Czech University of Life Sciences Prague

Faculty of Environmental Sciences

Department of Water Resources and Environmental Modeling

Bachelor Thesis

Validation of HEC-RAS model for Berounka catchment

Egor Suslin

Validation of HEC-RAS model for Berounka catchment

The objective of the thesis is:

- to review existing approaches to 1D hydraulic modeling
- to optimize the model setup to warrant stability under hydrological extremes
- to validate the simulation of the HEC-RAS model over the Berounka river basin with respect to observed runoff

1. Data preparation

The data from the HAMR system will be used to drive the HEC-RAS model. Therefore multi-day periods representing normal, drought and high flow conditions will be identified in the dataset and general statistics of the events will be provided. The events will be split into calibration and validation sets.

2. HEC-RAS simulation

The HEC-RAS model will be calibrated and validated with respect to observed data using standard hydrological model performance metrics with focus on extremes.

3. Python interface

The Python HEC-RAS interface will be used to efficiently enter the data into the model, for running the model and getting the simulated values. Consider solutions for running on a Linux server.

Keywords:

HEC-RAS, prediction

Recommended information sources:

1. Dysarz, T. (2018). Application of Python scripting Techniques for control and automation of HEC-RAS simulations. Water, 10(10), 1382.
2. Hanasaki, N., Kanae, S., & Oki, T. (2006). A reservoir operation scheme for global river routing models. Journal of Hydrology, 327(1-2), 22-41.
3. https://www.hec.usace.army.mil/software/hec-ras/
4. Pathan, A. I., & Agnihotri, P. G. (2020). Application of new HEC-RAS version 5 for 1D hydrodynamic flood modeling with special reference through geospatial techniques: a case of River Purna at Navsari, Gujarat, India. Modeling Earth Systems and Environment, 1-12.
5. Renssen, H., & Knoop, J. M. (2000). A global river routing network for use in hydrological modeling. Journal of Hydrology, 230(3-4), 230-243.
6. Thakur, B., Parajuli, R., Kalra, A., Ahmad, S., & Gupta, R. (2017). Coupling HEC-RAS and HEC-HMS in precipitation runoff modelling and evaluating flood plain inundation map. In World Environmental and Water Resources Congress 2017 (pp. 240-251).

7. Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., & Samaniego, L. (2019). The multiscale routing model mRM v1. 0: simple river routing at resolutions from 1 to 50 km. Geoscientific Model Development, 12(6), 2501-2521.

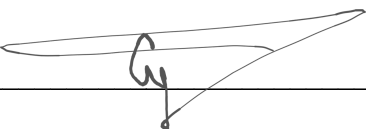Expected date of thesis defense: 2022/23 SS - FES

Advisor of thesis: Ing. Luděk Bureš, Ph.D.

Declaration

I declare that I have worked on my bachelor thesis titled "Validation of HEC-RAS model for Berounka catchment" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the bachelor thesis, I declare that the thesis does not break any copyrights.

In Prague on 28.02.2023

Validation of the HEC-RAS model for Berounka catchment

Abstract

Since its public release in 1995, HEC-RAS has become one of the most widely used software for performing hydraulic calculations for natural and constructed channels. It has proven to be a valuable tool for a wide range of hydrological analyses, but it also has some serious limitations. One of the challenges in using HEC-RAS is the manual updating of data, which can be time-consuming and prone to error. In this study, we aim to address this issue by developing and utilizing a Python interface for the HEC-RAS model to efficiently simulate runoff in the Berounka river basin. The study will utilize data from the HAMR system, a comprehensive dataset of hydrometeorological observations, and will focus on modeling normal, drought, and high-flow conditions. The HEC-RAS model will be calibrated and validated using standard hydrological performance metrics, with a particular emphasis on extreme events. The Python interface will be used to efficiently enter and process data and will be run on a Linux server. The ultimate goal of this study is to optimize the model setup to warrant stability under hydrological extremes and to accurately simulate the observed runoff in the Berounka river basin. This work aims to improve the efficiency and reliability of hydraulic modeling using HEC-RAS.

Table of Content

1. Introduction

Hydraulic modeling is crucial for understanding and predicting river and stream behavior. It is utilized to investigate many hydrological phenomena, including droughts, floods, and the quality of water. One hydraulic modeling tool that is frequently used is the River Analysis System of the Hydrologic Engineering Center (HEC-RAS). This software is used to simulate the flow of water in rivers and streams using variables like channel geometry, flow rate, and water depth.

However, the accuracy of HEC-RAS simulations depends on the quality of the input data and the suitability of the model setup. The results of the simulation can contain significant errors due to inaccurate input data or an inappropriate model setup. Therefore, it's essential to carefully optimize the model configuration and validate the simulation results with respect to the actual data that was observed.

This study aims to validate the simulation of the model over the Berounka river basin and optimize the setup of the HEC-RAS model in relation to observed runoff. In order to accomplish this goal, we will examine current methods for 1D hydraulic modeling and assess the HEC-RAS model's applicability to the Berounka catchment. We will also optimize the model setup to ensure stability in the presence of hydrological extremes like floods and droughts. Following that, the HEC-RAS model's simulation will be examined using standard hydrological model performance metrics, with a focus on extremes.

The Berounka catchment's HEC-RAS model must be tested because it will provide more information on the hydrological behavior of the catchment. The outcomes of the HEC-RAS simulations will be more reliable with a more accurate simulation. Additionally, it will be easier to enter data and analyze the results when Python HEC-RAS functionality is used. In conclusion, this research will expand our knowledge of hydraulic modeling with HEC-RAS and offer helpful information for Berounka catchment management.

2. Objective

The main objective of this thesis is to advance the field of hydraulic modeling by improving the accuracy and efficiency of HEC-RAS model simulations. Specifically, the following three objectives have been identified:

Development of Python Interface to Automatically Update Data in HEC-RAS Model:

The first objective of this study is to develop HEC-RAS Interface using Python functionality to automatically update the input data of the hydraulic model. This interface will enable the user to import, preprocess and integrate hydrological data into the model in a more efficient manner. By automating the data input process, the interface will reduce the potential for human-made errors and improve the speed and accuracy of the simulations.

Review of Existing Approaches to 1D Hydraulic Modeling:

The second objective is to provide review of the existing approach to hydraulic modeling. This review will examine the different methods, techniques and tools used to simulate river flows. The review will also highlight recent developments in 1D hydraulic modeling, with an emphasis on their applicability to real-world scenarios.

Optimization of Model Setup for Stability under Hydrological Extremes:

The third objective of this thesis is to optimize the setup of hydraulic models to ensure stability under hydrological extremes. To achieve this objective, the thesis will explore different methods of model calibration. The aim is to identify the perfect value of the model parameters to improve the accuracy and stability of the models.

Validation of HEC-RAS Model Simulation over Berounka River Basin:

The fourth objective is to validate the simulation of the HEC-RAS model over the Berounka river basin with respect to observed runoff. This will involve comparing the results

of the model simulations to observed hydrological data, and calculating the error metrics to evaluate the accuracy.

3. Literature Review

3.1 Introduction

Hydraulic modeling is a fundamental technique for comprehending and anticipating the actions of rivers and streams. It is applied to evaluate various hydrological conditions, including floods, droughts, and water quality. In his study, Vahid Gholami (2022) utilized hydraulic modeling to predict flood discharge and flood flow depth in ungauged forested watersheds. The hydraulic model used in the study was calibrated using the data from the gauged watersheds. The study concluded that hydraulic modeling can be an effective tool for predicting flood characteristics in ungauged forested watersheds and can assist in the development of flood forecasting models and flood hazard maps (Vahid Gholami, 2022). Also, the research by Vanessa C.F. Dias et al. (2017) aimed to predict the impact of district-metered area (DMA) implementation on water quality in a full-scale drinking water distribution system using hydraulic modeling. The study discovered that hydraulic modeling can be a helpful tool for predicting the effect of DMA implementation on water quality since it offers a better understanding of the flow and pressure distribution in the system (Vanessa C.F. Dias et al., 2017).

3.2 Hydraulic modeling

Hydraulic modeling is the process of simulating water flow in rivers, streams, lakes, and coastal areas to get a better understanding of water behavior and make a good simulation of real-world flow events. Hydraulic models can be classified into several categories: one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) models. There are advantages and disadvantages for each model and can be chosen based on the project requirements.

1D hydraulic models, which replicate water flow in a single direction, are frequently used to simulate flows in rivers and natural waterways. These models do not take into consideration lateral flow or variations in flow direction since they assume that the water flow is constant across the full width of the channel. Required input data for 1D models include flow discharge data, channel geometry, and roughness coefficient.

2D hydraulic models can simulate water flow in two dimensions and are used to model flow in areas where lateral flow and changes in the flow direction are significant, such as floodplains, urban areas, and wetlands. 2D models normally require discharge data, channel geometry, roughness coefficient, and topography.

3D hydraulic models are used to model water movement in coastal areas, harbors, and estuaries by simulating water flow in three dimensions. These models account for the vertical distribution of water velocity, density, and turbulence, which is critical for accurately modeling water movement in these complex environments. Flow data, channel geometry, roughness coefficients, topography, and hydrodynamic properties such as density, temperature, and salinity might be required as inputs for 3D models. However 3D models are better suited to localized hydraulic problems. They are not typically used for flood mapping because they require a great deal of detail, which results in lengthy simulation periods.

The model has to be chosen by the requirements of the research, as those models would perform differently, depending on the complexity of the problem being addressed. For example, Glock et al. (2019) compared hydrodynamics simulated by 1D, 2D, and 3D models focusing on bed shear stresses and discovered that 2D models were better suited for modeling water flows in rivers and floodplains with moderate complexity, whereas 3D models were better suited for modeling water movement in estuaries and coastal areas with high complexity (Glock et al., 2019).

Another example is from the study by Kesler (2023), where 1D, 2D, and 3D hydraulic models were used in urban flooding applications. In comparison to 1D and 3D models, Kesler discovered that 2D hydraulic models were the most effective for simulating urban flooding events. The study discovered that 2D models accurately represented urban

floodplains, including the flow of water over the surface and through street networks. Furthermore, the author stated that 3D models were computationally expensive and difficult to set up, whereas 1D models did not capture the full range of dynamics associated with urban flooding. However, the study also showed that 1D models can be more efficient and faster to run than 2D and 3D models, making them useful for large-scale projects (Kesler, 2023).

3.3 DEM and Roughness Coefficient

The accuracy of hydraulic modeling simulations can be due to problems in the resolution of the digital elevation models (DEMs) which are used as input data. Higher-resolution DEMs can provide more useful details with regard to the topography of the area of interest, which could positively affect the model's accuracy. Using low-quality DEMs simulated less accurate results and showed decreased performance in flood prediction. But it is still possible to use low-quality DEMs for the hydraulic models if you performed manual or programmatic data correction methods or used appropriate techniques to preprocess the data, such as interpolation, data filtering, and data smoothing. Additionally, you can combine DEMs with other data sources to improve their quality. It is possible to use field measurements or high-quality satellite data (Laks et al. 2017). In this study, high-resolution DEM will be used, however, there are examples of poor-quality geometry data of river reaches, as the model simulates the whole catchment water flow.

Additionally, the right definition of hydraulic roughness coefficients also has a significant impact on the hydraulic model performance. Meticulously selecting proper hydraulic roughness coefficients can significantly improve hydraulic simulation precision. Roughness coefficients represent the resistance to flow in channels and floodplains. Roughness is usually presented in the form of a Manning's n value for a hydraulic model (USACE, 2022). The roughness Coefficient is a significant parameter that has to be validated and calibrated along with the model's data. Error in Manning's roughness coefficient value may lead to problems with simulation results. (Kumar, M., 2019). It is important to carefully consider the characteristics of the study area when selecting the hydraulic roughness coefficients to ensure that they accurately reflect the conditions of the river or stream being

modeled. In the present study, the most appropriate hydraulic roughness coefficients were selected for each study area. That was achieved after 8 rounds of simulations.

3.4 Techniques to improve the accuracy of the model.

The reliability of the hydraulic model depends on the quality of the input data and the proper model setup. If the input data is not accurate enough for the analysis or the model setup is not suitable, these are the main factors leading to mistakes in simulation results. It is crucial to fine-tune model config setup and calibrate simulation results against observed data. For this work, I prepared data for two different years from the simulation of the conceptual Bilan model for each river reach. One is used to calibrate the model while the other is used for calibration. In this chapter, I am describing different methods of improving the model performance and how they can benefit working with hydraulic modeling.

An understanding of the hydrological characteristics of the study area will help improve the accuracy of the model simulation by optimizing the model setup. This includes climate, topography, use of land, and hydrological regimes of the water object. All above can heavily impact the flow of the water in the water body and need to be accurately represented in the model data. Taking all factors into the account will help you reach the required performance of the model, but it can verify for different analyses. Bravo et al. (2012) concluded that including hydrological characteristics significantly improves the accuracy of the model. Additionally, based on the article, the coupled hydrologic-hydraulic modeling approach can provide better predictions with complex imputed hydrological characteristics. Furthermore, gaining knowledge of those characteristics can influence decisions regarding water usage, irrigation, climate change, and other environmental and water-related matters. (Bravo et al., 2012)

It is worth mentioning several different ways to optimize the hydrological simulation setup. These include sensitivity analysis, uncertainty analysis, and model calibration. All of them have their distinct purpose. Sensitivity analysis entails slightly changing the input data or model setup to measure their impact on the simulation results. In contrast, uncertainty analysis aims at quantifying the variability of the output that is due to the variability of the input data and input parameters, following which the uncertainty is propagated in the

simulation procedure. That quantification is often performed by estimating input data's general statistical and populational quantities. In their study, Tate et al. (2014) discovered that the HAZUS-MH flood model's sensitivity to input parameters varied by location and that the results of the sensitivity analysis were helpful in identifying regions where more data collection or research could increase the model's accuracy. Additionally, the authors discovered that the uncertainty analysis offered crucial details on the range of model predictions and the probability of extreme events, which can be helpful for risk management and decision-making. (Tate et al., 2014).

Finally, the model calibration techniques use the approach of modifying input data or model parameters to minimize the difference between observed and predicted data. Error metrics are used to evaluate those differences and conclude if the model outputs are appropriate for the study area. The most suitable technique should be taken based on simulated analysis, parameters of the study area, and available data. For example, model calibration of parameters can improve model accuracy and the accurate incorporation of field data and calibration techniques into hydraulic modeling may enhance flood hazard assessment and management. (Ballesteros et al., 2011)

Kalman and particle filtering are two examples of data aggregation techniques that can be used with hydraulic modeling. These strategies can considerably enhance the correctness of model simulations by incorporating observed data in a systematic and statistically stringent manner. The use of Kalman filtering effectively corrects the forecast results of flood simulation in real-time, based on studies by Xiao-Ling Wu et al. (2013). It helped them to update the hydraulic model and make it even more efficient by incorporating real-time data (Xiao-Ling W et al., 2013). Particle Filtering techniques can also be beneficial when working with hydraulic models. Specifically, Nelly Jean-Baptiste in their work found that the particle filter was able to enhance the estimation of discharge and water level, even when there were significant disturbances to the model. It can be an essential tool for real-time monitoring and forecasting of river flows and have a significant benefit for flood management (Jean-Baptiste et al., 2010). Overall, Kalman and particle filtering can be the key techniques for improving accuracy of your model, if analysis requires real-time monitoring.

Besides the previously mentioned variables, many other matters can influence the precision of model simulation results. This includes representing the underground flow of water, which is referred to as subsurface flow. This kind of flow can have a major effect on the total flow pattern and hydraulic properties of the river or waterway, and it is fundamental to precisely simulate it inside the hydraulic model. This can be hard to do because of the necessity for meticulous data on the subsurface hydrology of the catchment.

Accurate modeling of channel beds and banks is critical to the reliability of simulations. These features determine a river or stream's shape, roughness, and other physical properties, which greatly impact the flow of water in terms of velocity, direction, and turbulence. Incorrectly entered geometry data in the hydraulic model could result in inaccurate flow simulations and incorrect predictions. To ensure accurate simulation results, it is imperative to ensure the correct loading of geometry data into the model, despite the challenge posed by the complexity of channel beds and banks and limited data availability. Casas et al. (2006) in their study mentioned the impact of topographic data sources on hydraulic flood modeling. They found that the quality of the digital terrain model (DTM) has a significant impact on the performance of the hydraulic model simulations, as higher resolution DTM resulted in more accurate models (Casas et al., 2006). It is important to consider only high-quality and high-resolution topographic data sources in hydraulic modeling for accurate predictions.

The representation of the vegetation is a key component that can influence how efficiently the model simulations perform. The analysis could generate inaccurate predictions if vegetation data were not represented precisely. It is crucial to accurately represent vegetation in the model simulation since it has the possibility of having a significant impact on the flow of water in waterways. Lama in their article examines the impact of vegetation on hydraulic modeling performance and the accuracy of the final results in a drainage channel by conducting field experiments and using a 1D hydraulic model. They stated that the presence of vegetation can increase the roughness of the channel and reduce the conveyance, which results in lower water level and velocity. They experimented with a model under different vegetation scenarios, suggesting that it can be used to evaluate the performance of riparian vegetation restoration measures (Lama, G et al., 2020). However,

vegetation has a complicated structural composition and available data is very limited, so it could be challenging to archive good flora representation results.

To sum up, that chapter, employing various strategies is crucial to enhancing the model's accuracy. Statistical analysis, a graphical depiction of actual and expected data, uncertainty analysis, sensitivity analysis, model calibration, and accurate model and configuration input of data are some of the techniques presented in this paper. Additionally, it is essential to ensure all input data and hydrological parameters. The model's performance can be considerably improved by combining these strategies, and they can also give you vital information for future model enhancements.

3.5 Use of Programming Languages (Python)

Using programming languages like Python, hydraulic modeling can be made more efficient and automated. The Python interface allows users to run hydraulic model simulations and can be used to analyze the results. Additionally, it can be of great assistance in updating model input data and can improve the accuracy of imported data while reducing human error. This can be very beneficial for challenging and large modeling projects that involve the analysis of a large amount of data. Especially if high-quality data input is required for the project. Tomasz Dysarz wrote in his article that Python scripting allows for the development of a custom workflow environment for the specific needs of the project. This can reduce the number of mistakes, time spent and effort required for modeling, allowing more flexibility in the model (Tomasz Dysarz, 2018). My goal was to develop a Python code that could quickly load or upgrade data for 1D hydraulic model simulations based on the user's configuration. The goal was to develop a dynamic tool that would assist scientists utilizing different modeling analyses.

3.6 Common Hydraulic Models

HEC-RAS

HEC-RAS (Hydrologic Engineering Center's River Analysis System) is a commonly employed hydraulic model and is used to reproduce the passage of water in rivers and

streams, taking into consideration factors such as flow rate, water depth, and channel geometry. Consequently, it is essential to thoroughly optimize the model setup and authenticate the simulation outcomes with regard to empirical data.

The HEC-RAS model has been successfully implemented for simulating extreme hydrological occurrences, such as floods and droughts, in numerous regions globally. For example, Hutanu et al. (2020) used HEC-RAS to simulate behavior during flood events in the Jijia River in Northern Romania. The authors note that the use of HEC-RAS modeling allowed for modeling detailed information about the river channel and floodplain, which enhanced the overall accuracy of the flood hazard map (Hutanu et al., 2020). Another study by Ikirri et al. (2021) describes the usage of HEC-RAS, along with the Watershed Modeling System (WMS), to model extreme hydrological conditions in the Ifni River watershed in Morocco. The authors tested the HEC-RAS model to simulate water flow and river flooding to estimate extreme events under different scenarios of climate change (Ikirri et al., 2021). This study will utilize the HEC-RAS model to replicate the hydrological behavior of the Berounka river during extreme events, such as floods, normal conditions, and drought, for better comprehension and prediction of the effects of these events on the catchment. Besides replicating extreme hydrological circumstances, the HEC-RAS model has been employed to assess the consequences of land use and land cover changes on hydrological processes. In 2002 an article was published by Jason J, Wiles, and Norman S. Levine, where the authors aimed to analyze the effect of urbanization on flooding in the Swan Creek watershed in Ohio. They successfully implemented HEC-RAS and Geographic Information System (GIS) to create a model to assess the potential impact of land use on flood risk and prevent it in that region. They simulated the flow of water under different land use scenarios, which helped them to create accurate results. (Wiles and Levine, 2002). So, the capacity of HEC-RAS to model the effects of different land use and land cover alterations makes it a practical instrument for recognizing and repairing the damage caused by human behavior on the hydrological system.

For the purposes of this thesis, HEC-RAS 6.3.1 will be employed, which is a widely recognized and reputable hydraulic modeling software in the field of water resources engineering. Its one of the latest versions includes new features that will further enhance the quality of the model simulations.

Approaches for hydraulic modeling

HEC-RAS 2D is a two-dimensional hydraulic modeling software developed by the US Army Corps of Engineers Hydrologic Engineering Center (HEC). It is intended to simulate the hydraulics of rivers and floodplains in intricate river systems using two-dimensional meshing techniques. HEC-RAS 2D is used for modeling flow behavior, water surface elevations, and inundation extents during flood events. The software can be used to analyze flood hazards, dam safety, and ecosystem restoration projects because it has extensive capabilities like flow routing, sediment transport, and water quality modeling. The hydraulic and hydrologic system analysis software HEC-RAS 2D is widely used by federal and state agencies, consultancies, and researchers. The software comes with a comprehensive user manual, tutorials, and technical support from the HEC team, making it accessible to users with different levels of experience. The HEC team also provides regular updates to the software to improve its capabilities and ensure its reliability (United States Army Corps of Engineers, 2023).

HEC-RAS is just one possible solution for 1D hydraulic modeling. There are some more potential approaches, which can be used for modeling the rivers or streams. For example, MIKE 11 is a one-dimensional hydraulic model developed by DHI Group for simulation water flow, sediment transport, and water quality processes in a river system. The model can simulate steady and unsteady flows, and can be used to predict water levels and discharges during flood events. It also includes a range of tools for data input, visualization, and analysis of simulation results. MIKE 11 is components of the MIKE software suite, which includes a wide range of different tools for modeling hydrology, hydraulics and water quality (DHI Group, n.d.).

Also, MIKE contains a 2D model, called MIKE 21 used for coastal engineering, marine resource management, and environmental impact assessment. It includes a range of tools for data input, visualization, and analysis of simulation results, and can be integrated with other models in the MIKE software suite for more comprehensive analyses.

Additionally, DHI Grouped developed a model like MIKE 21/3, which combines the capabilities of 2D and 3D modeling in MIKE for simulating complex hydrodynamic processes in marine and freshwater environments (DHI Group, n.d.). One of the disadvantages of using the MIKE model compared to HEC-RAS is the price of the subscription or licensing. It varies from 215 EUR to 1329 EUR per month, which might be quite expensive compared to free to use HEC-RAS.

Another approach for hydraulic modeling can be RiverFlow2D, which is a two-dimensional numerical model developed by Hydronia for simulating flow and sediment transport in rivers and estuaries. It is one of the most advanced models available for simulating these types of systems in its capability of handling a wide range of hydraulic conditions. It is a combined hydraulic and hydrologic flexible-mesh model offering a finite volume method to solve the depth-averaged Reynolds-averaged Navier-Stokes (RANS) for water flow and a sediment continuity equation for sediment transport. Also there are a lot of Add-On Modules to extend the functionality of the tool for different projects, including Sediment Transport (ST), Mud and Tailings Flow (MT), Pollutant Transport (PL), Water Quality (WQ), Urban Drainage (UD) and Plastic Transportation (PT). It is possible to use it for free during a free trial, but a basic one-user application would cost 3900 USD as one time payment (Hydronia, n.d).

Additionally to 1D, 2D and 3D hydraulic modeling approaches there are also simplified , conceptual representations of a part of the hydrological cycle, called conceptual hydraulic models. Conceptual hydraulic modeling uses a simplified representation of the underlying physical processes to simulate the hydrologic response of a watershed or drainage system. It is a mathematical model that describes the flow of water. Those models, which treat a hydrological system as a single unit with inputs and outputs also called lumped models. Those include models like precipitation-runoff, SCS Curve Number (SCS-CN), rainfall-runoff and many more. For example, Motyka (1998) in his study focused on development of the conceptual model for hydraulic networks in carbonate rocks based on field observations and geological information (J. Motyka , 1998).
While Alexandru Tatomir et al. (2018) was focusing on development and a new model based on a generic Features, Events, and Processes (FEP) database for assessing the potential impact of hydraulic fracturing on groundwater aquifers. The authors created a new approach

to building conceptual models by using a comprehensive and systematic inventory of features, events, and processes associated with hydraulic fracturing activities (Tatomir et al., 2018). To sum up, there are a lot of different types of conceptual hydraulic models used for different purposes in a wide range of applications and projects. It is possible to develop a unique mathematical model, which could be the best fit for your particular task, but performed worse in other projects or study areas. It is essential to ensure that the model you are using is suitable with the objectives of your study.

In conclusion to this chapter, there are a lot of different approaches for hydraulic modeling, including 1D, 2D, 3D and conceptual methods. It is essential to ensure what type of modeling is the best fit based on the objective of the research and specifications of the study area. In some cases, the development of the lumped conceptual models can significantly improve the accuracy of the predictions, although for some studies it is more efficient to use general approaches of hydraulic modeling.

1D hydraulic modeling approaches in HEC-RAS

One of the objectives of this study is to review existing approaches for 1D hydraulic modeling. During this chapter, I will describe different 1D models and will focus on HEC-RAS modeling. 1D modeling is a common method for simulating the flow of water in rivers or streams. HEC-RAS includes several approaches for hydraulic 1D modeling, including steady flow modeling, unsteady flow modeling, and mixed flow modeling.

The steady flow modeling approach assumes that the flow in the channel is with constant value over a period of time and Manning's equation to calculate flow velocity and depth. It is an empirical equation, which utilizes flow velocity, flow depth, and channel roughness to calculate the discharge. In HEC-RAS it is a default method for calculating channel friction and roughness in steady flow analysis. Users can define custom Manning's roughness values. (USACE, 2016).

Unsteady flow analysis is a more complex approach, as it takes into account the change in flow over time. It also utilizes Manning's equation to calculate the roughness coefficient for each computational cell for the model. The flow velocity and depth in the

river reach are both influenced by the roughness coefficient, which is a significant input parameter. However, in response to changes in flow parameters, channel geometry, and sediment transport, the roughness coefficient can also change over time and space. Then it uses the roughness coefficient from Manning's equation as a parameter in the full Saint-Venant equations of motion to calculate flow velocity and depth for each cell in the river over time (USACE, 2016).

There is one more approach to 1D modeling in HEC-RAS called mixed flow modeling. It is a combination of unsteady and steady flow modeling by assuming that the flow is steady in the main channel and unsteady in the floodplain. The simulation is based on a dynamic computational mesh that captures the transition from subcritical to supercritical flow, and the effects of lateral inflows, outflows, and boundary conditions are considered. The results obtained from mixed flow analysis are valuable for river engineering, hydraulic structure design, and floodplain management (USACE, 2016).

HEC-RAS also uses numerical methods for 1D modeling, including dynamic wave and kinematic wave models. They are utilized as tools to solve the full St. Venant equations. The dynamic model, which can simulate intricate flow patterns and hydraulic structures, takes into account the effects of inertia and pressure variations in the flow. The kinematic wave model, on the other hand, simplifies the equations for faster computations by assuming that pressure variations are insignificant in comparison to the friction losses in the flow, making it a better choice for large rivers or straightforward hydraulic structures (USACE, 2016).

3.7 HAMR

Hydrological, Agronomic, Meteorological and Retention (HAMR) is a system developed by the Czech Hydrometeorological Institute (CHMI) in collaboration with other research institutions and regional authorities. The system provides real-time information about the condition of water resources and potential drought events in the Czech Republic using a variety of data sources, including meteorological data, satellite imagery, and soil moisture sensors (MP, 2018). HAMR is a project that combines knowledge from hydrology,

agronomy, meteorology, and retention to better water management in agricultural areas. The project's primary goal is to create a web-based system that will track and forecast soil moisture, crop water needs, and runoff utilizing a variety of data sources and models. Farmers and water managers can also receive early warning notifications from HAMR for future droughts, floods, and other weather-related threats (Czech Hydrometeorological Institute, n.d.).

The HAMR system seeks to aid in reducing the effects of drought events on society, the economy, and the environment by assisting decision-makers and the general public in anticipating and responding to them (HAMR, n.d.). The European Union is funding the project, which aims to increase the nation's capacity for drought monitoring and warning (VUV, n.d.). It is integrated into the National Drought Monitoring System (NDMS) of the Czech Republic. The NDMS is a component of the CHMI-coordinated broader monitoring and warning system for hydrological and meteorological hazards in the Czech Republic. HAMR provides high-resolution data on precipitation, soil moisture, and other meteorological and hydrological parameters that are used to assess the drought situation and issue drought warnings. It is a part of the Hydrological Information System (HIS), which provides access to the data and monitoring tools.

The HAMR operates in the Czech Republic and does not cover the rest of the world; many countries have similar projects for water management. For example, the US Department of Agriculture's Agricultural Research Service has a research program called "Water Management and Conservation Research" that uses interdisciplinary approaches to optimize irrigation practices, manage soil moisture, and enhance water use efficiency in crops (USDA-ARS, 2021). Similarly, the Australian Bureau of Meteorology has a project called "Water Information Research and Development Alliance" that develops integrated water management tools using hydrological, meteorological, and other related data (CSIRO, n.d.).

For this study, the data from the HARM system was taken as input data into the conceptual Bilan model to simulate RM. That was used as input for the Berounka model into HEC-RAS. I will provide more information on the Bilan model in the next chapter.

To sum up, HARM is an essential and innovative approach to monitor, warm and forecast potential extreme meteorological and hydrological events. With its data integration from different sources, make the system an even more powerful project.

3.8 Bilan model

For simulating river runoff, the hydraulic conceptual Bilan model is an effective tool. The model is still in use and still accessible to both the general public and professionals in the Czech Republic, even though TGM WRI in Prague started working on it at the start of the 1990s. (Vodní zdroje, vodní hospodářství, 2022). A river basin's water flow can be simulated using the hydraulic Bilan model, which is also used to evaluate how flood control measures affect the management of water resources and the risk of flooding. By accounting for a wide range of factors, including precipitation, evapotranspiration, infiltration, runoff, and storage, the model seeks to identify the key factors affecting water flow and flood risk in the basin. The Bilan model is described as a semi-distributed model, as it divides a river basin into multiple sub-basins and considers the hydrological processes within each sub-basin separately. Additionally, the hydraulic Bilan model can also be used to assess the effects of climate change on flood risk and water resource management in the basin, as well as to gauge the efficiency of various flood protection techniques like levees, retention ponds, and floodplain restoration, as well as to assess the effectiveness of different flood protection measures (Hanel et al., 2014).

Precipitation, evapotranspiration, soil moisture, and runoff are the four main elements of the water cycle that the Bilan model can simulate. The model also includes various algorithms for simulating various elements of the water cycle, including snow accumulation and melting, groundwater recharge, and river routing. The model can be used at various spatial scales, from little catchments to big river basins (Bilan Model User Manual, 2015). However, the Bilan model provides reasonable estimates of hydrological variables, it is not without limitations. For example, the model does not consider lateral flow and assumes that all runoff is generated by excess precipitation (Hanel et al., 2014).

The Bilan model simulates runoff in a river basin using a simple water balance equation. The equation is based on the idea that the difference between precipitation input

and output from a basin in the form of evapotranspiration and infiltration equals the amount of runoff in that basin. The calculation looks like this:

Runoff = Precipitation - (Evapotranspiration + Infiltration)

The user manual for conceptual Bilan models described how to simulate runoff of river basins, as it uses the principle that runoff is equal to the difference between precipitation input and output from the basin (Evaporation) and infiltration. There are also instructions on how to specify input and data for simulation RM. (Bilan Model User Manual, 2015). Similarly, the report by Hanel et al. provides the case of estimating the hydrological impact due to climate change using a model and describes how Bilan used to simulate runoff in the river. It is stated that the model uses the water balance equation to simulate the main components of the water cycle including runoff (Hanel et al., 2014). The study by Melišová briefly discusses the Bilan model as one of the conceptual hydrological models used in the research. The paper mentions that the Bilan model has been used in previous studies to assess the impact of climate change on Czech river basins and that it has been calibrated and validated using observed hydrological data (Melišová et al., 2020).

A number of parameters that can be used to calibrate the Bilan model and increase its accuracy are also included in the model. These parameters include evapotranspiration rate, infiltration capacity, and soil moisture capacity. However, it's crucial to keep in mind that the specific characteristics of the river basin being modeled, as well as the caliber and accessibility of input data, can affect the accuracy of the Bilan model. Before using the model for water management or planning, it is advised to validate it against actual observations to ensure its accuracy.

In conclusion to this part, the conceptual Bilan model is a useful tool for simulating the main components of the water cycle, including runoff. After successful validation and calibration, it can produce accurate results, which can be used for further analysis or as input to other models.

4. Methodology

4.1 Methods

Data Preparation

The preparation of all necessary data for the Python Interface and the model as input is the first step in this study. To find multi-day periods that represent typical, drought-like, and high flow conditions, the dataset will be screened. The events will be divided into calibration and validation sets, and general statistics about the events will be presented. This step will guarantee that the data, whether measured or predicted, is accurate.

Python interface

After the data has been preprocessed, a Python interface will be created and used to load data into the HEC-RAS for various time frames. The interface will enable the model to automatically import, process, and incorporate hydrological data. The Python interface will be put to the test and evaluated to see if it can increase the precision and effectiveness of model simulations.

HEC-RAS simulation

The HEC-RAS model will be calibrated and validated with respect to observed data using standard hydrological model performance metrics with a focus on extremes. The model will be run for the selected events in the calibration and validation sets.

4.2 Study Area

Berounka basin

The Berounka catchment is located in the western part of the Czech Republic. primarily in the Pilsen Region (Plzeňský kraj) and partially in the Central Bohemian Region (Středočeský kraj) with a total study area of approximately 3,535 square km. The catchment area includes the Berounka River, as well as its tributaries, such as the Mže, Radbuza, Úhlava, and Úslava rivers. From Šumava mountains to its confluence with the Vltava river, the Berounka catchment length is around 139,1 km with a basin area of 8861 square km, but this measurement only includes the main stem the Berounka river itself not counting the length of its tributaries or smaller stems. The Mže River is approximately 106 km long with its source is in Germany's Upper Palatinate region, and it enters the Pilsen region of the Czech Republic before reaching the Radbuza River there. The length of the Radbuza River is roughly 111 kilometers. It starts in the Pilsen Region's Domažlice District and moves east across the region before merging with the Mže River close to Pilsen. The Uhlava River is around 108 kilometers long. It starts in the Pilsen Region's Klatovy District and travels northeast before merging with the Berounka River close to the town of Stříbro. The Uslava River covers roughly 94 km. It originates in the Tachov District of the Pilsen Region and flows northeastward through the Pilsen Region, where it joins the Mže River near the town of Stříbro (Infoglobe.cz, n.d.)

Berounka catchment plays an important role in the Czech Republic, as its temperate climate and varied topography make it a unique area. However, these conditions can lead to extreme hydrological events such as floods and droughts, which can have severe implications for local communities and their infrastructure. Thus, it is vital to understand how the catchment behaves under various hydrological conditions in order to prepare for and mitigate risks associated with extreme events. By studying the Berounka catchment, we can gain a comprehensive understanding of how best to protect our environment from the potential impacts of extreme weather.

Picture 1: Berounka catchment, main stem (Berounka river)

River Stations

Five datasets have been prepared for different locations of river stations situated in various rivers within the Berounka catchment. These include Berounka, Mže, Radbuza, Úhlava and Úslava.

River	Location	Reach	River Station
Berounka	Liblín	Berounka_6	5470
Mže	Stříbro	Mze_8	155.61
Radbuza	Staňkov	Radbuza_4	13130.7
Úhlava	Tajanov	Uhlava_6	2792
Úslava	Ždírec	Uslava_2	9902.51

Table 1: Study area river stations.

Datasets for each location will be utilized to calibrate and validate the model in this research. I will present general statistics about each river station in this part.

Liblín Station (Berounka):

The Berounka River, which is 139 km long and drains a basin that is 6455,795 square km in size, is located in the Plzeň region of Czechia. Station is located on the Berounka river above the bridge in Liblín. Data is collected by ČHMÚ Plzeň. Depending on the weather, its water temperature ranges from 6 to 25 °C and its average water flow rate is 27,6 m3/s. At the Liblin Station, the highest recorded water level was 7,03 meters in August 2002, and the lowest was 2,16 meters in March 1999. The annual discharge is 243 m3/s for Q1, 528 m3/s for Q5, 677 m3/s for Q10, 1080 m3/s for Q50, and 1270 m3/s for Q100, where the number is the amount of years. The station has three levels of flood activity, and the corresponding flow rates for each level are 141 m3/s (180 cm) , 263 m3/s (260 cm), and 420 m3/s (340 cm), respectively.

Stříbro Station (Mže):

The Stříbro station, located in the Plzeň region of the Czech Republic on the Mže river and operated by ČHMÚ Plzeň. The total length of the Mže river is 107 km. The station's catchment area is 1144,011 square km in size and is located at a distance of 41.8 km. For Q1, Q5, Q10, Q50, and Q100, the annual discharge for a particular recurrence interval is 51.2 m3/s, 102 m3/s, 130 m3/s, 212 m3/s, and 255 m3/s, respectively. The catchment's flow area represents 62.8 percent of the total catchment area. Each of the station's three flood activity levels corresponds to a different flow rate, which is 41.375 m3/s (150 cm), 58.628 m3/s (180 cm), and 83.769 m3/s (220 cm). With an average annual water level of 68 cm, the flow rate is 6,04 m3/s. On September 30, 1890, the station's water level reached its highest-ever recorded value of 450 cm.

Staňkov Station (Radbuza):

The Staňkov river station is located in the Plzeňský kraj, with a surface catchment area of 701.506 square km. The station has a recorded average annual water level of 43 cm and an average annual flow rate of 3.42 m3/s. The station has 3 levels of monitoring alerts for 160 cm is 25 m3/s (vigilance), 190 cm is 32.2 m3/s (emergency) and 220 cm with 40.1 m3/s (threat) of flow rate. The highest recorded water level of 360 cm was on 13th August

2002. Annual average water level is 43 cm with average flow rate around 3,42 m3/s. The annual discharge for a particular recurrence interval for Q1 is 31,1 m3/s, Q5 is 81,3 m3/s, Q10 is 110 m3/s, Q50 is 194 m3/s and Q100 is 238 m3/s. The data is owned and collected by ČHMÚ Plzeň.

Tajanov Station (Úhlava):

Tajanov Station is a hydrological monitoring station located in the Plzeňský kraj region of the Czech Republic. It is situated on the banks of the Uhlava River, which has a drainage area of 338.736 square kilometers. There are also 3 levels of monitoring alerts on 250 cm with 23 m3/s, 290 cm with 34,8 m3/s, and 300 cm with 49,9 m3/s. The average annual water level is 62 cm with an annual discharge rate of 3,23 m3/s. The annual discharge for a particular recurrence interval for Q1 is 20,2 m3/s, Q5 is 49,3 m3/s, Q10 is 65,7 m3/s, Q50 is 112 m3/s and Q100 is 137 m3/s. The highest recorded water level was 362 cm on the 13th of August 2002. Datacenter maintained by ČHMÚ Plzeň. The Tajanov Station's location is on the right bank of the Uhlava River, near the Povodí Vltavy závod Klatovy building.

Ždírec Station (Úslava):

The Ždírec station is a hydrological monitoring station located in the Plzeňský kraj region of the Czech Republic. It is situated on the left bank of the Uslava river, which has a drainage area of 375.807 square kilometers. The station is operated by the Czech Hydrometeorological Institute in Plzen and serves as a center for automated data collection. The station alerts on 150 cm with 14,689 m3/s flow rate, 190 cm with 28,214 m3/s and 220 cm with 54,305 m3/s. The annual average water level is 48 cm and annual discharge 1,51 m3/s. The annual discharge for a specific recurrence interval is 38,6 m3/s for Q1, 69,5 m3/s for Q5, 89,4 m3/s for Q10, 153 m3/s for Q50, and 189 m3/s for Q100. The Zdirec station has recorded several high water levels in the past, the highest being 217 centimeters in June 2013. The station is located on the left bank of the river, near a road bridge.

In summary, all the stations presented in the study area are located in the Plzen region and maintained by the Czech Hydrometeorological Institute in Plzen (ČHMÚ Plzeň). The

descriptive statistics for each river station and the corresponding river were provided, offering an overview of the study area. These rivers are significant contributors to the water resources of the Czech Republic and provide scenic beauty for tourists and locals. The data collected from these stations is essential for managing and mitigating the risks associated with flooding and managing water resources in the region.

4.3 Data Preparation

RDS data

The RDS file format is a binary file format used to store data objects in R, a statistical programming language. Currently, it is possible to process it in Python as well, using a specific library pyreadr, as it works by leveraging the R language to save an R RDS object as a dataframe.

One benefit of using the RDS format for hydraulic modeling is that it makes it simple to integrate with other R ecosystem tools and software programs. In other R-based software programs like "Shiny", "ggplot2", or "dplyr", for instance, data objects saved in the RDS format can be easily loaded, enabling seamless integration and analysis of hydraulic modeling data.

For hydraulic modeling, RDS files can be used to save, store, and share data objects for the modeling. HEC-RAS can be used to provide data for geometry and cross-section of the study area. In this study, I used RDS data as input for Unsteady Flow Analysis as separate RDS files for each river reach. Preprocessing of data is not needed to use the script, as you can provide the location of files and the code automatically reads it as an R object, stores it in the dataframe, and converts RM values into the discharge in meters per second (Q) using a file with the catch area of each river reach.

Area File

In order to convert RDS data values from RM to Q to use in the model, a data file with a catch area is required. Using the formula ((RM / 1000) * area) / (24 * 3600) is used

in the code. Area files have to be in excel format with two columns, including Reach Code (example BER0010) and the value of the area in square meters. It can be exported easily in any GIS software from the geometry data of the model using the attribute area. It is possible to calculate the area using a formula and export it with River Code in excel format.

Using area files to convert data in the script can make preprocessing way faster and more efficient, as it is not required to use any other tool to get Q values, instead of any GIS application. It was faster and more memory efficient to export area files and write functionality to convert data in Python, than use R to read all RDS, convert data there, and then save and export RDS for the whole time of the simulation.

Flow Condition

For the analysis, two years were chosen to validate and calibrate the Berounka model: 2015 and 2019. In 2019, the Berounka catchment area experienced flooding during the months of June and July. The heavy rainfall and thunderstorms during this period caused the river's water levels to rise significantly, leading to flooding in the surrounding areas. Numerous communities in the basin of the Berounka River had to be evacuated as a result of the flooding, which was among the worst to hit the nation in recent memory. There were reports of flooded streets, damaged infrastructure, and evacuations in some Berounka River communities. The report from the Czech Hydrometeorological Institute stated that the flow condition in 2019 was higher than normal, with locations slightly above normal (Czech Hydrometeorological Institute, 2019).

The flow condition in 2015 was the lowest. The Czech Hydrometeorological Institute reported that the water level and discharge of the Berounka catchment were at their lowest due to droughts in the area. In 2015, the Berounka catchment experienced drought conditions, which affected many parts of the Czech Republic. In May and June 2015, high temperatures and low precipitation caused droughts in some parts of the catchment. This resulted in the Berounka River reaching its lowest water levels and damaging the infrastructure. This year was chosen in contrast to the floods in 2019. According to the annual hydrological report for 2015, the condition of the flow was characterized by lower-than-normal precipitation in many parts of the Czech Republic during the summer months,

which led to decreased water levels and flow rates. Additionally, the report stated that there were a number of drought warnings and drought events during the spring and early summer of 2015 (Czech Hydrometeorological Institute, 2015).

Both reports for 2015 and 2019 provide general information about flow conditions in the Czech Republic, but they do not specify the Berounka catchment area. However, they provide information on the approximate locations of the extreme conditions for both years.

Measured water surface elevation (WSE) data

The following chapter presents the results of the data analysis collected from the Hydra2 resource, which stores, analyzes and provides data for different rivers in the Czech Republic about Depth, Runoff and Temperature. It is possible to choose any date between 29.07.2012 and up-to-date. In order to obtain more useful data about a specific river, it is possible to navigate into the Hydro Czech Hydrometeorological Institute via the link below the plot. All historical data was taken from the Hydra2 resource.

The goal of that study is to validate observed water surface elevation (WSE) with respect to predicted WSE by the model. To calculate observed data is it required to summarize Depth from Hydra2 resource with a Null Depth of the river. In the case below Hladina (Depth) in millimeters and Null (Null Depth) in meters. Final WSE is also in meters.

Datum	Hladina	Null	WSE
1/1/2015 0:00	56	269.75	270.31
1/1/2015 1:00	57	269.75	270.32
1/1/2015 2:00	59	269.75	270.34
1/1/2015 3:00	58	269.75	270.33
1/1/2015 4:00	57	269.75	270.32

Table 2: Sample of measured data

The predicted data of WSE was presented per data by the model, when the observed data from hydra2 resource is hourly. In order to get data for the same time period, it is important to find an average of hourly records to make it daily. I created Excel Pivot and group data by days with a mean of WSE. However, the real data contains missing values for some days, so it is important to fill all missing days manually. To fill it, I used the average of WSE from the first date before and after missing one. It is a variant of forward and backward fill method to enter values for missing dates. The technique of filling missing values with the average of the closest values is called "mean imputation". Note that all presented dataset contains less than 20% of missing data and can be used for validation and calibration.

Next, I want to present summary statistics of all original datasets exported from Hydra2 for 2015 and 2019. Statistics will be based on water surface elevation for hourly and daily versions.:

Statistics	Stribro	Liblín	Tajanov	Zdirec	Staňkov
STD	0,161	0,242	0,223	0,163	0,1371
MEAN	354,951	270,274	386,798	394,891	354,348
MIN	354,72	269,92	386,33	394,68	354,21
MAX	355,94	271,29	388,81	395,69	355,33
MEDIAN	354,93	270,24	386,78	394,84	354,32
MODE	354,93	269,98	386,76	394,79	354,25
Variance	0,0265	0,058	0,049	0,026	0,0187
SKEW	1,698	1,051	2,49	1,319	2,82
KURT	-10523,387	-3164,230	12969,194	-842,181	-10856,048
IQR	0,17	0,33	0,2	0,19	0,12
CV	0,045	0,089	0,057	0,041	0,038
2019					

Table 3: Measured hourly data before preprocessing for 2019

Statistics	Stribro	Liblín	Tajanov	Zdirec	Staňkov
STD	0,167	0,211	0,139	0,155	0,14208
MEAN	354,991	270,176	386,71	394,894	354,351
MIN	354,75	269,75	386,32	394,6	354,12
MAX	355,84	271,42	387,89	396,04	355,53

MEDIAN	354,95	270,12	386,67	394,89	354,33
MODE	354,83	269,99	386,62	394,97	354,36
Variance	0,028	0,045	0,019	0,024	0,02
SKEW	1,336	1,835	2,263	2,124	3,456
KURT	2092,617	2946,726	8159,140	7889,023	7268,857
IQR	0,24	0,26	0,14	0,2	0,13
CV	0,047	0,078	0,036	0,039	0,04
2015					

Table 4: Measured hourly data before preprocessing for 2015

Preprocessing of Real data and Summary Statistics

To preprocess the measured data to match the format of the model predicted data it is important to group the data by day period and find the average of WSE for each day. It was done using the pivot table in excel. Also, the missing data was filled in using mean imputation to get the full year period.

Statistics	Stribro	Liblín	Tajanov	Zdirec	Staňkov
STD	0,154	0,229	0,2143	0,154	0,121
MEAN	354,947	270,258	386,785	394,864	354,324
MIN	354,727	269,974	386,509	394,686	354,215
MAX	355,839	271,2277	388,265	395,545	355,129
MEDIAN	354,917	270,179	386,765	394,825	354,275
MODE	354,895	270,120	386,556	394,715	354,255
Variance	0,023	0,052	0,045	0,023	0,0146
SKEW	1,805	1,228	2,222	1,428	3,1613
KURT	5,391	1,6054	9,389	2,3353	13,5365
IQR	0,147	0,295	0,201	0,195	0,089
CV	0,043	0,084	0,055	0,039	0,034
2019					

Table 5: Measured daily data after preprocessing for 2019

Statistics	Stribro	Liblín	Tajanov	Zdirec	Staňkov
STD	0,155	0,201	0,142	0,144	0,1315
MEAN	354,979	270,150	386,6773	394,90	354,339
MIN	354,76	269,944	386,494	394,69	354,195
MAX	355,79	271,31	387,504	395,959	355,391
MEDIAN	354,92	270,06	386,64	394,95	354,299
MODE	354,92	270,0175	386,52	394,96	354,2775
Variance	0,024	0,040	0,020	0,020	0,017

SKEW	1,58	2,052	1,775	1,981	3,695
KURT	3,976	6,21	5,328	10,256	20,004
IQR	0,207	0,235	0,14	0,170	0,105
CV	0,043	0,0741	0,036	0,036	0,037
2015					

Table 6: Measured daily data after preprocessing for 2015

The tables shown here summarize the water surface elevation data for the Berounka catchment both before and after preprocessing. For the years 2015 and 2019, information was collected hourly, and then preprocessing was performed to group the data by day and apply mean imputation to fill in the missing values.

When we compare the statistics for 2015 before and after preprocessing, we can see that the standard deviation, skewness, and kurtosis values dropped after preprocessing, while the mean and median stayed almost the same. This shows that the data grew less spread out and skewed. The coefficient of variation (CV) also reduced, showing that the data's variability decreased following preprocessing. In general, following preprocessing, the data became more accurate and reliable.

Similarly, for 2019, the standard deviation, skewness, and kurtosis values fell following preprocessing, while the mean and median stayed nearly unchanged. The coefficient of variation (CV) also reduced, showing that the data's variability decreased following preprocessing. However, there were no significant differences in the before and after preprocessing statistics for the year 2019, and the data was already pretty consistent prior to the preprocessing.

Finally, the preprocessing strategy of grouping the data by day and utilizing mean imputation to fill in missing values has been found to be helpful in minimizing variability and making the data more reliable and consistent. The influence of preprocessing was more pronounced on data obtained in 2015, which was more dispersed and skewed. After the statistics we can conclude that the preprocessed measured data can be used efficiently to validate the model and accurately calculate the error metrics.
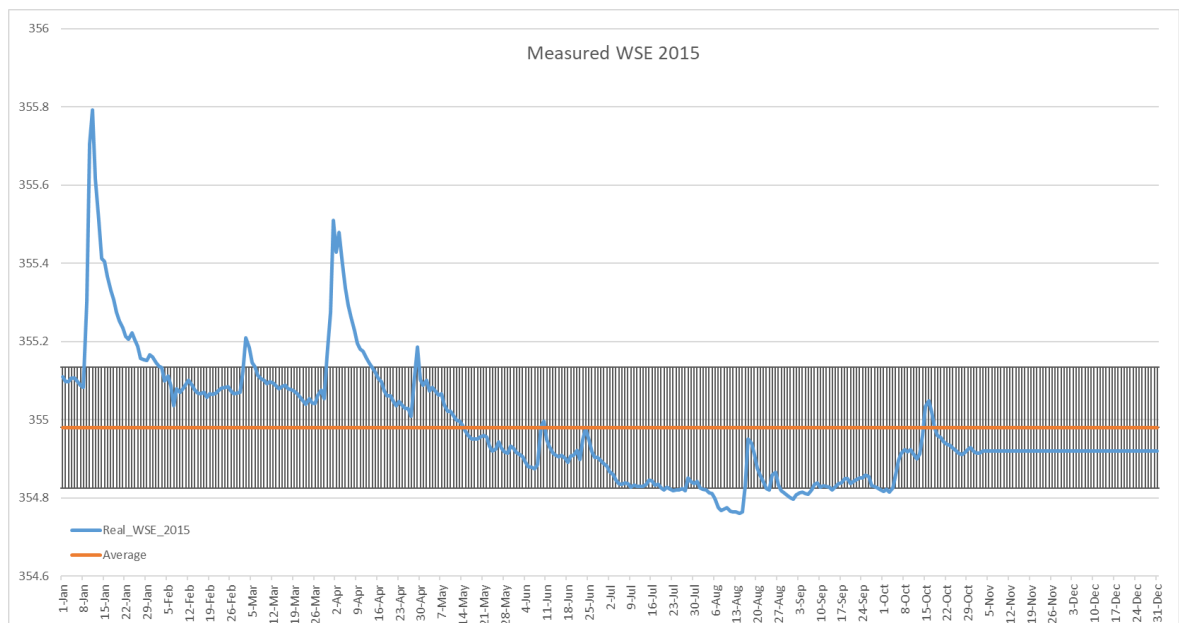
Graphical representation of the measured data.

In this section, I prepare the graphical representation of the measured data for each station for both 2015 and 2019 years. The crossing line in the Charts is a standard deviation.

Stribro:
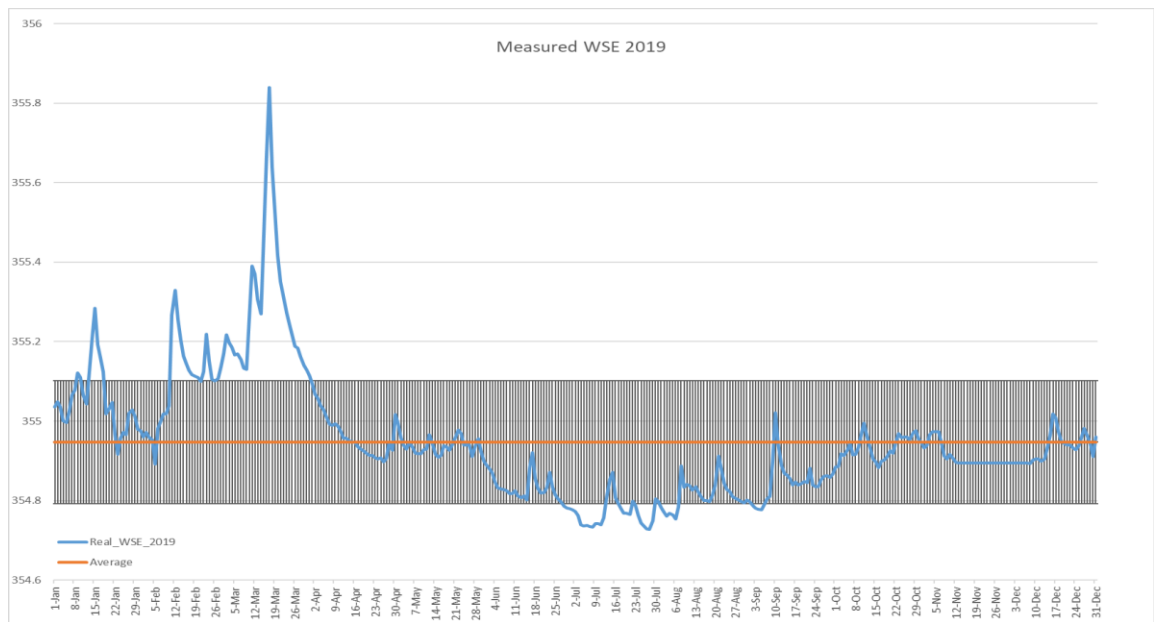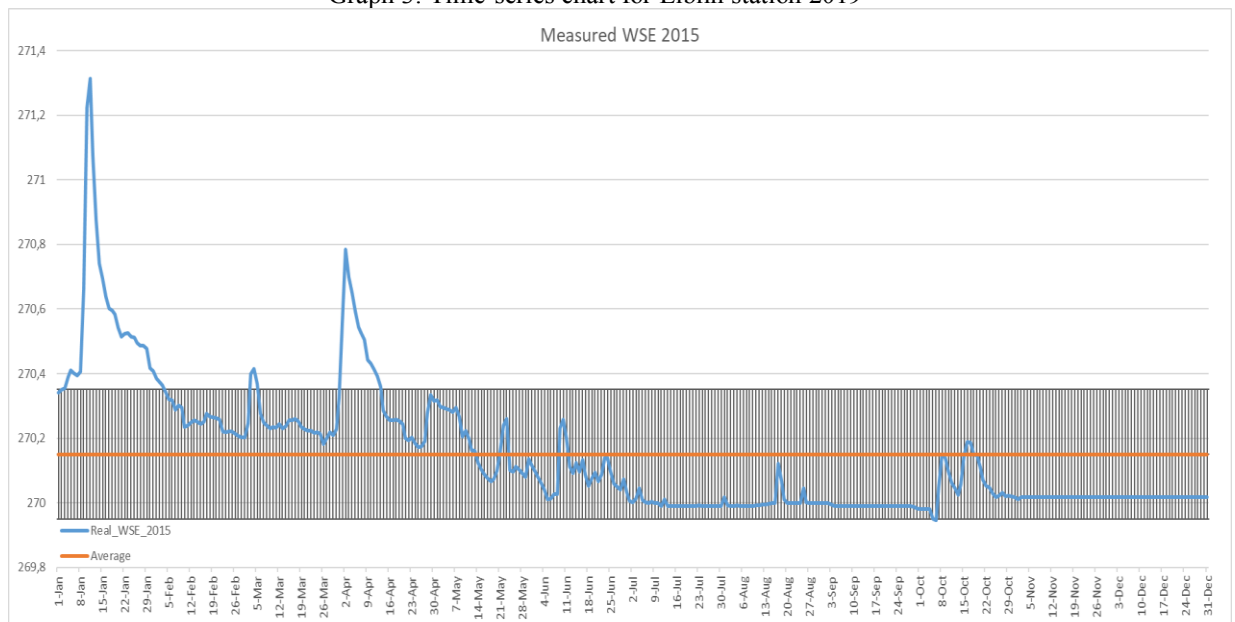

Graph 1: Time-series chart for Stribro station 2019

Graph 2: Time-series chart for Stribro station 2015

At station Stribro, the observed data show a clear seasonal pattern with high levels during January and March. Also, low levels of water surface elevation during July and August for 2019 and August and September for 2015. These seasonal patterns are likely driven by changes in precipitation and runoff in the catchment.

Liblín:
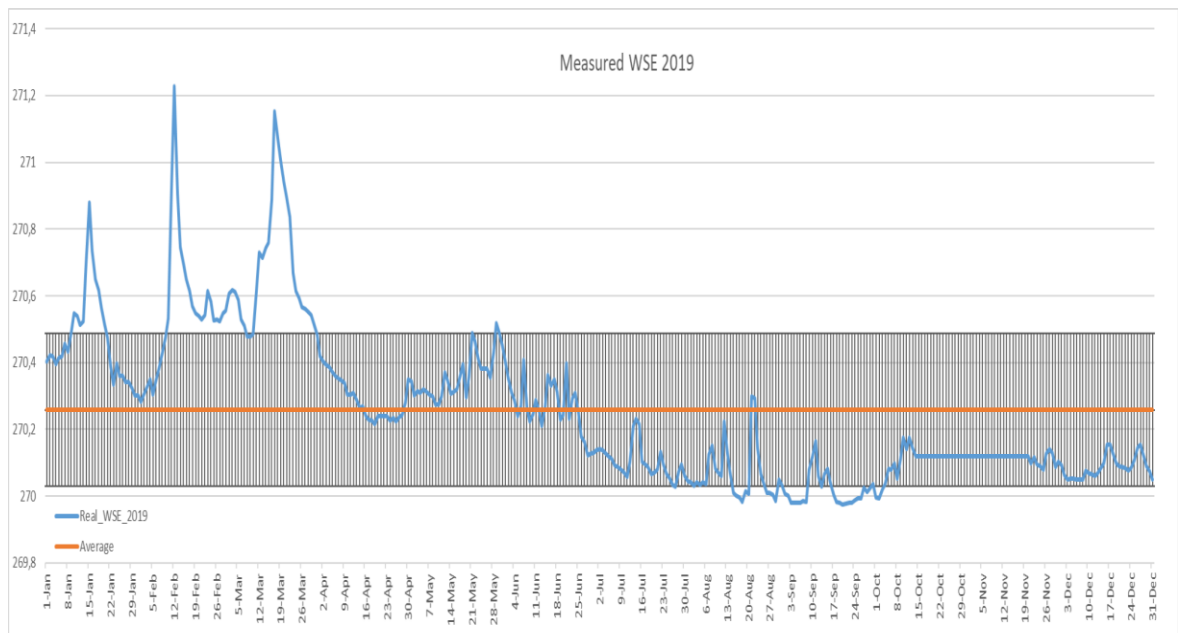
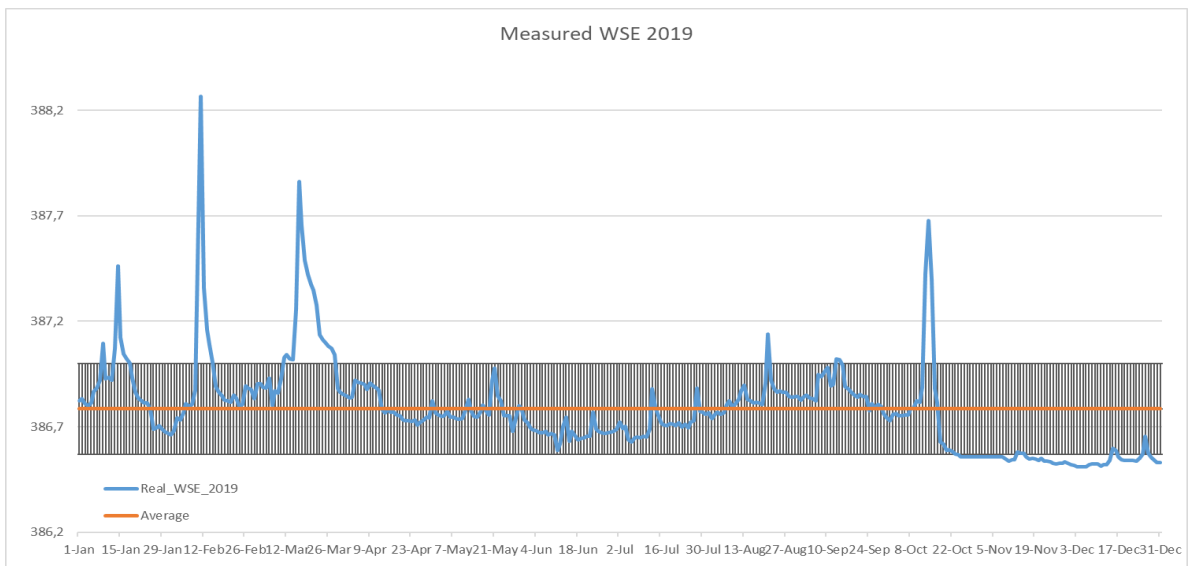Graph 3: Time-series chart for Liblin station 2019

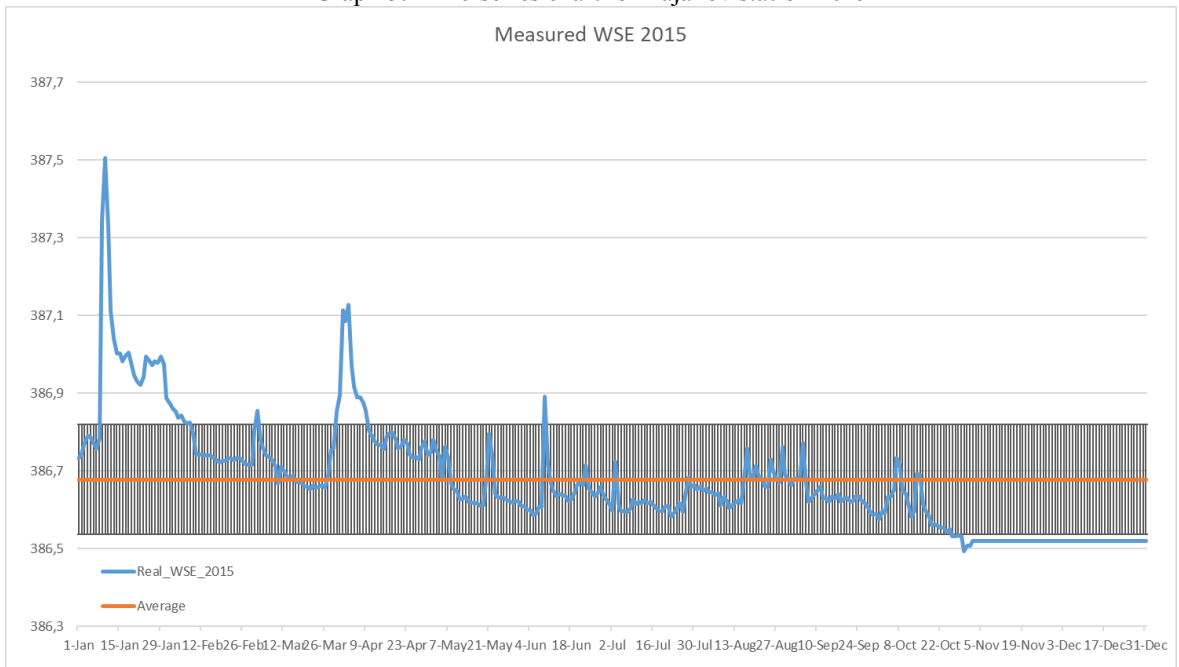Graph 4: Time-series chart for Liblin station 2015

From the charts for the station Liblin, it is clearly seen the peaks of the WSE during mid January, February and March for both years. Additionally, in 2019 WSE seems to be

lower than normal during August and September, while in 2015 it was within normal range for the whole period after the winter droughts. However, flow conditions in 2015 seem to be more extreme due to the higher WSE during the summer-spring and lower during the summer/autumn periods.

Tajanov:

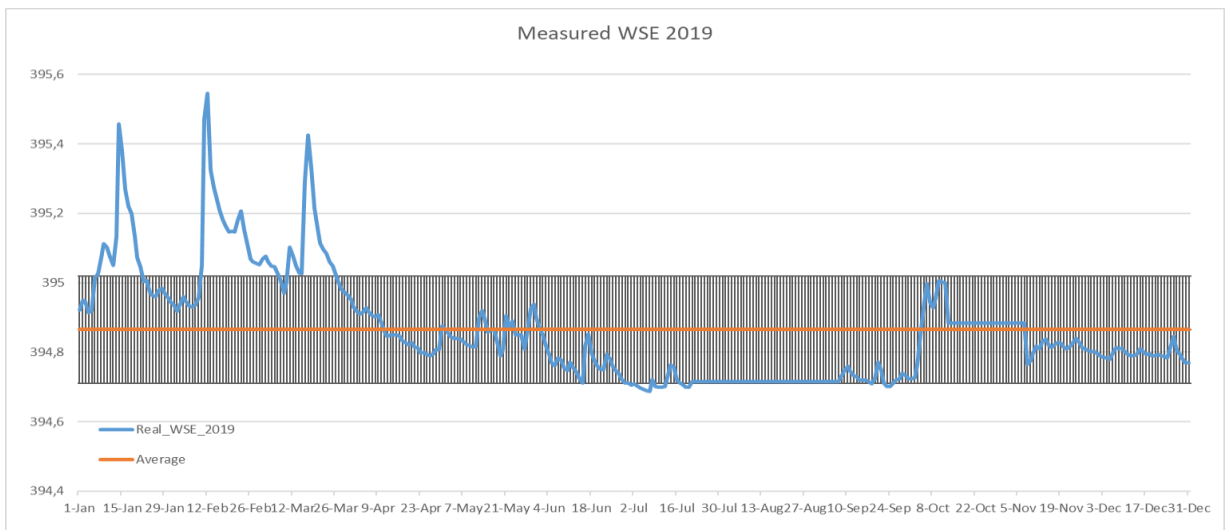Graph 5: Time-series chart for Tajanov station 2019
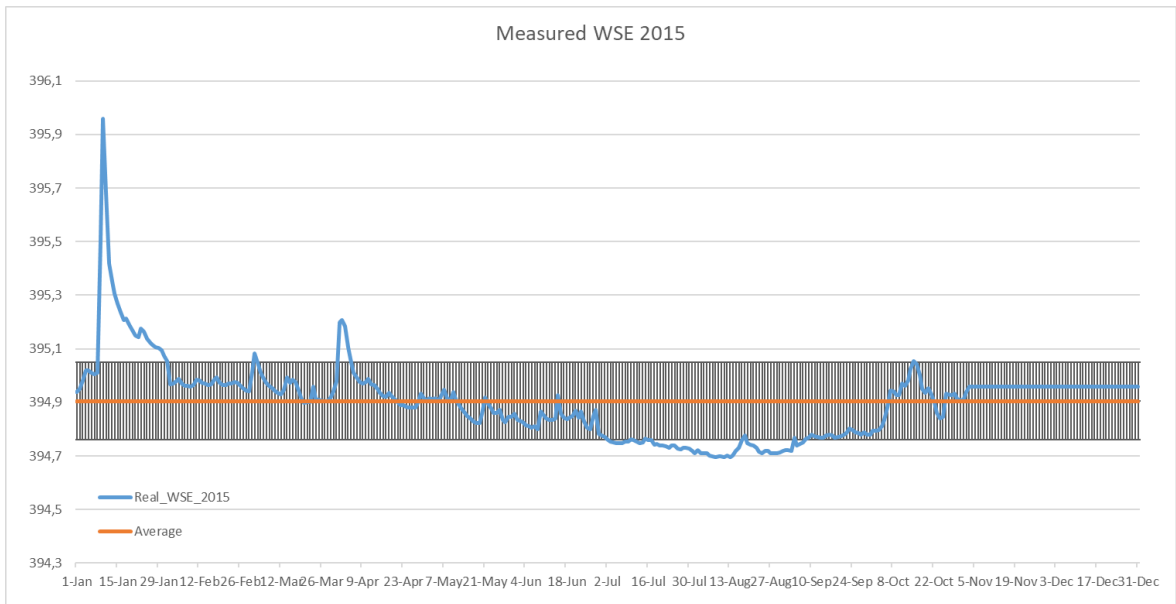

Graph 6: Time-series chart for Tajanov station 2015

For the Tajanov station in 2019 the measured data pattern is a bit different. With a WSE higher than normal January, February and March , there are two more peaks during

August and October. Also, WSE was lower than standard deviation in October, November and December. However for 2015 there are peaks in the charts during January, February and April with a lower WSE during October. There was missing data for November and December, so it was filled in with the last existing values.

Zdirec:

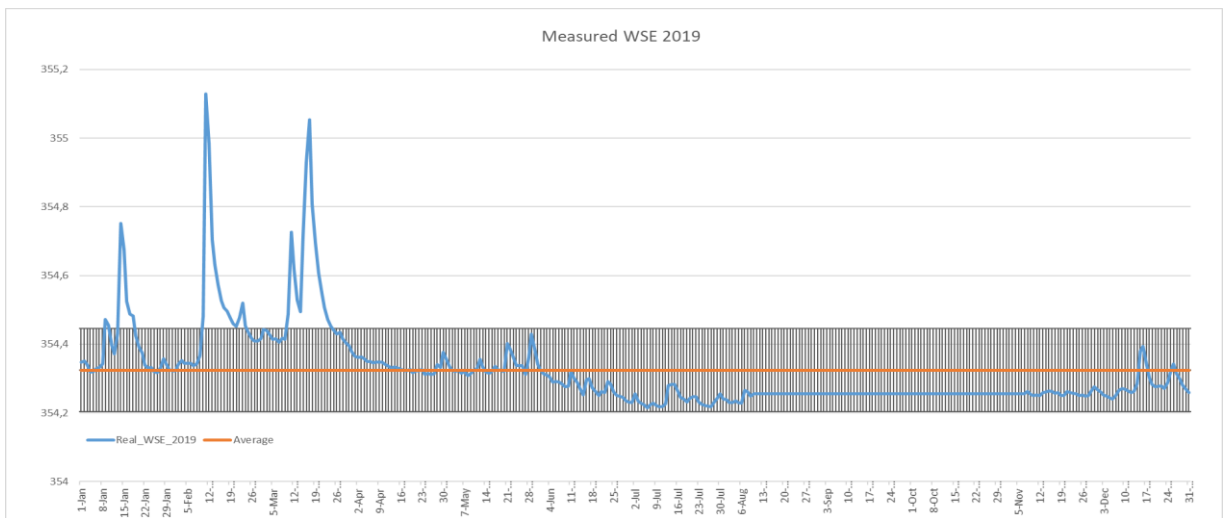Graph 7: Time-series chart for Zdirec station 2019
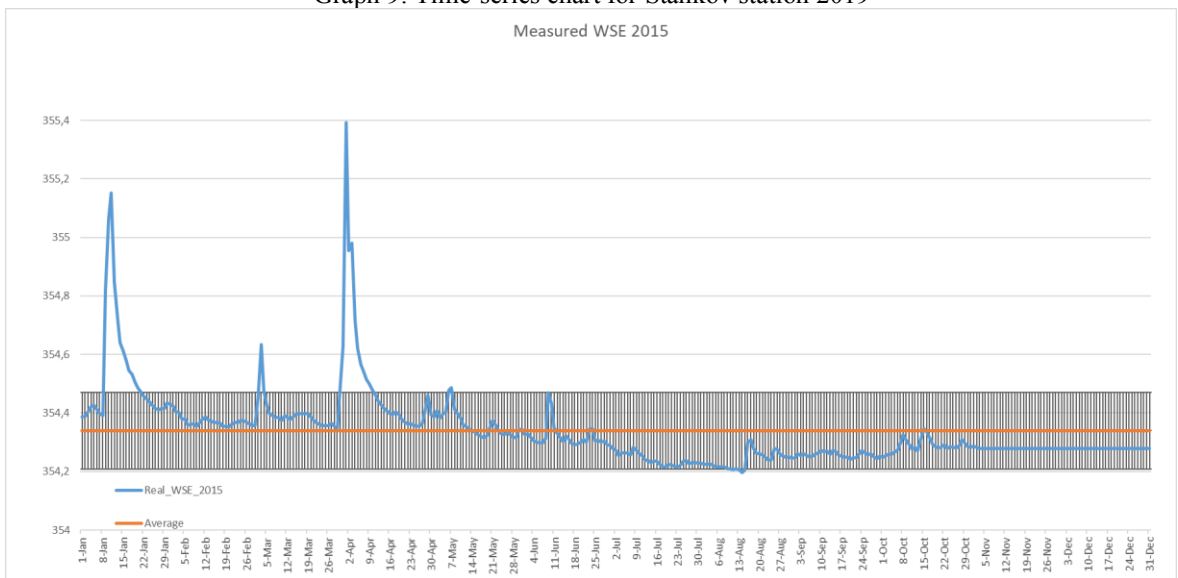

Graph 8: Time-series chart for Zdirec station 2015

On the station Zdirec high WSE was measured during the January, February and March in 2019 and during January and April in 2015. Also, WSE was slightly lower than

normal during Jaly in 2019 and in July and August during 2015. Upon analyzing the data sets, it is evident that the water surface elevation for the year 2019 exhibits a greater occurrence of high peaks, whereas the data for 2015 depicts a comparatively higher frequency of points below the standard deviation.

Staňkov:

Graph 9: Time-series chart for Stankov station 2019

Graph 10: Time-series chart for Stankov station 2015

At Stankov station there are high peaks of WSE during mid-January, March and April in both years. Data for the rest of the periods is in the range of standard deviation, which can be considered as normal condition.

The analyzed charts provided valuable information, which can help identify how the model would perform during different extreme flow events. Berounka catchments experience significant fluctuations in WSE, which can be a result of different climate events. The observed trends also indicate the impact of seasonal fluctuations and external causes on the local environment.

In hydraulic modeling, accurate data preparation is crucial for optimal performance. Inaccurate data can lead to unreliable results and inaccurate predictions, which can be detrimental to real-world conditions. The key to data preparation involves collecting trustworthy input data, accurately structuring and formatting the data, and validating its accuracy. By giving importance to accurate data preparation, hydraulic modelers can enhance the reliability and precision of their models, leading to improved decision-making and better management of water resources.

4.4 Python Interface

In this section, I'm excited to share my own Python Interface for the HEC-RAS model. The main point of this interface is to make the process of loading data into the model more efficient and less likely to contain errors. To do this, I've adopted the use of the pandas and pyreadr libraries in Python to read and process data from both Excel and RDS files. Additionally, the script takes advantage of regular expressions and various utility functions to accommodate a huge array of input parameters and data formats. The Python script I created aims to automate the process of updating data within the HEC-RAS model. This makes the script really flexible for various modeling scenarios, and it can really enhance the accuracy and productivity of 1D hydraulic modeling tasks. It is designed to handle a range of input parameters such as the path to the data files, the river reach code, the year for which the data needs to be updated, and the minimum acceptable value. Note that to use pyreadr 64-bit python is required and it might not work with 32-bit.

One of the functions in this script, the "process_rds" function, reads in area data of each river reach from an Excel file using the pandas library's read_excel function. It then uses regular expressions to replace the placeholder "XXX" in the file path with the actual river reach code, so that the appropriate RDS file containing the relevant data can be

accessed. The function then uses the pyreadr library's read_r function to read the data from the RDS file and convert it into a pandas dataframe. The dataframe is filtered to only include data for the specified year. The function then iterates through the rows of the dataframe and extracts the data values for that year, which are then converted from the rainfall-runoff model per second into the flow rate or discharge of water in a stream per day using another function called convert_rm_q. Finally, the function uses a list comprehension to round the converted values to three decimal places and return the resulting list.

The script's divide chunks function is essential because it divides the RDS data that has been processed into the required range of chunks. In order to comply with the HEC-RAS model's prerequisites, which consider the data to be divided into predetermined-sized chunks, this function is vital. The size of the chunks may vary depending on the specific analysis being performed, but in my case, the data was divided into 10 chunks as it was presented in the Unsteady Flow .u file. This process is necessary to format the RDS data for the HEC-RAS model. The function offers flexibility while ensuring adherence to the model's specifications by allowing the user to modify the chunk size in accordance with the demands of the particular analysis. Overall, the divide chunks function is an essential element that is required to make sure that the data is formatted correctly for the HEC-RAS model and is essential for the effective execution of the analysis.

The "if_num" function is a helper utility that checks if the first 8 characters of a line of text contain any numbers, to identify lines that should be ignored when adding new information to the HEC-RAS model document. On the other hand, the "return_spaces" function is used to ensure proper formatting of the data for the HEC-RAS model. The HEC-RAS model requires strict adherence to formatting guidelines, including a specific number of characters for each data cell. The "return_spaces" function calculates the necessary number of spaces required to meet these guidelines, based on the data length so that the data is correctly formatted for the HEC-RAS model. For example, if a data point with a value of 0.004 and a length of 5 is provided, the function will return 3 spaces to ensure compliance with the HEC-RAS formatting requirements. Those helper functions play a significant role in the main script correctly formatting out data in the right place.

Finally, the run function is used to execute the script and output the processed data to a specified file. This function takes in a number of parameters, including the output file name, the input file name, the lowest acceptable Q value, and the path to the RDS files. It also takes in the year for which data is to be updated, the names of the columns in the Excel file that contain the data values and area values, and the name of the Excel file containing the area data. All those parameters must be specified in Config.json before the execution of the main code.

In order to specify the setup for the Python tool, config.json was created to store that information. Users have to select a file with old data from HEC-RAS, output file, year of interest, and path to the RDS files, a column from RDS data containing loading data, the lowest value of Q assigned to the data, excel with area data for each reach and the column name of area data. We can parse RDS data without any preprocessing, but area data should contain one column with the name "UPOV_ID" and a column with the area in meters. You can get that data from Geometry data using any GIS software.

After specifying all required configurations, it is time to run a script. For Linux users, it is enough just to run the "run.sh" file, which will execute the script with configuration from Config.json. Installing the libraries is not needed, as it is included in the execution file. For Windows users the idea is the same, just run a .bat file in the terminal to utilize the tool. Note that both batch and sh scripts assume that 'pip' is on the system's PATH variables and can be run from the command prompt without specifying the full path to the executable.

The tool performs well in terms of speed and accuracy. The average execution time after running the script 10-12 times is 7 seconds, with the majority of that time spent on I/O operations: 6.2 seconds. Most of the code operates with the use of such operations. The total number of I/O operations is 298789. The memory usage is reasonable at 78.8 MB. CPU utilization is 0.0%, which indicates that the tool did not consume much of the CPU power while running 100 lines of code.

Despite this there are still a lot of limitations in my tool. It is not capable of creating a new data file from scratch, as it requires an input data file to iterate through. It is possible to add creation of the file from scratch based on users' configurations, but I need to investigate

more data files from different HEC-RAS models to code accurate, fast, and dynamic applications. Additionally, code contains few performance limitations, as it operates using RDS data files, which can be slow and consume a lot of memory. Also, the run function reads HEC-RAS data file line by line and can also decrease the performance for large files. Therefore, the tool can be enhanced even more and there are a lot of functionalities I will add in further development.

In conclusion to this part, Python can be a great tool for HEC-RAS modeling. Utilizing Python with HEC-RAS will save you a significant amount of time and enhance the accuracy of calculations. To work more efficiently with HEC-RAS and Python, the script should be thoroughly tested to reduce bugs and improve performance. To get the most out of it, Python code must be fast, as the main purpose of such tools is to perform manual work automatically, saving the scientist time.

4.5 HEC-RAS simulation

4.5.1 Berounka model

The objective of the present study is to validate the simulation of the HEC-RAS model over the Berounka river basin with respect to observed runoff. The data from the HAMR system was used in the Bilan model to simulate River Mean Runoff in millimeters (RM), which was preprocessed and used to drive the HEC-RAS model. Modeled runoff height [mm] was transformed using UPOV area to the discharge [m^3/s] as HEC-RAS requires it as input for the analysis. High flow conditions will be recognized in the dataset and general statistics of the events for the research region will be provided for a year period that represents normal and drought conditions. The dataset will be split into calibration and validation sets to improve the performance of the model. The HEC-RAS model will be calibrated and validated with respect to observed data using standard hydrological model performance metrics focusing on extremes. Python language will be used to load the data into the model efficiently based on the users' configuration. Results will be validated with respect to the real data to improve accuracy by changing the model setup.
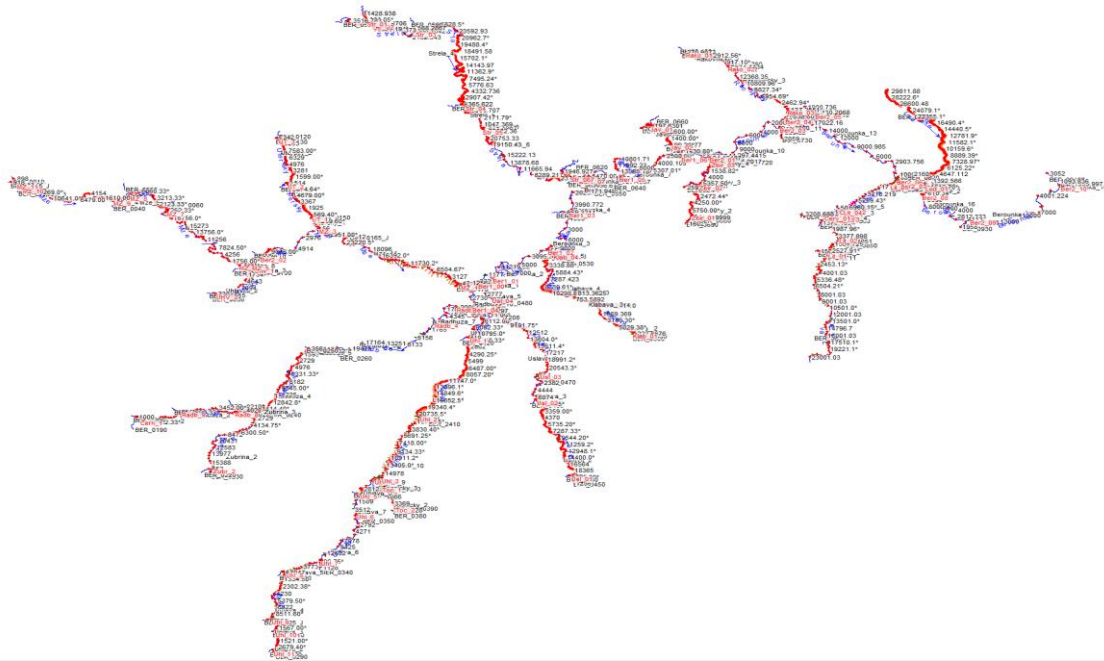
Picture 2: Berounka HEC-RAS model

4.5.2 Model Calibration

To ensure the accuracy of the HEC-RAS model used in this study, I conducted a comprehensive process that involved several essential steps. The calibration process of the HEC-RAS model for the Berounka catchment was done by testing the series of predicted data on the observed measurements under various flow conditions. After preprocessing the real data for each river station in the study area, the error metrics were calculated, including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), concerning predicted data from the model simulation. MAE is an absolute difference between predicted and observed data, while RMSE is a root of a squared difference. Both of these metrics are non-negative and in the same units as the data, in my case, it is in meters. To obtain model results from the simulation, open the attribute table tab in HEC-RAS and set up the reach and time range needed.

To successfully calibrate the HEC-RAS I choose datasets for 2019, as the flow conditions are varied within the whole year from normal to floods. The process of calibration contains a few key steps. First, the data for 2019 was loaded into HEC-RAS using the Python Interface, which I coded for this study. Firstly, the Unsteady Flow data file was added to the project as BER_2019. After loading the data, we manually checked it against the data from

the unsteady flow editor to ensure that it had been input correctly. Once we had verified the accuracy of the data, we proceeded to the next step in the calibration process. Working with hydraulic modeling it is important to check the data and model setup carefully.

To calculate the accuracy of the HEC-RAS model, we conducted a first round of error metrics calculation, specifically MAE and RMSE. it allowed us to understand if the model performed well and what steps have to be taken to minimize errors. After the first round of calculation, both errors were too big due to inaccurate input of the roughness coefficient. When the predicted data is predominantly higher than observed measurements, it indicates that the roughness coefficient of the channel might be too high, which would result in higher frictional resistance and a slower flow rate. And the opposite if the observed data is higher than predicted, the coefficient might be too low for that channel. We assumed now that DEM for that model was inputted correctly without any mistakes in it.

Before the next simulation, it is important to adjust the setting of the roughness coefficient in HEC-RAS in View/Edit Geometry data Manning's n table. The delta rate is different for each river reach and was chosen by performing more than 4 simulations, in a range of 0.01-0.1. In general 10 simulations were performed to ensure that chosen roughness coefficient results in accurate predictions. After each simulation I made a decision on the new delta for the roughness coefficient, based on the change in error metrics and how predicted values varied each time.

Finally, suitable Manning's n values were found for each river station, where the MAE and RMSE were at their minimum. Based on that I concluded that some of the river stations might have inaccuracies in their Geometry data or boundary conditions of a channel. River Station Zdirec showed the biggest errors from all 5 stations. Also, the station Stankov with errors quite higher than others might contain some issues with its geometry. However, stations Liblin, Tajanov, and Stribro performed well with suitable MAE and RMSE for this model. I will present the final results of each final metric for 2019. Also, I will provide here the final roughness coefficients, which will be used during the validation. Additionally, it is important to note that eros might be slightly higher due to missing data in original datasets, as they were filled in using mean imputation methods and are not 100% precise values.

Location	MAE	RMSE	Year	Station	River
Stříbro	0.143	0.272	2019	155.61	Mže
Libín	0.256	0.322	2019	5470	Berounka
Tajanov	0.172	0.249	2019	2792	Úhlava
Ždírec	0.867	0.882	2019	9902.51	Úslava
Staňkov	0.64	0.665	2019	13130.7	Radbuza

Table 7: Final MAE and RMSE for 2019

Final Manning's n values for the station for each stations:

Station	New Manning's n	Old Manning's n
155.61	0.038	0.03
5470	0.003	0.02
2792	0.048	0.035
9902.51	0.09	0.05
13130.7	0.01	0.03

Table 8: Manning's values after calibration

After calibrating the HEC-RAS model, we are now ready to move on to the validation process. For this next phase, we will be using the same dataset and study area as before, but focusing specifically on the year 2015, during which the flow conditions were known to be abnormal. By choosing this particular dataset, we can evaluate how well the model performs under conditions that are more challenging or differs, allowing us to better assess its accuracy and identify any potential areas for improvement. Through the calibration process, we adjusted various parameters in the model to minimize the error between the observed and predicted water surface elevations, and refine the model to more accurately represent the real-world conditions. Overall, by conducting a thorough validation and calibration process,

we aim to ensure the accuracy and reliability of the HEC-RAS model for use in future studies and engineering applications.

4.5.3 Model Validation

To validate the HEC-RAS model for a given study area, I followed an established process in order to produce accurate and reliable results. First, data sets were prepared for the year 2015. This specific year was chosen due to the presence of normal and extreme flow events, which differ from flow conditions in 2019 so we can see how the model could simulate results and error metrics under different flow conditions. The data set contained more than 85% non-missing values, which was deemed acceptable for the analysis.

Then, RDS data for 2015 was loaded into an HEC-RAS .u file using Python Interface. Additionally, observed data were preprocessed while missing values were filled in, in order to obtain a dataset for a full year. This was necessary in order to guarantee that the model provided accurate and reliable results. Then the predicted data from the simulation were compared against measured observations to calculate the MAE and RMSE ensuring that the predicted results can represent real-world scenarios. All results from the validation process will be presented in the following chapter.

The model highly depends on the accuracy of the geometry data and configuration parameters, like the roughness coefficient. After the calibration and validation process, the suitable roughness coefficients were found for each station and potential mistakes in data input were identified. The HEC-RAS model was successfully validated using observed data from multiple stations in the Berounka catchment, demonstrating the capability of the hydraulic modeling to accurately simulate water surface elevation in the river systems.

5. Results and Discussion

5.1 Python Interface

Overview and code statistics results:

Python is a popular and versatile programming language used for many tasks and applications. It is praised for simplicity, readability, and beginner-friendly, making it an excellent choice for a scientist working with hydraulic modeling. According to the TIOBE Index, Python has been among the top three most popular programming languages since 2018. Overall, Python language is one of the best options to automate monotonous and boring tasks for hydraulic modeling or simulation results analysis. However, it might be a bit challenging to use along with HEC-RAS.

Using Python language to load data into the HEC-RAS unsteady flow analysis automatically, performed successfully without any bags or errors. The script is fast and does not consume a lot of memory. It does not use any CPU, which makes the script a useful tool for HEC-RAS modeling. The average time of running the script is 7 seconds, which makes it a fast tool to update data for hydraulic modeling.

The Python Interface developed in this study provides a flexible and efficient way of updating data for the HEC-RAS modeling. It is specialized in unsteady flow modeling but can be enhanced to perform different analyses. There are only 4 external libraries utilized in the code, including pandas, pyreadr, numpy, and openpyxl. Also, it used regular expressions, along with utility functions to read and process different file formats, like RDS and excel. The script significantly reduces the likelihood of errors and can improve the accuracy of hydraulic modeling in HEC-RAS, as it is impossible to archive a good model with errors in inputted data.

The script uses RDS files as data input and does not require any preprocessing with them, which makes the script easy to use. To do so, I had to convert data from RM to Q right in the script. Converting RM into Q requires the area of each river reaches to use in the

formula, so the only preprocessing that has to be done is to create an Area file from the geometry data. However, it can be done easily using any GIS application, such as ArcGIS or QGIS.

Code uses several functions, which were described in the Practical Part of this study. Here I want to summarize the main functions, which are utilized by the Interface to load data into the model automatically. The "process_rds" function reads area data of each river reach from an Excel file, replaces the placeholder with the actual river reach code, and accesses the relevant RDS file using the pyreadr library's read_r function. The resulting data is filtered to include only the specified year, and the function iterates through the rows of the dataframe, extracts the data values for that year, and converts them into the flow rate per day. The "divide_chunks" function splits the RDS data into chunks of 10 as required by the HEC-RAS model. The "if_num" and "return_spaces" functions ensure proper formatting of data for the HEC-RAS model.

The "run" function executes the script and saves the processed data to a specified file. The tool's performance was satisfactory, and it was simple to set up and use using the "run.sh" or ".bat" file for Linux and Windows.

To sum up, Python Interface proves to be beneficial for 1D hydraulic modeling, as it speeds up the process of loading data into the model and improves the accuracy of inputted data. Loading data is only one approach that can be improved, as there are countless possibilities to automate working with hydraulic modeling.

Further Development:

While the developed Python interface has demonstrated its effectiveness and efficiency in automating data loading for HEC-RAS hydraulic modeling, there remains significant potential for further development in this area. In the future, there are several areas of development that could be explored to enhance the functionality and usability of the Python interface for HEC-RAS modeling. Next, I will describe some of the most important developments.

One of the most important features that can be added to the HEC-RAS Python Interface is the capability to load data into other models and analyses. It is possible to archive and make it an even more dynamic tool for updating data in HEC-RAS. In order to do so, it is important to analyze different hydraulic models from different analyses, including 1D, 2D, 3D and conceptual models. It will increase the number of parameters in configuration but might become an ultimate tool for hydraulic modeling in general.

To make it more user-friendly it is important to add a user-friendly interface. It might be a simple GUI, where you choose your configuration based on prompt requests from script or complete UX/UI interface for a python executable. It might reduce a wide range of potential mistakes during manual configuration and increase the possibilities of the code itself. Overall, any kind of interface can significantly improve user experience and make it the best tool for the HEC-RAS.

Next, it is essential to automate the calibration process of the HEC-RAS hydraulic modeling. That might include adjusting the model parameters, automate checking of the input and geometry data and many more. It can be achieved using genetic algorithm optimization and automated sensitivity analysis. Raspy-Cal tool utilizes genetic algorithms which mimic natural selection and genetic recombination to find the best set of model parameters that match the observed data. Also, it uses sensitivity analysis in order to find parameters which affect the errors the most and adjust them. The sensitivity analysis helps to narrow down the search space for the optimization algorithm, making the calibration process more efficient (Philippus, D., 2021). Similar approach can be implemented in my tool for HEC-RAS modeling, along with adding the functionality to validate geometry and boundary condition data, as it is also a part of calibrating the model.

Current version of the tool can handle unsteady flow models with flow hydrograph for a 365 year period. Next step of the development will be adding functionality to process different time periods and leap years. It is an important step to perform different analyses on the most suitable time stamp.

One more thing that is worth mentioning here is a functionality to export and analyze the results of the simulations. The analyzer could read and process the simulation results

generated by HEC-RAS, providing users with an easier and more efficient way to analyze and visualize the data. The analyzer could also offer customizable options for plotting and visualizing the results, allowing users to easily extract meaningful information from their simulations. Additionally, the analyzer could be designed to work seamlessly with the existing Python interface, streamlining the entire modeling process from data loading to output analysis. Tomasz Dysarz discusses in his article several approaches for result analyzer in Python for HEC-RAS using the HEC-DSSVue API to read and write data from the HEC-DSS file format, which is commonly used for storing model output data in HEC-RAS simulations. Additionally, he discusses the use of the Pandas library in Python for data analysis and visualization of the model results. The author demonstrates how to use Pandas to read the output data from HEC-DSS files, convert it into a Pandas dataframe, and perform various data analysis and visualization tasks. For example, the author shows how to plot the water level time series data from a model simulation using the matplotlib library in Python (Tomasz Dysarz, 2018).

To summarize, the potential use of Python with HEC-RAS is enormous and limited only by one's imagination, providing countless options for further growth and improvement that can enhance the work of hydraulic scientists and help minimize human errors, resulting in more accurate and efficient modeling and analysis. It is meaningless comparing Python Interface capabilities with manual work, as accurate automatization can only benefit every person of every profession.

Implementation of HEC-RAS Interface in different languages:

There are various other programming languages and technologies that may interface with HEC-RAS, including MATLAB, Fortran, R, and a Python library called PyRAS. PyRAS is a Python package that communicates with a HEC-RAS, allowing you to read and write HEC-RAS files, execute simulations, and extract results. Its capabilities are nevertheless constrained because it was created using an outdated version of HEC-RAS and does not provide you the ability to edit data. It can automate the process of running steady and unsteady models, producing cross sections, and analyzing the results.

Another technology, which can be used to automate HEC-RAS is MATLAB. It is possible to create a custom toolbox that connects with HEC-RAS to automate simulation

tasks and analyze the results of those simulations. While MATLAB can be a powerful tool for working with HEC-RAS models, it can also be expensive, need a lot of customization, and have a steep learning curve. Overall, it is not the best option in many circumstances, but worth investigating, if your tasks require working with large models. Additionally, the custom MATLAB toolbox described and developed by Lean et al. for their article is not publicly available, so you would need to develop your own implementation of the toolbox to use it in your own research (Lean et al. 2016).

It is also possible to use R and Fortran to handle automation of HEC-RAS simulations. It can be done using file preprocessing or HECRASController, which utilizes the HEC-RAS API. As HEC-RAS was developed in Fortran, it has functionality to communicate with it. Because the R programming language is not as well-established as other programming languages such as Python or Fortran, it may require more initial setup and customization to connect R with HEC-RAS. (Goodell, 2018)

5.2 Model Validation

The model's calibration and validation processes were generally successful, with all error metrics such as MAE and RMSE falling within the acceptable range. This demonstrates that the model is performing as expected.

Nonetheless, the findings revealed that the model data for two river reaches may be incorrect. It is possible that the errors are in the geometry data, and further investigation is required to determine the source of these errors and correct them in order to ensure the model's accuracy. To accomplish this, the geometry data should be checked to ensure that it is accurate and up to date, and any necessary adjustments made if errors are discovered.

After running the model and calculating MAE and RMSE, I got the following results:

Location	MAE	RMSE	Year	Station	River
Stříbro	0.16	0.336	2015	155.61	Mže

			54		
Libín	0.406	0.453	2015	5470	Berounka
Tajanov	0.175	0.235	2015	2792	Úhlava
Ždírec	0.858	0.875	2015	9902.51	Úslava
Staňkov	0.670	0.692	2015	13130.7	Radbuza

Table 9: Final MAE and RMSE for 2015

In 2015, the station Liblin had slightly higher errors. Model was calibrated based on extreme flow conditions, while most data for Liblin in 2015 were within a short normal range, except in mid-January. That led to an increase in error metrics. During November and December 2015, the model simulated results slightly worse for Liblin station, which led to bigger error metrics. That happened as data for November and December were missing in the measured dataset and filled in using mean imputation so the real data was not accurate enough. For the stations Tajanov and Stribro the HEC-RAS model simulated accurate predictions with low errors. However, the geometry data for stations Zdirec and Stankov were inputted with issues, so the model was validated with too big errors for those stations. From validation and calibration results of those stations, I found that DEM of Zdirec and Stankov have to be checked before simulating predictions in that study area for further research. In order to get accurate results, new geometry and/or boundary condition data have to be checked, prepared and updated to the HEC-RAS model, if necessary.

To make it more clear, the results of the validation process demonstrate the accuracy of the predictions for the Berounka catchment under all possible flow conditions. The model accurately predicted the water surface elevation (WSE) at the Tajanov and Stribro stations with low errors, while the Liblin river station performed slightly worse, but in the acceptable range. However, the model simulated bigger errors for river stations Zdirec and Stankov due to potential issues with the geometry data of the stations. Also, the real data for those stations contains more missing values for both years, than others. All those factors may lead to differences between real and predicted datasets.

Based on simulation results and error metrics MAE and RMSE calculation, adjustments to model configurations were made. In specific, the most suitable roughness coefficients of each river station were found, resulting in more accurate predictions of the real-world flow scenarios. It is important to carefully validate the model using full observed measurements and well-predicted values after the calibration of the model. Combination of those two processes can significantly improve the overall performance of the model and enhance the accuracy of the predicted values with different datasets.

The study demonstrates the effectiveness of using roughness coefficients for calibrating and validating hydraulic models in HEC-RAS. Validation metrics such as MAE and RMSE provide a quantitative measure of model accuracy, allowing for the identification of areas that require further calibration or improvement. However, it is important to consider the limitations of the model and the potential influence of factors such as input data accuracy and model assumptions, such as Manning's roughness coefficient.

To maintain accuracy, ongoing monitoring and calibration of the model may be necessary, particularly for systems that are subject to changes in hydrologic conditions or human activities. Additionally, it is important to consider the ability of the model to simulate extreme events or complex flow patterns, which may require additional model complexity or alternative approaches.

Selecting and validating the right hydraulic models is important for their usefulness and reliability in the management of water resources. Doing so provides insight into complicated hydraulic systems and helps in making informed decisions regarding flood control, infrastructure planning, and water allocation.

6. Conclusion

The main objective of this study was to develop a Python interface for the HEC-RAS model to automatically update the .u file for unsteady flow analysis, and to validate its accuracy for the Berounka catchment. With the assistance of the script, I was able to load data for several different years into the HEC-RAS fast and error-free to calibrate the model parameter, called the roughness coefficient. The Berounka 1D HEC-RAS model is an

abnormally large model that has almost 100 input reaches, so using an automotive Python Interface improved the time and accuracy of a data input. The Interface could become a unique and useful tool for large-scale hydraulic modeling projects in the HEC-RAS. Then the calibration process helped enhance the performance of the model and the accuracy of its simulation predictions for flow dynamics of the Berounka catchment, which has a significant effect on flood management in the area.

This study has been successful, but there are some limitations that need to be acknowledged. While the data for the model was fine and up-to-date, issues with the accuracy of the geometry or boundary condition data were found during the validation process. These inaccuracies led to bigger differences between the real and predicted data and led to higher values for MAE and RMSE. For future studies with this model for the Berounka catchment, geometry data have to be checked and updated, in order to get the best predictions from the simulation. Despite these limitations, my findings show how essential it is to utilize advanced modeling methods to better understand and manage natural resources.

Our findings show that the HEC-RAS model effectively predicted water surface elevation from flow discharge over the course of a year. The metrics used to assess the model's accuracy, such as MAE and RMSE, are within an acceptable range for the objectives of the analysis. Furthermore, the model's accuracy was increased by our calibration of the roughness coefficient, indicating that this is a method that might be utilized in the future to enhance other hydraulic models.

To sum up, this thesis has effectively demonstrated the importance of precise modeling of the study area's geometry and adjusting the model setup to enhance the accuracy of the HEC-RAS simulations. The precise loading of geometry data into the HEC-RAS model is crucial since it serves as the foundation for hydraulic simulations. Any inaccuracies or discrepancies in the data will have a significant impact on the reliability and validity of the simulation results. The use of programming languages like Python to perform HEC-RAS simulations quickly is a significant advantage as it helps to eliminate human errors and saves valuable time for scientists. These factors are essential for validating the HEC-RAS model of the Berounka catchment, resulting in simulation outcomes that more effectively and reliably represent the actual flow conditions in the real world.

References

1. Ballesteros, J.A., Bodoque, J.M., Díez-Herrero, A., Sanchez-Silva, M., & Stoffel, M. (2011). Calibration of floodplain roughness and estimation of flood discharge based on tree-ring evidence and hydraulic modeling. Journal of Hydrology, 403(1-2), 103-115. doi: 10.1016/j.jhydrol.2011.03.045

2. Bilan Model User Manual. (2015). Retrieved from http://bilan.vuv.cz/bilan/wp-content/uploads/2015/03/bilan_man_2015-06-23.pdf

3. Bravo, J. M., Allasia, D., Paz, A. R., & Collischonn, W. (2012). Coupled hydrologic-hydraulic modeling of the Upper Paraguay River Basin. Journal of Hydrologic Engineering, 17(5), 635 - 646. doi: https://doi.org/10.1061/(ASCE)HE.1943-5584.0000494

4. Casas, A., Benito, G., Thorndycraft, V.R., & Rico, M. (2006). The topographic data source of digital terrain models as a key element in the accuracy of hydraulic flood modelling. Earth Surface Processes and Landforms, 31(4), 444-456. https://doi.org/10.1002/esp.1278

5. CSIRO. (n.d.). Water Information Research and Development Alliance. Retrieved from https://www.csiro.au/en/research/natural-environment/water/water-resource-assessment/wirada

6. Czech Hydrometeorological Institute. (2015). Hydrological Yearbook 2015. https://www.chmi.cz/files/portal/docs/hydro/hydrologicke_rocenky/HR_2015.pdf

7. Czech Hydrometeorological Institute. (2019). Hydrological Yearbook 2019. https://www.chmi.cz/files/portal/docs/hydro/hydrologicke_rocenky/HR_2019.pdf

8. Czech Hydrometeorological Institute. (n.d.). O projektu [About the Project]. HAMR: Hydrologie x Agronomie x Meteorologie x Retence [Website]. Retrieved from https://hamr.chmi.cz/o-projektu.

9. Dias, V.C.F., Besner, M.C., & Prévost, M. (2017). Predicting Water Quality Impact After District Metered Area Implementation in a Full-Scale Drinking Water Distribution System. Journal - American Water Works Association, 109(9), E363-E380. https://doi.org/10.5942/jawwa.2017.109.0099

10. DHI Group. (n.d.). MIKE 11. Retrieved March 4, 2023, from https://www.mikepoweredbydhi.com/products/mike-11

11. DHI Group. (n.d.). MIKE 21/3. Retrieved March 4, 2023, from https://www.mikepoweredbydhi.com/products/mike-21

12. Dysarz, T. (2018). Application of Python scripting Techniques for control and automation of HEC-RAS simulations. Water, 10(10), 1382. https://doi.org/10.3390/w10101382

13. Gholami, V. (2022). Prediction of flood discharge and flood flow depth using a hydraulic model and flood marks on the trees in ungauged forested watersheds. Journal of Forest Science, 68, 190-198. https://doi.org/10.17221/6/2022-JFS

14. Glock, K., Tritthart, M., Habersack, H., & Hauer, C. (2019). Comparison of hydrodynamics simulated by 1D, 2D and 3D models focusing on bed shear stresses. Water, 11(2), 226. https://doi.org/10.3390/w11020226

15. Goodell, C. (2018, October 8). Using the HEC-RAS Controller with the R programming environment. Kleinschmidt Group. Retrieved from https://www.kleinschmidtgroup.com/ras-post/using-the-hec-ras-controller-with-the-r-programming-environment/#:~:text=While%20%E2%80%9CBreaking%20the%20HEC%2DRAS,for%20statistical%20computing%20and%20graphics.

16. HAMR. (n.d.). O projektu. Retrieved from https://hamr.chmi.cz/o-projektu

17. Hanel, M., Horáček, S., Daňhelka, J., Tomek, M., Hánová, K., Vizina, A., Ledvinka, O., Treml, P., & Melišová, E. (2014). Aktualizace odhadu hydrologických dopadů klimatických změn na povodích ČR. Vodohospodářské technicko-ekonomické informace, 52(1), 1-13. http://www.adaptacesidel.cz/data/upload/2015/07/028_aktualizace_odhadu_dopadu_kz_povodi.pdf

18. Hutanu, E., Mihu-Pintilie, A., Urzica, A., Paveluc, L. E., Stoleriu, C. C., & Grozavu, A. (2020). Using 1D HEC-RAS Modeling and LiDAR Data to Improve Flood Hazard Maps Accuracy: A Case Study from Jijia Floodplain (NE Romania). Water, 12(6), 1624. https://doi.org/10.3390/w12061624

19. Hydronia. (n.d.). RiverFlow2D: 2D River and Estuary Flow and Sediment Transport Model. Retrieved March 5, 2023, from https://www.hydronia.com/riverflow2d

20. Ikirri, M., Faik, F., Boutaleb, S., Echogdali, F. Z., Abioui, M., & Al-Ansari, N. (2021). Application of HEC-RAS/WMS and FHI models for extreme hydrological

events under climate change in the Ifni River arid watershed from Morocco. Climate and Land Use Impact on Natural and Artificial Systems, Elsevier, 2021, 251-270. https://doi.org/10.1016/B978-0-12-822184-6.00008-9

21. Infoglobe.cz. (n.d.). River Úhlava - Czech Republic - Europe. Retrieved from https://en.infoglobe.cz/rivers/czech-republic/europe/uhlava/

22. Infoglobe.cz. (n.d.). Berounka - Czech Republic - Europe. Retrieved from https://en.infoglobe.cz/rivers/europe/czech-republic/berounka/

23. Infoglobe.cz. (n.d.). River Uslava - Czech Republic - Europe. Retrieved from https://en.infoglobe.cz/rivers/europe/czech-republic/uslava/

24. Infoglobe.cz. (n.d.). Radbuza - Czech Republic - Europe. Retrieved from https://en.infoglobe.cz/rivers/europe/czech-republic/radbuza/

25. Infoglobe.cz. (n.d.). Mže - Czech Republic - Europe. Retrieved from https://en.infoglobe.cz/rivers/europe/czech-republic/mze/

26. Jean-Baptiste, N., Malaterre, P.-O., Dorée, C., & Sau, J. (2011). Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model. Mathematics and Computers in Simulation, 81(10), 2201-2215. doi: https://doi.org/10.1016/j.matcom.2010.12.021

27. Kesler, T. (2023). Comparing 1D, 2D, and 3D Hydraulic Models in Urban Flooding Applications [Master's thesis, Utah State University]. DigitalCommons@USU. (8691). https://digitalcommons.usu.edu/etd/8691

28. Kumar, M. (2019). Sensitivity analysis of Manning's roughness coefficient in MIKE 11 model for Mahanadi delta region. Annals of Plant and Soil Research, 21(1), 62-66. DOI: 10.13140/RG.2.2.25811.63523

29. Laks, I., Sojka, M., Walczak, Z., & Wróżyński, R. (2017). Possibilities of Using Low Quality Digital Elevation Models of Floodplains in Hydraulic Numerical Models. Water, 9(4), 283. https://doi.org/10.3390/w9040283

30. Lama, G. F. C., Errico, A., Francalanci, S., Chirico, G. B., Solari, L., & Preti, F. (2020). Hydraulic Modeling of Field Experiments in a Drainage Channel Under Different Riparian Vegetation Scenarios. In River Flow 2020 (pp. 347-354). Springer, Cham. https://doi.org/10.3390/geosciences10020047

31. Leon, A.S. and Goodell, C. (2016). Controlling HEC-RAS using MATLAB. Environmental Modelling & Software, 84, 339-348. doi: 10.1016/j.envsoft.2016.06.026.

32. Melišová, E., Vizina, A., Staponites, L. R., & Hanel, M. (2020). The Role of Hydrological Signatures in Calibration of Conceptual Hydrological Model. Water, 12(12), 3401. https://doi.org/10.3390/w12123401

33. Ministerstvo životního prostředí. (2018, December 13). MŽP financovalo modernizaci systému pro sledování sucha. Retrieved from https://www.mzp.cz/cz/news_181213_hamr

34. Motyka, J. (1998). A conceptual model of hydraulic networks in carbonate rocks, illustrated by examples from Poland. Hydrogeology Journal, 6, 469-482. doi: 10.1007/s100400050169

35. Philippus, D.; Wolfand, J.M.; Abdi, R.; Hogue, T.S. Raspy-Cal: A Genetic Algorithm-Based Automatic Calibration Tool for HEC-RAS Hydraulic Models. Water 2021, 13, 3061. https://doi.org/10.3390/w13213061.

36. Tate, E., Muñoz, C., & Suchan, J. (2014). Uncertainty and Sensitivity Analysis of the HAZUS-MH Flood Model. Natural Hazards Review, 15(1), 04013020. doi: 10.1061/(asce)nh.1527-6996.0000167.

37. Tatomir, A., McDermott, C., Bensabat, J., Class, H., Edlmann, K., Taherdangkoo, R., and Sauter, M.: Conceptual model development using a generic Features, Events, and Processes (FEP) database for assessing the potential impact of hydraulic fracturing on groundwater aquifers, Adv. Geosci., 45, 185–192, https://doi.org/10.5194/adgeo-45-185-2018, 2018.

38. United States Army Corps of Engineers. (2023). HEC-RAS 2D User's Manual. Retrieved from https://www.hec.usace.army.mil/confluence/rasdocs/r2dum/latest/

39. USACE. (2016). HEC-RAS 5.0 Reference Manual. Steady Flow Analysis. Washington, DC: U.S. Army Corps of Engineers. https://www.hec.usace.army.mil/software/hec-ras/documentation/HEC-RAS%205.0%20Reference%20Manual.pdf

40. USACE. (2022). Manning's roughness coefficients. HEC-RAS 1D Technical Reference Manual, version 6.2. Hydrologic Engineering Center, US Army Corps of Engineers. Retrieved from https://www.hec.usace.army.mil/confluence/rasdocs/ras1dtechref/6.2/performing-a-dam-break-study-with-hec-ras/mannings-roughness-coefficients

41. USDA-ARS. (2021). Water Management and Conservation Research. Retrieved from https://www.ars.usda.gov/pacific-west-area/maricopa-arizona/us-arid-land-agricultural-research-center/water-management-and-conservation-research/

42. Vodní zdroje, vodní hospodářství, vodní stavby. (2022, December 8). The current version of the Bilan model. Retrieved from https://www.vtei.cz/en/2022/12/the-current-version-of-the-bilan-model/

43. Výzkumný ústav vodohospodářský T. G. Masaryka. (n.d.). Projekt SUCHO HAMR. Retrieved from https://www.vuv.cz/projekt-sucho-hamr/o-projektu-10/

44. Wu, X., Xiang, X., Wang, C., & Chen, X. (2013). Coupled hydraulic and Kalman filter model for real-time correction of flood forecast in the Three Gorges interzone of Yangtze River, China. Journal of Hydrologic Engineering, 18(11), 1416-1425. doi: 10.1061/(ASCE)HE.1943-5584.0000473

List of pictures, tables, graphs, and abbreviations

1.   List of tables

2.   List of graphs

3.   List of abbreviations

HEC-RAS: Hydrologic Engineering Center's River Analysis System
DMA: District metered area
DEM: Digital Elevation Model
DTM: Digital Terrain Model
WMS: Watershed Modeling System
GIS: Geographic Information System
HEC: Hydrologic Engineering Center
RANS: Reynolds-averaged Navier-Stokes
ST: Sediment Transport
MT: Mud and Tailings Flow
PL: Pollutant Transport
WQ: Water Quality
UD: Urban Drainage
PT: Plastic Transportation
SCS-CN: Soil Conservation Service Curve Number

FEP: Features, Events, and Processes
HAMR: Hydrological, Agronomic, Meteorological and Retention
CHMI: Czech Hydrometeorological Institute
NDMS: National Drought Monitoring System
HIS: Hydrological Information System
RM: mean runoff
Q: discharge
WSE: water surface elevation
CV: coefficient of variation
MAE: Mean Absolute Error
RMSE: Root Mean Square Error

4. List of pictures

Appendix

Code.zip:

- all_stats_calculation\stats.excel: excel files with all preprocessing, validation, calibration and summary statistics with graphs
- data_hec: All required rds data is located there for each river reach (output from Bilan model)
- area_1.xlsx: area file extracted from geo data from data_hec\geo using QGIS
- output_files\Berounka.u03: input file with old unsteady flow data
- loader\config.json: file with user configuration for the script
- loader\loader.py: python script to handle and store user configurations
- performance\performance.py: script I used to measure the code performance
- performance\performance_logs_updated.txt: text file with performance information, output from performance.py
- requirements.txt: required libraries to run the script
- run.bat: Batch Script to install libraries and run the code on Windows
- run.sh: Shell Script to install libraries and run the script on Linux
- main.py: Main Script with all functionality
- output_files folder is for all generated by the tool files