

University of Stuttgart  
Germany



Department of Hydraulic Engineering and Water Resources Management  
Institute for Modelling Hydraulic and Environmental Systems

# **Spatially Explicit, Depth-dependent and Data-driven Roughness Calibration for Numerical Modelling**

Master's Thesis

Akash Jyoti Handique  
Water Resources Engineering and Management (WAREM)

Matriculation Number: 3503892

05.11.2022

**Examiner:** PD Dr. Stefan Haun

**Supervisors:** Dr. sc. (PhD) Sebastian Schwindt  
M.Sc. Beatriz Medeiros Fernandes Negreiros



# **Declaration**

I declare that I have developed and written the enclosed thesis completely by myself and that I have not used other sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. The thesis was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere. The enclosed electronic version is identical to the printed versions.

I agree that the present work is made available for scientific purposes in the libraries of the Institute for Water and Environmental Systems, University of Stuttgart (published according § 6 Abs. 1 UrhG (Copyright Act) and thereof can be cited under § 51 of the UrhG (Copyright Act).

Akash Jyoti Handique  
Stuttgart, 5 November 2022

# **Acknowledgments**

I would like to convey my gratitude to my supervisors Dr. sc. (PhD) Sebastian Schwindt and M.Sc. Beatriz Medeiros Fernandes Negreiros for their technical support in developing this master's thesis. Their guidance and encouragement were essential for the development of this project. I am thankful to all my friends and colleagues for making my master's journey a memorable one.

Finally, I would like to thank my parents for their continuous support.

# Abstract

Hydrodynamic numerical modelling lays the foundation for modelling studies to investigate sediment transport, water quality, ecological changes, waves, currents, and many more. Simulating the flow of fluid following the governing dynamics is a complicated task. This study explores spatially explicit, depth-dependent, and data-driven roughness calibration of a hydrodynamic numerical model for the Lower Yuba River, located in Marysville, California. The study implements Ferguson's approach for frictional computation and considers the river section's diverse morphological unit and roughness components. A workflow is developed to integrate Ferguson's law by altering the pre-existing subroutines of Telemac-2D software. Spatially explicit friction zones are demarcated to assign independent calibration parameters to different regions of the river section. Two constants of Ferguson's variable-power equation are considered as the calibration parameters for the model. The fully developed hydrodynamic numerical model becomes computationally expensive and manual calibration is inconvenient owing to the time factor. To overcome this challenge, the full complexity model is substituted by a surrogate model. Equivalency between the surrogate model and the full-complexity model is achieved by using Bayesian active learning to train the surrogate model. Based on the information theory score, Bayesian active learning improves the model in the regions of parameter space that are more relevant for Bayesian inference. The calibration results indicate that the surrogate model is able to predict the outcome of the full-complexity model with high precision. Bayesian active learning methodology enhances the predictive capacity of the surrogate model with increasing iterations. The simulated results of the hydrodynamic numerical model are inconsistent with the measurement values for flow velocity and water depth in the regions with higher grain sizes due to the lack of roughness attribute representation and computation of enhanced roughness in these regions. Ferguson's approach computes higher roughness in comparison to Manning's approach. The increased precision of the model in predicting the low-flow velocities is a significant observation from this study in contrast to this usually being a drawback of a hydrodynamic numerical model.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>viii</b>
<b>Nomenclature</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Objectives . . . . .	2
1.3 Outline . . . . .	2
<b>2 State-of-the-art</b>	<b>3</b>
2.1 Hydro-morphodynamic numerical modelling . . . . .	3
2.1.1 Numerical approaches . . . . .	3
2.1.2 Hydrodynamic modelling . . . . .	4
2.1.3 Morphodynamic modelling . . . . .	7
2.2 Model calibration and validation techniques . . . . .	9
2.2.1 Bayesian calibration techniques . . . . .	9
2.2.2 Calibration parameters of hydrodynamic models . . . . .	18
2.2.3 Validation approaches for hydrodynamic models . . . . .	18
<b>3 Materials and methods</b>	<b>21</b>
3.1 Study area . . . . .	21
3.1.1 The Lower Yuba River . . . . .	21
3.1.2 Existing models . . . . .	29
3.2 Adaptation of the hydrodynamic model . . . . .	31
3.2.1 Mesh construction . . . . .	31
3.2.2 Roughness zones . . . . .	33

3.2.3	Integration of roughness zones with the mesh geometry . . . . .	38
3.2.4	Implementation of Ferguson law . . . . .	39
3.3	Calibration and validation approach . . . . .	41
3.3.1	Measurement data . . . . .	41
3.3.2	Initial training data for the surrogate model . . . . .	42
3.3.3	Telemac-2D coupling with the Bayesian calibration process . . . . .	44
<b>4</b>	<b>Results</b>	<b>46</b>
4.1	Calibration parameters . . . . .	46
4.2	Evolution of information theory scores . . . . .	51
4.3	Surrogate and the calibrated model assessment . . . . .	53
4.3.1	Surrogate model . . . . .	53
4.3.2	Bayesian calibrated model . . . . .	54
4.4	Comparison of Manning's coefficients . . . . .	55
<b>5</b>	<b>Discussion</b>	<b>57</b>
5.1	Quality of the Bayesian calibration process . . . . .	57
5.2	Behaviour of the calibration parameters . . . . .	58
5.3	Physical interpretation of the calibration parameters . . . . .	59
5.4	Model evaluation . . . . .	59
<b>6</b>	<b>Conclusions and future research</b>	<b>61</b>
6.1	Conclusions . . . . .	61
6.2	Future work and recommendations . . . . .	62
<b>References</b>		<b>63</b>
<b>Appendices</b>		<b>66</b>
<b>A</b>	<b>The Lower Yuba River maps</b>	<b>67</b>
<b>B</b>	<b>Comaprison of Manning's coefficients (Extended)</b>	<b>70</b>

# List of Figures

2.1 Gaussian function.	14
2.2 Workflow for Surrogate-assisted Bayesian inversion.	17
3.1 Yuba River catchment (EPSG: 32610).	22
3.2 Lower Yuba River section (EPSG: 32610).	22
3.3 Yuba catchment land use (EPSG: 32610).	23
3.4 Yuba catchment land use percentage.	24
3.5 Lower Yuba River grain size distribution (EPSG: 32610).	26
3.6 Inflow hydrograph.	27
3.7 Stage-discharge curve.	28
3.8 Lower Yuba River mesh resolution (EPSG: 32610).	31
3.9 Lower Yuba River mesh with BOTTOM variable (EPSG: 32610).	32
3.10 Lower Yuba River channel (EPSG: 32610).	34
3.11 Lower Yuba River vegetation (EPSG: 32610).	35
3.12 Lower Yuba River floodplain (EPSG: 32610).	35
3.13 Lower Yuba River channel pixels (EPSG: 32610).	36
3.14 Pixels frequency for Friction ID.	37
3.15 Lower Yuba River model domain with random points (EPSG: 32610).	38
3.16 Lower Yuba River mesh with Friction ID (EPSG: 32610).	39
3.17 Location of the calibration points (EPSG: 32610).	41
3.18 Telemac-2D coupling with GPE-BAL.	45
4.1 Test-1 posterior calibration parameters.	49
4.2 Test-2 posterior calibration parameters.	50
4.3 BME score of Test-1.	51
4.4 RE score of Test-1.	51
4.5 BME score of Test-2.	52
4.6 RE score of Test-2.	52
A.1 Lower Yuba catchment.	68

A.2 Course of the river Yuba. . . . .	69
---------------------------------------	----

# List of Tables

3.1 Flow probabilities . . . . .	27
3.2 Tuflow model specifications . . . . .	30
3.3 Friction ID grouping . . . . .	34
3.4 Friction ID pixel grouping . . . . .	37
3.5 Parameter list . . . . .	42
3.6 Parameter results of PC1 (EPSG: 32610) . . . . .	43
4.1 Test-1 optimum calibration parameters . . . . .	46
4.2 Test-2 optimum calibration parameters . . . . .	46
4.3 Test-1 posterior calibration parameters . . . . .	47
4.4 Test-2 posterior calibration parameters . . . . .	48
4.5 Test-1 surrogate model verification . . . . .	53
4.6 Test-2 surrogate model verification . . . . .	53
4.7 Data sorting . . . . .	54
4.8 Test-1 RMSE value in quartiles . . . . .	54
4.9 Test-2 RMSE value in quartiles . . . . .	55
4.10 Test-1 Manning's comparison . . . . .	56
4.11 Test-2 Manning's comparison . . . . .	56
B.1 Test-1 Manning's value comparison (Extended) . . . . .	71
B.2 Test-2 Manning's value comparison (Extended) . . . . .	72

# Nomenclature

## Symbols

$\nu$ [m <sup>2</sup> s <sup>-1</sup> ]	kinematic viscosity
$\nu_t$ [m <sup>2</sup> s <sup>-1</sup> ]	turbulent eddy viscosity
$\rho_f$ [kg m <sup>-3</sup> ]	density of fluid
$\tau$ [N m <sup>-2</sup> ]	bed shear stress
$\tau_{*cr}$ [-]	critical dimensionless bed shear stress
$\tau_*$ [-]	dimensionless bed shear stress
$\tau_{ce}$ [N m <sup>-2</sup> ]	critical shear stress for erosion
$\tau_d$ [N m <sup>-2</sup> ]	critical shear stress for deposition
$a_1$ [-]	Ferguson's law constant
$a_2$ [-]	Ferguson's law constant
$C$ [-]	Chézy coefficient
$D$ [kg m <sup>-2</sup> s <sup>-1</sup> ]	diffusion flux
$D$ [m]	height of friction elements
$D_{84}$ [m]	grain diameter of which 84 % of grains are finer
$D_{pq}$ [m]	grain diameter of which $pq\%$ of grains are finer
$E$ [kg m <sup>-2</sup> s <sup>-1</sup> ]	erosion flux
$f$ [-]	Darcy-Weisbach friction coefficient
$F_x$ [ms <sup>-2</sup> ]	source term in x direction
$F_y$ [ms <sup>-2</sup> ]	source term in y direction

$g$ [ms $^{-2}$ ]	gravitational acceleration
$h$ [m]	water depth
$M$ [kg m $^{-2}s^{-1}$ ]	Krone-Partheniades constant
$n$ [m $^{-1/3}s$ ]	Manning's coefficient
$q_b$ [kg s $^{-1}m^{-1}$ ]	bed load transport parameter
$R_h$ [m]	hydraulic radius
$Re_*$ [-]	particle Reynolds number
$S$ [-]	gradient
$t$ [s]	time
$u$ [ms $^{-1}$ ]	water velocity in x direction
$v$ [ms $^{-1}$ ]	water velocity in y direction
$Z$ [m]	water surface elevation

### Probability symbols

$D_{KL}$	relative entropy between two probability distribution
$K$	covariance matrix of the Gaussian process
$n$	number of measured points
$N_{post}$	number of posterior samples
$N_{prior}$	number of prior samples
$p(w)$	prior distribution
$p(Z_{meas})$	Bayesian model evidence
$p(w   Z_{meas})$	posterior distribution
$p(Z_{meas}   w)$	likelihood
$R$	covariance matrix of measurement error
$w$	vector of calibration parameters
$x$	training data
$x_*$	test data
$Z_{meas}$	measured data

$Z_{mod}$  modelled data

## Abbreviations

$2D$	Two Dimensional
$BAL$	Bayesian Active Learning
$BME$	Bayesian Model Evidence
$DEM$	Digital Elevation Model
$GP$	Gaussian Process
$GPE$	Gaussian Process Emulators
$GPR$	Gaussian Process Regression
$LIDAR$	Light Detection and Ranging
$LYR$	Lower Yuba River
$MB$	Multibeam
$RE$	Relative Entropy
$RMSE$	Root Mean Square Error
$RTK$ -GPS	Real Time Kinematics - Global Positioning System
$VPE$	Variable Power Equation
$YCWA$	Yuba County Water Agency



# **Chapter 1**

## **Introduction**

### **1.1 Motivation**

Anthropogenic measures in the past have degraded the status of the Lower Yuba River. Hydraulic gold mining in the 19th century and rampant construction of hydraulic structures changed the river sediment dynamics creating navigation problems, flood level increment, and farmland destruction leading to a detrimental effect on riparian vegetation, and riverbed composition for spawning. Much research was carried out in the past to understand the hydrological and morphological characteristics of the river and contribute to the development of the ecosystem.

Past approaches in emulating the behaviour of the Lower Yuba River with modelling are conventional. With the emergence of new methods and equations which are more accurate in determining the river characteristics, incorporating state-of-the-art approaches will be beneficial in comprehending the natural phenomena of the river and further facilitating the process of the Lower Yuba River ecosystem rehabilitation. In the 2D hydrodynamic numerical model domain, a friction law is imperative in terms of determining the efficiency of the model. Capturing the varying physical characteristics of the river bed, representing the accurate velocity profiles, and roughness regions determine the usefulness of a friction law for different rivers. The inability of the existing friction laws to reproduce the hydraulic pattern for gravel and boulder bed rivers has subjected more significance in adapting a new friction law for 2D hydrodynamic models. Uncertainty in determining flow resistance usually arises due to ineffective characterisation of riverbed roughness. Ferguson (2007) considered alternative equations, capable of predicting hydraulic variables by taking into account the relative submergence experienced in gravel and boulder bed rivers. This study is motivated to improve the predicting capacity of the

Lower Yuba River 2D hydrodynamic numerical model by implementing Ferguson's law in spatially explicit friction zones and calibrating the model with a Surrogate-assisted Bayesian inversion process.

## 1.2 Objectives

The main purpose of this study is to run a hydrodynamic model for the Lower Yuba River using spatially explicit, depth-dependent and data-driven roughness calibration.

Specific objectives of this study deal with:

- Adapting Ferguson's approach for the existing Lower Yuba River hydrodynamic model in Telemac-2D.
- Categorisation of the spatially explicit roughness zones for the Lower Yuba River.
- Implementing the surrogate-assisted Bayesian inversion process for calibrating the Lower Yuba River hydrodynamic numerical model.
- Analysing the results and efficiency of the full complexity hydrodynamic numerical model and the surrogate model.

## 1.3 Outline

The content of this document is divided into 6 chapters which refer to the Bayesian calibration of the Lower Yuba River hydrodynamic numerical model with the adapted Ferguson's law in spatially explicit roughness zones.

Chapter 2 summarises the state-of-the-art concept and approaches of hydro-morphodynamic numerical modelling. The relevant Bayesian calibration techniques and parameters necessary to comprehend the content of this study are described elaborately.

Chapter 3 describes the hydrological and morphological characteristics of the study area. This chapter also delineates the underlying approaches of adapting the hydrodynamic model for the new friction law and spatially explicit roughness zones.

Chapter 4 showcases the results of the calibration process and checks the accuracy of the surrogate model and the full-complexity model.

Chapter 5 discusses the results of the study and states the reasoning for the observed behaviour.

Conclusions, future research, and recommendations are stated in Chapter 6 which also highlights the prominent findings of the study.

# **Chapter 2**

## **State-of-the-art**

### **2.1 Hydro-morphodynamic numerical modelling**

This section explains different approaches related to hydrodynamic and morphodynamic numerical modelling. Contemporary mathematical equations, concepts, and crucial steps corresponding to modelling are also addressed.

#### **2.1.1 Numerical approaches**

Numerical models are becoming important with advancements in the efficiency of simulating flow phenomena. In the field of hydraulics, models are widely used to solve various practical problems. By solving the Navier-Stokes equation through iteration, numerical models simulate flow characteristics. Choosing and understanding 1D, 2D and, 3D numerical models is a complex process. The inability of the 1D model to simulate the complex flow pattern of a river has put more significance to 2D and 3D models. With the advancement in the field of spatial data acquisition and accuracy, the 2D modelling domain has strengthened. Reducing Parameter sensitivity and inaccuracies made 2D modelling more efficient and easier to interpret. Higher computational time and sensitivity to boundary conditions impede the flawless execution of a 3D model. Telemac-2D solves the 2D shallow water equation, which is obtained from 3D averaged version of Reynold's Navier-Stokes equation(Wilson, Bates, & Hervouet, 2002). Navier-Stokes equation can describe the flow of a fluid with conservation of mass and momentum. This continuity equation along with a few assumptions like incompressibility and hydrostatic pressure distribution can illustrate the natural phenomena of a river. The mass and momentum conservation equations are as follows :

Continuity equation :

$$\frac{\partial h}{\partial t} + u \cdot \nabla h + h \cdot \nabla u = 0 \quad (2.1)$$

Momentum equation in x direction :

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = -g \frac{\partial Z}{\partial x} + F_x + \frac{1}{h} \cdot \nabla (h \cdot v_T \cdot \nabla u) \quad (2.2)$$

Momentum equation in y direction :

$$\frac{\partial v}{\partial t} + u \cdot \nabla v = -g \frac{\partial Z}{\partial y} + F_y + \frac{1}{h} \cdot \nabla (h \cdot v_T \cdot \nabla v) \quad (2.3)$$

where  $u$  and  $v$  are the velocity components,  $x$  and  $y$  are the space coordinates, and  $Z$  symbolises the free surface elevation,  $\mu$  is the viscosity,  $F_x$  and  $F_y$  represent the source terms,  $t$  is the time,  $g$  denotes the gravitational acceleration,  $h$  is the flow depth,  $v_T$  is the turbulent eddy viscosity.

Firstly, the model is being set up at steady state conditions i.e., the flow rate in the river is constant and does not change with temporal and spatial variations. In reality, a steady state condition is impossible to achieve in natural or even flow-controlled rivers. A steady state condition is executed to calibrate the model with field observations. After the calibration process, a model is used to simulate unsteady conditions. Unsteady condition is simulated by incorporating a time-dependent inflow hydrograph as a series of steady flow rates (Schwindt & Barros, 2022).

Hydrodynamic and Morphodynamic models are coupled to better replicate the riverbed parameters. The coupling process is flexible and a model can solve the Navier-Stokes equation along with the transport equations which leads to a longer computational time. In Coupling at an assigned frequency, transport equations are solved separately. The application ranges for both couplings are different. Section 2.1.3 discusses Morphodynamic modelling and equations associated with the modelling.

## 2.1.2 Hydrodynamic modelling

Hydrodynamic simulation deals with the motion of liquids. A robust hydrodynamical model is an amalgamation of various process that includes data pre-processing, Model preparation, grid setup, model running, calibration, and validation. Based on a topographic map, water depth, flow rate, surface roughness, boundary, and initial conditions with general, turbulence, and numerical parameterisation, the hydrodynamic model is capable of predicting the spatial and temporal patterns of hydraulic variables by solving the equations (2.1), (2.2), and (2.3).

### Mesh setup

An adequate mesh setup is a time-consuming and crucial step for hydrodynamic modelling. Mesh geometry reflects the original physical environment of the model location. Mesh comprised of triangular facets of various sizes and shapes, which facilitates the process of delineating model topography smoothly. Software like BlueKenue, QGIS provides specific tools for generating and amending 2D triangular meshes on Telemac-2D. BlueKenue provides a series of options to amend mesh parameters like resampling outline, edge length, and mesh density. Changing the density of the mesh in sensitive areas can result in accurate model output with numerical precision (Mensencal, 2012). Based on different morphological units within the model domain, mesh resolutions are adjusted. River channels are usually assigned with finer mesh resolution than floodplains and higher discharge areas. Various data can be inserted within the BlueKenue mesh. Data that can be inserted within the mesh in BlueKenue are :

- Friction coefficient
- Topography data
- Friction ID
- Free surface
- Water depth
- Wind data

### Liquid boundary condition

Boundary conditions are one of the most important factors that propagate within the model domain and model prediction capability depend on the boundary conditions. Assigning water depth, elevation, and discharge that reflect the physical conditions of the river hydrology on the model boundary is a sensitive process. In the upstream and downstream boundary conditions of the model, discharges, elevations, and velocity profiles can be prescribed (Schwindt & Barros, 2022). Telemac-2D provides a few options for assigning boundary conditions. For depth, the following conditions are assigned :

- Open boundary with prescribed depth
- Open boundary with free depth
- Closed boundary

For velocity and flowrate, following conditions are prescribed :

- Open boundary with prescribed flowrate
- Open boundary with prescribed velocity
- Open boundary with free velocity
- Close boundary with slip/friction
- Close boundary with one or two null velocity components

### **Friction boundary condition**

Telemac-2D enable the selection of 7 friction laws and the corresponding friction-coefficients. Following friction laws can be enabled :

- Haaland (Haaland, 1983)
- Chezy (Chézy, 1776)
- Strickler (Strickler, 1981)
- Mannings (Manning, 1891)
- Nikuradse (Nikuradse, 1933)
- Logarithmic law of the wall for turbulent flow (Karman, 1930)
- Colebrook and white (Colebrook & White, 1937)

Telemac-2D provides the flexibilty to change and add a new friction law whenever necessary. Independent friction law with coefficients can be allotted to designated friction zones in the model domain.

### **Initial condition**

Initial condition states the initial environment before the start of the simulation. Initial conditions involve assigning initial water elevation, depth, and dry or wet conditions for the model. Telemac-2D provides the following options for the initial condition :

- Zero elevation
- Constant elevation
- Zero depth
- Constant depth

Initializing the model with dry or wet conditions depends on the corresponding values of the initial condition.

Along with the above-mentioned factors, numerical parameters address the specific equations and approximations necessary for the simulation with the help of discretization methods. Turbulence parameterization considers the chaotic behavior of surface water. Formation of eddies, dissipation of energy is considered for better predicting the real-life hydraulic variables (Schwindt & Barros, 2022).

### 2.1.3 Morphodynamic modelling

Morphodynamical simulation emulates sediment processes in rivers and simultaneously considers riverbed formation, erosion, and deposition. Modelling of sediments and morphological evolution requires a spatial and temporal variation of grain size distributions and grading curves for simulating the alterations of alluvial bed in terms of bedload and suspended load. The cohesiveness of the sediments is also taken into account in morphodynamical modelling. Sediment transport in rivers originates from erosion, landslide, the flow of debris, and flooding. Sediment transport is impacted by river hydraulics, sediment characteristics, and availability. Transport capacity depends on the flow bed shear stress (du Boys, 1879). The equation of bed shear stress is stated below :

$$\tau = \rho_f \cdot g \cdot R_h \cdot S_e \quad (2.4)$$

where  $\rho_f$  represents the fluid density ( $1000 \text{ kgm}^{-3}$ ). The dimensionless expression of the bed shear stress is  $\tau_*$ , which refers to the particle density and the grain diameter. Dimensionless bed shear stress can be defined as follows :

$$\tau_* = \frac{R_h \cdot S_e}{(s - 1) \cdot D_{pq}} \quad (2.5)$$

where  $s = 2.68$ , it is the ratio of solid particle and water density, and  $D_{pq}$  is the grain diameter. Upon exceeding the  $\tau_*$  value, bed particles start to move. Known as critical dimensionless shear stress  $\tau_{*cr}$  or as the Shield's parameter, the grain mobility is influenced by  $\tau_{*cr}$ . The dimensionless particle diameter  $D_*$  or the particle Reynolds number  $Re_*$  (Einstein, 1950):

$$D_* = \left[ \frac{(s - 1) \cdot g}{\nu^2} \right]^{\frac{1}{3}} \quad (2.6)$$

$$Re_* = \frac{u_* \cdot D}{\nu} \quad (2.7)$$

This relationship results in the grain mobility curve and from the mobility curve, based on particle Reynolds number, dimensionless bed shear stress is determined.

Various processes like sediment transport plays a crucial role in river morphology. Sediments are fragmented by physical and chemical properties. Bedloads are the type of coarser sediments in the riverbed that slides, roll and jump. The hydraulic regime and transport rate are directly linked with bedload transport (Schwindt & Barros, 2022). Shear stress and flow rate Threshold defines the motion of the sediments. Meyer Peter Müller Equation 2.8 predicts the transportation of coarser bedload sediments (Meyer-Peter & Müller, 1948).

$$q_b = 8 \cdot (\tau_* - \tau_{*cr})^{\frac{3}{2}} \quad (2.8)$$

In the equation,  $\tau_*$  is the dimensionless bed shear stress,  $\tau_{*cr}$  is the Shield's parameter which value ranges from 0.047–0.07 and  $q_b$  is the dimensionless transport parameter.

For artificial and natural sediment supply limited rivers (Runoff in the river has potential of transporting more sediments than it is available in the riverbed) and for capacity-limited rivers (runoff in the river is not capable enough to transport more sediments that are available), different formulas are used for calculating the sediment supply. In rivers with infinite sediment supply, Meyer Peter Müller and its derivative equations are used (Schwindt & Barros, 2022).

Suspended loads are finer particles that are mobilized due to bulk flow and turbulence of the hydraulic regime. Advection -diffusion equations are helpful for simulating suspended load transport. The equations are as follows :

$$\frac{\partial(hC)}{\partial t} + \frac{\partial(huC)}{\partial x} + \frac{\partial(hvC)}{\partial y} = \frac{\partial}{\partial x}(h\varepsilon_s \frac{\partial C}{\partial x}) + \frac{\partial}{\partial y}(h\varepsilon_s \frac{\partial C}{\partial y}) + (E - D) \quad (2.9)$$

where  $u$  and  $v$  are horizontal and vertical velocity components,  $h$  denotes the water depth, and  $C$  is the depth-averaged concentration,  $\varepsilon_s$  is the turbulent diffusivity of the sediment, and  $E$  and  $D$  are Erosion and diffusion fluxes. The erosion and deposition fluxes are calculated as follows :

$$E = M \cdot \left( \frac{\tau}{\tau_{ce}} - 1 \right) \text{ if } \tau > \tau_{ce} \quad (2.10)$$

$$D = w_s \cdot C \cdot \left( \frac{\tau}{\tau_d} - 1 \right) \text{ if } \tau > \tau_d \quad (2.11)$$

where,  $M$  is the Krone-Partheniades erosion constant,  $\tau_{ce}$  and  $\tau_d$  are critical stress for erosion and deposition,  $w_s$  is the settling velocity and  $\tau$  is bed shear stress.

General and numerical parameters in morphodynamical modelling are quite similar to a hydrodynamic model. In the boundary of the model domain water depth, flow rate, and solid discharges can also be prescribed. To meet the objective of alluvial bed elevation change, water depth better not be prescribed in the upstream boundary. To attain the

flexibility of varying outflow according to the development of erosion and deposition, prescribing discharge at the downstream boundary is avoided.

On the liquid boundary, time-dependent sediment discharges are assigned. Suspended sediments are classified based on sediment density, diameters, and fractions. The concept of active layers in the morphodynamic model enables assigning of the erodible depth and mixing layers. Stratification and composition of riverbed layers based on characteristics of grain size of the sediment provide flexibility in terms of transportation, erosion, and deposition. Ensuring an efficient and timesaving simulation requires comprehending the type of dominant sediment classes and transport modes of the river (Schwindt & Barros, 2022).

## 2.2 Model calibration and validation techniques

Calibration of a numerical model is the adaptation of the numerical model to reality with the objective of minimizing the error between model results and observation data. Calibration process increases the likelihood of the numerical model to predict hydraulic variables.

### 2.2.1 Bayesian calibration techniques

Manual calibration of a hydrodynamic model with the input parameters can be computationally demanding. Therefore, the complexity of the model is replaced by a surrogate model with the help of Gaussian process regression. A good equivalence between the surrogate model and numerical model is achieved by training the surrogate model with Bayesian active learning based on information theory scores. Bayesian active learning improves the surrogate model iteratively in the regions of the parameter space which are more relevant correspond to Bayesian inference.

#### Bayesian inference

Proposed by Thomas Bayes in 1796, the Bayes theorem updates an initial prior hypothesis based on measured data and provide an accurate posterior hypothesis(Press, 2009). In the model calibration context, the Bayes theorem can be used to predict probable values for the calibrated parameters based on input data. Calibrated parameters are the parameters that are amended in the calibration process of a numerical model. Bayes Equation is as follows :

$$p(w | Z_{meas}) = \frac{p(Z_{meas} | w) \cdot p(w)}{p(Z_{meas})} \quad (2.12)$$

where  $w$  is the calibration parameters containing vector,  $p(w)$  is the initial prior probability of the parameters,  $p(Z_{meas} | w)$  is the likelihood which represents the probability of measured data  $Z_{meas}$  given the parameters  $w$ . Likelihood function aims to quantify the difference between the model results and measured values.  $p(w | Z_{meas})$  is the posterior probability of the parameters and  $p(Z_{meas})$  is Bayesian model evidence which acts as a normalizing factor to assure a valid probability distribution for the posterior. Prior distribution is the initial assumed distribution of the designated calibration parameters and posterior distribution is acquired from prior distribution utilising the Bayes Equation 2.12.

The following equation is used to calculate the likelihood function assuming the difference between the measured and modelled values are normally distributed :

$$p(Z_{meas} | w) = (2\pi)^{-\frac{n}{2}} \cdot \det(R)^{-\frac{1}{2}} \cdot \exp(-.5(Z_{meas} - Z_{mod})^T R^{-1} (Z_{meas} - Z_{mod})) \quad (2.13)$$

In the equation  $n$  is the number of the measurement points, and  $R$  is the covariance matrix of the measurement error,  $Z_{meas}$ , and  $Z_{mod}$  are measured and modelled values.

The Equation (2.12) is solved using Monte-Carlo simulation and repeated random sampling of the calibration parameters (Raychaudhuri, 2008).

Following steps are necessary for executing the Bayes theorem :

- Assume prior probability distribution for the calibration parameters
- Prior samples are generated by random sampling from the prior distribution
- Prior samples are evaluated in the model
- Model results are compared to the measurement values through a likelihood function
- Posterior probability distribution is generated from the prior distribution, likelihood function and Bayesian model evidence using equation (2.12).

solving (2.12) with Monte-Carlo and repeated random sampling becomes inconvenient and time-consuming for expensive models. In order to get rid of the process of running the expensive model rigorously, the concept of a surrogate-assisted model is applied. Accommodating the surrogate-assisted Bayesian inversion techniques save enough computational time and the calibration process becomes feasible (Espinoza, 2021).

### **Likelihood function**

The likelihood function aims to quantify the discrepancy between the modelled results and the measurement values. In the Equation 2.13, the term  $(2\pi)^{-\frac{n}{2}} \cdot \det(R)^{-\frac{1}{2}}$  is a nor-

malizing factor to make sure that the Probability density function integrates to one. the similarity between the model and the measured values are accounted by the term  $\exp(-.5(Z_{meas} - Z_{mod})^T R^{-1}(Z_{meas} - Z_{mod}))$ . If the model results and the measured values are similar, the exponential term value is 1 and on the other hand, the value is 0 if the difference is extremely high.  $R$  term denotes the measurement error and takes a diagonal matrix form.

$$R = \begin{bmatrix} e_1^2 & 0 & 0 \\ 0 & e_2^2 & 0 \\ 0 & 0 & e_3^2 \end{bmatrix} \quad (2.14)$$

In the above  $3 * 3$  matrix,  $e_i$  represents the measurement error. Influence of the difference between the model and the measured values decreases with significant measurement error as the measurement values are not plausible (Espinoza, 2021).

### Bayesian model evidence

The probability of the measured data is termed as Bayesian model evidence. BME is calculated by arranging the Bayes theorem and integrating over parameter space. The following Equation 2.15 shows the final stage of the integration procedure :

$$p(Z_{meas}) = \int_{\Omega} p(Z_{meas} | w) p(w) dw \quad (2.15)$$

The equation represents the expected value and the likelihood  $p(Z_{meas} | w)$  is supported by the prior probability  $p(w)$  as weight. A big distribution sample tend to have average value resembling the expected one. Using this concept, BME is approximated by the following equation :

$$BME = \int_{\Omega} p(Z_{meas}|w) p(w) dw \approx \frac{1}{N_{prior}} \sum_{i=1}^{N_{prior}} p(Z_{meas}|w_i) \quad (2.16)$$

BME is the quantification of how good the model can reproduce the measurement data with the provided prior distribution. Prior sample size plays a crucial role in determining a credible BME value (Espinoza, 2021).

### Surrogate-assisted Bayesian inversion

The basic idea of a surrogate model can be boiled down to a regression problem. Using Calibration parameter combinations and model results as input data, a surrogate model can resemble the full-complexity model and produce results in a similar manner. This strategy saves the computational time of running the model several times with multiple combinations of calibrated parameters. The input calibration parameters and model

results generate a response surface, which is represented by one or many axes. The input parameter combinations and corresponding result data are termed training data and the data used to validate the accuracy of the model are known as test data in the language of machine learning (Espinoza, 2021).

Gaussian process regression is a non-linear regression that uses Gaussian functions and weightage factors to adjust the input and output points. GPR comes in the process of building the surrogate model. As GPR requires fewer points to converge, it is suitable for computationally expensive models. Using basis functions and non-linear system of equations, GPR performs regression to interpolate in parameter space(Rasmussen & Williams, 2006).

The formal definition of GPR explores the idea of GP as a random variable, which has joint Gaussian distribution. Gaussian distribution can be defined as the following notation :

$$f(x) \sim GP(M, K) \quad (2.17)$$

where  $M$  is the mean and  $K$  is the Covariance matrix of the Gaussian distribution.

The covariance matrix in the process can be updated to accommodate newly added test data points. A reduced formulation of the covariance matrix is presented in the Matrix 2.18, where  $K(x, x)$  is the covariance matrix between training data,  $K(x, x_*)$  and  $K(x_*, x)$  are the covariance matrix between training and test data and  $K(x_*, x_*)$  is the covariance matrix within test data. In the matrix,  $x$  symbolises training data and  $x_*$  represents test data.

$$\begin{bmatrix} K(x, x) & K(x, x_*) \\ K(x_*, x) & K(x_*, x_*) \end{bmatrix} \quad (2.18)$$

The output prediction value for the test data points can be calculated using conditional probability and can be expressed as follows:

$$y_* | X_*, X, y \sim N(K(x_*, x) \cdot K(x, x)^{-1} \cdot y, K(x_*, x_*) - K(x_*, x) \cdot K(x, x)^{-1} \cdot K(x, x_*)) \quad (2.19)$$

Equation (2.19) indicates the value of  $y_*$ , the output space of the test data when the information of  $X, X_*, y$  are given. The equation is based on the principle that the conditional probability of a multivariate Gaussian is a Gaussian distribution with the term  $K(x_*, x) \cdot K(x, x)^{-1} \cdot y$  as mean and  $K(x_*, x_*) - K(x_*, x) \cdot K(x, x)^{-1} \cdot K(x, x_*)$  as variance. The Equation 2.19 is valid for multi-dimensional parameters as well (Espinoza, 2021).

Hundreds of thousands of prior calibration samples can be tested in the surrogate model in a fraction of a second. To incorporate surrogate-assisted Bayesian inversion with

Bayesian inference, model is run several times with various calibration parameter combinations and associated results are evaluated as training data for the GPR. GPR generates a response surface that is capable of approximating the full-complexity numerical model results.

### Gaussian process regression

Gaussian process regression is similar to the concept of non-linear regression. The only difference is the use of Gaussian function instead of a polynomial basis. Weight factors are used to adjust the Gaussian function to the available points. Figure 2.1 shows the fitting of a Gaussian function to a few random points.

While computing basis function at each points, the function influences all the other nearby points. The following matrix represents the influence of the polynomial basis function :

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{bmatrix} \cdot \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \quad (2.20)$$

The Matrix 2.20 shows the polynomial equation of  $y = \theta_0 + \theta_1.x + \theta_2.x^2 + \theta_3.x^3$  with  $\theta$  as the weight factor,  $x$  denotes the data and  $y$  is the function value.

The influence of Gaussian function increases to the points that are nearby and as the distance between the points increases, the influence of the Gaussian function decreases.

To execute a regression with GP, a covariance matrix is defined. In the covariance matrix, a chosen Gaussian base function is calculated at each training data point and each function value will represent the influence of data points in pairs (Espinoza, 2021). The concept can be comprehended by the following matrix :

$$\begin{bmatrix} f(x_1, x_1) & f(x_1, x_2) & f(x_1, x_3) \\ f(x_2, x_1) & f(x_2, x_2) & f(x_2, x_3) \\ f(x_3, x_1) & f(x_3, x_2) & f(x_3, x_3) \end{bmatrix} \quad (2.21)$$

In the matrix,  $f(x_1, x_2)$  is a Gaussian base function that represents the influence of point  $x_1$  on  $x_2$ . The same concept applies to the other members of the matrix. The Matrix (2.21) is a co-variance matrix with 3 training data points.

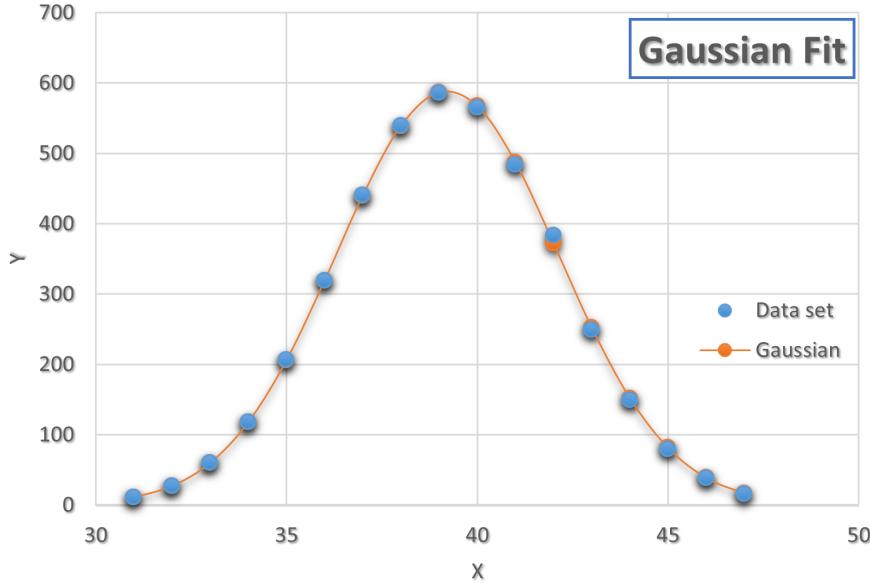


Figure 2.1: Gaussian function.

### Bayesian active learning

The accuracy and efficiency of a surrogate model depend on the equivalency between the numerical model and the surrogate-assisted model. BAL method improves the predictive capability of the surrogate-assisted model(Oladyskin, Mohammadi, Kroeker, & Nowak, 2020). Oladyskin et al. (2020) Tested various information theory scores and found that relative entropy is the best component in terms of quantifying the difference between two distributions. BAL methodology compares the relative entropy between prior and posterior distribution using the following Equation :

$$D_{KL}(p(w|Z_{meas}), p(w)) = \int_{\Omega} \ln \left[ \frac{p(w|Z_{meas})}{p(w)} \right] p(w|Z_{meas}) dw \quad (2.22)$$

where  $p(w)$  is prior distribution,  $p(w|Z_{meas})$  is posterior distribution and  $D_{KL}$  denotes the Kullback-Leibler divergence which quantifies the scores of information theory.

In Bayesian context, relative entropy is represented as the information gained while moving from prior to posterior(Oladyskin & Nowak, 2019). According to Oladyskin and Nowak (2019), Equation (2.22) is unfeasible in many applications to solve as it requires multidimensional integration. Instead, prior, and posterior sample techniques should be used for solving the Equation (2.22) and calculating the information theory score. The following equation is suggested to calculate the relative entropy :

$$D_{KL} \approx -\ln \left[ \frac{1}{N_{prior}} \sum_{i=1}^{N_{prior}} P(Z_{meas}|w_i) \right] + \frac{1}{N_{post}} \sum_{i=1}^{N_{post}} \ln [P(Z_{meas}|w_i)] \quad (2.23)$$

In the equation  $N_{prior}$  and  $N_{post}$  are number of prior and posterior samples. The second term in the equation signifies the gained information from prior to likelihood and third term denotes the gained information from likelihood to posterior. The sum of both the term provides the reduction in uncertainty from prior to posterior.

The number of training points signifies the accuracy of the surrogate-assisted model. With more training points, the predictive capacity of the surrogate model increases. BAL method localizes the optimum regions in the parameter domain and based on maximum relative entropy value, new optimum posterior values are selected. Newly selected parameter values will add to the list of training data and will reinforce the response surface of the surrogate model and reduce the number of points necessary for convergence. This way BAL method improves the surrogate model iteratively. Figure 2.2 guides through the workflow of surrogate-assisted Bayesian inversion.

### Entropy

Definition of entropy is based on its application. In the context of the study, entropy is the expected value of uncertainty related to random variables. Inverse of a probability is uncertainty. With increasing probability, uncertainty gradually decreases. uncertainty can be defined by the Equation 2.24 :

$$U = \ln\left(\frac{1}{P(x)}\right) \quad (2.24)$$

Entropy can also be computed as the expected value of the uncertainty. Expected value of a data set can be expressed as :

$$E(x) = \int x p(x) dx \quad (2.25)$$

Where  $x$  is the data value,  $p(x)$  is the probability density function and  $E(x)$  is the expected value. Combining Equation 2.24 and 2.25, entropy is quantified as :

$$H(x) = \sum_i \ln\left(\frac{1}{P(x_i)}\right) P(x_i) = -\sum_i [P(x_i)] P(x_i) \quad (2.26)$$

As defined by Shannon (1948), the continuous definition of the entropy is as follows :

$$H(x) = \int \ln\left(\frac{1}{P(x)}\right) P(x) dx = - \int \ln[P(x)] P(x) dx \quad (2.27)$$

### Relative entropy

Relative entropy is a information theory score and usually termed as Kullback-Leibler divergence. As Equation 2.22 mentioned in the previous section is unsuitable for practical

applications, Oladyshkin and Nowak (2019) avoid the multidimensional approach. The authors apply natural logarithm to the arranged Bayes theorem :

$$\ln \left[ \frac{p(w|Z_{meas})}{p(w)} \right] = \ln \left[ \frac{p(Z_{meas}|w)}{p(Z_{meas})} \right] \quad (2.28)$$

Both side of the Equation 2.28 is multiplied by the posterior distribution and integration is done over the parameter space :

$$\int_{\Omega} \ln \left[ \frac{p(w|Z_{meas})}{p(w)} \right] p(w|Z_{meas}) dw = \int_{\Omega} \ln \left[ \frac{p(Z_{meas}|w)}{p(Z_{meas})} \right] p(w|Z_{meas}) dw \quad (2.29)$$

Exploring the logarithmic properties in the right side of the Equation 2.29, following expression is achieved :

$$D_{KL} = \int_{\Omega} \ln [p(Z_{meas}|w)] p(w|Z_{meas}) dw - \int_{\Omega} \ln [p(Z_{meas})] p(w|Z_{meas}) dw \quad (2.30)$$

Inserting the expression for the BME, Equation 2.31 is generated :

$$D_{KL} = -\ln [BME] + \int_{\Omega} \ln [p(Z_{meas}|w)] p(w|Z_{meas}) dw \quad (2.31)$$

The authors refer to the second term as non-normalized cross-entropy.

By substituting Equation 2.16 in Equation 2.31, Equation 2.23 is attained, which is the final expression of relative entropy.

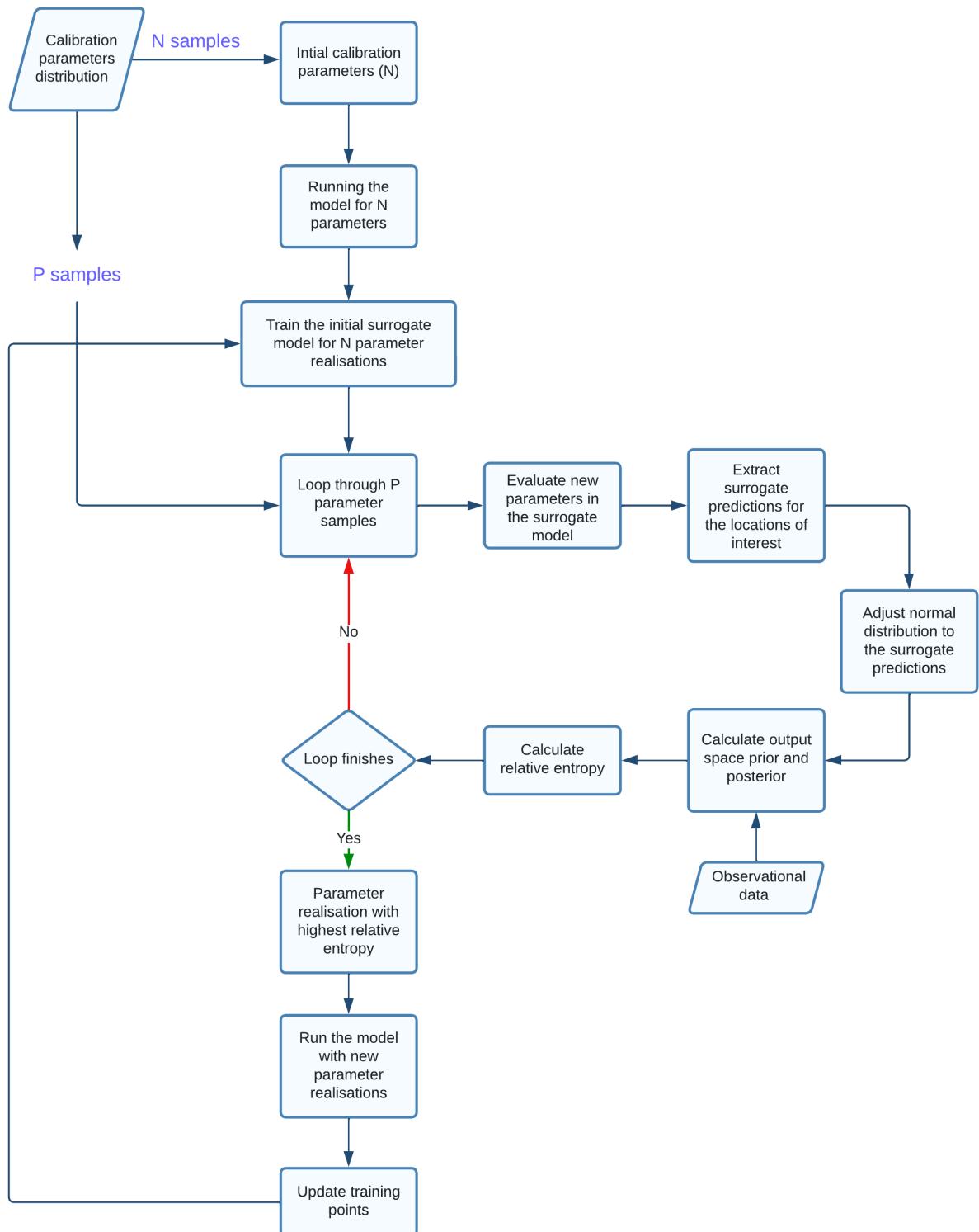


Figure 2.2: Workflow for Surrogate-assisted Bayesian inversion.

### 2.2.2 Calibration parameters of hydrodynamic models

The discrepancy between the model results and observation data emanates from the model uncertainties. Boundary conditions, topography, discretisation methods, time steps, and computational error are the main sources of uncertainty in a numerical model. sometimes the governing formulas failed replicate the complex physical process of the river flow and their interaction. Simplifying terms and conditions, various assumptions along the way contribute to the uncertainty. Improving the accuracy of the model involves tuning the model input parameters in a systematic way while comparing with measurement data. The restraining factor in the process is the availability of the measurement data. Spatial and temporal variability in the measurement data will improve the confidence interval of the calibration. Variation of model output concerning variation of the input parameter provides an understanding of the model sensitivity towards the specific input parameters. In 2D hydrodynamical model, following parameters are used for calibration:

- Friction coefficient
- Numerical parameters
- Model initialization
- Turbulence model and parameters

### 2.2.3 Validation approaches for hydrodynamic models

Validation of a numerical model assess the accuracy or goodness of a model without adapting the model parameters. This is the fundamental difference between the calibration and the validation process. Calibration process amend the model parameters based on reliable data. The main aim of the validation process is to verify whether the model predictions represents the observed behaviour of the hydraulic variables. According to Landry, Malouin, and Oral (1983), model validation is not only the representativeness of the model in the real field but the usefulness, effectiveness, time considerations are also prominent attributes.

Model validation step is carried after the calibration process. Validation is achieved by the observational points within the same or different time period than the period it is calibrated for. However, the observational points which are subjected for validation has to be different to the observational points analysed for the calibration. The general objectives of the validation process is decided by the modeler beforehand. Based on the quality of the data, criteria of validation varies. Validation process is carry forward to achieve the intended task for the specific model. The same model can not be counted valid for other tasks.

## Evaluation methods

Model evaluation methods involves comparing the model predictions with observational data through visualisation and/or statistical approach. The first step in the evaluation process is the visualisation analysis. Visualisation provides the modeler a general overview of the model performance. In this domain, analysing hydrograph, exceedance probability curves and colour gradient plots of the maps are good practice. The limitation surrounding these approaches is the subjectivity of the modeler. The similar visuals can be interpreted differently based on modelers skills.

The lack of automated approach in the visualisation has culminated more significance on the statistical methods. In the statistical approach, using mathematical functions and formulations model performance is quantified. Model validation statistical tools are able to quantify the accuracy and biasing of the model as well.

According to Sutherland, Peet, and Soulsby (2004), for accuracy purpose, mean absolute error (MAE), mean square error (MSE) are common measures. MAE and MSE are defined in Equation 2.32 and 2.33:

$$MAE(Y, X) = \frac{1}{j} \sum_{j=1}^j |y_j - x_j| = \langle |Y - X| \rangle \quad (2.32)$$

$$MSE(Y, X) = \frac{1}{j} \sum_{j=1}^j (y_j - x_j)^2 = \langle (Y - X)^2 \rangle \quad (2.33)$$

In the above expressions,  $Y$  and  $X$  are the two dataset.  $j$  is the number of data in both the sets.  $\langle \rangle$  denotes the mean.

RMSE is the square root of MSE. Due to unit similarity with the measured values, RMSE and MAE is preferred in the evaluation process than MSE. The Equation 2.32 and 2.33 can be extended by inclusion of weight factors. The choice between two evaluation measures is considered based on importance of the outliers. The outliers are the points in which large differences are observed between the measured and the predicted values. As the difference between the dataset are squared in RMSE, it is more sensitive to outliers than MAE. This makes the RMSE a conservative predictor of error. Difference of RMSE and MAE depends on the shape of the  $Y$  and  $X$  distribution. Modelling exercise and evaluation can decide the importance of outliers and influence the choice between RMSE and MAE. Both the evaluation methods consider no error in the measurement data. This assumption is not valid in practical cases.

Linear correlation, coefficient of determination are well known methods of determining correlation between two series.

Linear correlation is a well-known method of determining linear relation between two series. The expression is as follows:

$$r_{XY} = \frac{\langle (Y - \langle Y \rangle)(X - \langle X \rangle) \rangle}{\sigma_X \sigma_Y} = \frac{s_{XY}}{\sigma_X \sigma_Y} \quad (2.34)$$

where  $\sigma_X$  and  $\sigma_Y$  are measured and prediction standard deviation,  $s_{XY}$  symbolises the covariance of the series. The value of  $r_{XY}$  ranges from [0,1] where 1 is the highest correlation and 0 is the lowest.

Pearson's correlation coefficient ( $r$ ) and coefficient of determination ( $R^2$ ) methods is capable of describing the collinearity between two series. The correlation coefficient range is [-1,1]. If  $r = 0$ , the relationship does not exist. Positive and negative linear relationship are addressed with  $r = 1$  and -1.  $R^2$  expound the variance proportion of the model, ranging from [0,1]. Higher  $R^2$  values represent less error variance and values greater than 0.5 are considered as acceptable. This statistical measure is over sensitive to outliers and insensitive to proportional and additive discrepancies between the model and the measured values (Moriasi et al., 2007). The Equation is shown below:

$$r = \frac{\sum_{j=1}^J (x_j - \langle X \rangle)(y_j - \langle Y \rangle)}{\sqrt{\sum_{j=1}^J (x_i - \langle X \rangle)^2 \sum_{j=1}^J (y_i - \langle Y \rangle)^2}} \quad (2.35)$$

where  $j$  is the number of data in the series.  $\langle \rangle$  signifies the mean and  $X$  is the observed data set and  $Y$  is the model predicted data set.

Model and data uncertainties are not accounted in the statistical evaluation of a numerical model. Map comparison based on fuzzy theory can enhance the efficiency of the evaluating process. The ability of capturing similarity between the prediction and the observational behaviour while considering uncertainties can realistically evaluate a numerical model (B. Negreiros, Schwindt, Haun, & Wieprecht, 2022).

# **Chapter 3**

## **Materials and methods**

### **3.1 Study area**

This section address the specific area based on which the Lower Yuba river hydrodynamic model is prepared in Telemac-2D. The section states various hydrological and morphological features of the study area.

#### **3.1.1 The Lower Yuba River**

##### **Watershed**

The Yuba River is a Sacramento River's tributary. The Yuba River flows from the Sierra Nevada to the confluence with the Feather River. The river is located in North-Central California, United States of America(Hopkins & Pasternack, 2018).

The Yuba River is flowing from East to West passing through the forest, and agricultural lands and joining the Feather River in the end. The yuba river has three main branches North, Middle, and South branch as well as one dry and deer creek. The Yuba catchment is mostly consisting of Granite soils, such as glaciated granodiorites and the riverbed is mostly Gravel and Cobble (James, 2005). Yuba catchment encompasses an area of 3500 km<sup>2</sup>(Hopkins & Pasternack, 2018).

The LYR model domain is defined as the river segment between the Highway 20 bridge and Daguerre Point Dam that stretches 11.2 km. Daguerre Point Dam with two fish passes in North, South direction and an arch bridge named Parks Bar Bridge are prominent structures within the LYR. Figure 3.1 shows the Yuba River catchment and Figure 3.2 displays the Lower Yuba River section.

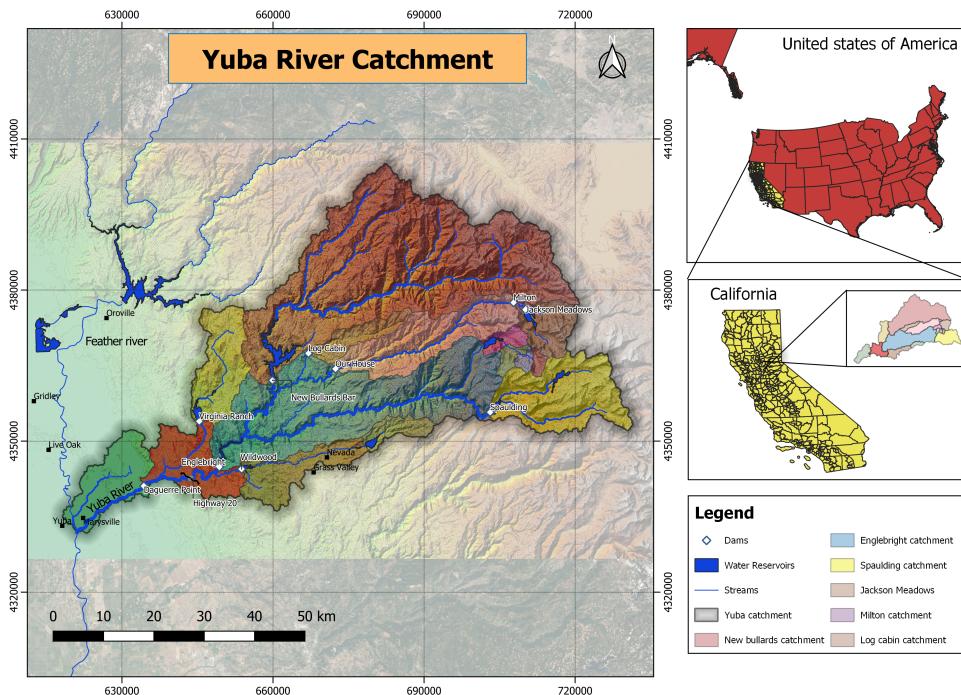


Figure 3.1: Yuba River catchment (EPSG: 32610).

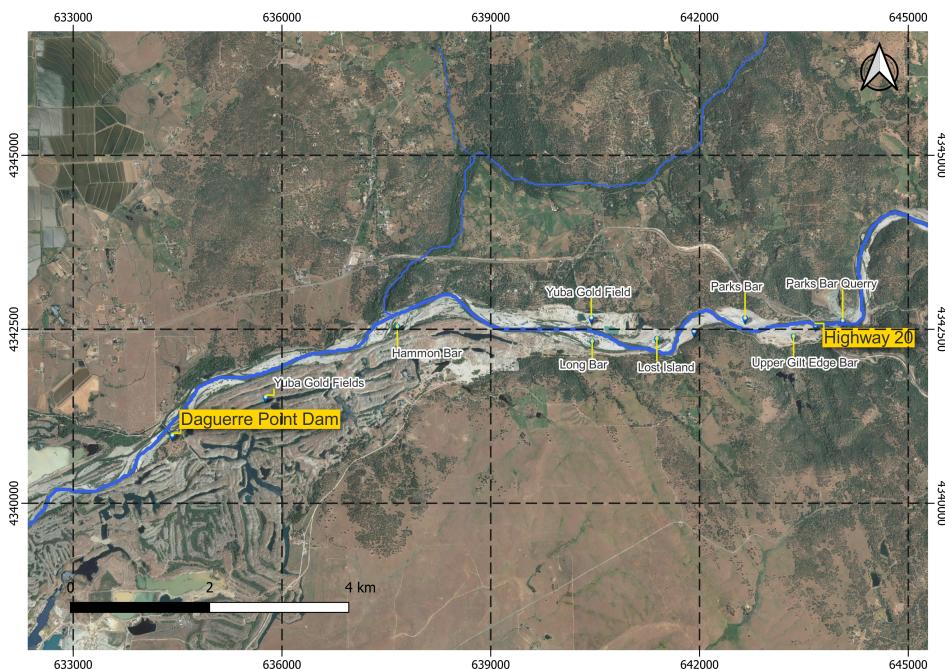


Figure 3.2: Lower Yuba River section (EPSG: 32610).

### Land use and cover

The Yuba River land use patterns are quite distributed and diverse. Dominating land coverage are mostly water, trees, grass, flooded vegetation, crops, scrub, built area, and bare ground. The abundance of forest and crops in the low land areas to scattered rural settlements and mining areas are located surrounding the river. Figure 3.3 exhibit the land use pattern in the Yuba River catchment.

Significant land coverage are trees representing more than 74 % of land area, consisting of dispersed clusters of plants with occasional exposure to rock/soil. Scrubs and build areas are also predominant in the LYR section covering area(Karra et al., 2021). Anthropogenic activities are facilitating the degradation process of riparian vegetation, affecting the aquatic species due to a lack of nutrients and shelter. This has culminated in a greater threat to migratory birds as well(US Army Corps of Engineers, 2019). Pie chart 3.4 shows the percentage of land use for different catagories in the Yuba catchment.

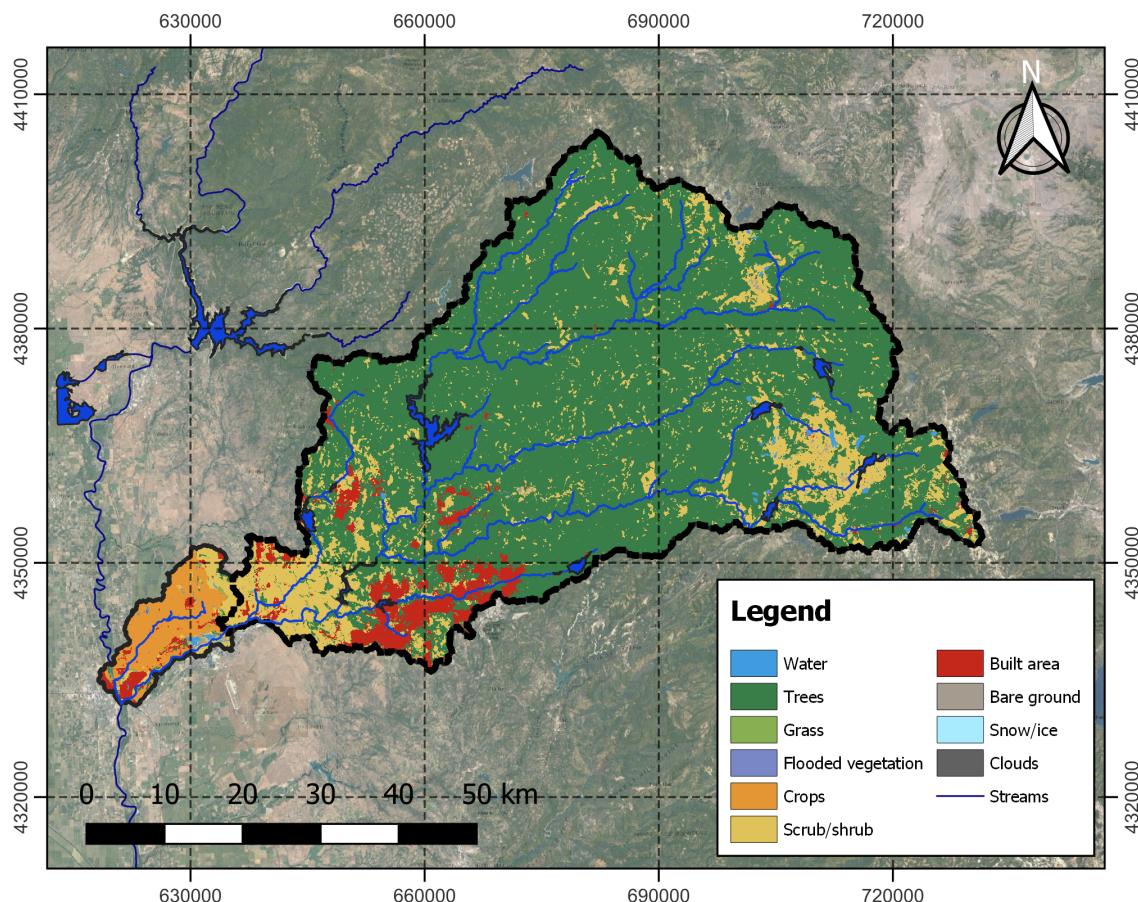


Figure 3.3: Yuba catchment land use (EPSG: 32610).

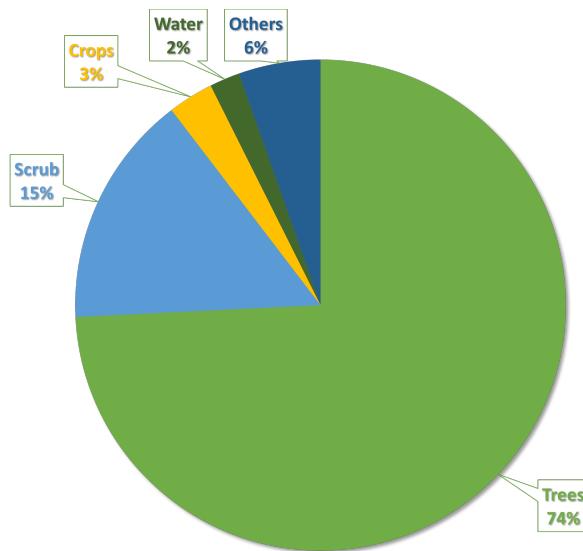


Figure 3.4: Yuba catchment land use percentage.

### Landscape pattern

The Cordillera of Sierra Nevada is the most prominent feature stretching more than 640 km to eastern California(US Army Corps of Engineers, 2019). In the LYR section, the elevation ranges from 31 m to 141 m. Drainage channels are mostly made of metamorphic rocks and igneous rocks. The catchment is developing through canyons along alluvial sections with steep, high ridges. sediment flux from mining activity buried historic terraces(Curtis, Flint, Alpers, Wright, & Snyder, 2006).

### Hydroclimate

The Mediterranean climate in the catchment is characterized by wet, cold winters and dry summers. 80 % of the precipitation occurs between October to April. Accumulation of snow is evident at higher elevations during winter. The occurrence of flooding can happen in Winter due to warm rain on snow. During April-June, warm temperature prevails in the region, and spring-runoff becomes the dominating process. Annual precipitation varies from Sierra Nevada (1500 mm) to Marysville (500 mm)(Hopkins & Pasternack, 2018). Major runoff is generated by snowmelt in the spring, and storms during late summer and early autumn. Peak flows can be as intense as more than three times the base flow. Constructed water reservoir structures control the intensity of the flood water(Curtis et al., 2006).

### Morphodynamical features

Morphological features are functions of the channel response to various evolutionary processes. The LYR section is a single-thread structure channel, and the section is entrenched. The ratio of width and depth ranges from moderate to high and the sinuosity is low. Sinuosity is a physical character of a channel, represented by the ratio of baseflow centreline length and valley corridor length. With a channel slope of 0.0016 and 97 mm substrate diameter, the LYR section is classified as a C3 stream(Wyrick & Pasternack, 2012). According to Rosgen and Silvey (1996), C3 type stream can be described as a slightly meandering, entrenched channel with a substantially developed floodplain. The C3 stream type is cobble-dominated, generally seen in glacier valleys with a u-shape. The composition of the bank is heterogeneous and unconsolidated alluvial materials which are finer than the bed material. Channels are vulnerable to the erosion of the bank.

Sediment supply is high if the erodibility condition of the bank is high. While comparing different C3-type rivers all around the world, it is found that the LYR exhibit similar characteristics to other C3-type rivers. Classification of a heterogenous river is a complicated task and Parker (1976) derived a quantitative classification procedure to recognize channel patterns. His methodology analyses the ratio of slope and Froude number along with the inclusion of width and depth. With an average Froude number of 0.26, the LYR section can be classified as Being in a transitional stage between straight and meandering.

The LYR geo-morphic characteristics are transitioning from upstream to downstream. Near the upstream, channel bed slope is high, and reach is short with an increase in valley width and baseflow width. At the confluence with deer creek, the channel becomes narrower. Surpassing the deer creek, the valley becomes wider and forms noticeable meandering. Increasing channel width continues and enables wider floodplains and terraces. In the flowing path, with decreasing slope few braided sections are also created with multiple paths of flow. The end of the downstream is characterized by a meandering channel with several floodplains, paths, and ponds(Wyrick & Pasternack, 2012).

### Sediment characteristics

In different parts of the LYR section, sediment characteristics vary. Grain size distribution changes as the river flow from upstream to downstream. Mean grain size decreases towards the downstream. At the advent of the LYR section, Daguerre Point dam is located as a sediment barrier. There are no river-rounded, but angular gravel-cobble sediments in this section. The same characteristics prevail from Deer creek as few cobbles bar can be seen. Section until Highway 20 is wider and provides room for a river meandering alluvial system. Surrounding Highway 20 bridge, a valley constriction exists. As a result of this

valley constriction, deposition of gravel occurs in a backwater zone, yielding a significant braided channel. As the river is flowing toward the Dry creek confluence, the valley width doubles and enables wider sedimentary bars(Wyrick & Pasternack, 2012). The mean grain size of vegetated areas is 81 mm and the bank full segment has a mean grain size of 97.4 mm (Wyrick, Gonzalez, & Pasternack, 2013). Evolution of grain size distribution in the LYR is shown in the Figure 3.5.

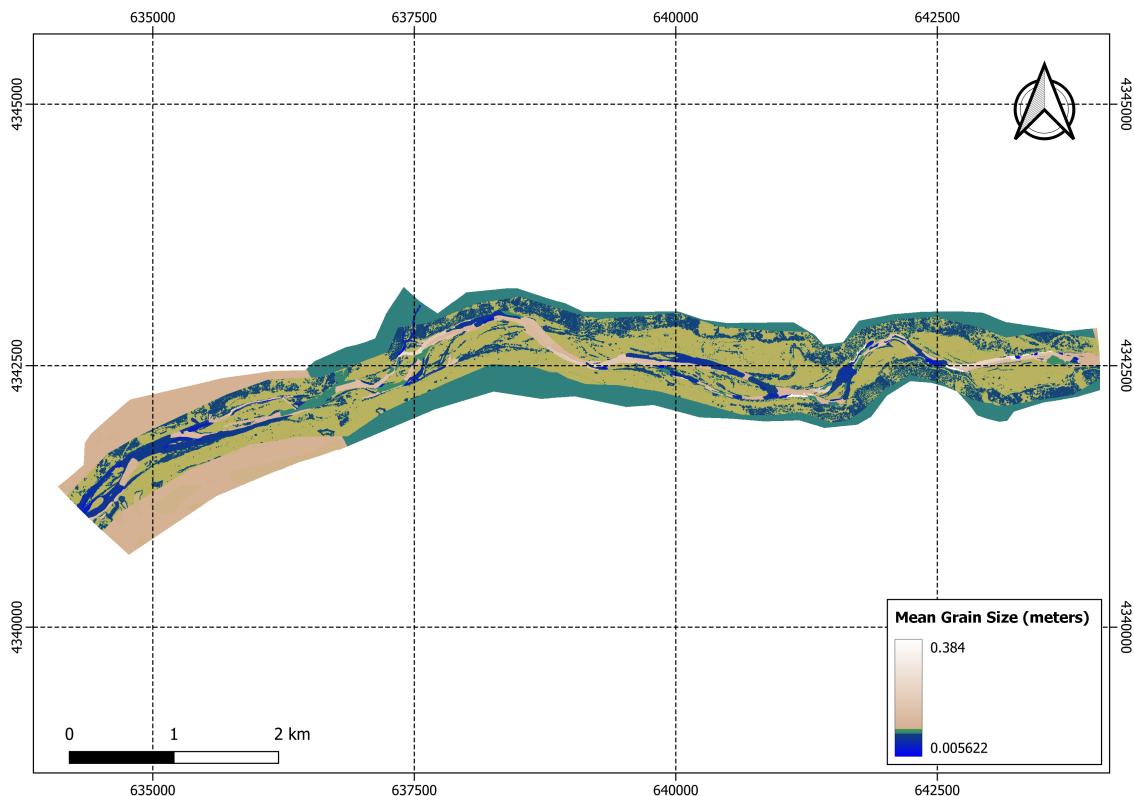


Figure 3.5: Lower Yuba River grain size distribution (EPSG: 32610).

## Hydrology

The Lower Yuba river watershed is controlled by different reservoirs which control water and trap sediments to various degrees. Prominent water control structures include the Daguerre dam and New Bullards reservoir. New Bullards bar reservoir is the biggest water reservoir in the Yuba watershed. The reservoir provides flows for recreation, environmental protection, irrigation, and flood control purpose. Prior to the season of precipitation, most of the flow in the LYR is regulated by New Bullards reservoir. In the basin, there is only limited provision for retaining snow-melt runoff. Released water from New Ballards reservoir passes downstream where most of the flow is released from hydroelectric power generation. Flow in the Yuba River is measured in two gauging stations. One at Smartsville

Gage, United states geological Survey Station 11418000 and another at Marysville Gage, United States Geological Survey station 11421500. Average annual discharge of 73.62 m<sup>3</sup>/s is indicated from these gaging stations. As a result of the velocity gauging campaign in December 2009 and January 2010, inflow hydrograph data was collected. Repeating discharge values in the interval of seconds enable the hydrodynamic simulation to achieve a steady state. The inflow hydrograph is shown in the Figure 3.6.

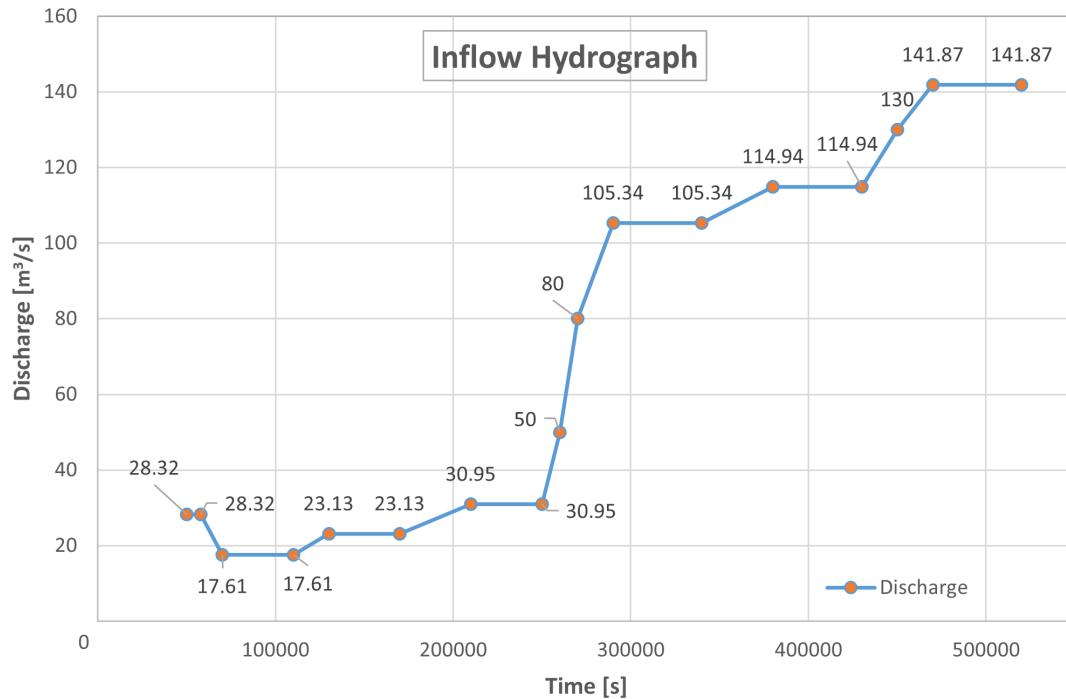


Figure 3.6: Inflow hydrograph.

Based on a 100-year simulation developed by US Army Corps, flow probabilities for 10, 25, 50, 75, and 100 years are determined (US Army Corps of Engineers, 2010). The Table 3.1 contains the probability discharge values for the respective years.

Table 3.1: Flow probabilities

Flood event	Discharge [m <sup>3</sup> /s]
1 in 10 years	651.29
1 in 25 years	1444.16
1 in 50 years	2406.93
1 in 75 years	3228.12
1 in 100 years	3822.77

In the LYR, bank full discharge is reached once every 1.25 years with an 82 % probability of filling the riverbank channel every year. Floodplain-filled flow has a return period of 2.5 years with nearly 40 % exceedance probability annually. The field determined bank full discharge in the LYR is  $141.6 \text{ m}^3/\text{s}$  and floodplain-filling flow is  $597.5 \text{ m}^3/\text{s}$  (Wyrick & Pasternack, 2012). From the stage-discharge Figure 3.7, it is evident that small changes in the height correspond to larger transitions in the river discharge. This phenomenon suggests that increased discharge is more likely to create new flow paths as well as wetted areas.

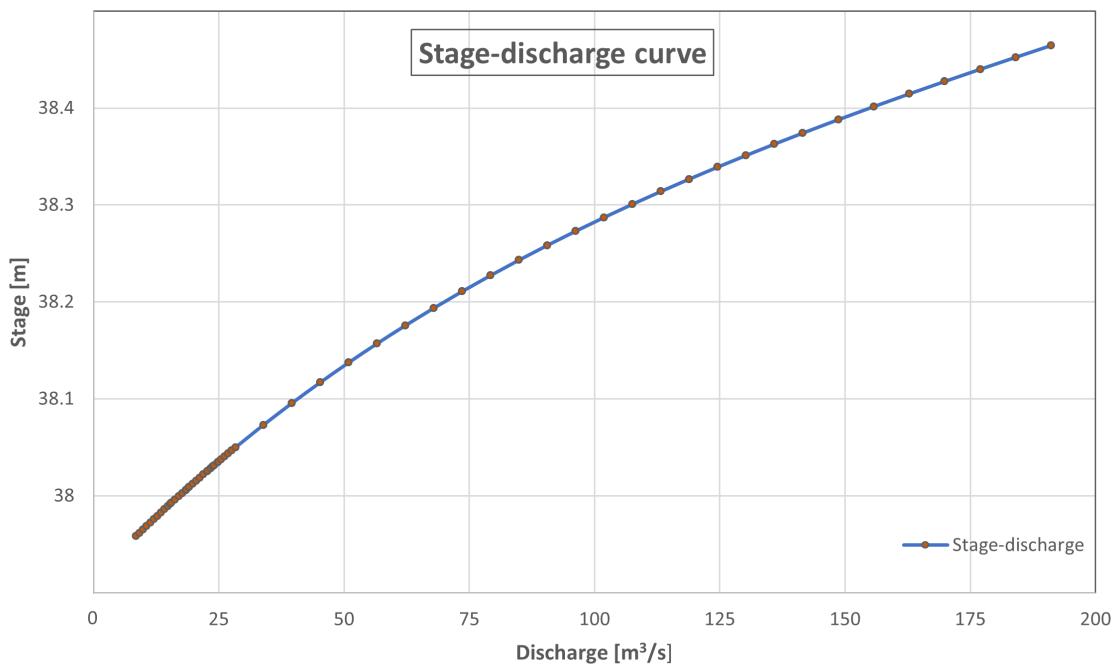


Figure 3.7: Stage-discharge curve.

### 3.1.2 Existing models

Yuba river rehabilitation is a concern for many agencies. Yuba County Water Agency (YCWA) is developing comprehensive spatial datasets to facilitate the analysis of the Yuba River characteristics. The first comprehensive and high-resolution 2D model of the LYR was developed in 2008. Pasternack, Tu, and Wyrick (2014) expounded the 2008 model meticulously. 2008 model was thoroughly validated and cited in many published papers. The 2008 model was used in many application purposes like river development, fisheries restoration, pre-feasibility study, ecosystem restoration, irrigation, and many more. In 2014, the whole LYR section has been mapped with utmost precision for ecological and geo-morphological functions. A new 2D hydrodynamic model was developed using TUFLOW GPU covering the LYR with a 3\*3 computational grid. The new model allows for predicting velocity, water depth, shear stress, and associated other variables. Using emerging technologies in the model development field, the model was calibrated and validated extensively using map and other associated data. 2014 model encompasses the LYR section from Englebright dam to the confluence of Feather River. For the model, topography and bathymetry data were collected across the LYR. MB sonar data, and RTK-GPS data were collected for the purpose of higher resolution. Rating curves were obtained from Smartsville and Marysville gaging stations. Rating curves were extrapolated in order to run the 2D model lower and higher than the observed flow rate. To validate the model after simulation, water surface elevation, depth elevation, and velocity vector validation measurements were carried out on special discharge rates. The advantage of the 2014 model above the 2008 model is the use of TUFLOW instead of SRH-2D. TUFLOW-GPU computation is substantially faster than SRH-2D in the same resolution. Post-processing steps were also reduced(Hopkins & Pasternack, 2018). Special attention was paid to friction parameterisation. Calibration was done with global Manning's value and spatially distributing Manning's values. For surface roughness characterization, different approaches were tried for vegetated and unvegetated terrains. In the roughness allocation, a base set of values for Manning's were selected, and based on the yielded results, friction values were adjusted.

As the model over predicted lower velocities, which is a common occurrence with 2D models, 2014 model adopted bank slope parameter inclusion. The idea is that the flow within the channel is impacted by the bank slope and bed. With increasing bed slope, systematic roughness was added to achieve desirable outcome. Another approach was to include spatially distributed roughness using vegetated depth raster and LIDAR point heights and substituting the base Manning's values in locations where the anticipated Manning's is higher than the base values. Bank slope and spatially distributed method yielded improved results in terms of simulating lower velocities. In 2014 LYR model

spatially distributed method was incorporated. It was concluded after comparing the results with observational data that the model exceeded the expected performance of a 2D model(Hopkins & Pasternack, 2018).

Following table highlights the specifications of the TUFLOW model.

Table 3.2: Tuflow model specifications

<b>Model parameter</b>	<b>Specifications</b>
Projection	NAD_1983_2011_StatePlane_California_II-_FIPS_0402_Ft_US
Units	US Customary (ft, cfs)
Mesh Resolution/Cell Size	3'* for flows < 10,000 cfs 3' for flows > 10,000 cfs
Cell Wet/Dry Depth	0.3'
Timestep	Adaptive timestepping
Eddy Viscosity Formulation	Smagorinsky
Eddy Viscosity Coefficient	0.5
Manning's n < 1000 cfs	Spatially distributed surface roughness
Manning's n > 1000 cfs	Spatially distributed vegetation roughness

In 2021, Hydrodynamic 2D model is developed for the LYR using Telemac-2D. Water depths and flow velocities were calculated with a 2-meter resolution channel mesh structure, produced in the software BlueKenu. The 2021 model explored two Manning friction schemes with grain size  $D_{84}$ . For the model,  $D_{mean}$  data was available from which  $D_{84}$  values were calculated using the Formula :

$$D_{mean} = D_{84}/2.5 \quad (3.1)$$

As proposed by Rickenmann and Recking (2011), Manning's coefficient was calculated from  $D_{84}$  grain size :

$$1/n = 20.4/D_{84}^{0.167} \quad (3.2)$$

In the above equation  $n$  is the Manning's coefficient. Explicit spatial roughness zones are implemented based on Manning's  $n$ . Calibration was executed according to the trial-and-error method. Two simulation conditions are analysed: a steady state and a quasi steady state condition employing measurement data and TUFLOW results. Telemac-2D model yielded similar results to the TUFLOW model.

## 3.2 Adaptation of the hydrodynamic model

This section meticulously explains the process of adapting the Lower Yuba River hydrodynamical model to accommodate the new friction law as well as the categorisation of the roughness zones.

### 3.2.1 Mesh construction

To generate a suitable mesh for the LYR model, the unstructured mesh is constructed using the software BlueKenu. Unstructured mesh is a combination of triangular facets of different size and shapes. Various sizes of the triangular facets allow the adequate representation of the complex topography of the LYR section. 3 mesh resolutions are constructed:

- 2-meter mesh resolution in the river channel and dry creek
- 5-meter mesh resolution in the high discharge flow areas
- 30-meter mesh resolution in the floodplains

The LYR mesh section consists of 817485 nodes and 1634026 elements. Mesh resolution in different locations of the model domain is shown in the Figure 3.8.

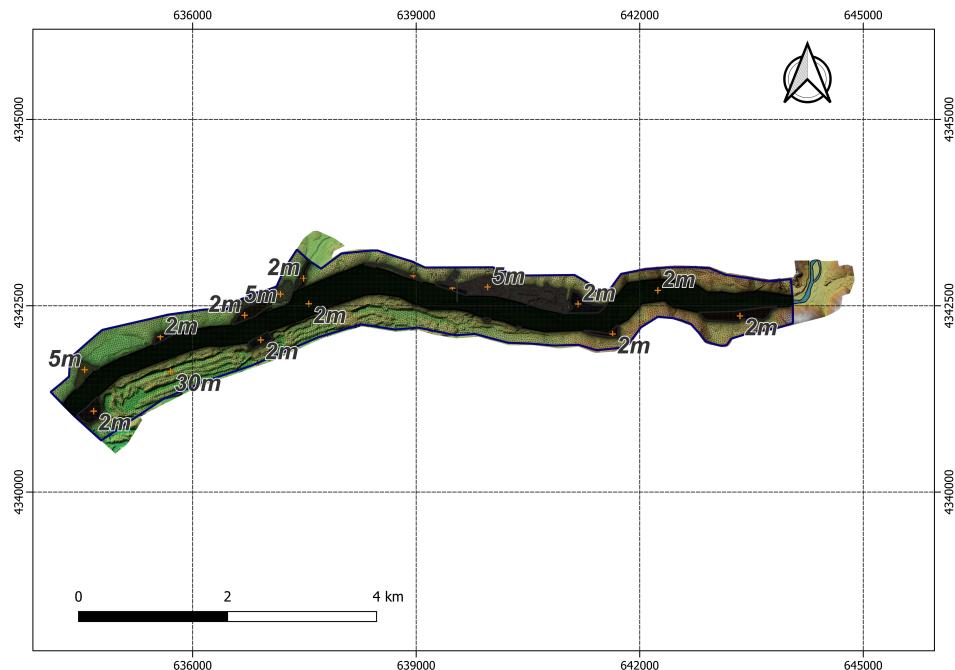


Figure 3.8: Lower Yuba River mesh resolution (EPSG: 32610).

After the mesh generation, LYR mesh is incorporated with DEM data. DEM is the elevation data for the LYR section. Following procedures are followed for incorporating DEM data with the LYR mesh :

- DEM data is exported to BlueKenu in a (.xyz) point cloud file
- A new variable BOTTOM is created and linked to the mesh
- Individual point values in the (.xyz) file are interpolated for the whole LYR domain using 2D interpolator tool in BlueKenu
- A selafin (.slf) file is generated with model geometry and BOTTOM variable with DEM data

Following the mentioned procedures, Figure 3.9 is generated in the BlueKenu.

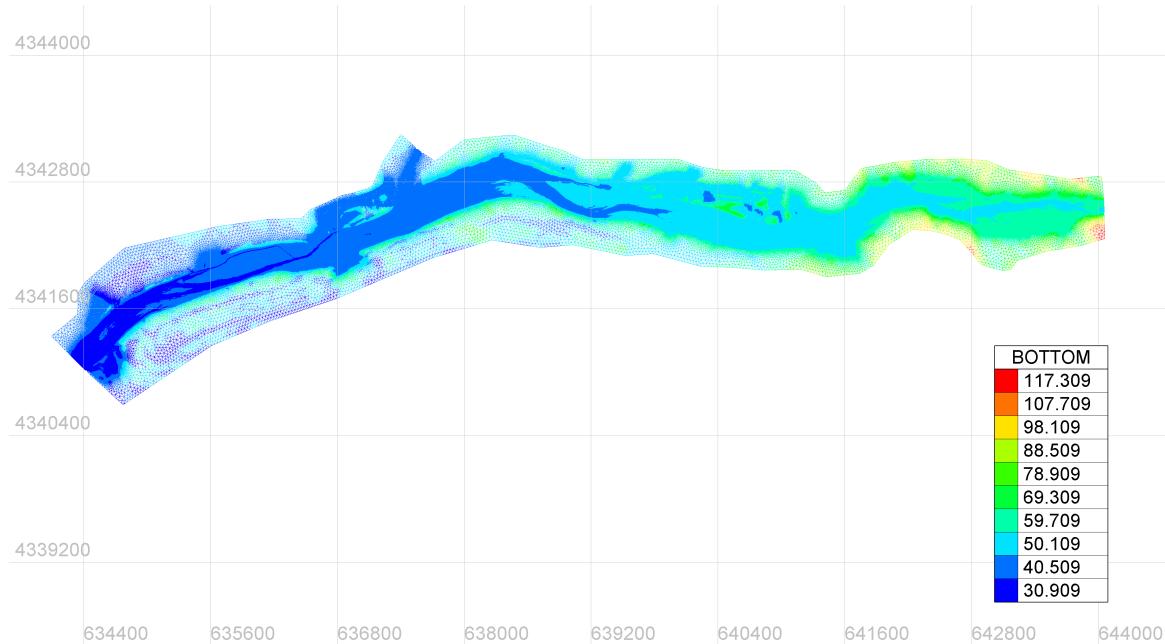


Figure 3.9: Lower Yuba River mesh with BOTTOM variable (EPSG: 32610).

### 3.2.2 Roughness zones

Spatially explicit roughness zones are demarcated in the LYR 2D model. In order to delineate the LYR extent, various zones are created based on similarity of characteristics. Three major categories of division are:

- Channel bed
- Floodplain
- Vegetation

Due to homogeneous features across the floodplain and vegetation, both the categories contain one roughness zone only. The channel bed is divided into 246 roughness zones throughout the LYR section based on:

- Topographical features
- Sediment characteristics
- Patches of bedrock
- Vegetation bank
- Presence of obstruction elements

Friction zones are implemented as Friction ID in Telemac-2D and the limit of Friction ID that can be inserted is 80. Therefore, friction zones are categorised based on  $D_{84}$  values.  $D_{84}$  grain sizes are calculated from  $D_{mean}$  using the Equation 3.1.

After Segregating and re-organizing the friction zones, 246 explicit friction zones are reduced to 15 friction zones and each zones are assigned a Friction ID. Table 3.3 illustrate the Friction ID and corresponding  $D_{84}$  values. Vegetation and floodplain are represented as Friction ID 14 and 15 consecutively. To assign separate calibration parameter pair for different Friction ID, Frictions ID are divided into 4 groups. Firstly, all the 4 groups are assigned serially to the Friction ID. Figures 3.10,3.11 and 3.12 shows the boundary for the LYR channel, vegetation, and floodplain.

Table 3.3: Friction ID grouping

Friction ID	$D_{84}[m]$	Group
1	0.013	B
2	0.018	A
3	0.038	C
4	0.053	D
5	0.077	B
6	0.111	A
7	0.162	C
8	0.220	D
9	0.298	B
10	0.401	A
11	0.436	C
12	0.660	D
13	0.564	B
14	0.202	A
15	0.240	C

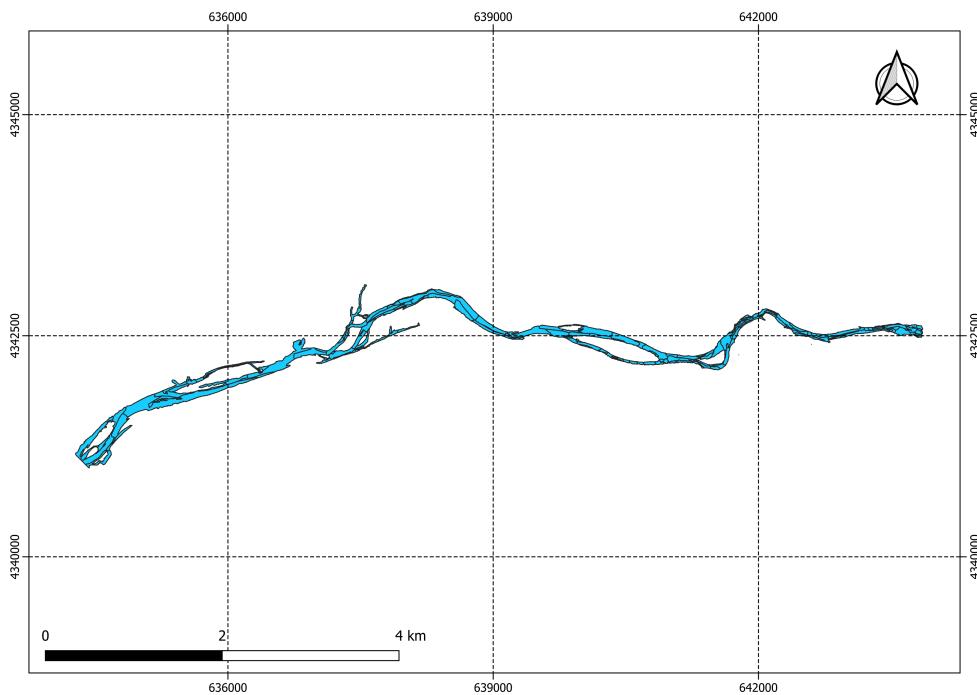


Figure 3.10: Lower Yuba River channel (EPSG: 32610).

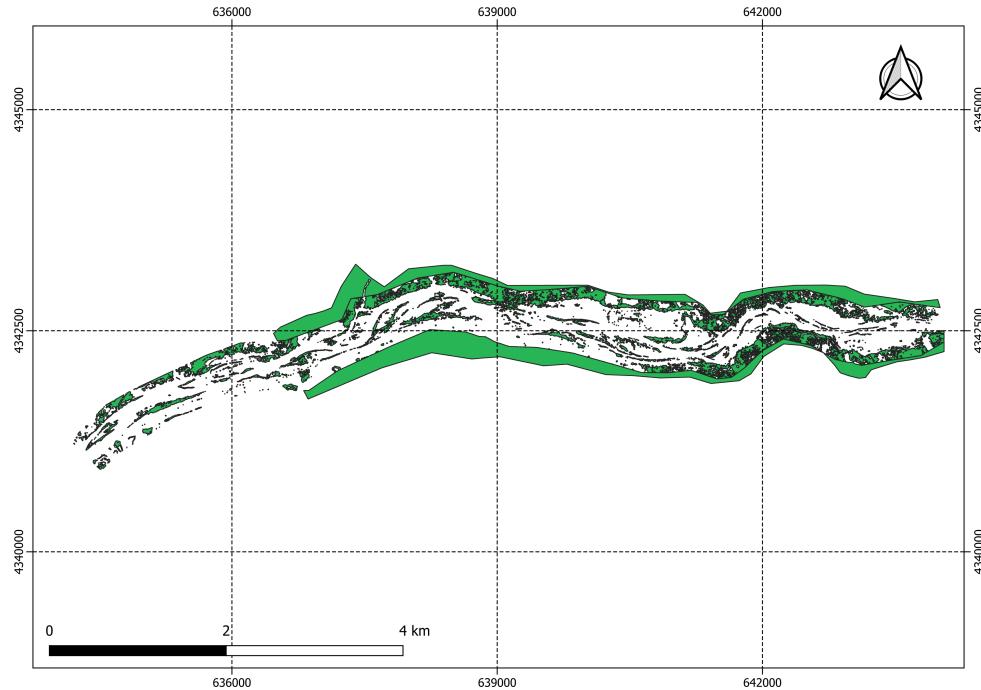


Figure 3.11: Lower Yuba River vegetation (EPSG: 32610).

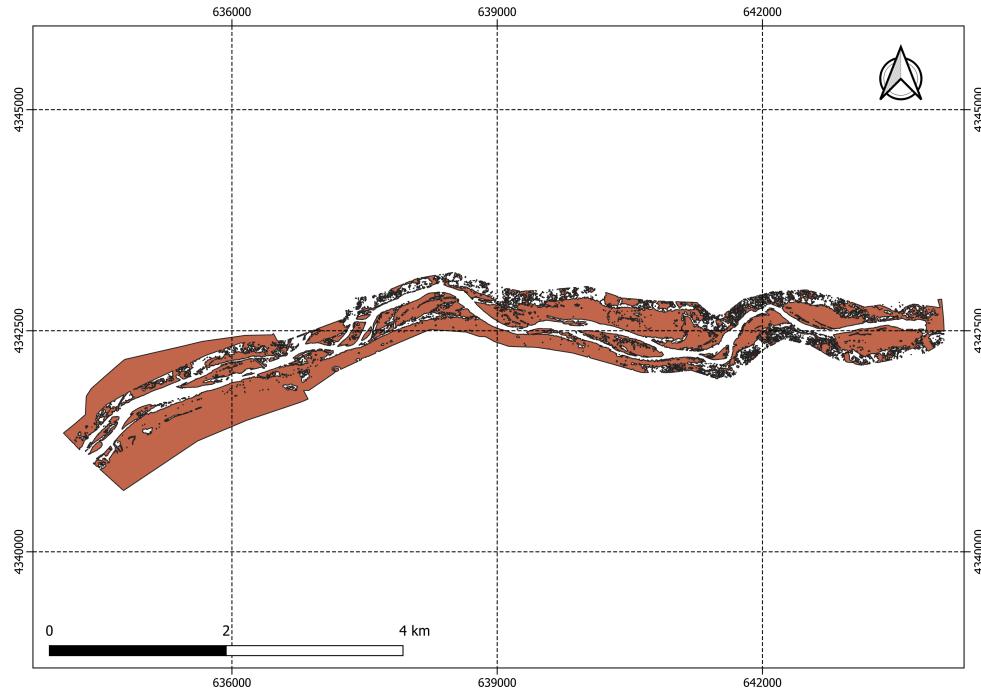


Figure 3.12: Lower Yuba River floodplain (EPSG: 32610).

In order to assign the 4 groups to the 15 Friction IDs, a few options are explored.

In the first option, the physical location of the 15 Friction IDs are analysed in the LYR model domain to find any similarities between them and representation of the morphological features corresponding to each Friction ID are assessed to meaningfully carry the grouping process. However, the Friction IDs are scattered throughout the model domain without any discernible pattern. therefore, grouping based on resembling features and closeness of the Friction ID is abandoned.

In the second option, weightage of the Friction IDs are evaluated and pixels occupied by the Friction IDs are taken into account in the grouping process. Finally, the Friction IDs are categorised based on the weightage of the pixels occupied by each Friction ID in the LYR model domain.

To implement the grouping process, the whole river channel is divided into 15 Friction IDs based on  $D_{84}$  value and using QGIS, points are generated in pixel centroids within the river channel of the LYR. Figure 3.13 shows the generated pixels in the river channel.

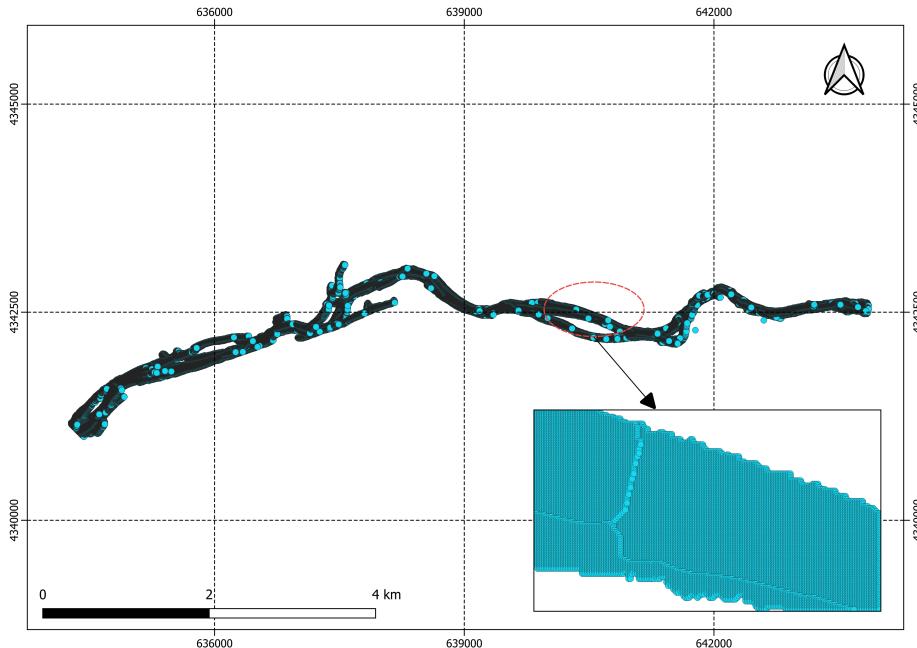


Figure 3.13: Lower Yuba River channel pixels (EPSG: 32610).

Number of pixels associated with each Friction ID are extracted and it is found that vegetation and floodplain representation in terms of pixels within the LYR domain is very high in comparison to the remaining Friction IDs. Figure 3.14 exhibits the frequency of pixels associated with 13 remaining Friction IDs.

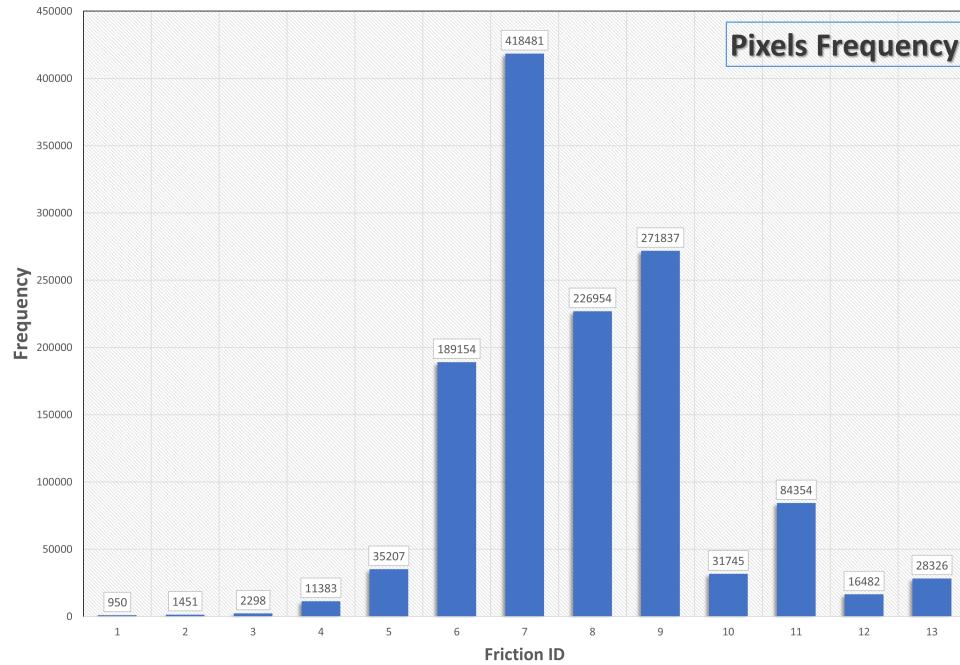


Figure 3.14: Pixels frequency for Friction ID.

Table 3.4: Friction ID pixel grouping

Friction ID	$D_{84}[m]$	Group
1	0.013	A
2	0.018	A
3	0.038	A
4	0.053	A
5	0.077	A
6	0.111	A
7	0.162	A
8	0.220	B
9	0.298	B
10	0.401	B
11	0.436	B
12	0.660	B
13	0.564	B
14	0.202	C
15	0.240	D

Table 3.4 contains the grouping of the Friction IDs. Total number of pixels under group A and B is 658924 and 659698. Representation of Group A and B in the LYR in terms of pixels are similar. Group A symbolises the Friction ID zones with  $D_{84}$  grain size lower than 0.220 m and Group B denotes the Friction ID zones with  $D_{84}$  grain size between 0.220 m to 0.660 m. Vegetation and floodplain fall under Group C and D. Both the Friction IDs are considered under individual group due to the wide representation of these categorises in the pixels of the LYR model.

### 3.2.3 Integration of roughness zones with the mesh geometry

Mesh geometry for the LYR model contains the DEM data in the form of the BOTTOM variable. In order to add a new Friction ID variable in the mesh geometry,  $D_{84}$  grain size raster data for the LYR section is used. Within the LYR section extent, a plethora of random points are generated inside the designated boundaries of the channel bed, floodplain, and vegetation. The random points are filled with latitude and longitude attributes based on the location of the points. Figure 3.15 contains the LYR section with randomly generated points. With an overlapping  $D_{84}$  particle size raster, the value of  $D_{84}$  in the assigned point locations are extracted and a (.xyz) file is generated containing the latitude, longitude, and Friction ID of the allocated points. Further, (.xyz) file data are imported to the mesh geometry and inserted in the form of a Friction ID variable. The final geometry file consists of two variables with elevation and Friction ID data on each node. Figure 3.16 showcase the mesh geometry with Friction ID.

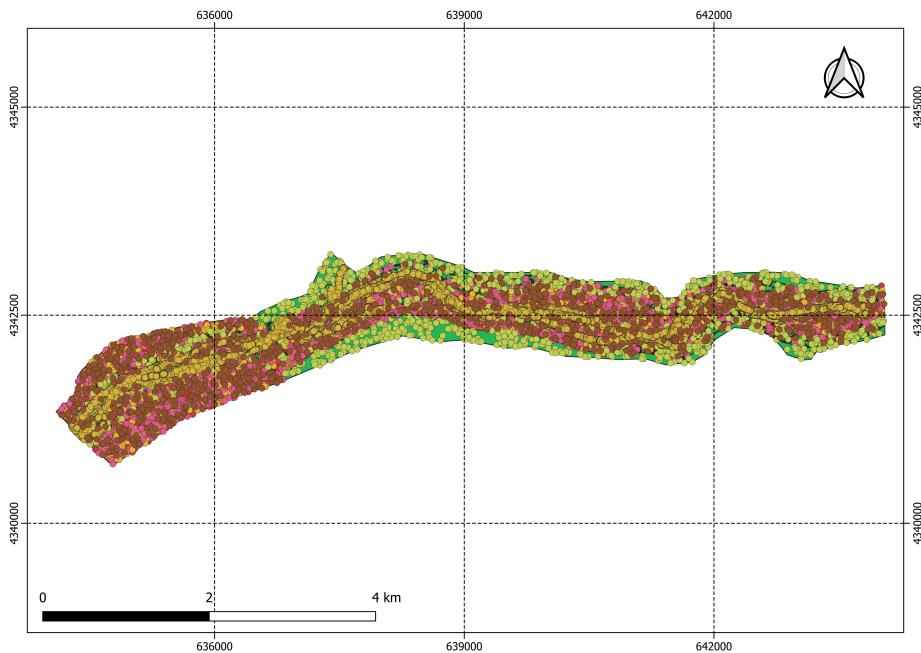


Figure 3.15: Lower Yuba River model domain with random points (EPSG: 32610).

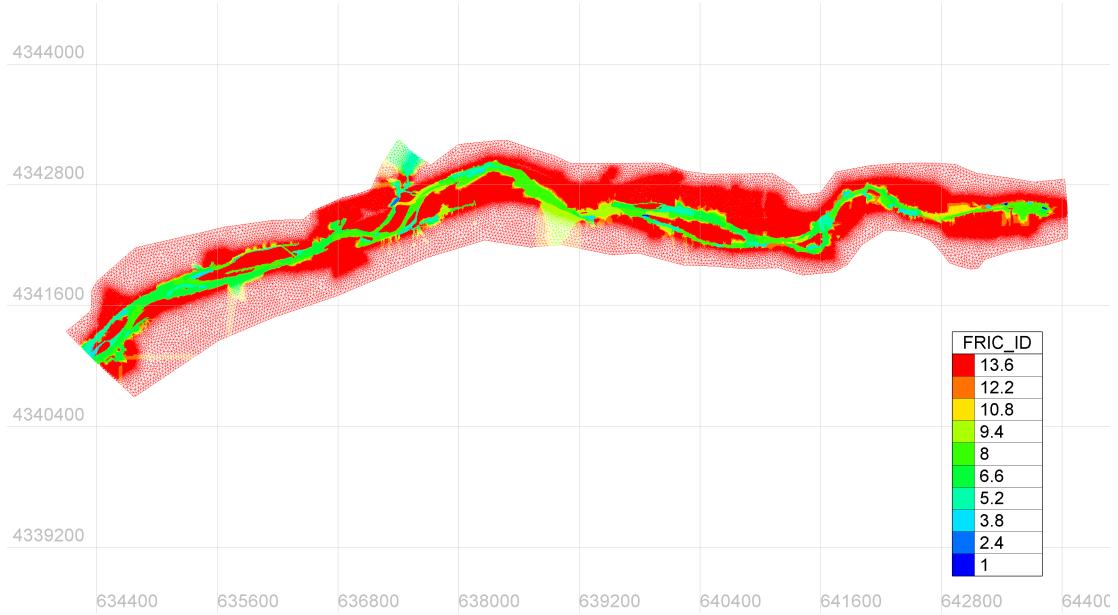


Figure 3.16: Lower Yuba River mesh with Friction ID (EPSG: 32610).

### 3.2.4 Implementation of Ferguson law

Calibration parameters and the range of the designated parameters are selected before running GPE-BAL with Telemac-2D. For this study, two constants of applied roughness law are selected as calibration parameter. Ferguson (2007) variable-power equation is used as friction law in Telemac-2D to calculate the hydraulic variables. Flow resistance is generally quantified by equation coefficients which merge velocity, water depth, and gradient based on the assumption that retardation of flow is tantamount to the downslope water weight component:

$$V = C(dS)^{\frac{1}{2}} = (8gdS/f)^{\frac{1}{2}} = d^{\frac{2}{3}}S^{\frac{1}{2}}/n \quad (3.3)$$

The above classic equations are used to relate water depth, velocity and, gradient where  $g$  is gravitational acceleration,  $d$  is the mean depth of water,  $S$  is the gradient,  $V$  is the velocity,  $C$  is Chézy coefficient,  $f$  is Darcy-Weisbach friction and  $n$  is Manning's coefficient in the equation. Variable-power equation treats the Darcy-Weisbach coefficient as a combination of bedforms and skin friction. The Variable-power equation is as follows :

$$(8/f)^{\frac{1}{2}} = a_1 a_2 (d/D) / \left[ a_1^2 + a_2^2 (d/D)^{\frac{5}{3}} \right]^{\frac{1}{2}} \quad (3.4)$$

where  $D$  is the height of friction elements and  $d$  is the water depth.  $a_1$  and  $a_2$  are equation constants taken as calibration parameters. According to Ferguson (2007), the Variable-power equation is a friction equation capable of quantifying resistance for shallow and

deep-water flows in coarse-river bed. The equation represents the concept of relative submergence and altering dominant forces as the flow becomes shallower and avoids implementing logarithmic profiles of velocity in situations where velocity is not logarithmic. The desired values for the constant  $a_2$  and  $a_1$  are in the range [1,4] and [6,8]. However, Ferguson pointed out that the optimum values are  $a_1=6.5$  and  $a_2 = 2.5$ . The performance of the variable-power equation is tested in comparison with field data over a wide range of conditions in previous studies.

Different roughness laws are implemented using the Fortran subroutine in Telemac-2D configuration. The (friction\_calc.f) Fortran subroutine in Telemac-2D is consist of 7 friction law formulas and their parameters. The Value of friction coefficients for the laws can be represented by assigning an individual value to each Friction ID.

The process of calculating Darcy-Weisbach friction coefficients is different for various formulas available in Telemac-2D. For example, in the Manning's formula, based on Manning's coefficient value, Darcy-Weisbach friction coefficient is calculated and for Ferguson law,  $D_{84}$  grain size distribution and water depth is employed for this calculation.  $D_{84}$  grain size is an approximation for the height of the roughness elements.

Adding a new law in the (friction\_calc.f) subroutine with the existing 7 laws will alter the framework of the Telemac-2D subroutine configuration. Therefore, to conveniently add the VPE 3.4 in the (friction\_calc.f) subroutine, the VPE(3.4) formula is inserted by replacing (Strickler, 1981),(Manning, 1891), (Nikuradse, 1933) and (Chézy, 1776) law. Effort is made to change The remaining 3 laws to the Equation 3.4, but the input conditions and number of variables assigned to these remaining laws are different and this impede the process modifying the laws. Each modified law contains 2 equation constants which are considered as calibration parameter for this study. (friction\_calc.f) subroutine gives the flexibility to assign variables globally. For 4 modified laws, a total of 8 calibration parameters are assigned globally in the Fortran subroutine. The calibration parameter values assigned in the (friction\_calc.f) can be subjected to change as per the new posterior calibration parameter generated by the GPE-BAL calibration process.

After assigning the calibration parameters in the (friction\_calc.f) subroutine and modifying the laws, assignment of 4 groups are executed in the (roughness.tbl) file provided by Telemac-2D. (roughness.tbl) allows to define the Friction ID and their associated  $D_{84}$  value and the law to be implemented for the Friction ID. Using (friction\_calc.f) and (roughness.tbl) in the Telemac-2D simulation, the concept of explicit roughness zones are performed.

## 3.3 Calibration and validation approach

### 3.3.1 Measurement data

Measurement data are crucial to evaluate the quality of the LYR model. To validate the hydrodynamical LYR 2D model, velocity measurements over the model extent are used. water depth values are obtained from 2014 TUFLOW model. 849 Velocity and water depth values are available for unsteady simulation discharge of  $23.13 \text{ m}^3/\text{s}$  on different point locations throughout the model domain.

Utilising 849 measurement points for calibration is not a plausible case due to the issue called dimensionality curse. Dimensionality curse refers to the problem of increased dimension of Gaussian multivariate and disruption in the process of getting a representative sample in the likelihood calculation of the Equation(2.13).

100 measurement points are chosen from different morphological unit. Figure 3.17 shows the locations of 100 measurement points selected for calibration. For validation, remaining measurement data are assessed.

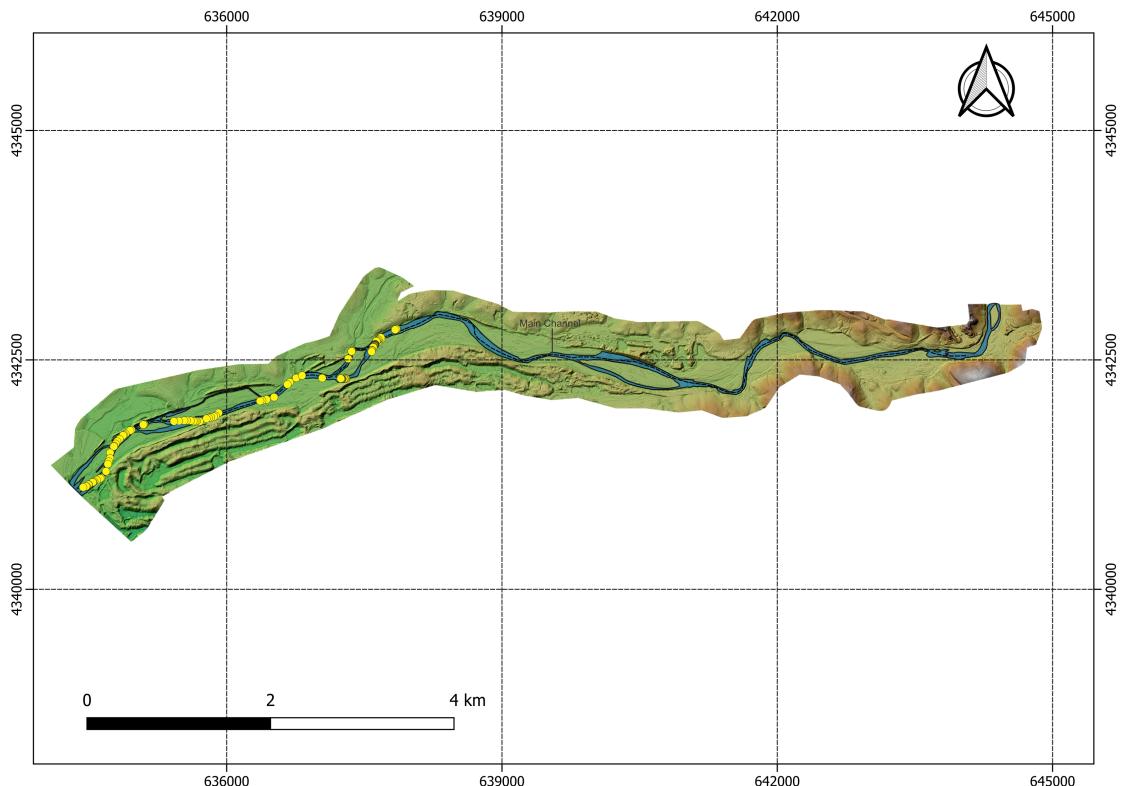


Figure 3.17: Location of the calibration points (EPSG: 32610).

### 3.3.2 Initial training data for the surrogate model

The efficiency and accuracy of the calibration process depend on the number and quality of the provided input data. Equivalency between the surrogate model and the full complex hydrodynamic model can be achieved by the chosen input training points. To predict the optimum combination of calibration parameters using a surrogate model, training data are provided in two ways.

- Parameter list
- Parameter results

Parameter lists consist of the 15 calibration parameter combinations which are available in the Table 3.5. Each parameter combinations are of 4 pairs of the constant  $a_1$  and  $a_2$  of the Ferguson law (3.4). Calibration parameters are chosen as input for the initial surrogate model based on the relevant boundary limitations. Parameter combination values scatter within the designated range for both constants.

Table 3.5: Parameter list

<b>Simulation name</b>	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$
	<b>Group A</b>		<b>Group B</b>		<b>Group C</b>		<b>Group D</b>	
PC1	7.9	3.3	6.7	1.8	7.8	1.5	7.1	1.9
PC2	6.2	2.0	7.8	2.9	6.4	1.0	7.7	3.6
PC3	6.2	2.0	6.7	1.8	6.1	3.0	6.4	2.6
PC4	6.3	3.4	6.6	2.6	6.4	1.3	7.0	3.0
PC5	7.9	1.1	7.9	3.5	6.3	3.8	7.3	3.6
PC6	7.9	3.9	6.6	2.2	7.2	3.9	7.4	1.3
PC7	7.8	3.9	7.7	1.9	7.4	3.2	7.4	3.7
PC8	7.0	2.5	6.4	2.0	6.7	2.3	7.3	1.2
PC9	6.4	3.4	6.2	3.9	6.7	2.2	6.2	1.8
PC10	6.8	3.2	7.7	1.9	7.8	3.0	7.5	2.8
PC11	7.3	1.7	7.3	2.7	7.2	2.8	6.1	2.5
PC12	7.7	3.1	6.7	1.1	7.6	3.5	6.9	1.0
PC13	7.6	3.4	7.4	2.2	7.0	3.9	6.8	3.2
PC14	7.8	3.2	7.8	1.8	6.0	1.2	7.9	2.5
PC15	7.6	1.0	6.0	1.4	6.2	3.4	6.8	2.9

Each Calibration parameter combination contains a unique name to discern from one another. For each of the 15 calibration parameter combinations, the LYR 2D hydrodynamic model is being run and flow velocity and water depth are extracted from the result of the

simulation. The location of the extracted results for velocity and water depth aligns with the calibration points location. 20 sample of the parameter results of PC1 provided for the surrogate model are shown in the Table 3.6.

Table 3.6: Parameter results of PC1 (EPSG: 32610)

<b>Latitude</b>	<b>Longitude</b>	<b>Velocity [m/s]</b>	<b>Water depth [m]</b>
637847.6583	4342828.999	0.41178	1.41194
637689.6481	4342733.228	0.31054	1.18169
637650.2832	4342696.645	0.25604	0.97937
637364.5438	4342601.289	0.15731	1.33206
637361.6536	4342593.204	0.09349	0.8547
637361.1888	4342591.879	0.08793	0.76001
637324.6094	4342512.573	0.34696	0.86205
637041.282	4342301.379	0.3451	1.05283
636768.1024	4342312.723	0.79179	1.49794
636760.7767	4342304.922	0.65627	1.0522
636757.7032	4342302.357	0.56553	1.17908
636688.593	4342256.795	0.54201	2.14199
636662.2029	4342232.225	0.6088	1.87145
636516.9753	4342092.996	0.36618	1.77365
636435.4555	4342068.851	0.32312	1.33182
636386.4465	4342057.233	0.40357	1.36899
636364.6312	4342051.068	0.4429	1.4369
635916.51	4341923.129	0.25861	1.76567
635914.2684	4341922.836	0.25587	1.74379
635912.2125	4341922.323	0.24796	1.70027

Parameter file and parameter results are provided for the surrogate model to minimize the discrepancy of prediction between the surrogate model and the full complexity hydrodynamic model.

To calculate the relative entropy for BAL process, likelihood function need to be calculated and used in the Equation 2.23 as mentioned in the chapter 2. To substitute the likelihood function values, individual errors are assigned to all the 100 measurement velocity and water depth values. The error value represent the difference between the modelled and the measurement values. 11.53 % average error difference is assigned to all measurement data based on previous analysis of the LYR section.

### 3.3.3 Telemac-2D coupling with the Bayesian calibration process

The idea of integrating the BAL method with the surrogate model is to reduce the number of input data necessary for an accurate prediction. For this workflow, the parameter file and parameter results need to be updated during the process. This update requires:

- Generating a new calibration parameter combination from the BAL method
- Running Telemac-2D simulation
- Extracting results in the measurement locations for the new calibration parameter combination
- Adding newly extracted results to the previous model results

Integration of Telemac-2D simulations files and post-processing of the results is a necessary steps to successfully execute the GPE-BAL calibration for LYR 2D model. Telemac-2D is coupled with GPE-BAL following the workflow of the Figure 3.18.

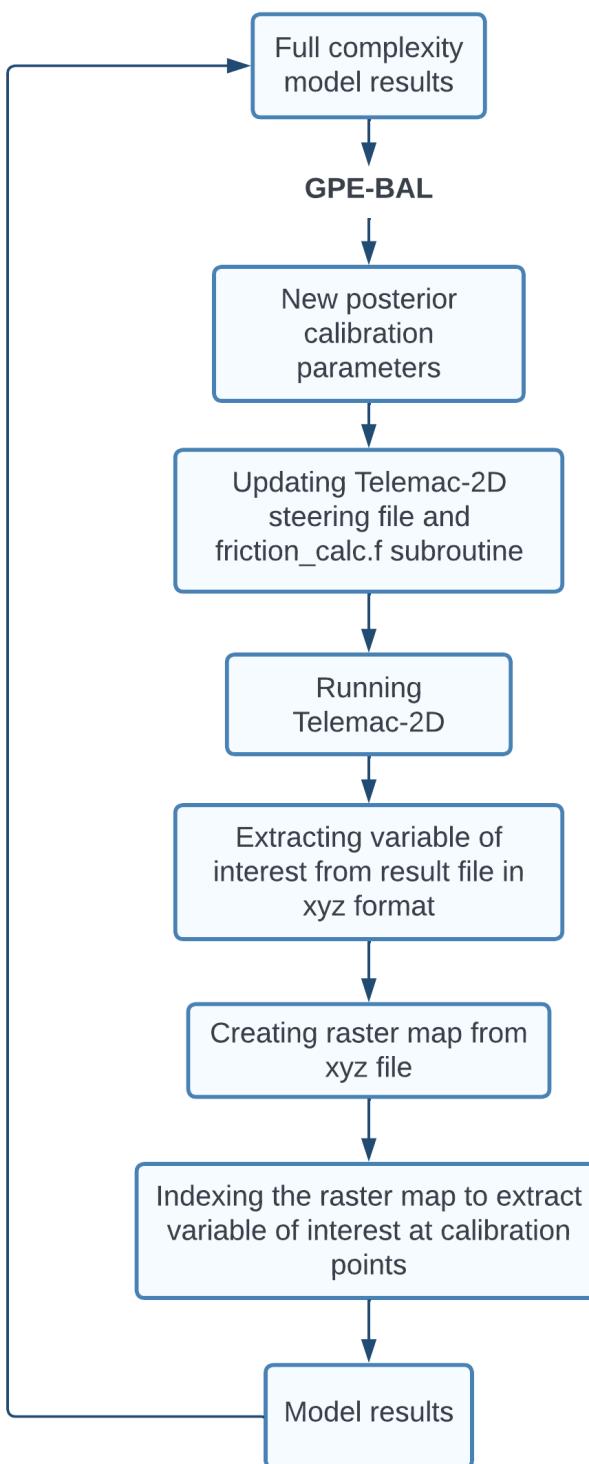


Figure 3.18: Telemac-2D coupling with GPE-BAL.

# Chapter 4

## Results

### 4.1 Calibration parameters

As mentioned in the Materials and Methods chapter, the surrogate model is constructed using the 15 calibration parameter combinations of Ferguson (2007) VPE constants and the corresponding results of the LYR 2D model for the combinations. The 15 parameter combinations are shown in the Table 3.5. The GPE-BAL calibration process is conducted for 2 Telemac-2D unsteady simulations. In Test-1, calibration process is performed for unsteady simulation with  $23.13 \text{ m}^3/\text{s}$  discharge taking into consider the Friction ID grouping of Table 3.3 and Test-2 calibration is performed with similar discharge in unsteady condition but following the grouping of the Table 3.4. Within the limited time frame, 30 iterations of GPE-BAL process for Test-1 and 7 iterations for Test-2 are successful. Based on relative entropy, BAL method is capable of selecting new parameter combinations with each iterations. After running GPE-BAL for LYR 2D model considering both the tests, final calibration parameter combinations are shown in Table 4.1 and 4.2.

Table 4.1: Test-1 optimum calibration parameters

<b>Simulation name</b>	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$
	<b>Group A</b>		<b>Group B</b>		<b>Group C</b>		<b>Group D</b>	
PC45	7.2	2.2	6.4	1.1	7.4	3.1	6.4	3.9

Table 4.2: Test-2 optimum calibration parameters

<b>Simulation name</b>	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$
	<b>Group A</b>		<b>Group B</b>		<b>Group C</b>		<b>Group D</b>	
PC22	6.2	1.6	7.8	1.2	6.1	3.9	6.3	1.9

BAL selected posterior calibration parameter combinations are showcased in Table 4.3 and 4.4.

Table 4.3: Test-1 posterior calibration parameters

<b>Simulation name</b>	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$
	<b>Group A</b>		<b>Group B</b>		<b>Group C</b>		<b>Group D</b>	
PC16	7.5	1.2	6.5	3.6	7.3	3.9	6.1	1.0
PC17	6.0	1.2	7.2	3.6	7.9	1.4	6.7	1.3
PC18	6.3	2.4	7.8	1.1	7.7	1.1	6.4	3.9
PC19	7.6	1.2	6.2	1.0	6.7	1.2	6.2	2.0
PC20	6.0	3.9	6.0	3.9	6.0	3.4	6.8	3.9
PC21	7.3	3.0	7.7	2.7	6.2	1.1	6.1	1.0
PC22	7.8	3.7	6.3	3.8	6.2	1.1	7.1	3.7
PC23	7.7	1.0	7.9	1.1	6.0	2.2	7.6	3.9
PC24	6.6	3.6	6.6	3.9	7.9	1.0	7.7	1.6
PC25	6.0	1.0	6.4	1.3	7.6	3.3	7.2	1.1
PC26	7.9	1.2	6.6	2.8	6.1	1.1	7.6	1.2
PC27	7.9	1.5	6.1	2.5	7.2	1.2	6.1	3.9
PC28	6.1	2.0	7.8	3.5	6.7	3.9	7.9	1.0
PC29	6.5	2.1	6.3	1.5	7.8	1.1	6.9	1.0
PC30	7.8	1.4	7.1	3.8	7.9	2.2	7.4	3.7
PC31	7.5	3.9	6.0	3.7	7.4	3.3	6.2	1.1
PC32	7.9	3.9	6.8	3.7	6.6	2.1	6.3	3.9
PC33	6.3	1.0	7.0	1.0	6.9	3.6	7.6	3.7
PC34	6.4	3.9	7.3	1.0	6.0	1.3	6.4	2.1
PC35	6.6	1.1	6.3	3.8	7.2	3.5	7.7	2.0
PC36	6.9	3.6	6.2	1.0	6.6	1.0	6.7	3.3
PC37	7.5	3.1	7.8	3.6	7.7	1.9	7.8	1.0
PC38	6.1	3.3	6.1	3.7	7.8	1.4	6.1	3.9
PC39	6.7	1.3	7.6	1.1	7.5	1.3	7.8	1.4
PC40	6.8	1.2	7.6	3.9	7.1	1.5	6.0	3.0
PC41	7.3	2.1	7.4	1.0	6.5	3.8	7.7	1.0
PC42	6.1	1.0	6.2	2.2	6.2	1.2	7.3	2.7
PC43	7.4	3.9	7.4	3.4	6.1	3.9	7.9	1.8
PC44	7.5	3.8	6.1	1.5	6.0	2.0	7.8	1.0
PC45	7.2	2.2	6.4	1.1	7.4	3.1	6.4	3.9

Table 4.4: Test-2 posterior calibration parameters

<b>Simulation name</b>	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$	$a_1$	$a_2$
	<b>Group A</b>		<b>Group B</b>		<b>Group C</b>		<b>Group D</b>	
PC16	6.6	1.1	7.5	1.0	6.5	1.1	6.5	2.5
PC17	6.5	1.2	6.3	3.9	6.7	1.8	7.9	1.5
PC18	7.8	3.7	7.1	1.0	7.0	1.1	7.9	3.6
PC19	6.8	2.9	6.3	1.1	7.0	3.1	6.0	3.8
PC20	6.4	3.5	7.1	3.6	7.6	3.9	7.1	3.8
PC21	7.7	3.2	7.4	3.8	7.7	1.0	6.3	3.7
PC22	6.2	1.6	7.8	1.1	6.0	3.9	6.3	1.9

For Test-1, The distribution of the calibration parameters concerning the 4 groups are devoid of any discernible patterns. BAL process is generating the posterior parameters within the designated boundaries and higher percentage of the parameter values for  $a_1$  and  $a_2$  are closer to the lower and upper boundaries of the corresponding parameter. Smaller sample size of the posterior calibration parameter combinations obtained for Test-2 hinder the process of distinguishing any tangible pattern.

In order to visualise the distribution of the posterior calibration parameters, scatter plots in Figure 4.1 and 4.2 are generated for Test-1 and Test-2.

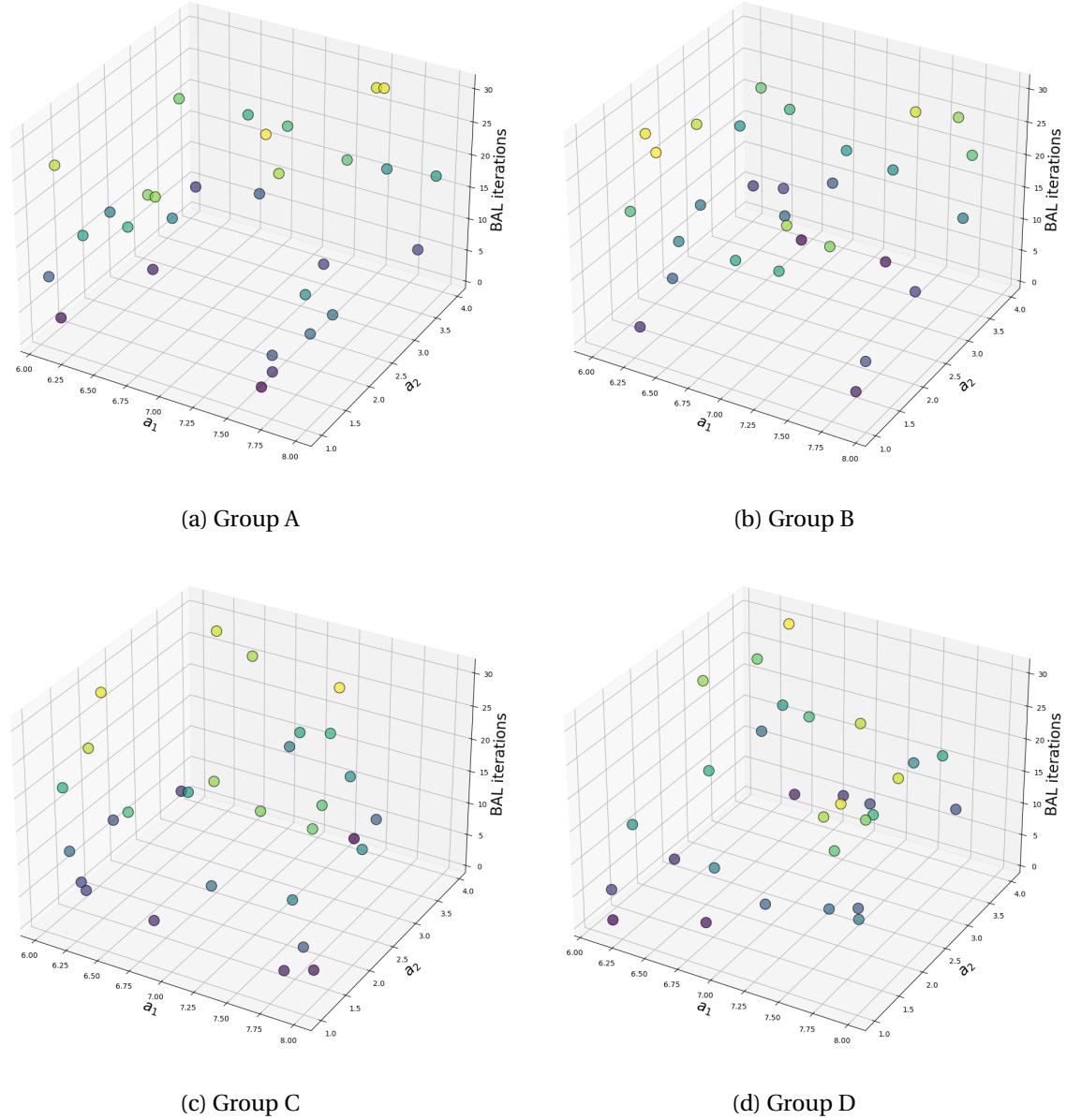


Figure 4.1: Test-1 posterior calibration parameters.

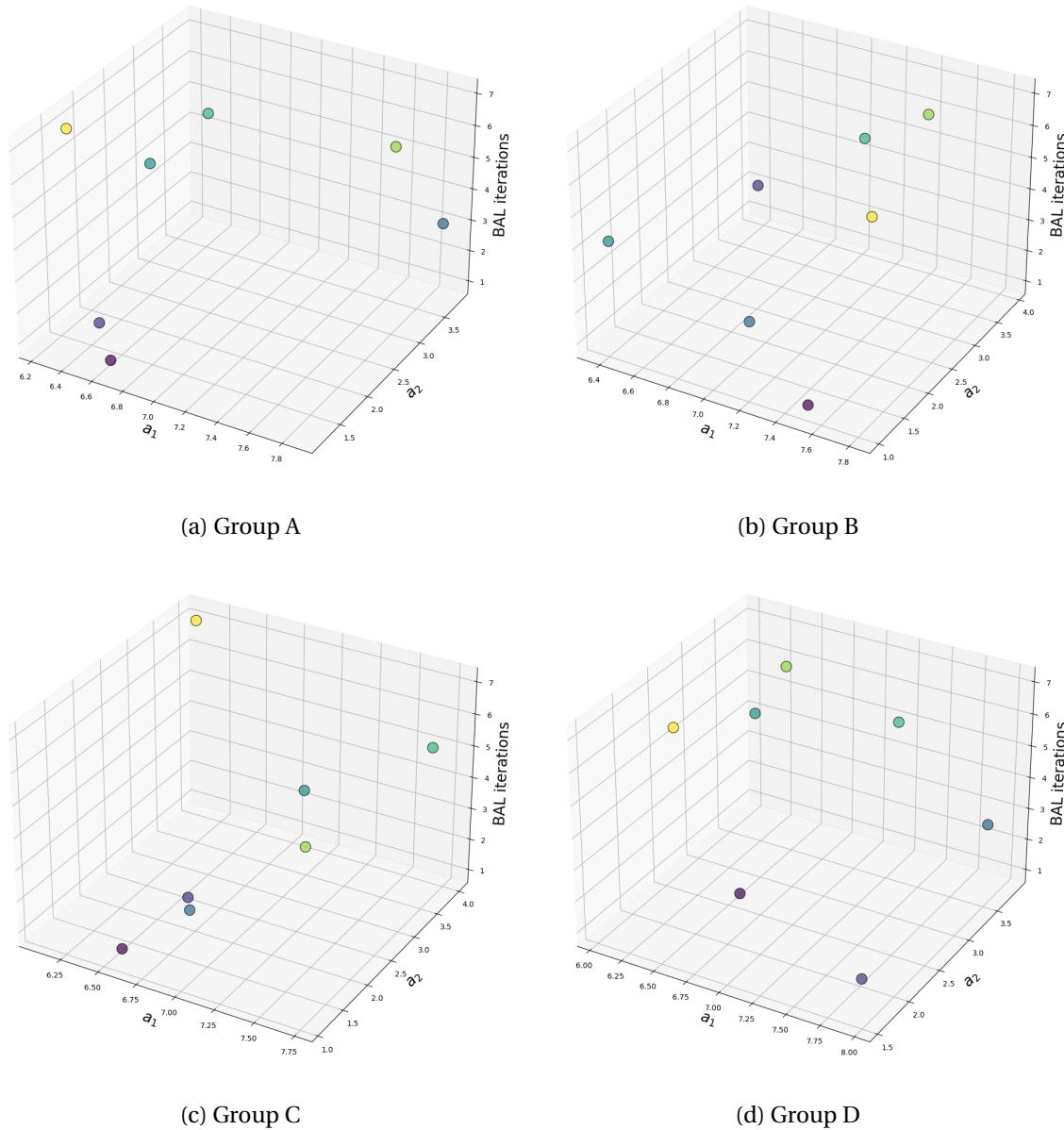


Figure 4.2: Test-2 posterior calibration parameters.

## 4.2 Evolution of information theory scores

The evolution of the Bayesian model evidence and relative entropy throughout the iterations for the two tests are plotted in Figure 4.3, 4.4, 4.5 and 4.6.

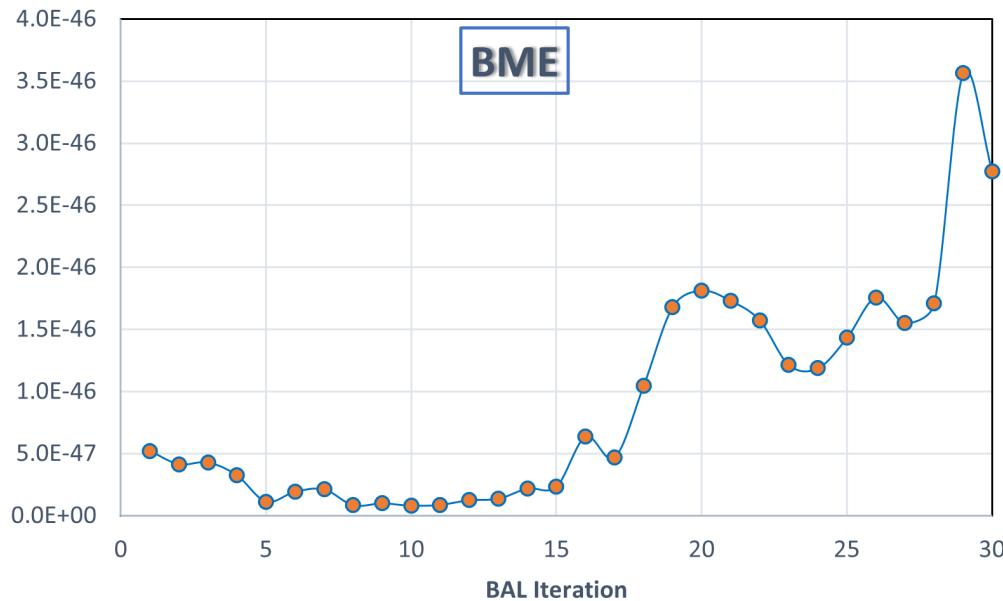


Figure 4.3: BME score of Test-1.

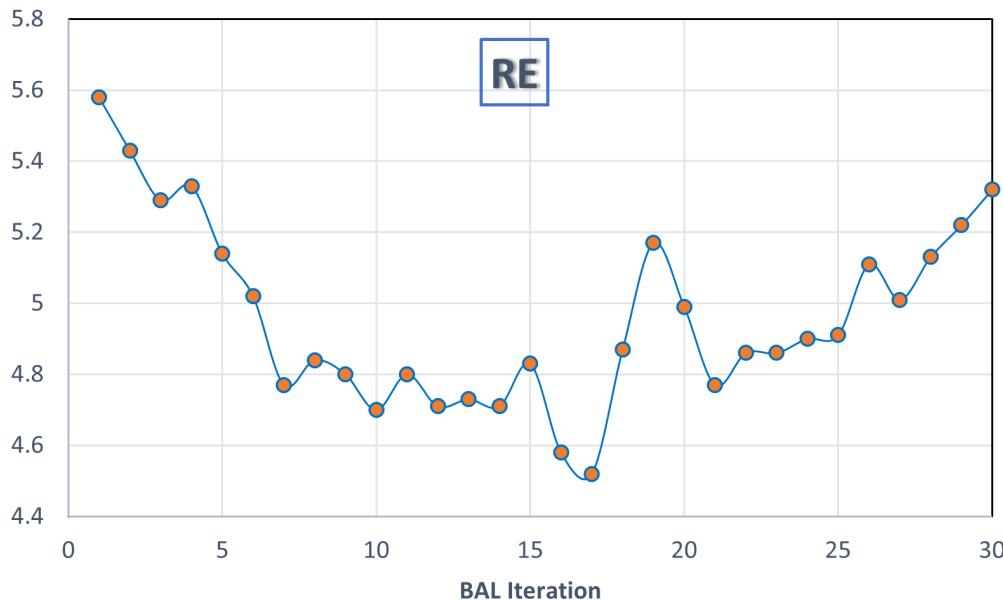


Figure 4.4: RE score of Test-1.

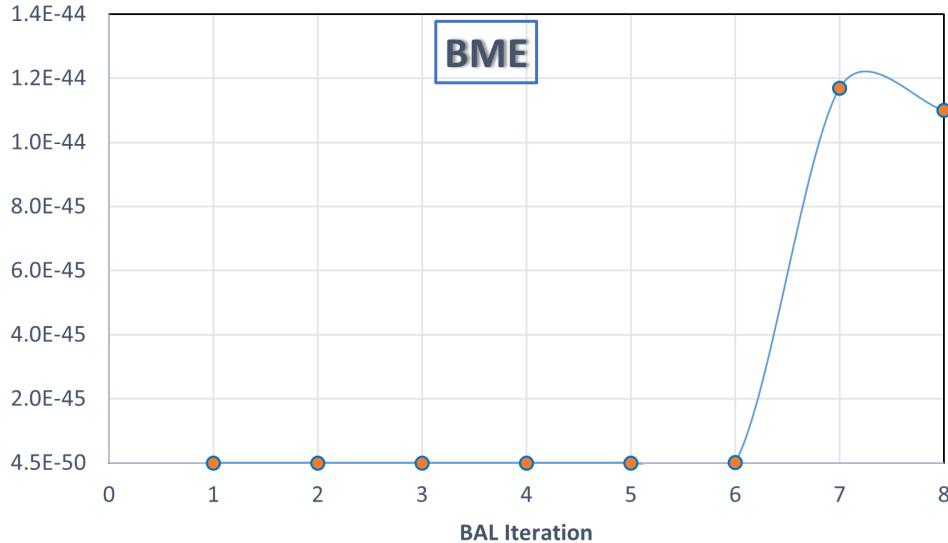


Figure 4.5: BME score of Test-2.

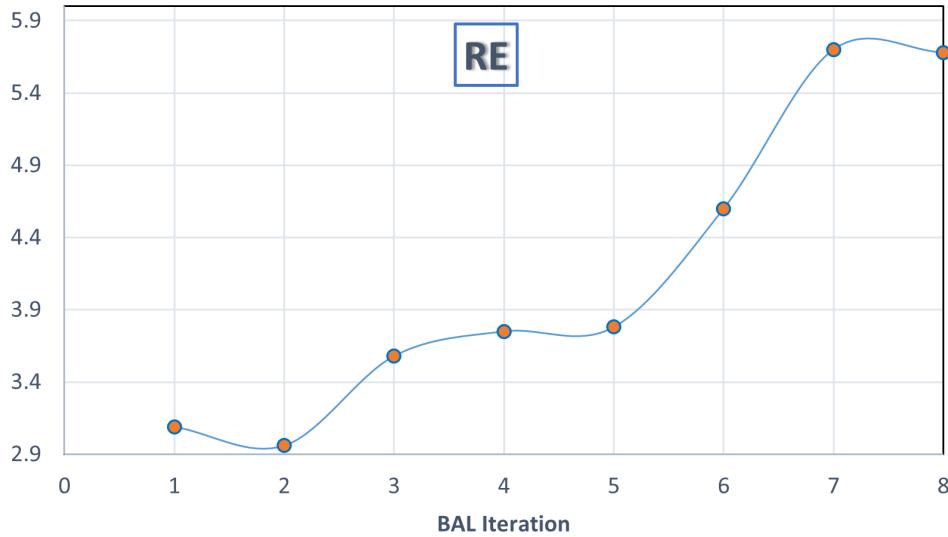


Figure 4.6: RE score of Test-2.

Test-1 BME score tend to converge till 15 iterations. After 15 iterations, BME score is rising consistently. Occasional downtrend in iterations 20 and 26 are noticeable. RE score is manifesting and verifying the behaviour of the BME score. Following a dip untill 15 iterations, RE score increases to the 30 iterations. In Test-2, BME value shows a sharp increase in 6 iterations. RE score also escalate to 7 iterations.

## 4.3 Surrogate and the calibrated model assessment

### 4.3.1 Surrogate model

In order to verify the surrogate model predictive capacity, set of calibration parameter combinations obtained after the calibration process are applied to both the surrogate and the LYR 2D model. To assess the BAL method meticulously, RMSE between the surrogate and LYR 2D model is documented for the 4 lower and higher likelihood posterior calibration parameter samples generated by GPE-BAL. PC17-20 combinations are selected under lower likelihood and PC42-45 are selected under higher likelihood for Test-1. For Test-2, PC16-22 are selected for higher likelihood. RMSE value between the LYR and the surrogate model for flow velocity and water depth in measurement locations are shown in the Table 4.5 and 4.6.

Table 4.5: Test-1 surrogate model verification

<b>Simulation name</b>	<b>RMSE</b>	
	<b>Flow velocity [m/s]</b>	<b>Water depth [m]</b>
PC17	0.016	0.050
PC18	0.036	0.053
PC19	0.014	0.046
PC20	0.015	0.056
PC42	0.011	0.033
PC43	0.017	0.024
PC44	0.023	0.051
PC45	0.017	0.020

Table 4.6: Test-2 surrogate model verification

<b>Simulation name</b>	<b>RMSE</b>	
	<b>Flow velocity [m/s]</b>	<b>Water depth [m]</b>
PC16	0.038	0.087
PC17	0.020	0.056
PC18	0.023	0.045
PC19	0.010	0.028
PC20	0.016	0.060
PC21	0.033	0.073
PC22	0.012	0.045

A striking difference in RMSE is observed in Test-1 for the higher and the lower likelihood parameter combinations. For PC17-20, the surrogate model predictions are more accurate for velocity than the water depth. There is an improvement in RMSE from PC42-46 for water depth in comparison to the lower likelihood parameter combinations. Atypical behaviour is shown by PC 44 and PC 18 combination, where the RMSE for both velocity and water depth increases instantly. Test-2 surrogate model performance is not progressing than Test-1 as far as lower likelihood parameter combinations are concerned. For flow velocity and water depth, RMSE values are higher than Test-1.

### 4.3.2 Bayesian calibrated model

To assess the efficiency of the Bayesian calibrated LYR 2D model, the optimum calibration parameter combination obtained for Test-1 and 2 are evaluated in the LYR 2D set up, and water depth and velocity variables are extracted at the measurement locations. For Test-1, the RMSE value of 0.25 m/s is calculated for velocity and 0.38 m for water depth. For Test-2, RMSE values of 0.27 m/s and 0.49 m are acquired for velocity and water depth. To test the models for different ranges of the hydraulic variables, measurement velocity values and TUFLOW 2014 water depth values are divided into quartiles. The calculated quartile values are shown in the Table 4.7.

Table 4.7: Data sorting

Hydraulic Variable	1 <sup>st</sup> quartile	2 <sup>nd</sup> quartile	3 <sup>rd</sup> quartile	4 <sup>th</sup> quartile
Velocity [m/s]	[0-0.32)	[0.32-0.53)	[0.53-0.60)	[0.60-2.43]
Water depth [m]	[0-0.58)	[0.58-0.88)	[0.88-1.16)	[1.16-2.22]

For different quartile, measurement and LYR 2-D model results are analysed and RMSE values are showcase in Table 4.8 and 4.9.

Table 4.8: Test-1 RMSE value in quartiles

Hydraulic variable	RMSE			
	1 <sup>st</sup> quartile	2 <sup>nd</sup> quartile	3 <sup>rd</sup> quartile	4 <sup>th</sup> quartile
Velocity [m/s]	0.07	0.16	0.20	0.44
Water depth [m]	0.35	0.40	0.43	0.31

Table 4.9: Test-2 RMSE value in quartiles

<b>Hydraulic variable</b>	<b>RMSE</b>			
	<b>1<sup>st</sup> quartile</b>	<b>2<sup>nd</sup> quartile</b>	<b>3<sup>rd</sup> quartile</b>	<b>4<sup>th</sup> quartile</b>
Velocity [m/s]	0.08	0.17	0.21	0.47
Water depth [m]	0.50	0.50	0.54	0.39

For the calibrated model, similar patterns are observed in Test-1 and Test-2. For velocity, RMSE gradually increases from 1<sup>st</sup> quartile to 4<sup>th</sup> quartile. In case of water depth, RMSE in 2<sup>nd</sup> and 3<sup>rd</sup> quartile is higher than 1<sup>st</sup> and 4<sup>th</sup> quartile.

## 4.4 Comparison of Manning's coefficients

The resultant 4 pairs of calibration parameters obtained from Test-1 and 2 owing to the GPE-BAL process represents 15 Friction IDs including floodplain and vegetation. Friction IDs are categorised from 246 friction zone divisions of the LYR model based on  $D_{84}$  grain size. Each friction zones occupy a specific tangible area in the river section and has unique morphological properties based on its location.

To verify the physical representation of the calibration parameters, Darcy-Weisbach friction coefficient is determined from the calibration pairs using the Equation 3.4. From average water depth of each Friction ID and calculated Darcy-Weisbach friction coefficient, Manning's coefficient is calculated utilising Equation 3.3. An analogy is made between the Manning's coefficient value obtained from the Ferguson calibration parameters and the Manning's value determined from  $D_{84}$  grain size using the Equation 3.2.

In the Table 4.10 and 4.11, Mann. coefficient(1) denotes the Manning's values obtained from the back-calculation of the calibration parameters and Mann. coefficient (2) signifies the Manning's values used in 2021 LYR Telemac-2D model obtained from  $D_{84}$  grain size.

From the comparison, it is clear that there is equivalency between the Manning's coefficients for smaller  $D_{84}$  grain sizes. In Test-1, apart from Friction ID 5, all the remaining Friction ID smaller than grain size 0.298 m are similar. From 0.298 m grain size, both the Manning's values are dissimilar. Disparity between the Manning's values rise for Group B in Test-2. The discrepancy escalates significantly for Friction ID 10, 11, 12, and 13.

Table 4.10: Test-1 Manning's comparison

<b>Friction ID</b>	<b>Group</b>	<b><math>D_{84}[m]</math></b>	<b>Mann. coefficient (1)</b>	<b>Mann. coefficient (2)</b>
1	B	0.013	0.025	0.024
2	A	0.018	0.020	0.025
3	C	0.038	0.022	0.028
4	D	0.053	0.028	0.030
5	B	0.077	0.045	0.032
6	A	0.111	0.033	0.034
7	C	0.162	0.034	0.037
8	D	0.220	0.042	0.038
9	B	0.298	0.089	0.041
10	A	0.401	0.074	0.043
11	C	0.436	0.055	0.044
12	D	0.660	0.065	0.047
13	B	0.564	0.270	0.048

Table 4.11: Test-2 Manning's comparison

<b>Friction ID</b>	<b>Group</b>	<b><math>D_{84}[m]</math></b>	<b>Mann. coefficient (1)</b>	<b>Mann. coefficient (2)</b>
1	A	0.013	0.025	0.024
2	A	0.018	0.023	0.025
3	A	0.038	0.028	0.028
4	A	0.053	0.031	0.030
5	A	0.077	0.042	0.032
6	A	0.111	0.042	0.034
7	A	0.162	0.049	0.037
8	B	0.220	0.089	0.038
9	B	0.298	0.087	0.041
10	B	0.401	0.126	0.043
11	B	0.436	0.119	0.044
12	B	0.660	0.181	0.047
13	B	0.564	0.268	0.048

# **Chapter 5**

## **Discussion**

### **5.1 Quality of the Bayesian calibration process**

The surrogate-assisted Bayesian calibration of the LYR 2D hydrodynamical model has produced posterior calibration parameter combinations along with BME and RE for each iteration.

BME score represents the ability of the model to reproduce the measured values based on prior assumptions. Oladyshkin et al. (2020) proposed to stop the BAL iterations based on information theory scores. In an ideal situation, convergence of the BME index is the decisive factor for the quality of the calibration process. This state is hard to achieve in computationally expensive model like LYR 2D. For this study, 30 BAL iterations of Test-1 and 7 iterations of Test-2 are possible with the calibration within the time frame. From Figure 4.3, it is clear that in Test-1 after 17 BAL iterations, BME becomes unstable and is not converging till 30 BAL iterations. BME pattern suggests that the model has not reached its full potential with the provided BAL iterations and the calibration process is capable of generating better posterior calibration parameter combinations with further iterations. The time constraint for this study propelled us to take PC45 as the optimum calibration parameter combination for Test-1. Verifying the BME pattern, RE score in Figure 4.4 demonstrate that after 17 iterations, the discrepancy between the modelled results and measurement values are increasing contrary to the pattern exhibited for iterations 1 to 16. With additional BAL iterations, RE score will reduce following the convergence of BME. Test-2 BME and RE score are in the preliminary state of the calibration process where any drawn pattern will not hold significance in representing the quality of the Bayesian calibration process.

To verify the BAL process proposed by Oladyshkin et al. (2020), posterior calibration pa-

parameter combinations associated with each iterations are tested on the LYR 2D model and analysed with the surrogate model predictions. For Test-1, RMSE between the surrogate model and the LYR 2D model shows a distinct pattern. The pattern observed from assessing the lower likelihood and higher likelihood of the calibration parameter combinations shows the effectiveness of the GPE-BAL process with increasing iterations. However, a detour from this striking pattern is noticeable in case of PC44 where the RMSE for both velocity and water depth surged before reaching the minimum value at PC45.

## 5.2 Behaviour of the calibration parameters

Figure 4.1 and 4.2 illustrates the distribution of the posterior calibration parameters in groups for Test-1 and Test-2. The idea of separating the parameter combinations into 4 group is to represent the spatially explicit roughness zones by different calibration parameters and utilise the Telemac-2D subroutine which give an edge to assign 4 different calibration parameter pair to 4 separate Friction ID zones instead of assigning 1 set of calibration parameter for all the Friction IDs. In Test-1, 4 groups are assigned to 15 Friction ID serially. However, in Test-2 the groups represent the physical characteristics of the river section based on  $D_{84}$  grain size and weightage of pixels occupied by Friction ID in the model domain. Group A representation is  $D_{84}$  grain size smaller than 0.220 m and Group B represent the larger grain sizes upto 0.660 m. In Test-1, 75-80% of the posterior calibration parameter combinations are closer or on the defined boundary of the two constants of the Ferguson (2007) VPE law. Most of the posterior calibration parameter combinations are distributed surrounding the upper and lower boundary. This state is not optimal as the distribution infer that the boundaries for both the constants should have been more extensive to achieve the goal of minimising the variance between the full complexity model results and the calibration point values. The degree of influence of the altering calibration parameters on the LYR 2D model outcome is low. The observation is verified by tuning the calibration parameter  $a_1$  and  $a_2$  to lowest and highest value and documenting the changes in the flow velocity and water depth variables. Minimising a large variance between the measurement point values and the model results by amending the calibration parameters within the designated boundary of this study is not plausible.

### 5.3 Physical interpretation of the calibration parameters

Verification of the optimum calibration parameter is implemented by producing the equivalent Manning's coefficient from the calibration parameters and comparing it with the Manning's values obtained from the grain size  $D_{84}$ . Notable outcome is the similarity of the Manning's coefficients for smaller  $D_{84}$  grain sizes. For Friction IDs 1, 2, 3, 4, 5, 6, and 7 the Manning's values are alike, suggesting the authenticity of the calibration parameter combinations for these Friction IDs. Considering the Manning's coefficients generated from  $D_{84}$  grain size as credible (Rickenmann & Recking, 2011), the calibration parameter values failed to represent the roughness of the LYR model domain in higher  $D_{84}$  Friction ID zones. Specially in Test-2, the calibration parameter representation of the roughness zones are particularly substandard for the larger grain sizes. This observation is occurring because of the assignment of only one calibration pair group to the  $D_{84}$  grain size from 0.220 to 0.660 m. whereas, in Test-1, the groups are diverse than Test-2, resulting in better representation of the roughness zones by the calibration parameters. Test-2 grouping is also not optimal as the implementation of only 4 calibration parameter pair do not allow the calibration parameter pairs to represent each of the 15 Friction IDs in full potential. As one calibration parameter pair is representing 3 or more Friction IDs, altering morphological units and roughness elements of the various roughness zones are not addressed properly. The higher equivalent Manning's coefficients in larger grain size locations infer that the Ferguson implemented friction law will underestimate the flow velocity and miscalculate the water depth as well. In large  $D_{84}$  grain sizes, calibration parameter  $a_1$  and  $a_2$  value combinations influences the Darcy-Weisbach friction coefficient to a higher degree. Higher Darcy-Weisbach friction coefficient corresponds to increased roughness of the friction zones. This phenomena reduces the velocity value of the respective roughness zones and increase the water depth.

### 5.4 Model evaluation

The optimum posterior calibration parameter combination is found to be PC45 for Test-1 after 30 iteration of the GPE-BAL. PC45 combinations are implemented on the LYR 2D model. Analysing the calibrated model with the measurement values for velocity and 2014 TUFLOW water depth values, it found that the model can simulate velocity better than it can simulate water depth for the LYR model domain. Ferguson (2007) law implemented calibrated model underestimate the high flow velocities and overestimate water depth in most of the regions. However, the calibrated model is accurate in terms of predicting the low flow velocities with RMSE value of 0.07 m/s. With increasing measurement velocity value, the accuracy of prediction decreases drastically and reaches a RMSE value of 0.44

m/s in high flow velocity regions. In case of water depth, model is capable of predicting low and high water depth in a better way than it can predict the average water depth in the  $2^{nd}$  quartile and  $3^{rd}$  quartile of the measurement data. For test 2, optimum calibration parameter combination is PC22 after 7 iterations. RMSE values of similar fashion is also observed in the Test-2 calibrated LYR 2D model. However, the accuracy of the Test-2 model is less than Test-1 due to limited BAL iterations in the calibration process and different assignment of Friction ID grouping. The limited measurement points for LYR 2D model occupy Friction IDs 7, 8, 9, and 10. Measurement points values in remaining Friction IDs are absent. The ability of the model to predict the lower velocities can be explained by the fact that the higher roughness is executed by the Ferguson approach in these Friction IDs as the  $D_{84}$  grain sizes are large. Same reasoning is valid to explain the inaccuracy of the model in simulating the high-flow velocities in the mentioned Friction IDs.

# Chapter 6

## Conclusions and future research

### 6.1 Conclusions

The results associated with the surrogate-assisted Bayesian calibration of the LYR 2D hydrodynamical model with spatially explicit roughness zones lead to the conclusion that the calibration process can be implemented for numerically expensive models. As proposed by Oladyshkin et al. (2020), Bayesian active learning generates a plausible set of parameters based on information theory scores. The plausible set of parameters is able to reinforce the surrogate model and successfully substitute the full-complexity model with fewer training data. While improving the surrogate model with Bayesian active learning, the number of iterations play a crucial role in terms of bringing the calibration process to an end.

The implemented workflow for the LYR 2D hydrodynamical model is successful in adapting Ferguson's law in Telemac-2D and representing the spatially explicit roughness zones with independent calibration parameters in groups. However, this workflow is limited as there is a lack of individual representation of 15 Friction IDs with 15 separate calibration parameters. The calibration parameters for the Friction ID grouping of Test-1 exhibit a better depiction of the roughness zones than Test-2.

The optimum calibration parameters acquired from the calibrated LYR 2D hydrodynamical model with Ferguson (2007) approach is ineffective in delineating the physical roughness of the spatially explicit friction zones with larger  $D_{84}$  grain sizes, which are tantamount to the size of boulders. The model executes an enhanced Darcy-Weisbach roughness coefficient. Hence, the LYR 2D model underestimates the flow velocity and overestimates the water depth in these locations. However, the representation is accurate in friction zones with smaller  $D_{84}$  grain sizes.

## 6.2 Future work and recommendations

In Future, it is recommended to implement or amend the current workflow so that it considers assigning individual calibration parameters to all the spatially explicit roughness zones associated with the model domain. The attributes of the altering roughness zones will be delineated in a better way in the model than assigning the calibration parameters in groups of a few roughness zones. As Ferguson (2007) suggested the optimum values for the two constants as the considered calibration parameters in this study, employing  $D_{84}$  as a calibration parameter in the model can produce an insightful outcome.

The number of roughness zones is also imperative in representing the frictional elements of the river channel. Increasing the number of roughness zones is also suggested for improving the model accuracy. Special care should be taken while dealing with a larger  $D_{84}$  grain size roughness zones. It is suggested to check the credibility of the Ferguson law's calibration parameters with the equivalent Manning or Darcy-Weisbach coefficients so that the model does not execute unrealistic frictional calculations in the roughness zones.

Telemac-2D provides the flexibility to assign different friction laws to different roughness zones. Utilising the methodology and diversifying the frictional calculation by assigning Manning's law for higher velocity zones and Ferguson's law for lower velocity zones can potentially depict real-life hydraulic variables.

In the Bayesian calibration context, for likelihood function calculations, measurement uncertainty and model uncertainty should be addressed properly. Bayesian model evidence convergence is crucial for the full utilisation of Bayesian active learning.

Investigation of the Ferguson approach sensitivity corresponding to different  $D_{84}$  grain sizes will unravel the feasibility of the approach in different conditions.

Surrogate-assisted Bayesian calibration process is coupled with Telemac-2D<sup>1</sup> by referring the work of (B. M. F. Negreiros, 2020) and (Espionaza, 2021).

---

<sup>1</sup>[https://github.com/akashHandique/GPE-BAL\\_Calibration](https://github.com/akashHandique/GPE-BAL_Calibration)

# References

- Chézy, A. (1776). Formula to find the uniform velocity that the water will have in a ditch or in a canal of which the slope is known. *Journal Association of Engineering Societies*, 18.
- Colebrook, C. F., & White, C. M. (1937). Experiments with fluid friction in roughened pipes. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 161(906), 367–381.
- Curtis, J. A., Flint, L. E., Alpers, C. N., Wright, S. A., & Snyder, N. P. (2006). Use of sediment rating curves and optical backscatter data to characterize sediment transport in the upper yuba river watershed, California, 2001–03.
- du Boys, M. (1879). Étude du régime du rhône et de l'action exercée par les eaux sur un lit à fond de graviers indéfiniment affouillable. vol. 18 of 5. In *Annales des ponts et chaussées*.
- Einstein, H. A. (1950). *The bed-load function for sediment transportation in open channel flows* (No. 1026). US Department of Agriculture.
- Espinoza, E. A. (2021). *Stochastic Calibration of Sediment Transport Models Using Surrogate-Assisted Bayesian Inversion* (Unpublished master's thesis). University of Stuttgart, Stuttgart.
- Espinaza, E. A. (2021, December). *surrogated\_assisted\_bayesian\_calibration*. Retrieved from [https://github.com/eduardoAcunaEspinaza/surrogated\\_assisted\\_bayesian\\_calibration](https://github.com/eduardoAcunaEspinaza/surrogated_assisted_bayesian_calibration) (GitHub Repository)
- Ferguson, R. (2007). Flow resistance equations for gravel-and boulder-bed streams. *Water resources research*, 43(5).
- Haaland, S. E. (1983). Simple and explicit formulas for the friction factor in turbulent pipe flow.
- Hopkins, C., & Pasternack, G. (2018). Autumn 2014 lower yuba river tuflow gpu 2d model description and validation. *Davis (CA): University of California at Davis. Prepared for Yuba County Water Agency*.
- James, L. A. (2005). Sediment from hydraulic mining detained by Englebright and small dams in the yuba basin. *Geomorphology*, 71(1-2), 202–226.

- Karman, T. (1930). Von, "mechanical similitude and turbulence", nachr. ges. wiss. *Göttingen, Math-Phys. Class.*
- Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., & Brumby, S. P. (2021). Global land use/land cover with sentinel 2 and deep learning. In *2021 ieee international geoscience and remote sensing symposium igarss* (pp. 4704–4707).
- Landry, M., Malouin, J.-L., & Oral, M. (1983). Model validation in operations research. *European journal of operational research*, 14(3), 207–220.
- Manning, R. (1891). On the flow of water in open channels and pipes: Institute of civil engineers of ireland transactions, v. 20.
- Mensencal, Y. (2012). Use of telemac software system as a technical modelling tool for coastal zone development studies. *SOGREAH Eau-Energie-Environnement-SPICOSA project*.
- Meyer-Peter, E., & Müller, R. (1948). Formulas for bed-load transport. In *Iahsr 2nd meeting, stockholm, appendix 2*.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885–900.
- Negreiros, B., Schwindt, S., Haun, S., & Wieprecht, S. (2022). Fuzzy map comparisons enable objective hydro-morphodynamic model validation. *Earth Surface Processes and Landforms*, 47(3), 793–806.
- Negreiros, B. M. F. (2020, October). *fuzzycorr*. GitHub repository (Version 0.0.1). Retrieved from <https://github.com/beatriznegreiros/fuzzycorr>
- Nikuradse, J. (1933). Stromungsgesetze in rauhen rohren. *vdi-forschungsheft*, 361, 1.
- Oladyshevkin, S., Mohammadi, F., Kroeker, I., & Nowak, W. (2020). Bayesian active learning for the gaussian process emulator using information theory. *Entropy*, 22(8), 890.
- Oladyshevkin, S., & Nowak, W. (2019). The connection between bayesian inference and information theory for model selection, information gain and experimental design. *Entropy*, 21(11), 1081.
- Parker, G. (1976). On the cause and characteristic scales of meandering and braiding in rivers. *Journal of fluid mechanics*, 76(3), 457–480.
- Pasternack, G., Tu, D., & Wyrick, J. (2014). Chinook adult spawning physical habitat of the lower yuba river. *Prepared for the Yuba Accord River Management Team, University of California, Davis, CA*.
- Press, S. J. (2009). *Subjective and objective bayesian statistics: principles, models, and applications*. John Wiley & Sons.
- Rasmussen, C., & Williams, C. (2006). *Gaussian processes for machine learning. the mit press, cambridge, massachusetts, usa, london*. England UK.
- Raychaudhuri, S. (2008). *Introduction to monte carlo simulation,[in:] 2008 winter simula-*

- tion conference. *ieee*, 91–100.
- Rickenmann, D., & Recking, A. (2011). Evaluation of flow resistance in gravel-bed rivers through a large field data set. *Water Resources Research*, 47(7).
- Rosgen, D. L., & Silvey, H. L. (1996). *Applied river morphology* (Vol. 1481). Wildland Hydrology Pagosa Springs, CO.
- Schwindt, S., & Barros, R. (2022). *Hydro-Informatics*. Retrieved 2022-10-17, from <https://hydro-informatics.com/>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379–423.
- Strickler, A. (1981). Contributions to the question of a velocity formula and roughness data for streams, channels and closed pipelines.
- Sutherland, J., Peet, A., & Soulsby, R. (2004). Evaluating the performance of morphological models. *Coastal engineering*, 51(8-9), 917–939.
- US Army Corps of Engineers, Y. (2010). *Lower yuba river gravel augmentation project yuba and nevada counties, california. technical report november, california*.
- US Army Corps of Engineers, Y. (2019). *Yuba river ecosystem restoration feasibility study california, technical report*.
- Wilson, C., Bates, P., & Hervouet, J.-M. (2002). Comparison of turbulence models for stage-discharge rating curve prediction in reach-scale compound channel flows using two-dimensional finite element methods. *Journal of Hydrology*, 257(1-4), 42–58.
- Wyrick, J., Gonzalez, R., & Pasternack, G. (2013). Hydraulic geometry of the lower yuba river: Depth-discharge relationships.
- Wyrick, J., & Pasternack, G. (2012). Landforms of the lower yuba river.

# **Appendices**

## **Appendix A**

### **Maps of the Lower Yuba river**

The Lower Yuba River additional maps are shown below:

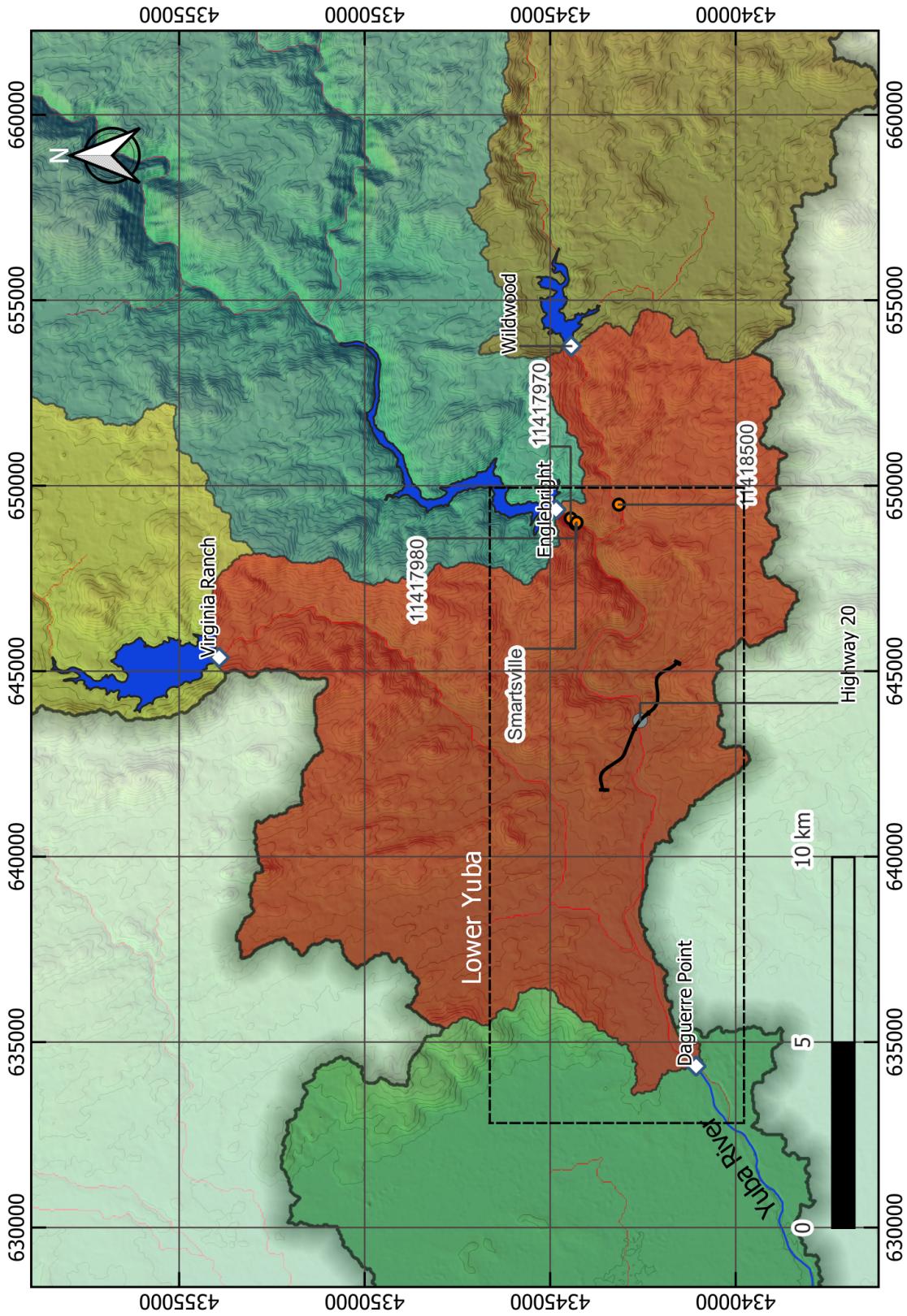


Figure A.1: Lower Yuba catchment.

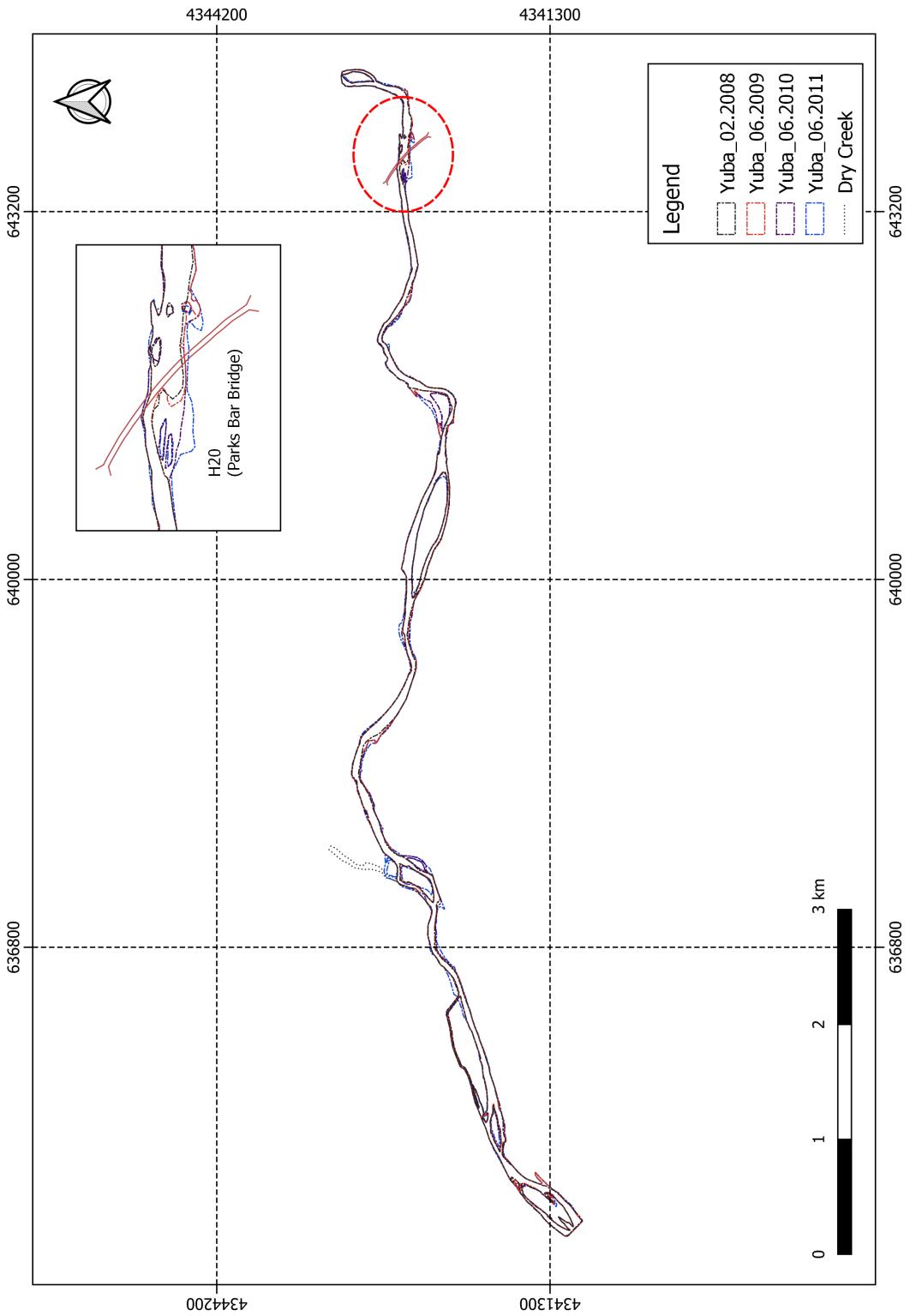


Figure A.2: Course of the river Yuba.

## **Appendix B**

### **Manning's comparison tables**

Table B.1: Test-1 Manning's value comparison (Extended)

Friction ID	Group	$D_{84}[m]$	Mann. coefficient (1)	Mann. coefficient (2)	Darcy-Weisbach coefficient	Water depth [m]
1	B	0.013	0.025	0.024	0.082	0.227
2	A	0.018	0.020	0.025	0.034	0.809
3	C	0.038	0.022	0.028	0.043	0.851
4	D	0.053	0.028	0.030	0.067	0.803
5	B	0.077	0.045	0.032	0.199	0.568
6	A	0.111	0.033	0.034	0.089	0.904
7	C	0.162	0.034	0.037	0.090	0.991
8	D	0.220	0.042	0.038	0.155	0.797
9	B	0.298	0.089	0.041	0.616	1.04
10	A	0.401	0.074	0.043	0.459	0.849
11	C	0.436	0.055	0.044	0.240	1.04
12	D	0.660	0.065	0.047	0.323	1.15
13	B	0.564	0.270	0.048	7.07	0.530

Table B.2: Test-2 Manning's value comparison (Extended)

Friction ID	Group	$D_{84}[m]$	Mann. coefficient (1)	Mann. coefficient (2)	Darcy-Weisbach coefficient	Water depth [m]
1	A	0.013	0.025	0.024	0.086	0.186
2	A	0.018	0.023	0.025	0.045	0.844
3	A	0.038	0.028	0.028	0.066	0.829
4	A	0.053	0.031	0.030	0.080	0.832
5	A	0.077	0.042	0.032	0.185	0.465
6	A	0.111	0.042	0.034	0.148	0.826
7	A	0.162	0.049	0.037	0.195	0.937
8	B	0.220	0.089	0.038	0.718	0.657
9	B	0.298	0.087	0.041	0.611	0.973
10	B	0.401	0.126	0.043	1.31	0.864
11	B	0.436	0.119	0.044	1.10	1.03
12	B	0.660	0.181	0.047	2.57	1
13	B	0.564	0.268	0.048	7.04	0.510