results of the seasonality component derived from seasonal decomposition method on river depth data at the Piacenza shallow point.



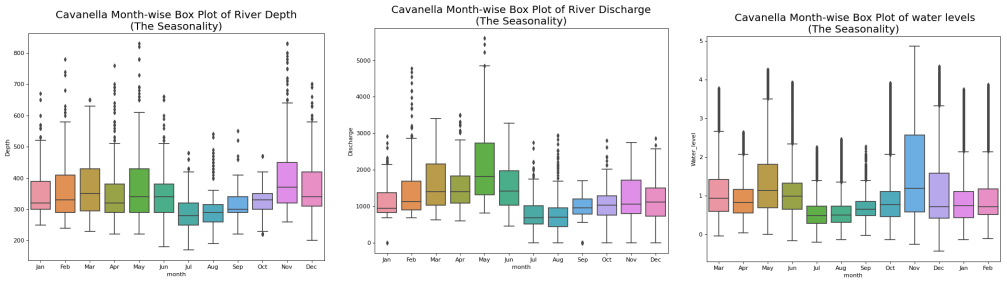
(a) River Depth (b) River Discharge rates (c) Water level

Figure 20: Box Plots show the consolidated monthly distribution of river depth, river discharge rates and water levels at Piacenza shallow point.



(a) River Depth (b) River Discharge rates (c) Water level

Figure 21: Box Plots show the consolidated monthly distribution of river depth, river discharge rates and water levels at Monte P.Te Revere shallow point.

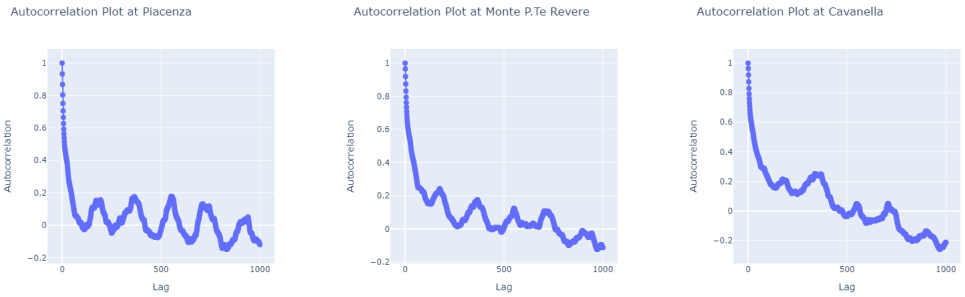


(a) River Depth (b) River Discharge rates (c) Water level

Figure 22: Box Plots show the consolidated monthly distribution of river depth, river discharge rates and water levels at Cavanella shallow point.

7.4. ACF and PACF Plots

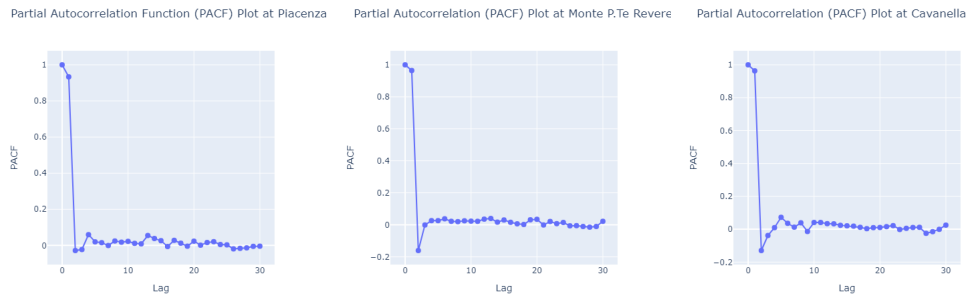
The Auto Correlation Function (ACF) and the Partial Auto Correlation Function (PACF), as described in sections 5.1.4 and 5.1.5, provides valuable insights into the temporal dependencies within a time series dataset comprising depth, discharge and water level data from all three shallow points. The ACF method is applied with a lag limit of 500 to observe long-term temporal patterns over the years. On the other hand, the PACF method is used with a lag limit of 30 to reflect short-term temporal patterns and consider optimal sequence length as input to train RNN models. The graphical representation of the ACF plots (see figure 23) generated for river depth data reveals engaging patterns. A gradual decline in correlation with increasing lag is observed until approximately 105 lags, followed by a slight increase and cyclic behaviour every 180 lags, suggesting a seasonal pattern from all three shallow points.



(a) Piacenza (b) Monte P.Te Revere (c) Cavanella

Figure 23: Auto Correlation Function (ACF) from river depth data at all three shallow points.

Conversely, the partial auto correlation graphs show the direct relationship between observations at specific lags after removing the influence of intermediate observations. Notable spikes in partial auto correlation plots indicate direct dependencies between observations. For instance, in the graph for Piacenza (see figure 24a), a slight spike at a lag of 13 is observed. In contrast, the PACF values are nearly close to zero and minor spikes are noted at lag 6 for Monte P.Te Revere (see figure 24b) and Cavanella (see figure 24c), suggesting the chance of cyclic behaviour and temporal dependencies. These lag patterns are valuable for training RNN models, as they help specify the sequence number for forecasting future values based on the cyclic behaviour exhibited by the data.



(a) Piacenza (b) Monte P.Te Revere (c) Cavanella

Figure 24: Partial Auto Correlation Function (PACF) from river depth data at all three shallow points.

8. Machine Learning Approach for River Depth Forecasting

This section addresses RQ2 by developing machine learning forecasting models based on their functionality described in section 5.2 for predicting water depth levels at critical shallow points. The models leverage LSTM, GRU, LSTNet and VAR algorithms and are structured from two temporal perspectives: daily and hourly. Strategies for designing models are explained based on data characteristics, followed by dividing data into training and testing sets.

8.1. Workflow

According to the problem statement described in section 1.1, in response to the AIPo authority's requirement, a strategic noble approach is proposed in this work involving the development of two distinct forecasting models for each shallow point serving daily and hourly predictions. The historical dataset compiled by the AIPo, encompassing river depth, discharge rates and upstream water levels, serves as the foundational data for training these models. By leveraging this data, the models can be finely trained to forecast the next occurrence precisely.

Moreover, to achieve the desired forecasting horizon of 14 days and 24 hours, the model can be augmented by incorporating forthcoming river discharge rates from the EFAS (Copernicus Climate Change Service, 2019) and the AIPo's existing water level estimation technology (Interregional Agency of the Po River, 2011). The AIPo holds access to EFAS data and monitors data populated from the EFAS model and their own estimation system. This integration complements the AIPo's existing monitoring system, offering a comprehensive solution for predicting water depth levels with heightened accuracy and reliability.

Two models are developed to meet the forecasting requirements set by the AIPo. The daily forecasting model will utilise future river discharge rates from the EFAS and upstream water levels from the AIPo's estimation system to predict river depth levels at specific shallow points over the next 14 days. This model will be trained using historical data on river depth, discharge rates and water levels to ensure accurate predictions.

Given the lack of data regarding sand sedimentation at shallow points, along with incomplete information on river depth and river discharge rates at hourly intervals, the hourly forecasting model focuses on predicting water fluctuations in upstream water level measurement stations. This strategy operates under the assumption that ships can

navigate safely when water levels are elevated above ship draft length, particularly given the proximity of the water level station to shallow points. The hourly forecasting model also leverages the interdependence between river depth, discharge rates and upstream water levels. By training on historical data, the model forecasts water levels for the next 24 hours using interpolated river depth from the daily model and river discharge rates from EFAS as input. These models collectively aim to provide comprehensive and timely forecasts, contributing to the navigation scope.

In this thesis work, as a proof of concept, the forecasting models will exclusively utilise historical data provided by the AIPo authority. Based on the forecast horizon, the dataset is divided into training and testing subsets, assuming that future river discharge rates from the EFAS and the AIPo's estimated upstream water levels align with the testing dataset. This approach ensures the models are trained and evaluated using relevant historical data aligned with the anticipated forecast scenarios.

Daily Forecasting Model

The first model (see figure 25) is designed to forecast daily river depth using preprocessed historical data of river discharge rates and upstream water levels as input variables and river depth as the target variable. This model is trained on the input variables to learn the behaviour of the target variable over time.

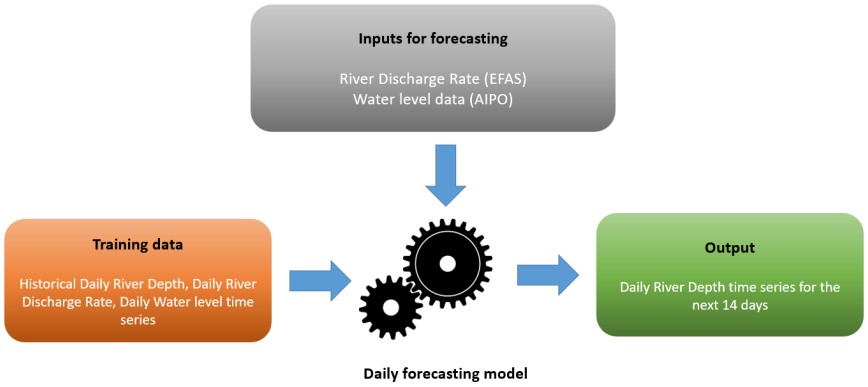


Figure 25: Strategy employed for daily river depth forecasting, encompassing input variables, target variable and inputs for forecast.

The model employs an iterative approach to predict daily river depth for the next 14 days. Initially, the model generates predictions based on the last sequence of input variables.

The resulting prediction is then combined with input variables from the testing dataset to forecast the next occurrence. Subsequently, this process is repeated by incorporating the latest sequence of data to predict the river depth for the next day. This iterative method continues until the 14-day forecasting horizon is achieved, allowing the model to refine its predictions iteratively based on evolving input data. Using historical data and resampled water levels enhances the model's capability to capture variations in river depth over time accurately.

Hourly Forecasting Model

The second model (see figure 26) is dedicated to forecasting upstream hourly water levels. It involves training with preprocessed historical river depth data and river discharge rates at hourly intervals as input variables, with upstream water levels as the target variable.

The hourly forecasting model is structured to predict upstream water levels on an hourly basis by utilising inputs such as river discharge rates from the EFAS forecast model and river depth (derived from the outcomes of the first model and interpolated to hourly intervals). Similar to the first model, the hourly model adopts an iterative approach to forecast hourly upstream water levels for a 24-hour forecast horizon. This iterative method involves generating predictions based on sequential input data, refining the forecasts iteratively to account for changing conditions and improving predictive accuracy over the forecast period.

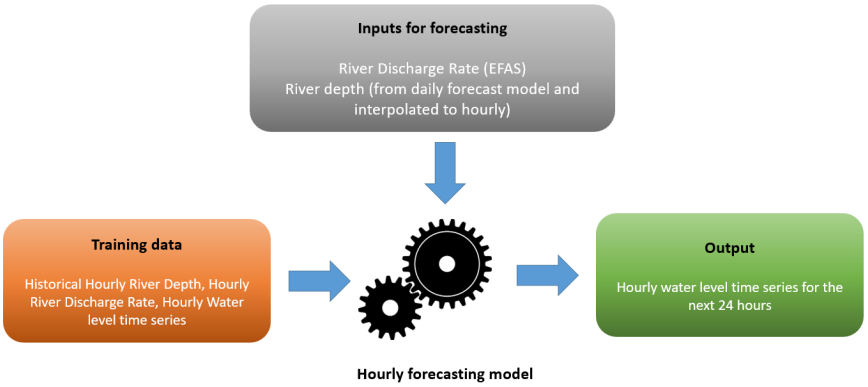
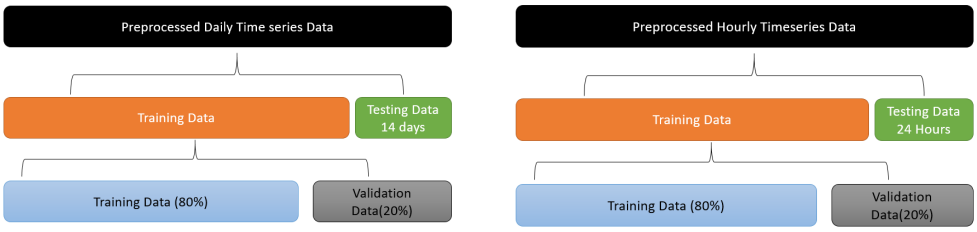


Figure 26: Strategy employed for hourly upstream water level forecasting, encompassing input variables, target variable and inputs for forecast.

This approach is chosen based on the observation that utilising the original 30-minute raw water level data frequency yields superior performance compared to interpolated river depth data. Moreover, acknowledging the strong correlation between river depth and water level, leveraging both variables enhances the model's predictive capabilities and ensures robust forecasting outcomes. By integrating interpolated data and input parameters from the first model along with the existing forecast model for river discharge rates, this second model aims to deliver detailed forecasts of water levels at shorter intervals.

8.2. Data Preparation for Modeling

In this analysis, significant emphasis has been placed on partitioning historical data to assess the performance of forecasting models, as described in figure 27. The daily and hourly datasets are segmented into training (80%) and validation (20%) sets. This partitioning strategy facilitates the development and evaluation of forecasting models specific to different temporal resolutions, ensuring robustness and accuracy in model training and validation processes.



(a) Daily data (b) Hourly Data

Figure 27: Train-test split strategy for the preprocessed daily and hourly dataset.

The training set is the foundation for model training, allowing algorithms to learn patterns and relationships from historical data. Conversely, the validation set provides an independent dataset to evaluate model performance, offering insights into how well the trained models generalise to unseen data. Following this partitioning strategy, a subsequent split as testing data reflected input from EFAS and the AIPo's estimation model to compare forecasts over 14 days and 24 hours for daily and hourly perspectives, respectively.

Shallow Point	Training Data Range	Total	Training (80%)	Validation (20%)	Testing Data Range	Testing (forecast in days)
Piacenza	2010-01-22 to 2021-10-23	4293	3435	858	2021-10-24 to 2021-11-06	14
Monte P.Te Revere	2014-03-28 to 2022-10-28	2954	2363	591	2022-04-29 to 2022-05-12	14
Cavanella	2010-03-19 to 2019-01-13	3223	2578	645	2019-01-14 to 2019-01-27	14

Table 17: Summary of training and testing data split for shallow points with a daily forecasting model.

Shallow Point	Training Data Range	Total	Training (80%)	Validation (20%)	Testing Data Range	Testing (forecast in hours)
Piacenza	2010-01-22 to 2021-10-23	103330	82664	20666	2021-11-05 to 2021-11-06	24
Monte P.Te Revere	2014-03-28 to 2022-10-28	71185	56948	14237	2022-05-11 to 2022-05-12	24
Cavanella	2010-03-19 to 2019-01-13	77641	62112	15529	2019-01-26 to 2019-01-27	24

Table 18: Summary of training and testing data split for shallow points with hourly forecasting model

Tables 17 and 18 detail the data split across the three shallow points, outlining the specific training and testing data ranges for both daily and hourly time series. This rigorous evaluation period facilitates a forward looking assessment of the model's predictive capabilities across defined time horizons, enabling a thorough examination of forecast accuracy and providing inputs for probability calculations in navigation planning.

A normalisation step is applied to prepare the data for model training using the Standard Scaler function from the `scikit-learn` library (Pedregosa et al., 2011). The process ensures that input features and the target variable have a mean of 0 and a standard deviation of 1, facilitating stable and consistent training across different datasets.

The sequences to train the model are generated using a sliding window approach, where a fixed sequence length is defined. This sequence length is determined based on PACF plots (see figure 24), which indicate the appropriate number of time steps needed to capture relevant temporal patterns. For Piacenza, a sequence length of 13 is observed, while Monte P.Te revere and Cavanella have a sequence length of 6 for daily points. For hourly water level forecasting, a sequence length of 24 is considered. By extracting sequences of this length from the dataset at each time step, input-output pairs are formed, enabling the model to learn temporal dependencies effectively. Overall, defined steps ensure the dataset is structured into sequences suitable for training sequence-based models like LSTM, GRU and LSTNet.

8.3. Model Training

The thesis considers various time series forecasting approaches, as described in section 5.2, encompassing different methodologies, including RNN such as LSTM (see section 5.2.2) and GRU (see section 5.2.3), along with a hybrid neural network named LSTNet (see section 5.2.4) and a statistical approach known as VAR (see section 5.2.5). These models are thoroughly investigated for both daily and hourly forecasting tasks and their performance is evaluated using RMSE (see section 5.4), allowing for a comparative analysis of their forecasting accuracy.

During the model development process, various combinations of parameter values are explored, including the number of layers, units and epochs, which are trained and evaluated. Through this exploration, it is determined that the architecture and training configuration described for Vanilla LSTM (see table 19), GRU (see table 22), LSTNet (see table 23) and VAR (see table 24) achieved better and performance for the task.

The described architectures for LSTM, GRU and LSTNet models are designed using `keras` library of Tensor flow framework (Abadi et al., 2015) and executed using Python 3.11.8 (Python Software Foundation, 2023).

Parameter	**Value**
Input	River Depth, Water level and River Discharge rate
Output	Daily: River Depth; Hourly : Water level
Sequence number	Daily: Piacenza (13), Monte P.Te Revere, Cavanella (6); Hourly: 24 (all three stations)
Dataset Distribution Structure	Training set (80%), Validation set (20%), Testing set (last 14 days or 24 hours)
Activation Function	Rectified linear unit (ReLU)
Loss Function	Mean Square Error (MSE)
Number of Epochs	50
Batch Size	default (32)
Optimizer	Adam
Model structure	One bidirectional layer of LSTM layer with 50 units followed by a dense layer with 1 unit.

Table 19: Model configuration details of the Vanilla LSTM network for time series forecasting.

Due to the higher training time for the Vanilla LSTM model, it is fine-tuned using FLAML tool as described in section 5.3 for each shallow point on the daily dataset to enhance its predictive performance. For the tuning objective, the minimum RMSE value is considered as the tuning objective and the search space described in table 20 are considered:

Tuning Objective	**Range**
Number of LSTM Units	10 to 128
Number of layers	1 to 5
Batch Size	6 to 256
Activation Function	Rectified linear unit (ReLU), Hyperbolic tangent function (Tanh)
Number of Epochs	10 to 500

Table 20: Tuning variables and search space for optimising the Vanilla LSTM model.

After exhaustive experimentation and parameter tuning and based on the results (see tables 36,37,38 in the appendix), the best configuration model with the least RMSE for the respective dataset is optimised with the following hyper-parameters:

Parameter	**Piacenza**	**Monte P.Te Revere**	**Cavanella**
Activation Function	Tanh	ReLU	ReLU
LSTM Units	64	109	122
Number of Layers	3	3	3
Number of Epochs	339	11	192
Batch Size	6	179	53
Model structure	Bi directional Three layers of LSTM based on above units with single dense layers as final Output		

Table 21: Model configuration details of the tuned LSTM network for time series forecasting.

Due to limited computation resources, only the LSTM model is tuned using the FLAML tool. In the perspective of GRU (see table 22), LSTNet (see table 23) and VAR (see table 24) models, standard model architecture with custom configuration are considered, which are inspired by works of Lin et al. (2022), Yang and Zhang (2022) and Hartini et al. (2015) respectively.

Parameter	**Value**
Input	River Depth, Water level and River Discharge rate
Output	Daily: River Depth; Hourly : Water level
Sequence number	Daily: Piacenza (13), Monte P.Te Revere, Cavanella (6); Hourly: 24 (all three stations)
Dataset Distribution Structure	Training set (80%), Validation set (20%), Testing set (last 14 days or 24 hours)
Activation Function	Rectified linear unit (ReLU)
Loss Function	Mean Square Error (MSE)
Number of Epochs	50
Batch Size	default (32)
Optimizer	Adam
Model structure	One bidirectional layer of GRU layer with 50 units followed by 1 Dimensional Global Average Pooling layer and Output layer by a Dense layer with 1 unit.

Table 22: Model configuration details of the GRU network for time series forecasting.

Parameter	Value
Input	River Depth, Water level and River Discharge rate
Output	Daily: River Depth; Hourly : Water level
Sequence number	Daily: Piacenza (13), Monte P.Te Revere, Cavanella (6); Hourly: 24 (all three stations)
Dataset Distribution Structure	Training set (80%), Validation set (20%), Testing set (last 14 days or 24 hours)
Activation Function	Rectified linear unit (ReLU)
Loss Function	Mean Square Error (MSE)
Number of Epochs	50
Batch Size	default (32)
Optimizer	Adam
Model structure	One bidirectional layer of CNN branch with kernel size 3, filter size 128, LSTM layer with 100 units, joined using concatenation layer and final output layer by a Dense layer with 1 unit.

Table 23: Model configuration details of the hybrid approach of LSTNet network for time series forecasting.

To train the VAR model, complete historical data is utilized, with the last 14 records reserved for testing the daily model and the last 24 records reserved for testing the hourly model. For the current analysis, `tsa.vector_ar` model from `statsmodels` library (Seabold and Perktold, 2010) is used and executed in Python 3.11.8 (Python Software Foundation, 2023).

Parameter	Value
Input	River Depth, Water level and River Discharge rate
Output	River Depth, Water level and River Discharge rate (Both Daily and Hourly)
Dataset Distribution Structure	Training set, Testing set (last 14 days or 24 hours)
freq	Daily (default: D), Hourly (default: H)

Table 24: Model configuration details of the statistical approach Vector Auto Regression(VAR) for time series forecasting.

8.4. Cross Validation

To further evaluate the model's performance and robustness, a 20 fold cross validation procedure is employed. This approach divides the dataset into 20 equally-sized folds, each serving as a validation set while the remaining folds are used for training. The process is repeated 20 times with a different fold designated as the validation set. By averaging the evaluation metrics, such as RMSE, across the folds, a more reliable estimate of the model's performance is obtained. This accounts for variability in the data and reduces the risk of overfitting.

After the best-performing model is identified, its forecasting results are used to calculate the probability of navigability risk. The model's predictive capabilities are further leveraged to assess the likelihood of potential navigational challenges or hazards based on the forecasted river depth.

9. Findings and Discussion

The subsequent section presents the outcomes of an extensive investigation using machine learning algorithms to deepen the understanding of river depth forecasting at critical points along the River Po. This analysis addresses the research question RQ2, as outlined in section 1.2. The section showcases the results of daily river depth forecasting and hourly water level forecasting using the machine learning approaches discussed in section 8. These forecasts are conducted at pivotal locations such as Piacenza, Monte P.Te Revere and Cavanella, providing valuable insights into the predictive capabilities of the developed models and their implications for navigational planning and risk assessment along the River Po.

The aim is to compare the effectiveness of machine learning algorithms, including LSTM, GRU, LSTNet and VAR models, in predicting water depth levels at critical points. Additionally, the research explores the probabilities of navigational risk for various ship classes based on forecasted water depth levels. This section entails a structured presentation of comparative discussions through findings, detailed analysis and considerations of limitations and validity.

9.1. Forecasting Results

The forecasted daily river depth values at Piacenza (see table 39 in the appendix), Monte P.Te Revere (see table 40 in the appendix) and Cavanella (see table 41 in the appendix) shallow points are presented over distinct date ranges, utilising a variety of predictive models such as LSTM (Vanilla), LSTM (tuned), GRU, LSTNet and VAR. Based on the available data, the forecasting period of daily models is 14 days from 10 October 2021 to 06 November 2021 for Piacenza, 29 April 2022 to 05 May 2022 for Monte P.Te Revere and 14 January 2019 to 27 January 2019 for Cavanella.

The daily forecasting results for the specified dates across the shallow points can be found in tables 39, 40 and 41 in the appendix, which present the outcomes obtained

from the considered models using testing data. Similarly, the hourly forecasting results produced by the considered models can be viewed in figures 44 and 45 in the appendix.

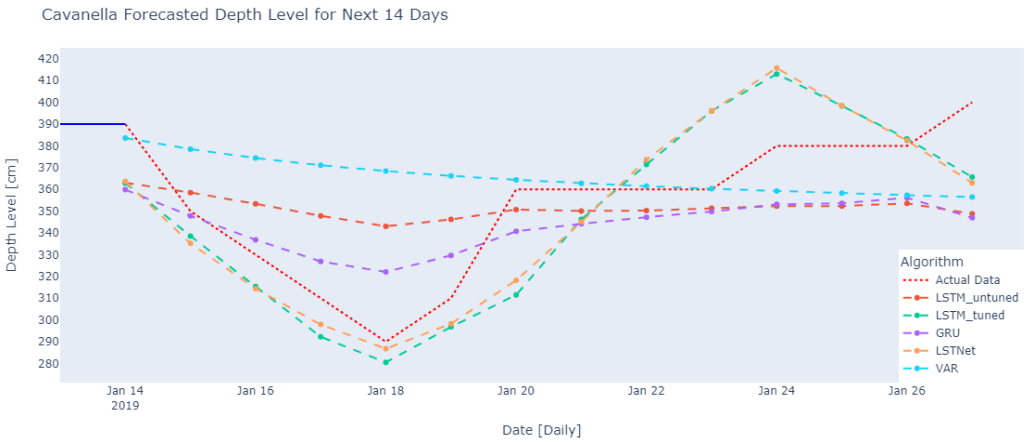


Figure 28: Results from daily river depth forecasts for next 14 days at Cavanella by corresponding models.

In examining the daily forecasted river depth values across different locations and date ranges, the LSTNet hybrid model emerges as a consistent performer, showcasing stable and accurate predictions with low variability compared to the other models (see figures 28, 40 and 41 in the appendix). Specifically, at Piacenza, the LSTNet model, alongside LSTM (tuned) and GRU, effectively captures the pattern in river depth fluctuations, exhibiting closely aligned forecasts. Similarly, at Monte P.Te Revere and Cavanella, LSTNet consistently delivers competitive performance in capturing observed river depth trends. Although GRU and LSTM (tuned) also offer good predictions, they may display more variability, particularly in response to changing input data. Conversely, the LSTM (Vanilla) with standard parameters and VAR model shows less distinct performance, characterised by higher forecast variability than the neural network counterparts. Overall, the LSTNet model is reliable and robust in river depth forecasting.

Regarding the hourly forecasting results at Monte P.Te Revere, the LSTNet model demonstrates superior prediction performance compared to the other models, closely aligning with the actual testing data (see to figure 44 in the appendix). Conversely, the results from Cavanella (see figure 45 in the appendix) show similar forecasting capabilities among LSTM, GRU and LSTNet models, with minor variations from the actual testing data. This behaviour is expected because the actual testing data at Cavanella

exhibits variability within a few hours, making it challenging for models to precisely predict every fluctuation.

9.2. Comparative Analysis

The comparison analysis allows for assessing each model's predictive capability in capturing daily variations in water depth and hourly water levels at the shallow points. The RMSE metric provides insights into the magnitude of prediction errors, with lower RMSE values indicating better predictive accuracy. Additionally, training times highlight the computational efficiency of each model, which is crucial for real-time or operational applications. These performance metrics collectively inform the selection of the most effective machine learning approach for river depth forecasting in navigational contexts along the River Po.

Daily Forecasting Models

The table 25 compares the performance of various machine learning models on daily time series data of Piacenza, Monte P.Te Revere and Cavanella. The evaluation metrics include RMSE in centimetres of river depth and training time in seconds. RMSE measures the accuracy of forecasted results against actual observations of the validation data set.

	Piacenza		**Monte P.Te Revere**		**Cavanella**	
Model	**RMSE**	**Training Time**	**RMSE**	**Training Time**	**RMSE**	**Training Time**
LSTM (Vanilla)	28.055	341.532	27.132	128.186	29.408	203.641
LSTM (tuned)	14.475	41.305	19.686	13.007	24.261	24.173
GRU	14.667	96.184	26.987	31.214	24.424	47.173
LSTNET	13.778	36.474	25.211	20.330	23.753	21.154
VAR	26.769	0.009	36.090	0.008	37.030	0.008

Table 25: Performance comparison of different models on daily time series data at all three shallow points.

The results show that the Recurrent neural network LSTM (Vanilla) model, despite its extended training time exceeding 341 sec (approx. 5 minutes) for Piacenza, exhibits moderate performance with RMSE values ranging from 27.132 to 29.408 across the shallow points. In contrast, the tuned LSTM model demonstrates improved performance with lower RMSE values ranging from 14.475 to 24.261, with still considerable training times. The GRU model achieves comparable performance with RMSE values ranging from 14.667 to 26.987 and moderate training times. Notably, the hybrid approach with

the LSTNet model stands out with the lowest RMSE values ranging from 13.778 to 25.211 and relatively shorter training times, balancing performance and efficiency. In contrast, the statistical approach of the VAR model, while efficient in training time, shows higher RMSE values ranging from 26.769 to 37.030, indicating less accurate predictions than the neural network models.

Overall, the findings suggest that the hybrid approach combining CNN layer technique with RNN architecture, specifically LSTNet, is a promising model for daily water depth prediction at critical points along the River Po, providing both accuracy and efficiency. The results highlight the trade-offs between model complexity, prediction performance and training time, offering valuable insights for decision-making in navigational planning and risk assessment scenarios.

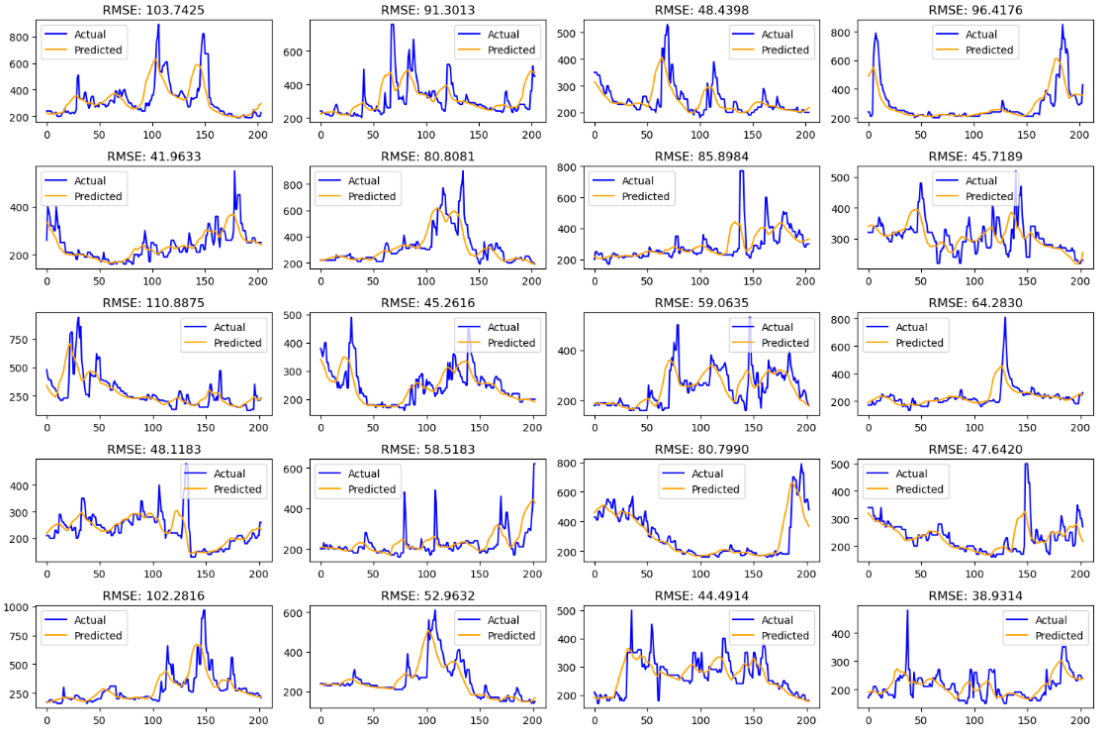


Figure 29: Performance of the LSTNet model on the Piacenza dataset at each cross-validation fold, comparing actual and predicted values using the RMSE metric.

To assess the robustness of the LSTNet model, a 20-fold cross-validation method is employed. Figure 29 compares actual and predicted river depth for each fold within the cross-validation process at Piacenza critical point. The model's performance across all

folds is evaluated using RMSE values, ranging from the most favourable result of 38.93 cm to 103.74 cm. Similarly, the results from cross-validation on Monte P.Te Revere (see figure 42) and Cavanella (see figure 43) show consistent outcomes. Table 42 in the appendix refers to average RMSE values across all folds at the Piacenza, Monte P.Te Revere and Cavanella shallow points reflecting the model's robustness irrespective of input data.

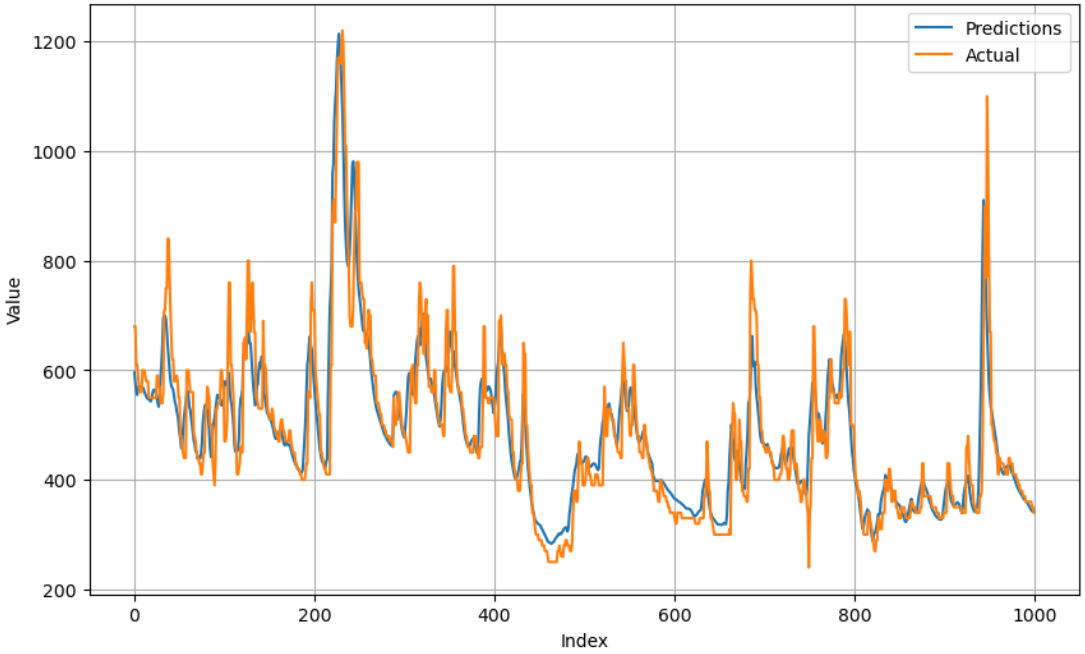


Figure 30: Line graph comparing LSTNet model predictions with actual values of the cross validation data for daily river depth levels in Monte P.Te Revere.

The results from the unseen validation data are presented in figure 30, illustrating the model's performance in predicting river depth over the last 1000 records of cross validation data. This figure highlights the model's ability to capture significant temporal dependencies, including short-term and long-term patterns. This graph underscores the model's effectiveness in leveraging historical data to make accurate predictions, emphasising its capability to capture and adapt to temporal variations in river depth dynamics at critical locations at the selected points along the River Po.

Hourly Forecasting Models

Similar to the daily forecasting model, the hourly time series model presents table 26, which compares the performance of various machine learning models in forecasting water levels from critical points. RMSE measures the accuracy of forecasted results against actual observations of the validation data set.

	Piacenza		Monte P.Te Revere		Cavanella	
Model	**RMSE**	**Training Time**	**RMSE**	**Training Time**	**RMSE**	**Training Time**
LSTM (Vanilla)	26.903	1452.514	19.203	1087.333	13.128	784.173
GRU	24.727	394.214	17.369	278.260	12.327	228.440
LSTNeT	22.801	695.727	10.515	496.175	14.470	363.490
VAR	45.002	0.166	48.701	0.118	47.087	0.117

Table 26: Performance comparison of different models on hourly time series data at all three shallow points.

The results show notable differences in model performance across different locations and algorithms. The LSTM (Vanilla) model exhibits moderate performance with RMSE values ranging from 13.128 cm to 26.903 cm across the three locations, accompanied by relatively longer training times ranging from about 784 to 1452 seconds. In comparison, the GRU model performs competitively with lower RMSE values ranging from 12.327 cm to 24.727 cm and relatively shorter training times ranging from approximately 228 to 394 seconds.

LSTNet emerges as the top-performing model, achieving notably lower RMSE values, particularly reaching 10.515 cm for Monte P.Te Revere. However, this superior performance takes longer training times, ranging from about 363 to 695 seconds. In contrast to that, the VAR model exhibits higher RMSE values ranging from 45.002 cm to 48.701 cm, indicating less accurate predictions than neural network-based models. Despite its simplicity, the VAR model maintains minimal training times.

In summary, LSTNet demonstrates exceptional accuracy in hourly water depth prediction due to its higher training set size compared to daily training set size, particularly at Monte P.Te Revere, although with longer training times. The trade-offs between prediction accuracy and computational efficiency outline the practical considerations in selecting the most suitable model for operational use in river management and navigational planning scenarios along the River Po. These findings provide valuable insights into the comparative performance of machine learning approaches and inform decision-making processes for calculating navigational risk at these points.

9.3. Probability of Navigational Risk at Shallow Points

The navigability assessment of the Po River incorporates a statistical method recommended by the AIPo authority. To address the research question RQ3 stated in section 1.2, the time series of historical daily water depth levels available at each monitoring station can be used to define a reasonable estimation of probability along Piacenza, Monte P.Te Revere and Cavanella. Combined with the forecasted daily water depth level data by the best machine learning model in section 9.2, those probabilities provide a primary method to compute the likelihood of navigability for each ship class for 14 days. The probability is calculated based on the percentage of occurrences of the event where water depth levels are greater than the minimum ship draft in the historical data.

$$P_{cr.point}(nav) = \frac{\text{Number of days where depth} \geq \text{minimum depth class and } Q \leq Q_{\text{obs}}}{\text{Total number of days where } Q \leq Q_{\text{obs}}}$$

The formula determines the likelihood of the observed depth (Q_obs) being less than or equal to a specific threshold depth (Q) while the water depth exceeds a minimum class depth. This probability is derived by dividing the days where the water depth surpasses the minimum class threshold and is less than or equal to the observed depth by the total number of days where the depth is less than or equal to the observed depth.

	2021-10-24	2021-10-25	2021-10-26	2021-10-27	2021-10-28	2021-10-29	2021-10-30	2021-10-31	2021-11-01	2021-11-02	2021-11-03	2021-11-04	2021-11-05	2021-11-06
140	99.258392	99.258392	99.185598	99.185598	99.061265	99.061265	98.905530	98.905530	98.905530	98.674110	98.674110	98.905530	99.061265	99.258392
160	96.370023	96.370023	96.013716	96.013716	95.405138	95.405138	94.642857	94.642857	94.642857	93.510119	93.510119	94.642857	95.405138	96.370023
180	85.284934	85.284934	83.840549	83.840549	81.373518	81.373518	78.283410	78.283410	78.283410	73.691556	73.691556	78.283410	81.373518	85.284934
200	69.398907	69.398907	66.395199	66.395199	61.264822	61.264822	54.838710	54.838710	54.838710	45.289602	45.289602	54.838710	61.264822	69.398907
220	44.067135	44.067135	38.576940	38.576940	29.199605	29.199605	17.453917	17.453917	17.453917	0.000000	0.000000	17.453917	29.199605	44.067135
250	8.938329	8.938329	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	8.938329

Figure 31: Probabilities for safe navigation (in green) and risk (in red) at the Piacenza shallow point based on forecasted river depth by the LSTNet model for the upcoming 14 days.

Figure 31 provides a detailed outlook of how ship draft length and forecasted river depth influence the probability of navigation risk across various dates from 24 October to 06 November 2021 at Piacenza. The data reveals a consistent trend, where the probability of navigation risk increases when ship draft length increases. For instance, on 24 October 2021, the probability values for ship drafts of 140, 160, 180, 200, 220 and 250 centimetres were 99.25, 96.37, 85.28, 69.39, 44.06 and 8.93, respectively. These values suggest that ships with shallower drafts face significantly lower risks compared to those with deeper drafts under similar river conditions. Furthermore, the table highlights fluctuations

in risk probabilities across different dates, indicating varying river depths impacting navigation safety. Days with higher probability values (coloured in green) likely signify safer river conditions. In contrast, lower values (coloured in red) indicate a heightened risk for navigation chances due to shallower river depth levels.

	2022-04-29	2022-04-30	2022-05-01	2022-05-02	2022-05-03	2022-05-04	2022-05-05	2022-05-06	2022-05-07	2022-05-08	2022-05-09	2022-05-10	2022-05-11	2022-05-12
140	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
160	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
180	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
200	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
220	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
250	87.431694	96.354992	98.008658	98.161471	98.161471	98.161471	98.008658	96.354992	94.320988	92.857143	95.490196	98.008658	98.687215	98.959276

Figure 32: Probabilities for safe navigation (in green) and risk (in red) at the Monte P.Te Revere based on forecasted river depth by the LSTNet model for the upcoming 14 days.

Similarly, projections from Monte P.Te Revere (see figure 32) and Cavanella (see figure 33) indicate higher probabilities (highlighted in green) of safe navigation across the forecasted period for all draft lengths except for the 250 class. The noticeable reduction in probability (highlighted in red) for the 250 class suggests potential caution for navigation in this specific draft category.

	2019-01-14	2019-01-15	2019-01-16	2019-01-17	2019-01-18	2019-01-19	2019-01-20	2019-01-21	2019-01-22	2019-01-23	2019-01-24	2019-01-25	2019-01-26	2019-01-27
140	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
160	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
180	99.916037	99.900200	99.874055	99.830221	99.798995	99.830221	99.874055	99.907149	99.919743	99.924214	99.927589	99.924214	99.922058	99.916037
200	99.328296	99.201597	98.992443	98.641766	98.391960	98.641766	98.992443	99.257196	99.357945	99.393710	99.420710	99.393710	99.376461	99.328296
220	96.473552	95.808383	94.710327	92.869270	91.557789	92.869270	94.710327	96.100279	96.629213	96.816976	96.958726	96.816976	96.726422	96.473552
250	87.699412	85.379242	81.549118	75.127334	70.552764	75.127334	81.549118	86.397400	88.242376	88.897310	89.391745	88.897310	88.581450	87.699412

Figure 33: Probabilities for safe navigation (in green) and risk (in red) at the Cavanella shallow point based on forecasted river depth by the LSTNet model for the upcoming 14 days.

Understanding these temporal patterns is crucial for informed navigation planning, enabling ships to assess risk levels and make informed decisions to ensure safe passage through river environments. By leveraging results from hourly water level forecasts (see figure 34) and visualising trends in water levels near shallow points, more profound insights can be gained to enhance planning for upcoming hours. As described in the feature engineering section 6, the strong correlation between river depth levels and water levels underscores the importance of such analyses in providing actionable information for refining navigation risk management strategies.

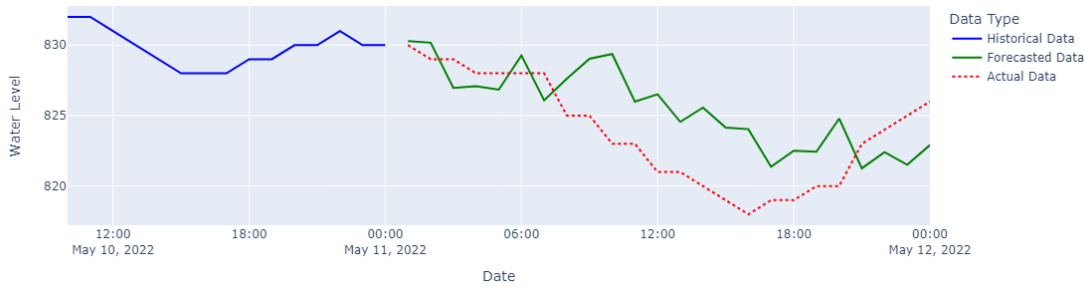


Figure 34: Plot of forecasted hourly water levels at the Monte P.Te Revere shallow point for the next 24 hours using the LSTNet model.

For example, the forecast from Monte P.Te Revere (see Figure 34) for the next 24 hours illustrates a favorable trend of improved navigation safety, with water levels ranging from 820 to 840 cm. These levels exceed the maximum ship draft length of 250 cm over the forecasted time frame, providing safe passage to ships. This information is crucial for understanding the relative risks associated with different ship draft lengths and guiding navigation decisions accordingly.

9.4. Limitations and Validity

When considering the limitations and validation aspects of this thesis, several critical points emerge concerning the forecasting of river depth and the reliability of the available data.

The primary limitation is reliance on daily data for river depth forecasting, which inherently restricts the ability to capture hourly fluctuations and variations that could significantly impact navigational planning. The original daily river depth measurements are recorded and stored based on 10-centimetre differences rather than capturing data at a finer granularity of each centimetre. The RMSE values presented in table 25 reflect the differences between actual and predicted river depth, emphasising the challenges posed by this granularity limitation. Notably, results from the table 42 in the appendix suggest that rounding forecasted river depth values to the nearest 10-centimetre increment reduces error rates when compared against actual values.

The absence of finer-grained hourly river depth data significantly limits the precision and responsiveness of forecasting models, particularly in capturing rapid changes and significant events at the shallow points along the River Po. Current hourly forecasting relies on water level stations geographically proximate to these points, which may

not fully capture the variations in river depth. Additionally, the lack of information about sand dunes at shallow points further contributes to forecasting limitations, as these features play a crucial role in shaping river dynamics and navigational conditions (Whitmeyer and FitzGerald, 2006). This data constraint poses challenges in accurately representing short-term fluctuations and rapid changes, potentially leading to inaccuracies in forecasting outcomes. Enhancing data resolution by integrating finer-grained hourly river depth measurements is crucial for improving the reliability and effectiveness of river depth forecasting models. Additionally, incorporating detailed information on environmental features like sand dunes can enhance maritime safety and risk assessment capabilities along the River Po.

Regarding climate data, the observed low correlation values based on findings in section 6.2 may indicate limitations in capturing the full spectrum of climate impacts on shallow points along the River Po. The reliance on daily data for climate variables may obscure finer-scale patterns and associations, limiting the scope of insights into the direct influence of climate factors on river depth variations. However, despite the challenges posed by low correlation values from daily climate data, the current analysis has successfully identified geographical points where climate factors have a noticeable impact on river depths. This positive perspective highlights the practical utility of findings, emphasising the importance of understanding localised climate impacts even with limited data resolution. By identifying specific geographical areas of influence, future research can prioritise targeted interventions and further investigate to enhance the understanding of climate-river dynamics interactions along the River Po.

Furthermore, the presence of missing data, regardless of imputation techniques, poses a substantial challenge, affecting the completeness and accuracy of the datasets used for model training and validation. Missing data points can introduce biases and uncertainties, potentially influencing the performance and reliability of the forecasting models.

Future research and data collection efforts should prioritise addressing these limitations to advance the understanding and predictive capabilities in river management and navigation.

9.5. Future Scope

By extending this thesis, future directions include integrating the developed forecasting model with the river discharge values from EFAS (Copernicus Climate Change Service, 2019) to enhance river depth predictions. Additionally, leveraging the Coperni-

cus weather forecasting model (European Centre for Medium-Range Weather Forecasts (ECMWF), 2022), post-training with historical climate data, can significantly improve predictive accuracy by providing advanced weather predictions. Incorporating data on extreme weather events from historical climate records into the model would further enhance its resilience and capacity to anticipate disruptions such as storms, heavy rainfall or droughts.

From a logistics perspective, the model can be applied to synchronise various supply chain modes, including road, rail and inland transportation. This helps to optimise corridor logistics and enhance resilience in managing river systems. These advancements collectively contribute to more robust and adaptive forecasting systems, vital for effective water management and logistics planning amidst evolving environmental challenges and extreme weather conditions.

9.6. Publication

The substantive content of this research is developed as part of a research publication titled "Towards an Automatic Tool for Resilient Waterway Transport: The Case of the Italian River Po," which has been accepted by MDPI Proceedings and is pending publication. The work will also be presented at the ITISE 2024 (10th International Conference on Time Series and Forecasting), followed by selected contributions that will be considered to be published in the book series of "Springer: Contributions to Statistics". This work represents a significant contribution to the field of statistics and logistics, particularly inland waterway transport resilience. It highlights the importance of leveraging advanced machine learning forecasting techniques for effective navigational planning along the River Po. This research aims to enhance the resilience and efficiency of waterway transport systems through data-driven insights and predictive modelling.

10. Summary

The analysis conducted within the framework of CRISTAL project focuses on developing data-driven solutions for effective navigational planning in inland waterways. The River Po in northern Italy offers significant potential for navigation by enhancing supply chain management and corridor logistics. The navigable stretch of the Po River, while offering strategic advantages, presents challenges due to shallow points that can disrupt navigation under specific hydro-metric conditions. The presence of shallow points within distinct branches is part of the navigational complexity. Developing a machine learning prediction model capable of providing timely and forecasted information on river depth levels at these shallow points was essential for overcoming logistical challenges and ensuring accurate planning.

This thesis has addressed three key research questions to enhance the understanding of river depth fluctuations and navigational risk along the River Po. Firstly, the investigation into key hydrological components revealed significant influences of river discharge rates, upstream water levels, total precipitation, temperature and snow depth on daily river depth fluctuations at shallow points (Piacenza, Monte P.Te Revere, Cavanella) within the River Po. Through experiments conducted during feature engineering of hydrological components, it was concluded that river discharge rates (correlation coefficients: 0.85, 0.92, 0.88) and upstream water levels (correlation coefficients: 0.89, 0.95, 0.87) exhibit a robust correlation, significantly influencing river depth at shallow points (Piacenza, Monte P.Te Revere, Cavanella).

From the perspective of climate attributes, like total precipitation, temperature and snow depth, show weaker correlation values (maximum is 0.46) than other hydrological attributes. However, total precipitation was found to have a notable effect on river depth fluctuations when considering geographically influencing regions (within the river catchment area). For instance, the Piacenza depth point showed the highest correlation value of 0.46 with a lag of 3 days relative to precipitation in the Graian Alps mountains, situated 135 km from the critical point. Similarly, despite the less evident correlation value, results indicate the direction of the relationship between river depth levels with snow depth and temperature over the river basin across all three shallow points, with an average lag of 359 days (based on lag starting from March).

Secondly, comparing daily and hourly forecasting abilities of machine learning algorithms (LSTM, GRU, LSTNet, VAR) revealed varying performance in predicting water depth levels at critical points along the River Po. Among these models, the hybrid

approach using the LSTNet model demonstrated notable forecasting accuracy with the lowest RMSE values ranging from 13.778 cm to 25.211 cm for daily forecasts and 10.515 cm to 22.801 cm for hourly forecasts across the critical points. Additionally, the LSTNet model exhibited relatively shorter training times than other recurrent networks like LSTM. This combination of strong predictive performance and efficiency underlines the effectiveness of the LSTNet model in achieving a favourable accuracy, outperforming traditional recurrent networks and statistical approaches like VAR.

Lastly, assessing navigational risk probabilities for different ship classes at critical shallow points using the best-performing machine learning algorithm (LSTNet) provided valuable insights into safety considerations and navigational planning along the River Po. Projections from Monte P.Te Revere and Cavanella demonstrate higher probabilities of safe navigation across the forecasted period for all draft lengths except for the 250 class. The noticeable reduction in probability for the 250 class suggests caution may be warranted for navigation within this specific draft category. By leveraging results from hourly water level forecasts and visualising trends in water levels near shallow points, more profound insights can be gained to enhance planning for upcoming hours.

The findings from this study highlight the critical need to enhance data resolution and refine predictive modelling techniques to strengthen waterway transport resilience. Despite notable advancements, limitations arise from factors such as data granularity, sample size, absence of sand dunes information and climate data, which constrain the accuracy and resolution of current forecasting models. Future research efforts prioritise exploring fine granular data collection methods and integrating additional environmental variables to improve forecasting model precision.

In conclusion, this thesis provides valuable insights into the complex relationships among hydrology, machine learning and logistics along the River Po. This study also helps future research and practical applications in waterway transport resilience and management, especially in the Po River. This analysis acts as a foundation step for advancing research in hydrology, digital twin technology and synchronous logistics models that integrate diverse transportation methods to achieve environmental objectives. Furthermore, in the domain of data science, this work demonstrates the effectiveness of various machine learning models through comparison. It also demonstrates applying advanced machine-learning techniques to address real-time challenges in the logistics domain. This integrated approach holds significant promise for enhancing waterway transport resilience and sustainability

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Ajewole, K., samuel oluwaseun, A., and Jemilohun, V. (2020). Test for stationarity on inflation rates in nigeria using augmented dickey fuller test and phillips-persons test. Vol. 16:11–14.

Ambra, T., Caris, and Macharis, C. (2019). Towards freight transport system unification: reviewing and combining the advancements in the physical internet and synchromodal transport research. *International Journal of Production Research*, 57(6):1606–1623.

Arttna, S., Ferraresl, M., Todini, E., and Zoccoli, F. (1990). Water resources planning of the river po basin: a surface and groundwater hydrology study.

Atashi, V., Gorji, H. T., Shahabi, S. M., Kardan, R., and Lim, Y. H. (2022). Water level forecasting using deep learning time-series analysis: A case study of red river of the north. *Water*, 14(12):1971.

Beniston, M. (2012). Impacts of climatic change on water and associated economic activities in the swiss alps. *Journal of Hydrology*, 412:291–296.

Borovykh, A., Bohte, S., and Oosterlee, C. W. (2018). Conditional time series forecasting with convolutional neural networks.

Borwarnginn, P., Haga, J. H., and Kusakunniran, W. (2022). Predicting river water height using deep learning-based features. *ICT Express*, 8(4):588–594.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer New York, New York, NY.

Castellarin, A., Di Baldassarre, G., and Brath, A. (2011). Floodplain management strategies for flood attenuation in the river po. *River Research and Applications*, 27(8):1037–1047.

Chatfield, C. (1986). Exploratory data analysis. *European Journal of Operational Research*, 23(1):5–13.

Cho, K., van Merrienboer, B., Bahdanau, D., and Bengio, Y. (2014a). On the properties of neural machine translation: Encoder-decoder approaches. *CoRR*, abs/1409.1259.

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014b). Learning phrase representations using RNN encoder–decoder for statistical machine translation. In Moschitti, A., Pang, B., and Daelemans, W., editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics.

Chow, V. T. (2017). *Handbook of applied hydrology*. Mcgraw-Hill Education, New York, second edition edition.

Copernicus Climate Change Service (2019). River discharge and related forecasted data from the European Flood Awareness System. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed on 24-Feb-2024.

Copernicus Climate Change Service (C3S) (2023). Era5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed on 26-Jun-2023.

CRISTAL project team (2022). Climate resilient and environmentally sustainable transport infrastructure, with a focus on inland waterways – CRISTAL Project. `http://www.cristal-project.eu/`. European Union's Horizon Europe research and innovation program under grant agreement no. 101069838. Accessed on 26-Jun-2023.

Cuppini, N., Frapporti, M., and Pirone, M. (2015). Logistics struggles in the po valley region: Territorial transformations and processes of antagonistic subjectivation. *South Atlantic Quarterly*, 114(1):119–134.

Dettinger, M. D. and Cayan, D. R. (1995). Large-scale atmospheric forcing of recent trends toward early snowmelt runoff in california. *Journal of Climate*, 8(3):606–623.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., and Bargellini, P. (2012). Sentinel-2: Esa's optical high-resolution mission for gmes operational services. *Remote Sensing of Environment*, 120:25–36. The Sentinel Missions - New Opportunities for Science.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2):179–211.

European Centre for Medium-Range Weather Forecasts (ECMWF) (2022). Ensemble Forecasts. *ECMWF Forecast Products*. Accessed on 24-Feb-2024.

Federal Institute of Hydrology (BfG) (2020). Major River Basins of the World / Global Runoff Data Centre, GRDC. Accessed on 26-Jun-2023.

Fiseha, B. M., Setegn, S. G., Melesse, A. M., Volpi, E., and Fiori, A. (2014). Impact of climate change on the hydrology of upper tiber river basin using bias corrected regional climate model. *Water Resources Management*, 28(5):1327–1343.

Ghimire, B. N. S. (2017). Application of arima model for river discharges analysis.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Greunen, J. V., Heymans, A., Heerden, C. V., and Vuuren, G. V. (2014). The prominence of stationarity in time series forecasting. *Studies in Economics and Econometrics*, 38(1):1–16.

Hamzaçebi, C., Akay, D., and Kutay, F. (2009). Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting. *Expert Systems with Applications*, 36(2, Part 2):3839–3844.

Hartini, S., Hadi, M. P., Sudibyakto, S., and Poniman, A. (2015). Application of vector auto regression model for rainfall-river discharge analysis. In *Forum Geografi*, volume 29.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2023). ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Accessed on 26-Jun-2023.

Hewamalage, H., Bergmeir, C., and Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1):388–427.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9:1735–80.

Hoyer, S. and Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1).

Interregional Agency of the Po River (2011). AIPO - Agenzia Interregionale del fiume Po. Accessed on 26-Jun-2023.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., et al. (2014). The climate of daily precipitation in the alps: development and analysis of a high-resolution grid dataset from pan-alpine rain-gauge data. *International Journal of Climatology*, 34(5):1657–1675.

Jordahl, K., den Bossche, J. V., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., Rey, S., Ren, C., Arribas-Bel, D., Wasser, L., Wolf, L. J., Journois, M., Wilson, J., Greenhall, A., Holdgraf, C., Filipe, and Leblanc, F. (2020). geopandas/geopandas: v0.8.1.

Kazimierski, L., Re, M., and Menéndez, A. (2012). Impacts of changes in river discharge on sedimentation of the paranÁ navigation route.

Kirilenko, Y. and Epifantsev, I. (2023). Automatic recognition of draft marks on a ship's board using deep learning system. In Guda, A., editor, *Networked Control Systems for Connected and Automated Vehicles*, pages 1393–1401, Cham. Springer International Publishing.

Krizhevsky, A., Sutskever, I., and Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25.

Lai, G., Chang, W.-C., Yang, Y., and Liu, H. (2018). Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 95–104.

Lin, H., Amin Gharehbaghi, Q. Z., Shahab S. Band, Hao Ting Pai, K.-W. C., and Mosavi, A. (2022). Time series-based groundwater level forecasting using gated recurrent unit deep neural networks. *Engineering Applications of Computational Fluid Mechanics*, 16(1):1655–1672.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Mazzetti, C., Carton de Wiart, C., Gomes, G., Russo, C., Decremer, D., Ramos, A., Grimaldi, S., Disperati, J., Ziese, M., Schweim, C., Sanchez Garcia, R., Jacobson, T., Salamon, P., and Prudhomme, C. (2023). River discharge and related historical data from the european flood awareness system, v5.0. European Commission, Joint Research Centre (JRC). Accessed on 22-Jul-2023.

Montanari, A. (2012). Hydrology of the po river: looking for changing patterns in river discharge. *Hydrology and Earth System Sciences*, 16(10):3739–3747.

Montanari, A., Nguyen, H., Rubinetti, S., Ceola, S., Galelli, S., Rubino, A., and Zanchettin, D. (2023). Why the 2022 po river drought is the worst in the past two centuries. *Science Advances*, 9(32):eadg8304.

Montgomery, D., Jennings, C., and Kulahci, M. (2011). *Introduction to Time Series Analysis and Forecasting*. Wiley Series in Probability and Statistics. Wiley.

Mugade, U. and Sapkale, J. (2015). Influence of aggradation and degradation on river channels : A review. *International Journal of Engineering and Technical Research (IJETR)*, 3:209–212.

Mushtaq, R. (2011). Augmented dickey fuller test. Available at SSRN: `https://ssrn.com/abstract=1911068` or `http://dx.doi.org/10.2139/ssrn.1911068`.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Plotly Technologies Inc. (2015). Collaborative Data Science. Plotly Technologies Inc., Montreal, QC.

Python Software Foundation (2023). Python 3.11.8. `https://www.python.org/downloads/release/python-3118/`.

Qiusheng Wu (2021). Leafmap: A python package for interactive mapping and geospatial analysis with minimal coding in a jupyter environment. *Journal of Open Source Software*, 6(63):3414.

Ravazzani, G., Barbero, S., Salandin, A., Senatore, A., and Mancini, M. (2015). An integrated hydrological model for assessing climate change impacts on water resources of the upper po river basin. *Water Resources Management*, 29(4):1193–1215.

Seabold, S. and Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

Shumway, R. H. and Stoffer, D. S. (2017). *Characteristics of Time Series*, pages 1–44. Springer International Publishing, Cham.

Staudemeyer, R. C. and Morris, E. R. (2019). Understanding LSTM - a tutorial into long short-term memory recurrent neural networks. *CoRR*, abs/1909.09586.

Vezzoli, R., Mercogliano, P., Pecora, S., Zollo, A. L., and Cacciamani, C. (2015). Hydrological simulation of po river (north italy) discharge under climate change scenarios using the rcm cosmo-clm. *The Science of the total environment*, 521-522:346–358.

Wang, C., Wu, Q., Weimer, M., and Zhu, E. (2021). Flaml: A fast and lightweight automl library. In Smola, A., Dimakis, A., and Stoica, I., editors, *Proceedings of Machine Learning and Systems*, volume 3, pages 434–447.

Wang, Z., Yan, W., and Oates, T. (2016). Time series classification from scratch with deep neural networks: A strong baseline.

Whitmeyer, S. and FitzGerald, D. (2006). Sand waves that impede navigation of coastal inlet navigation channels. page 33.

Witt, A., Kurths, J., and Pikovsky, A. (1998). Testing stationarity in time series. *Phys. Rev. E*, 58:1800–1810.

Yang, X. and Zhang, Z. (2022). A cnn-lstm model based on a meta-learning algorithm to predict groundwater level in the middle and lower reaches of the heihe river, china. *Water*, 14(15).

Yao, S., Hu, S., Zhao, Y., Zhang, A., and Abdelzaher, T. (2017). Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In *Proceedings of the 26th international conference on world wide web*, pages 351–360.

Yin, J., Rao, W., Yuan, M., Zeng, J., Zhao, K., Zhang, C., Li, J., and Zhao, Q. (2019). Experimental study of multivariate time series forecasting models. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, page 2833–2839, New York, NY, USA. Association for Computing Machinery.

Zwicklhuber, T. and Kaufmann, M. (2023). Euris (european river information services system) – the central european ris platform. In *Proceedings of the International Conference on River Navigation and Information Systems*, pages 850–856. Springer, Singapore.

Appendices

A. Additional Figures

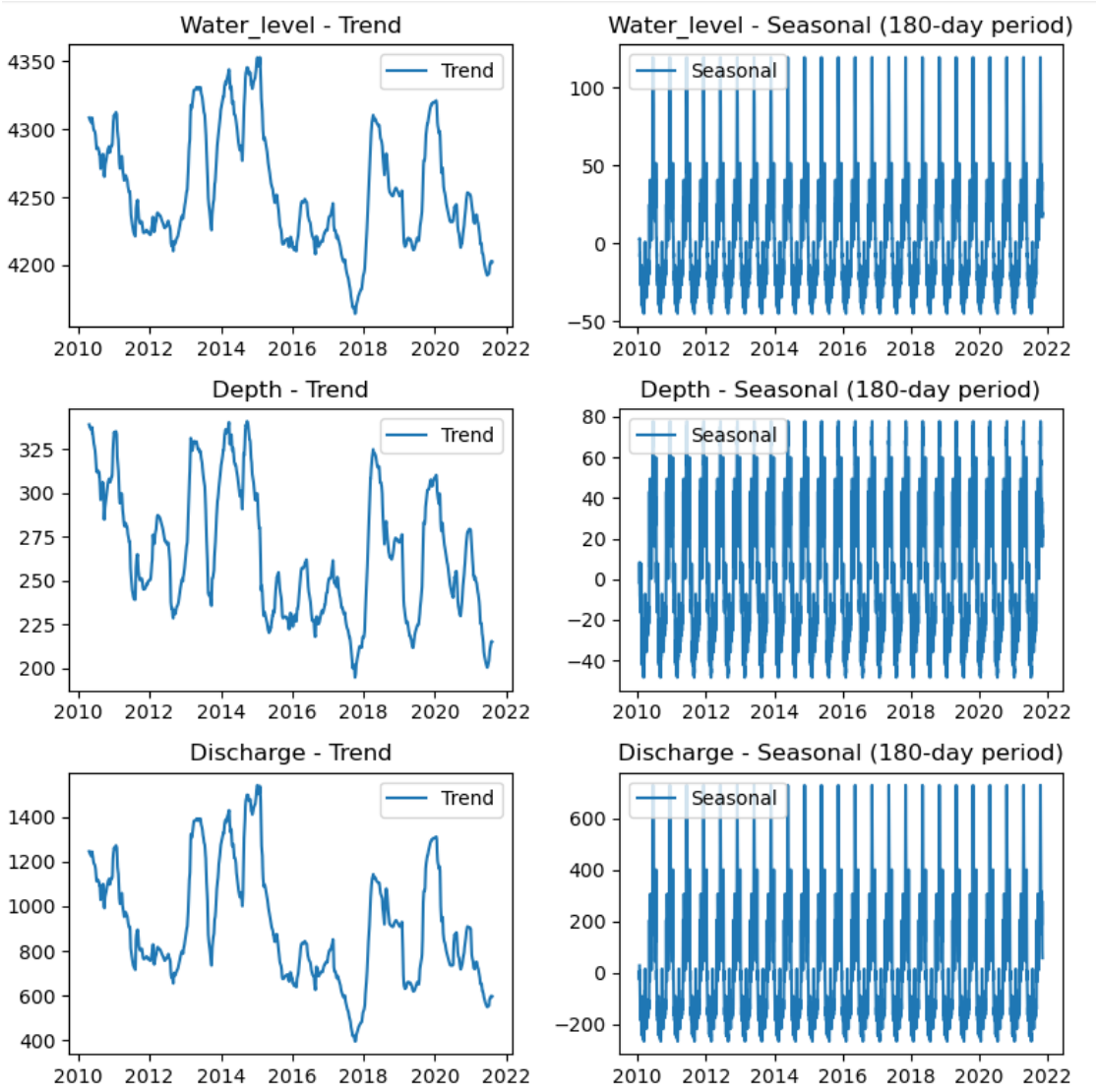


Figure 35: Graphs from seasonal decompose method using the additive model on Piacenza dataset.

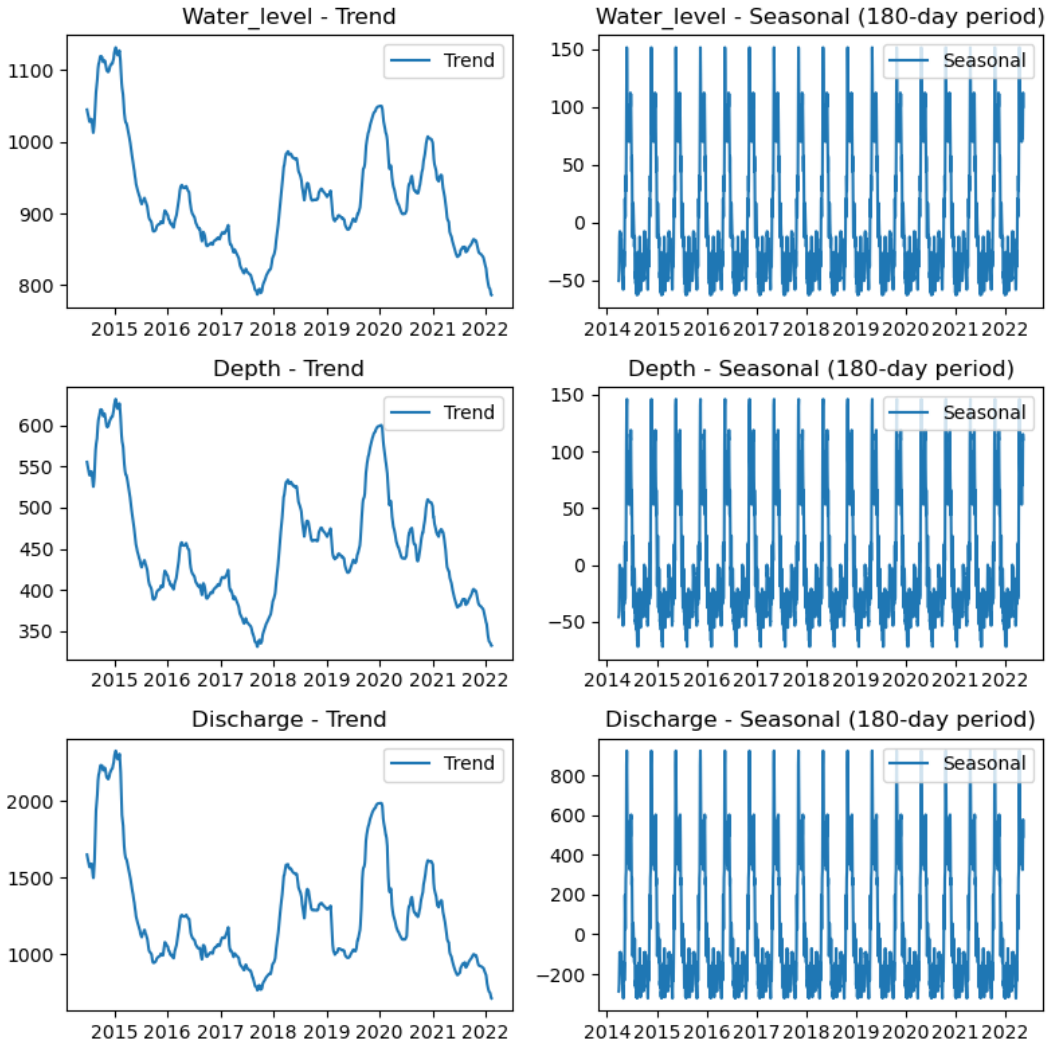


Figure 36: Graphs from seasonal decompose method using the additive model on Monte P.Te Revere dataset.

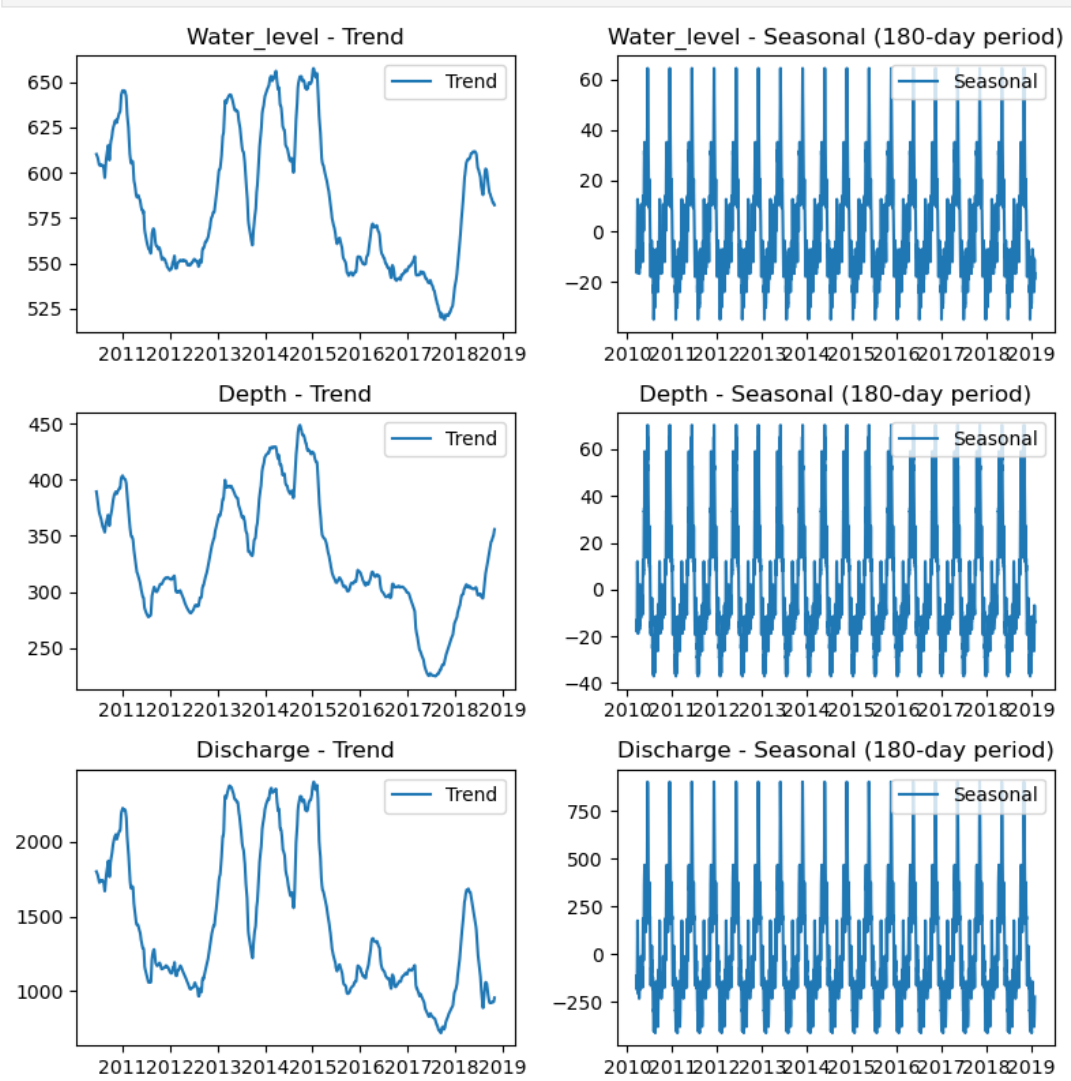


Figure 37: Graphs from seasonal decompose method using the additive model on Cavanella dataset.

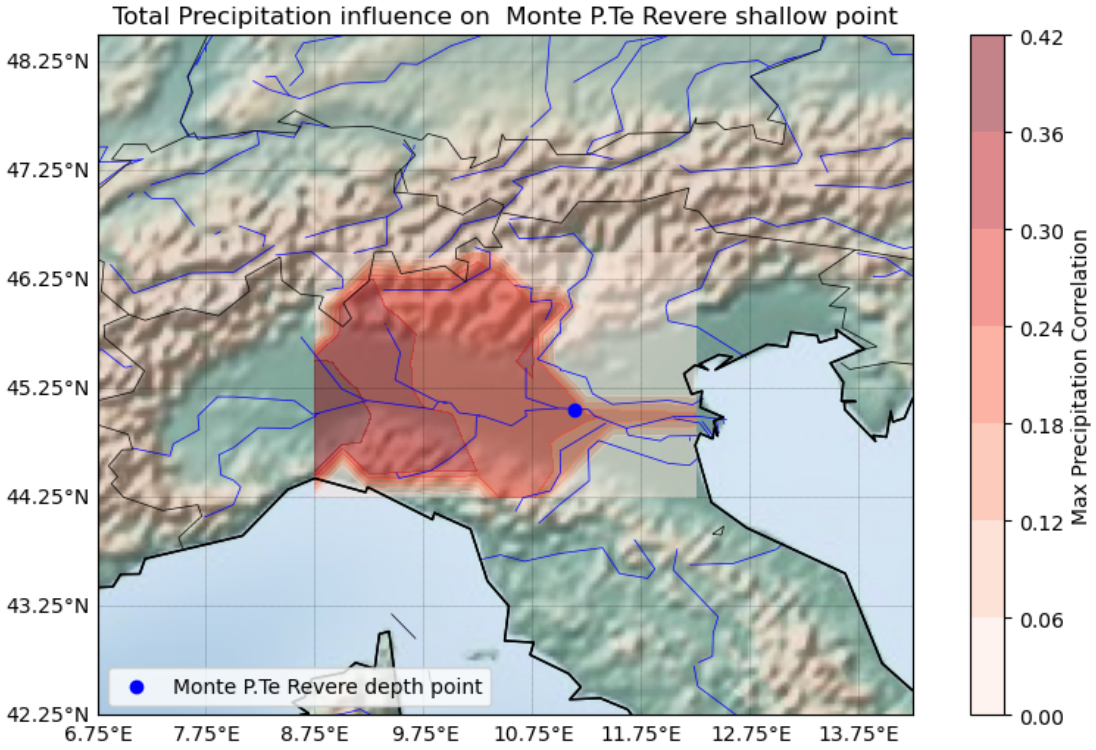


Figure 38: Map highlighting regions (in red) with maximum cross-correlation between total precipitation and river depth at Monte P.Te Revere (blue point).

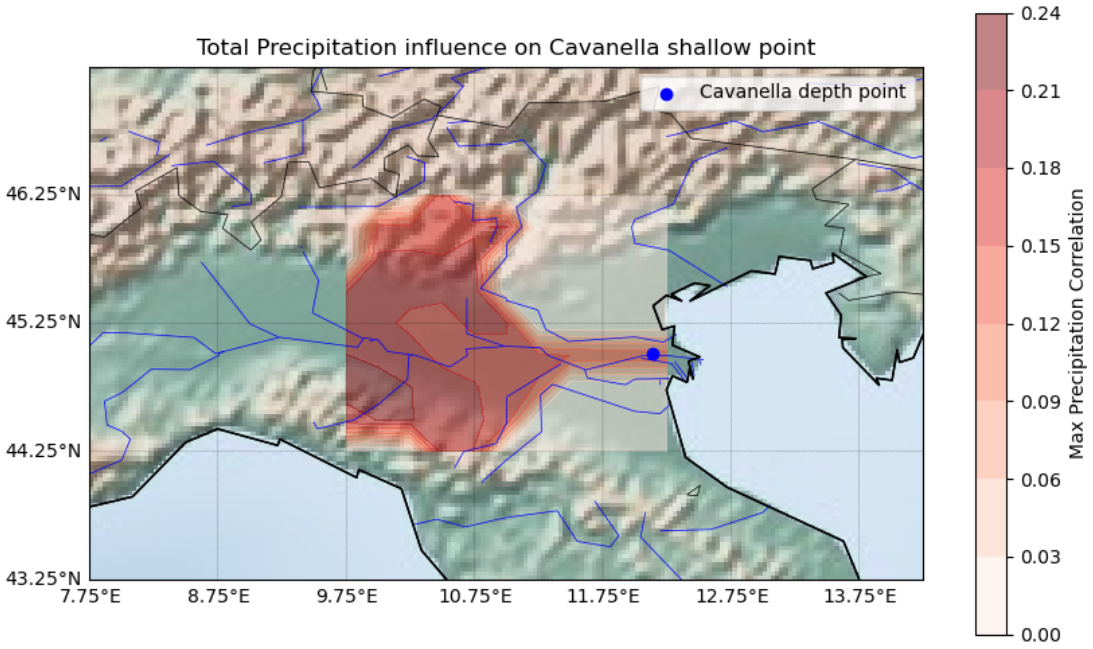


Figure 39: Map highlighting regions (in red) with maximum cross-correlation between total precipitation and river depth at Cavanella (blue point).

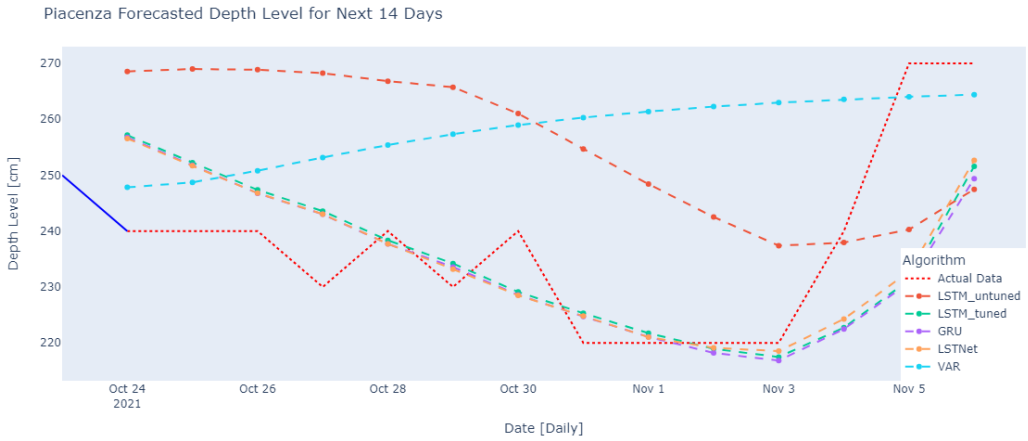


Figure 40: Plot of forecasted daily depth levels at the Piacenza shallow point for the next 14 days by selected models.

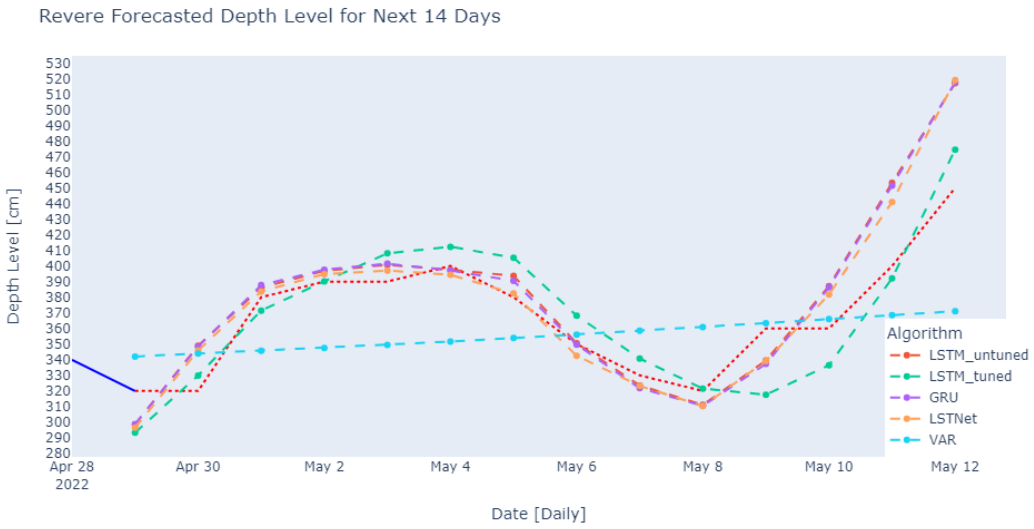


Figure 41: Plot of forecasted daily depth levels at the Monte P.Te Revere shallow point for the next 14 days by selected models.

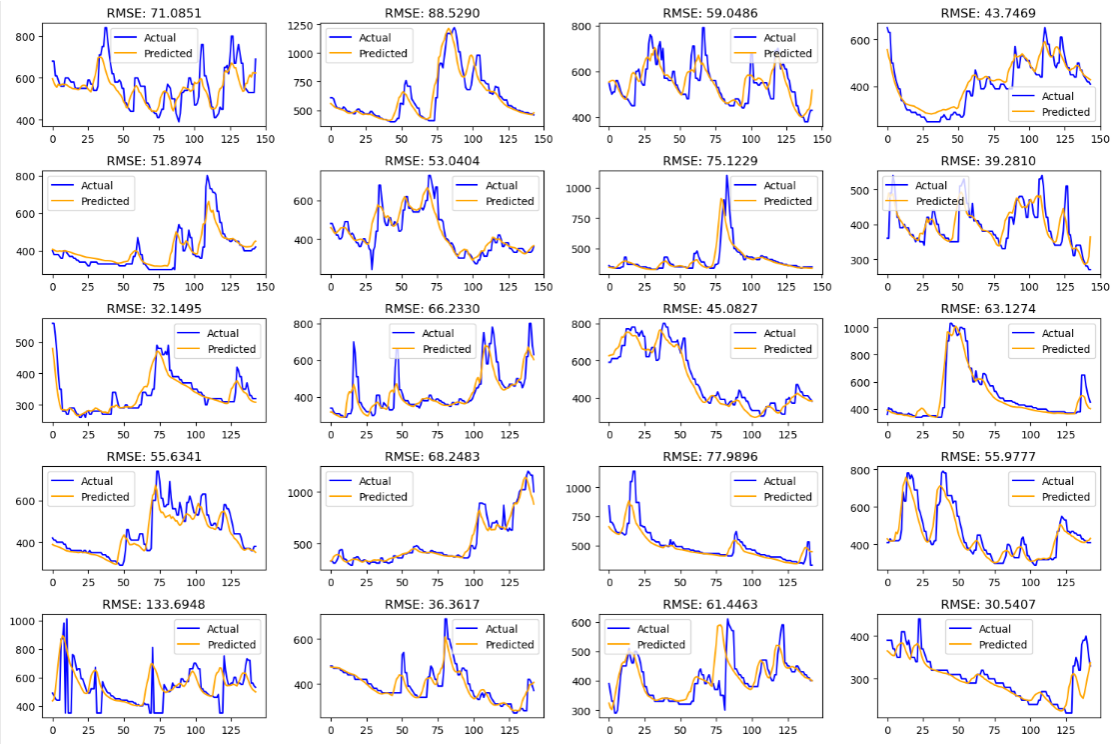


Figure 42: Performance of the LSTnet model on the Monte P.Te Revere dataset at each cross- validation fold, comparing differences among actual and predicted values using the RMSE metric.

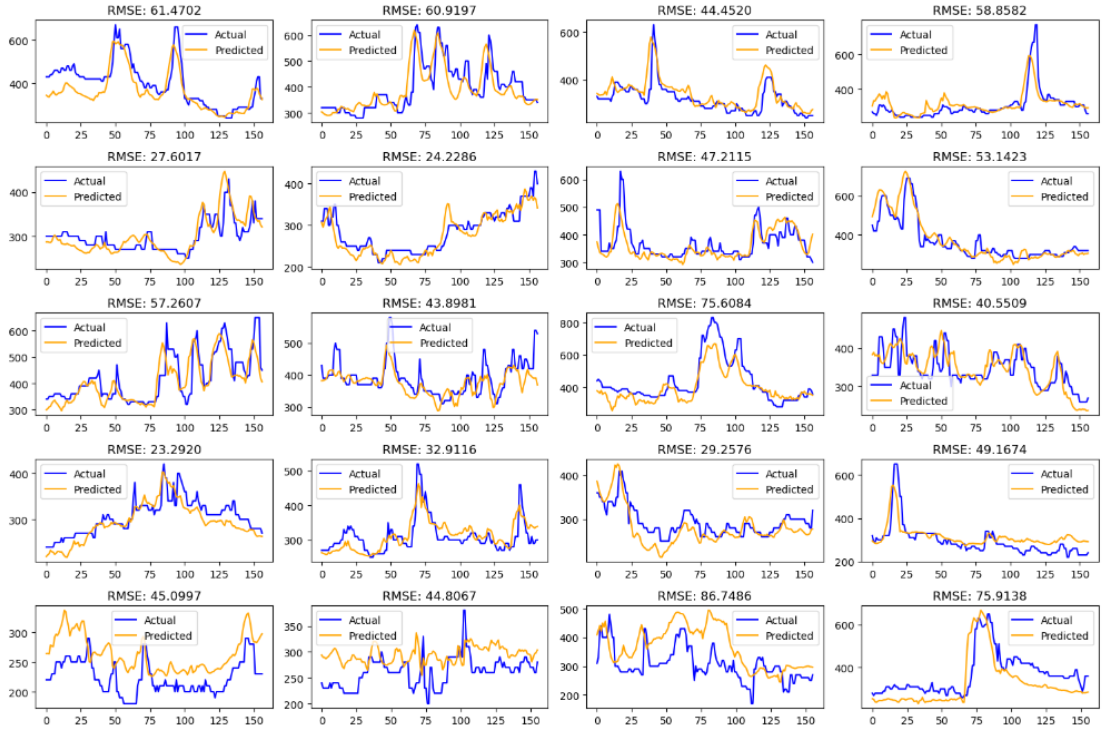


Figure 43: Performance of the LSTnet model on the Cavanella dataset at each cross- validation fold, comparing differences among actual and predicted values using the RMSE metric.

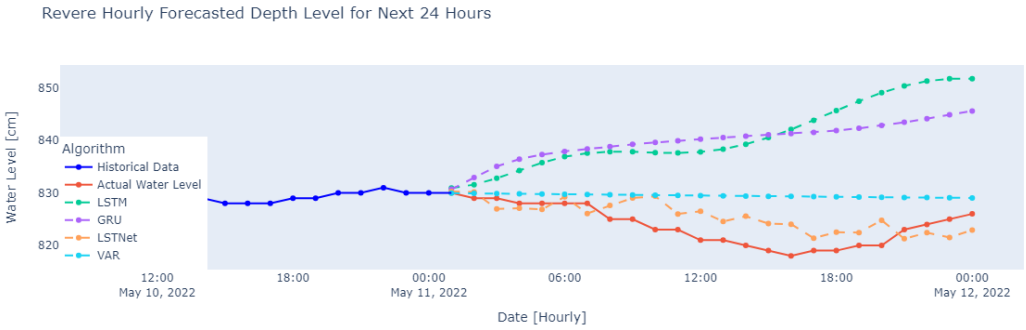


Figure 44: Plot of forecasted hourly upstream water levels at the Monte P.Te Revere shallow point for the next 24 hours by selected models.

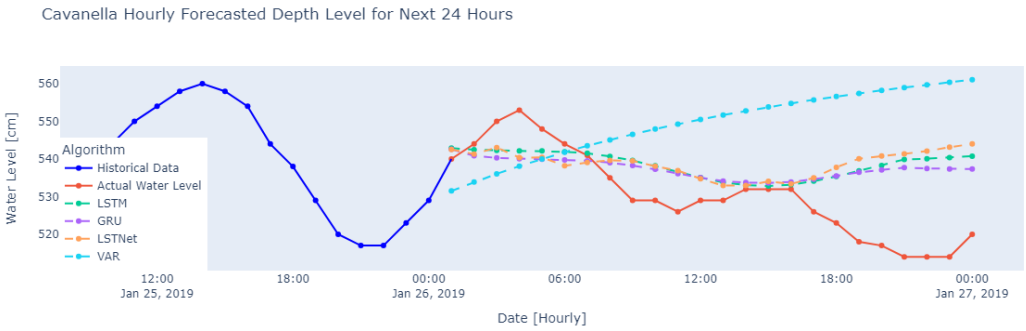


Figure 45: Plot of forecasted hourly upstream water levels at the Cavanella shallow point for the next 24 hours by selected models.

B. Additional Tables

Location	Distance	Lag	Correlation value
45.25, 8.0	135.312866	3	0.465442
45.25, 7.75	154.701999	3	0.459719
45.5, 7.75	160.577976	3	0.452488
45.0, 7.75	153.705467	3	0.449365
45.5, 8.0	142.080649	3	0.448031

Table 27: Top 5 locations exhibiting the highest cross-correlation with total precipitation at the Piacenza shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
45.0, 8.0	134.080365	359	−0.127683
44.5, 8.75	97.620648	359	−0.127646
44.25, 7.5	196.158305	2	−0.127481
44.75, 8.0	138.530966	359	−0.127070
44.75, 7.75	157.680921	359	−0.125815

Table 28: Top 5 locations exhibiting the highest cross-correlation with temperature at the Piacenza shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
46.25, 10.0	134.456731	360.0	−0.260711
46.25, 9.75	132.524294	360.0	−0.246129
46.25, 10.25	139.102622	361.0	−0.240595
46.0, 10.5	121.655949	362.0	−0.201759
46.0, 10.75	132.640987	365.0	−0.200127

Table 29: Top 5 locations exhibiting the highest cross-correlation with snow depth at the Piacenza shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
45.25,8.75	188.201536	4	0.404489
45.0,8.75	187.462622	4	0.390845
44.75,8.75	190.813378	4	0.383074
45.5,8.75	192.983149	4	0.380024
45.0,9.0	167.829150	4	0.379251

Table 30: Top 5 locations exhibiting the highest cross-correlation with total precipitation at the Monte P.Te Revere shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
44.5,8.75	198.046331	361	-0.209718
44.5,9.25	161.022213	361	-0.207072
44.75,8.75	190.813378	361	-0.204479
44.5,9.5	142.999393	359	-0.204303
44.5,10.0	108.747295	359	-0.203626

Table 31: Top 5 locations exhibiting the highest cross-correlation with temperature at the Monte P.Te Revere shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
46.25,10.0	159.430942	175.0	-0.141639
46.5,10.25	174.664635	161.0	-0.141414
46.25,10.25	149.567416	167.0	-0.136123
45.5,10.75	57.869833	210.0	-0.133654
45.5,10.5	70.053507	210.0	-0.131260

Table 32: Top 5 locations exhibiting the highest cross-correlation with snow depth at the Monte P.te Revere shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
44.75, 10.0	170.882605	4	0.218575
44.75, 9.75	190.303266	4	0.217492
44.5, 10.25	159.672820	4	0.211367
44.75, 10.25	151.532181	4	0.210845
44.5, 10.0	178.216901	4	0.210784

Table 33: Top 5 locations exhibiting the highest cross-correlation with total precipitation at the Cavanella shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
44.25, 10.5	155.124716	360	−0.280212
44.5, 10.0	178.216901	360	−0.279988
44.5, 9.75	196.992590	361	−0.279535
44.25, 10.75	139.052858	359	−0.276898
44.5, 10.25	159.672820	360	−0.276556

Table 34: Top 5 locations exhibiting the highest cross-correlation with temperature at the Cavanella shallow point over a specified number of days (lag).

Location	Distance	Lag	Correlation value
45.5, 10.75	120.893230	216.0	−0.176565
45.5, 10.25	156.950121	168.0	−0.171089
45.5, 10.5	138.713599	168.0	−0.168527
45.5, 10.0	175.472948	176.0	−0.124840
46.25, 10.5	186.794230	166.0	−0.119593

Table 35: Top 5 locations exhibiting the highest cross-correlation with snow depth at the Cavanella shallow point over a specified number of days (lag).

Layers	Activation Function	Units	Epochs	Batch Size	RMSE
3	tanh	64	339	6	13.985152
3	tanh	70	339	6	14.079039
3	tanh	68	322	6	14.308261
3	tanh	60	343	18	14.309830
3	tanh	68	358	6	14.310603

Table 36: Results of the top 5 model configurations tuned using the FLAML tool on Piacenza daily data.

Layers	Activation Function	Units	Epochs	Batch Size	RMSE
3	tanh	109	11	179	18.904749
3	tanh	93	19	251	18.999330
3	tanh	96	10	246	19.803955
3	tanh	96	10	239	19.914644
3	tanh	93	18	242	20.095669

Table 37: Results of the top 5 model configurations tuned using the FLAML tool on Monte P.Te Revere daily data.

Layers	Activation Function	Units	Epochs	Batch Size	RMSE
3	relu	122	192	53	19.687378
3	relu	36	382	55	19.704619
3	relu	118	206	92	19.718627
3	relu	102	262	51	19.914872
3	relu	104	276	60	19.974990

Table 38: Results of the top 5 model configurations tuned using the FLAML tool on Cavanella daily data.

Forecast	Depth	LSTM (Vanilla)	LSTM (tuned)	GRU	LSTNet	VAR
24-10-2021	240.00	268.56	257.16	256.75	256.56	247.82
25-10-2021	240.00	268.99	252.24	251.79	251.69	248.72
26-10-2021	240.00	268.86	247.36	246.77	246.78	250.80
27-10-2021	230.00	268.25	243.58	243.00	243.04	253.16
28-10-2021	240.00	266.82	238.35	237.69	237.71	255.39
29-10-2021	230.00	265.74	234.20	233.57	233.18	257.34
30-10-2021	240.00	261.02	229.10	228.58	228.49	258.97
31-10-2021	220.00	254.67	225.30	224.73	224.79	260.30
01-11-2021	220.00	248.41	221.73	221.06	221.03	261.39
02-11-2021	220.00	242.52	218.93	218.20	219.10	262.26
03-11-2021	220.00	237.38	217.47	216.83	218.52	262.97
04-11-2021	240.00	237.94	222.73	222.46	224.26	263.55
05-11-2021	270.00	240.27	231.23	230.80	232.77	264.02
06-11-2021	270.00	247.48	251.57	249.37	252.66	264.41

Table 39: Forecast of daily river depth over a 14-day period at the Piacenza shallow point for selected models.

Forecast	Depth	LSTM (Vanilla)	LSTM (tuned)	GRU	LSTNet	VAR
29-04-2022	320.00	298.62	293.31	298.73	296.55	342.07
30-04-2022	320.00	348.94	329.92	348.46	345.94	343.91
01-05-2022	380.00	386.19	371.48	387.76	383.76	345.72
02-05-2022	390.00	396.90	390.20	397.66	394.80	347.59
03-05-2022	390.00	400.75	408.08	401.50	397.14	349.56
04-05-2022	400.00	397.63	412.32	397.11	394.38	351.65
05-05-2022	380.00	393.71	405.36	390.48	382.46	353.85
06-05-2022	350.00	350.56	368.23	349.74	342.53	356.15
07-05-2022	330.00	323.52	340.73	321.79	323.26	358.53
08-05-2022	320.00	311.08	321.52	310.52	310.37	360.97
09-05-2022	360.00	339.46	317.58	337.28	339.15	363.46
10-05-2022	360.00	387.14	336.50	386.32	381.73	365.98
11-05-2022	400.00	453.28	392.11	451.52	440.99	368.53
12-05-2022	450.00	517.02	474.54	517.39	519.04	371.08

Table 40: Forecast of daily river depth over a 14-day period at the Monte P.Te Revere shallow point for selected models.

Forecast	Depth	LSTM (Vanilla)	LSTM (tuned)	GRU	LSTNet	VAR
14-01-2019	390.00	363.02	362.92	359.84	363.62	383.58
15-01-2019	350.00	358.54	338.57	347.81	335.19	378.47
16-01-2019	330.00	353.37	315.36	336.84	314.40	374.38
17-01-2019	310.00	347.84	292.25	326.88	297.91	371.09
18-01-2019	290.00	343.06	280.53	322.05	286.80	368.40
19-01-2019	310.00	346.16	296.82	329.67	298.28	366.20
20-01-2019	360.00	350.72	311.51	340.73	318.21	364.36
21-01-2019	360.00	350.10	346.23	344.18	345.08	362.80
22-01-2019	360.00	350.24	371.40	347.22	373.68	361.46
23-01-2019	360.00	351.29	396.00	349.81	395.98	360.28
24-01-2019	380.00	352.32	413.00	353.12	415.74	359.22
25-01-2019	380.00	352.33	398.32	353.60	398.47	358.25
26-01-2019	380.00	353.57	383.21	356.16	382.38	357.35
27-01-2019	400.00	348.80	365.62	346.95	362.96	356.51

Table 41: Forecast of daily river depth over a 14-day period at the Cavanella shallow point for selected models.

RMSE	Piacenza	Monte P.Te Revere	Cavanella
Original	71.3509	64.5289	52.001
Floored	51.5148	43.1486	36.2442

Table 42: Comparison of average original versus floored RMSE values obtained from 20-fold cross-validation of the LSTNet model across different shallow points.

Eidesstattliche Versicherung

(Affidavit)

Dhavaleswarapu, Sohith

Name, Vorname
(surname, first name)

224839

Matrikelnummer
(student ID number)

☐ Bachelorarbeit
(Bachelor's thesis)

■ Masterarbeit
(Master's thesis)

Titel
(Title)

Advancing Inland Waterways Logistics: Monitoring Navigability Through Water Level Forecasting, Predictive Analysis and Feature Engineering of Hydrological Data

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem oben genannten Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present thesis with the above-mentioned title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution before.

Dortmund, 15-05-2024

Ort, Datum
(place, date)

Unterschrift
(signature)

Belehrung:
Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin") zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:
Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the Chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, Section 63 (5) North Rhine-Westphalia Higher Education Act (*Hochschulgesetz, HG*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund University will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:*

Dortmund, 15-05-2024

Ort, Datum
(place, date)

Unterschrift
(signature)

***Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**