
Fourier Neural Operator-Based Modelling for Solving the Shallow Water Equations in Spherical Coordinates

Melissa Ulsøe Jessen (s194322)

Supervisor: Allan Peter Engsig-Karup, DTU Compute



Master Thesis
Mathematical Modelling and Computation

DTU Compute
Technical University of Denmark
February 2, 2025

Preface

This thesis has been prepared over five months at the Department of Applied Mathematics and Computer Science, at the Technical University of Denmark, DTU, in fulfillment of the requirements for the degree of Master of Science in Mathematical Modelling and Computation. The project was carried out from September 2024 to February 2025 and is equivalent to 30 ECTS points.

Melissa Ulsøe Jessen (s194322)

Abstract

This Master's thesis explores the integration of numerical and data-driven methodologies to solve the shallow water equations (SWE), a cornerstone in computational fluid dynamics. Accurate and efficient simulation of shallow water dynamics is essential for critical applications such as flood forecasting, tsunami modeling, and coastal management. Traditional numerical approaches, such as the finite volume method (FVM), solve partial differential equations (PDEs) by discretizing the domain into grids. While finer grids enhance accuracy, they substantially increase computational cost, posing a trade-off between accuracy and efficiency.

To address this challenge, this study focuses on solving the SWE in 1D, 2D, and 1D spherical coordinates while comparing Fourier neural operators (FNOs) and convolutional neural networks (CNNs) as data-driven approaches. The generated data from the FVM is used to train these models. FNOs learn mappings between function spaces, enabling grid-independent solution transfer and efficient handling of non-linearities. This allows FNOs to achieve zero-shot super-resolution by transferring solutions seamlessly between coarse and fine meshes. In contrast, CNNs operate on fixed grids, offering a complementary perspective on spatially localized features.

The thesis evaluates the performance of FNOs and CNNs in predicting the evolution of water levels and assesses their computational efficiency and accuracy. By training on coarse grids and inferring solutions on finer grids, both approaches demonstrate potential scalability and adaptability for real-time applications. The results highlight the strengths and limitations of each method, with FNOs excelling in grid-independent operations and CNNs providing robust localized feature extraction.

This work contributes a comprehensive computational toolkit for solving SWE, combining traditional numerical techniques with advanced neural network methodologies. The findings emphasize the potential of integrating FNOs and CNNs into hydrodynamic modeling frameworks, paving the way for enhanced disaster preparedness and water resource management.

Acknowledgements

Acronyms

- ANN** Artifical Neural Network
- CNN** Convolutional Neural Network
- ERK4** Explicit Runge-Kutta 4th Order
- FDM** Finite Difference Method
- FEM** Finite Element Method
- FNN** Feedforward Neural Network
- FNO** Fourier Neural Operator
- FORCE** First Order Centred
- FVM** Finite Volume Method
- HLL** Harten, Lax and van Leer
- HLLC** Harten, Lax, van Leer, Contact
- IC** Initial Condition
- IVP** Initial Value Problem
- LSWE** Linearized Shallow Water Equations
- LSTM** Long Short-Term Memory
- MOL** Method of Lines
- MUSCL** Monotonic Upstream-centered Scheme for Conservation Laws
- ODE** Ordinary Differential Equation
- PDE** Partial Differential Equation
- PINN** Physics-Informed Neural Network
- SFNO** Spherical Fourier Neural Operator
- SWE** Shallow Water Equations
- TVD** Total Variation Diminishing
- WAF** Weighted Average Flux

Contents

Preface	i
Abstract	ii
Acknowledgements	iii
Acronyms	iv
Contents	vii
1 Introduction	1
1.1 Motivation	1
1.2 Literature	3
1.3 Thesis overview	4
2 The Shallow Water Equations	5
2.1 Notation	5
2.2 Derivation of the Shallow Water Equations	6
2.3 The SWE in vector form	11
2.4 The 1D SWE in integral form	11
2.5 The Shallow Water Equations in Spherical Coordinates	12
2.6 The Linearized Shallow Water Equations in Spherical Coordinates	18
3 The Finite Volume Method	20
3.1 The Finite Volume Method for the 1D Shallow Water Equations	21
3.2 The Finite Volume Method for the 2D Shallow Water Equations	23
3.3 The Finite Volume Method for the spherical 1D Linearized Shallow Water Equations	24
3.4 The Riemann problem	26

3.4.1	The Dam-Break problem	27
3.4.2	Wave decomposition in the Riemann problem	27
3.5	Numerical fluxes	29
4	Neural Networks and Fourier Neural Operators	32
4.1	Convolutional Neural Networks	32
4.2	Fourier Neural Operators	33
5	Data generation	37
5.1	Data generation using the Finite Volume Method	37
5.2	Data generation for the sphere	42
6	Numerical results	46
6.1	The 1D Dam Break Problem	46
6.2	Toro test cases	47
6.3	1D linearized Shallow Water Equations on a sphere	53
6.4	2D idealised Circular Dam Break Problem	53
6.5	Scalability	54
6.6	Animations	55
7	Data-driven results	57
7.1	1D Shallow Water Equations with Gaussian initial conditions	57
7.2	1D linearized Shallow Water Equations in Spherical Coordinates	63
7.3	2D Shallow Water Equations with Gaussian initial conditions	67
7.3.1	Long-term predictions for the 2D Shallow Water Equations	71
8	Discussion	75
9	Conclusion	76
9.1	Further work	76
Bibliography		78

A Appendix	79
A.1 Additional figures from 1D SWE spherical CNN and FNO	79
A.2 2D SWE long term prediction	82

Chapter 1

Introduction

This chapter provides the motivation for the project and explains the background and key objectives of the thesis. It also includes a literature review, where we outline the most important sources relevant to this work. Finally, this chapter offers an overview of the thesis and its structure.

1.1 Motivation

Natural disasters such as floods and tsunamis are becoming increasingly frequent, causing devastating impacts on human life and property [1]. Populations in coastal areas, river basins and flood-prone regions are particularly at risk. A recent example is the catastrophic flooding in Valencia, Spain, in late October 2024. On October 29, 2024, Valencia received a year's worth of rain in just eight hours, leading to flash floods that devastated the area, resulting in significant loss of life and property damage [2]. Satellite images highlight the extent of the flooding, comparing the region before the event on October 8, 2024, and after on October 30, 2024. These images are presented in Figure 1.1.

While it is impossible to prevent such disasters from occurring, trustworthy forecasts delivered in sufficiently short time can help in emergency response and disaster management. By having efficient tools to simulate flood scenarios it would be possible to reduce the time needed for decision-making in critical situations, potentially saving lives and minimizing the impact of such disasters. During a catastrophe, fast simulations are particularly valuable, even if some accuracy is sacrificed, as long as the results are accurate enough to make decisions. This is one aspect where fast simulations could be useful.

Another important aspect is that, in non-catastrophic situations, running multiple simulations of different scenarios can provide a better understanding of water dynamics and a stronger data foundation. In this case it may not be needed to use a fast simulation, as we can prioritize accuracy to provide information of potential flood scenarios. For instance, the flooding in Germany during July 2021 was unexpected in several ways. Meteorological forecasts predicted heavy rain in the affected regions, and there were warnings of the potential for flooding [3]. Despite these forecasts, the intensity and scale of the floods surpassed expectations, and climate scientist expressed shock of the extent of the flooding [4]. This is a good example of how a deeper understanding of water dynamics, supported by simulations, could have helped to predict the water flow and its spread.

This motivates the study of the shallow water equations (SWE), a set of hyperbolic partial differential equations that describe the motion of a fluid in a shallow layer of water. These equations are essential for understanding and simulating water dynamics in shallow water regions, such as coastal areas, rivers, and flood-prone regions. A wide range of problems can be modeled by the SWE, such as flooding, tsunamis and dam break scenarios. By

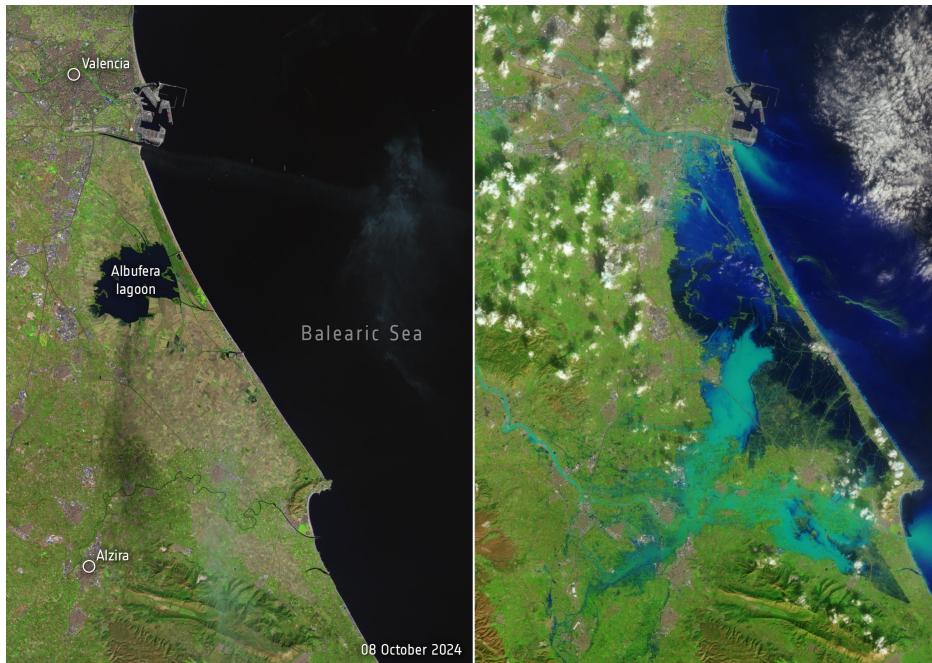


Figure 1.1: Before and after the floods in Valencia, Spain, October 2024.

Source: https://www.esa.int/ESA_Multimedia/Images/2024/10/Valencia_flood_disaster.

solving the shallow water equations, we can simulate and predict the behavior of water in these regions, providing valuable insights for disaster management and emergency response. In this work, we will derive the SWE with one spatial dimension (1D), two spatial dimensions (2D), and in spherical coordinates, which are particularly useful for modeling water flow on the Earth's surface. We will also derive the linearized shallow water equations (LSWE) in spherical coordinates for one spatial dimension, providing a more simple framework for implementation.

The SWE can be solved using numerical methods, such as the finite volume method (FVM), which is a numerical technique for solving partial differential equations, by discretizing the domain into small control volumes and integrating the equations over these volumes. This method is widely used in computational fluid dynamics to model fluid behavior. In this work, we will implement the FVM in 1D and use it to solve the SWE for several scenarios, including the dam break problem. We will extend the implementation of the FVM to 2D and solve the idealized circular dam break problem. These FVM solvers will be validated against known test cases, as this validation is critical for future work. We will also implement the FVM to solve the LSWE on a sphere. For the FVM, it holds that a finer grid resolution improves solution accuracy but comes at the cost of an increased computational run time, resulting in slower simulations. This means that while high-resolution grids can provide more accurate results, they may not be practical for real-time simulations, such as during a flood event, where fast simulations are crucial for decision-making. This limitation motivates the investigation of data-driven approaches to solve the shallow water equations, which may offer a faster and more efficient way to simulate water dynamics while maintaining an acceptable level of accuracy.

This project investigates whether data-driven methods can provide a faster and more efficient alternative to traditional numerical methods for solving the shallow water equations. Specifically, we will explore training a convolutional neural network (CNN) and a Fourier neural operator (FNO) to solve the SWE. The data-driven models will be trained on the data generated from the FVM solvers, and evaluated on performance metrics such as run time, accuracy, grid transferability and their response to new initial conditions. These initial conditions could include varying water heights, velocities, and other environmental factors that may change in real-world flood scenarios. We will compare the advantages and limitations of the data-driven approaches against numerical methods, focusing on their potential to improve disaster responder and flood prediction. This comparison will allow

us to assess whether data-driven approaches offer a practical solution to efficiently solve the shallow water equations. We expect, that data-driven models may be preferred for applications requiring fast simulations, while numerical methods may be more suitable for scenarios demanding high accuracy. An interesting aspect is to investigate how much more efficient the data-driven methods may be, compared to the traditional numerical methods, to determine to which applications they are best suited.

Additionally, we will examine the different models abilities to make long-term predictions, meaning predicting many time steps ahead. Numerical methods, including the FVM, solve the SWE one time step at a time, which can be computationally expensive, particularly for long-term predictions. In contrast, data-driven models also predict one time step ahead at a time, but they may not need to be retrained for each time step. Once trained, these models can make predictions for all time steps, potentially making them faster than numerical methods.

By analysing and comparing these approaches, this project aims to identify the context in which data-driven models can complement or even replace traditional numerical methods for solving the shallow water equations.

1.2 Literature

When working in this area it is inevitable to mention the work of E. F. Toro, who has written several books on the topic of Riemann solvers and the finite volume method, specifically for the shallow water equations. In this project, we will use the books *Shock-Capturing Methods for Free-Surface Shallow Flows* from 2001 [5], *Riemann Solvers and Numerical Methods for Fluid Dynamics* from 2009 [6] and the rather new book from 2024 *Computational Algorithms for Shallow Water Equations* [7] as references. The books have been especially useful when deriving the shallow water equations as well as understanding and describing the finite volume method, including the Riemann solvers used in this project. The course *Advanced Numerical Methods for Environmental Models* at the University of Trento, has provided a good foundation for the numerical methods used in this project, both in terms of lecture notes and exercises [8].

Working with the SWE in spherical coordinates, the papers *Well-balanced methods for the shallow water equations in spherical coordinates* by Castro et al. [9] and *Physics-informed neural networks for the shallow-water equations on the sphere* by Bihlo et al. [10] are references for deriving the SWE in spherical coordinates. Additionally, the lecture notes *Shallow water on a sphere* by Raymond from New Mexico Tech [11] and the notes from Geophysical Fluid Dynamics Laboratory [12] have been valuable in this derivation. Furthermore, the papers by Gavete [13] and Galewsky [14] also provide important insights into the spherical shallow water equations. Implementing the FVM to solve the shallow water equations in spherical coordinates is a challenging task, as exact literature on the topic is limited. However, some sources discussing the discontinuous Galerkin scheme [15] for solving the spherical shallow water equations have been useful in this context.

Convolutional neural networks (CNNs) are a well-researched area, with numerous sources available on the topic. For this project, the paper *An Introduction to Convolutional Neural Networks* [16] and the blog post *A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way* [17] have been particularly helpful in explaining the theoretical foundations of the methods. The field of Fourier neural operators (FNO) is relatively new, and as a result, there is limited literature on the topic. However, the paper *Fourier Neural Operator for Parametric Partial Differential Equations* [18], written by several authors, is a key reference. The paper suggests that FNOs have shown great potential in being an efficient and resolution independent operator in the emerging field of scientific machine learning to solve partial differential equations. A key reason of their success is their ability to generalize to unseen data. For more on the topic of scientific machine learning, consider the tech report [19]. In the last years the company Nvidia has done some very interesting work on the topic of FNO, and they have published several blog posts on the topic. One of the posts consider the use of spherical Fourier neural operators (SFNO) to generate weather forecasts around the globe [20]. Another paper regarding SFNOs is [21], which generalizes FNOs on the sphere.

1.3 Thesis overview

The rest of the thesis is structured as follows. In chapter 2, we derive the shallow water equations (SWE) in 1D, 2D, and spherical coordinates. In chapter 3, we present the finite volume method (FVM) used to solve the shallow water equations, including the Riemann problem and the numerical fluxes essential for the FVM. These chapters form the theoretical and methodological foundation for the numerical methods employed in this project, leading into the discussion of data-driven methods.

In chapter 4, we introduce the concepts of convolutional neural networks (CNNs) and Fourier neural operators (FNOs), explaining how these methods can be applied to solve the shallow water equations. In chapter 5, we describe the process of generating the data required for the data-driven methods, as we produce all the data used in this project ourselves.

Chapter 6 focus on the numerical results for the 1D and 2D SWE using the FVM. To validate these results, we compare the FVM outputs with test cases from the literature. Validation is crucial since the data-driven models are trained on the FVM data. We also present the numerical results of solving the 1D LSWE on a sphere using the FVM. In chapter 7, we present the results for the shallow water equations using the data-driven models. We analyze the outcomes, discuss the performance of the data-driven models, and compare them to the numerical results.

Finally, in chapter 8, we discuss the findings, and in chapter 9, we conclude the thesis and propose directions for future work.

Chapter 2

The Shallow Water Equations

In this chapter, we explore the theory behind the shallow water equations (SWE). We begin by introducing the relevant notation used throughout the report. Next, we derive the SWE in conservative form and present them in both vector and integral form for both the 1D- and 2D-case in cartesian coordinates. We extend the derivation to spherical coordinates, where we derive the nonlinear SWE and afterwards linearize the equations in one spatial dimension.

2.1 Notation

Before deriving the SWE, we will introduce the notation that will be used throughout this report. In both the 1D case and the 2D case of the SWE, we use cartesian coordinates (x, y, z) with time denoted by t . Given that linear algebra is a fundamental tool used in this report, we first establish the relevant notation. Lowercase bold letters represent vectors, while uppercase bold letters represent matrices. For instance, \mathbf{a} is a vector of size $r \times 1$, where $r \in \mathbb{N}$, and \mathbf{A} is a matrix of size $m \times n$ with $m, n \in \mathbb{N}$. The identity matrix, denoted by \mathbf{I} , is a square matrix with ones along the diagonal and zeros elsewhere. For example, the 3×3 identity matrix is given by:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In this project, differential calculus plays a significant role. We denote partial derivatives using the following notation:

$$f_x = \frac{\partial f}{\partial x}, \quad f_y = \frac{\partial f}{\partial y}, \quad f_z = \frac{\partial f}{\partial z}. \quad (2.1.1)$$

The gradient operator, denoted by ∇ , gives the gradient of a scalar function $f(x, y, z)$ as a vector:

$$\nabla f = [f_x \quad f_y \quad f_z].$$

Given two vectors $\mathbf{a} = [a_1 \quad a_2 \quad a_3]^\top$ and $\mathbf{b} = [b_1 \quad b_2 \quad b_3]^\top$, the dot product of \mathbf{a} and \mathbf{b} is given by:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

The dot product can also be written as a matrix product:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^\top \mathbf{b}.$$

The divergence operator, represented as $\nabla \cdot$, gives the divergence of a vector \mathbf{a} as:

$$\nabla \cdot \mathbf{a} = \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z} = a_{1x} + a_{2y} + a_{3z},$$

using the notation for partial derivatives introduced in (2.1.1). The tensor product of two vectors \mathbf{a} and \mathbf{b} , denoted by $\mathbf{a} \otimes \mathbf{b}$, is a matrix where each element is the product of the elements of \mathbf{a} and \mathbf{b} , i.e.,

$$\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{bmatrix}.$$

Establishing this relevant notation, we can now derive the shallow water equations.

2.2 Derivation of the Shallow Water Equations

The shallow water equations are derived from the conservation laws for mass and momentum, and are fundamental in fluid dynamics. The derivation follows four steps: First we consider the conservation laws for mass and momentum, and then we consider the boundary conditions for a free surface problem. Afterwards we make some necessary assumptions and finally we use the boundary conditions to integrate the conservation laws over depth. The derivation follows the methods outlined in [5] and [22].

Conservation laws

The conservation laws for mass and momentum state that the total mass and momentum are conserved over time. These laws can be generally expressed as [5, eq.'s (2.1) and (2.2)]:

$$\rho_t + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (2.2.1a)$$

$$(\rho \mathbf{v})_t + \nabla \cdot (\rho \mathbf{v} \otimes \mathbf{v} + p \mathbf{I} - \mathbf{T}) = \rho \mathbf{g}, \quad (2.2.1b)$$

where ρ is the fluid density, $\mathbf{v} = [u \ v \ w]^\top$ is the fluid velocity in the x, y and z -direction respectively, p is the pressure, \mathbf{I} is the identity matrix, and the vector $\mathbf{g} = [g_1 \ g_2 \ g_3]^\top$ represents body forces including gravity. In these equations, the density ρ and the pressure p are dependent of x, y, z and t , but later we will introduce some assumptions that simplify the equations. The matrix \mathbf{T} is the viscous stress tensor, given by

$$\mathbf{T} = \begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{bmatrix},$$

which accounts for the viscous forces in the fluid. However, in this project the viscous stress tensor \mathbf{T} is neglected, since we assume the function $\tau(x, y, z)$ is constant. The matrix $\mathbf{v} \otimes \mathbf{v}$ represents the tensor product of the velocity vector \mathbf{v} with itself, i.e.,

$$\mathbf{v} \otimes \mathbf{v} = \begin{bmatrix} u^2 & uv & uw \\ vu & v^2 & vw \\ wu & wv & w^2 \end{bmatrix}.$$

Note that $\mathbf{v} \otimes \mathbf{v} = \mathbf{v}\mathbf{v}^\top$. Putting this together, we can rewrite the momentum equation (2.2.1b) as

$$(\rho \mathbf{v})_t + \nabla \cdot (\rho \mathbf{v}\mathbf{v}^\top + p \mathbf{I}) - \rho \mathbf{g} = 0. \quad (2.2.2)$$

Applying the divergence operator and the product rule for differentiation, we can write out the mass equation (2.2.1a) as

$$\rho_t + \rho(u_x + v_y + w_z) + u\rho_x + v\rho_y + w\rho_z = 0. \quad (2.2.3)$$

In this project we consider incompressible fluids, meaning that the fluid density ρ is independent of the pressure p . We also assume that the fluid density only depends on temperature and salinity, and thus is independent of t, x, y and z . Additionally, we assume ρ is nonzero. Hence from (2.2.3) we obtain

$$u_x + v_y + w_z = 0, \quad (2.2.4)$$

also referred to as the mass conservation equation. Applying the divergence operator $\nabla \cdot$, the momentum conservation equation (2.2.2) can be written out as:

$$\rho_t \mathbf{v} + \rho \mathbf{v}_t + \rho \begin{bmatrix} (u^2 + p)_x + (uv)_y + (uw)_z \\ (vu)_x + (v^2 + p)_y + (vw)_z \\ (wu)_x + (wv)_y + (w^2 + p)_z \end{bmatrix} - \rho \mathbf{g} = 0. \quad (2.2.5)$$

We neglect all body forces in \mathbf{g} , except the gravitational force in the z -direction, i.e., $\mathbf{g} = [0 \ 0 \ -g]$, where g is the gravity acceleration, which we assume to be constant. Hence, by using the product rule in (2.2.5) and that $\rho_t = 0$ we obtain

$$\rho \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} + \rho \begin{bmatrix} p_x + uu_x + vu_y + wu_z + u(u_x + v_y + w_z) \\ p_y + uv_x + vv_y + wv_z + v(u_x + v_y + w_z) \\ p_z + uw_x + vw_y + ww_z + w(u_x + v_y + w_z) \end{bmatrix} - \rho \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} = 0. \quad (2.2.6)$$

We apply (2.2.4) to (2.2.6) to remove terms, we move the pressure terms to the right hand side, and we divide by ρ . Putting it all together, the mass equation (2.2.1a) and the momentum equation (2.2.2), split in x -, y - and z -directions, simplify to

$$u_x + v_y + w_z = 0, \quad (2.2.7a)$$

$$u_t + uu_x + vu_y + wu_z = -\frac{1}{\rho} p_x, \quad (2.2.7b)$$

$$v_t + uv_x + vv_y + wv_z = -\frac{1}{\rho} p_y, \quad (2.2.7c)$$

$$w_t + uw_x + vw_y + ww_z = -\frac{1}{\rho} p_z - g. \quad (2.2.7d)$$

Boundary conditions

In this project, the focus is on the flow of water with a free surface, meaning that the surface is not fixed and can move or change over time. To solve the SWE, it is essential to impose boundary conditions at both the bottom of the water column and at the free surface. We assume the bottom b is defined by a function

$$z = b(x, y),$$

meaning that the bottom is dependent on x and y , but not on the time t . Since the bottom is not moving over time, we refer to it as fixed. The free surface is defined by

$$z = s(x, y, t) \equiv b(x, y) + h(x, y, t),$$

where $h(x, y, t)$ is the water depth at time t . The following illustration helps to visualize the setup:

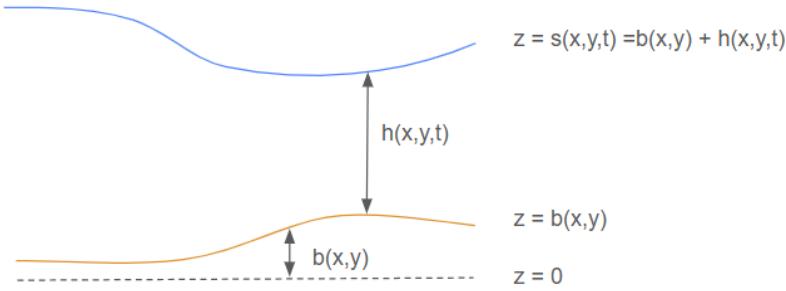


Figure 2.1: Illustration of a water column with a free surface.

We impose boundary conditions at the bottom and at the free surface, addressing both kinematic and dynamical conditions. To describe the boundaries mathematically, we introduce a boundary function $\psi(x, y, z, t)$ that is zero on the boundaries:

$$\psi(x, y, z, t) = 0.$$

For the free surface, this boundary is given by

$$\psi|_{z=s} = z - s(x, y, t) = 0, \quad (2.2.8)$$

and for the bottom, it is described by

$$\psi|_{z=b} = z - b(x, y) = 0. \quad (2.2.9)$$

In deriving a kinematic condition, we assume that fluid particles on the boundary remain on the boundary over time. Mathematically this is expressed as

$$\frac{d}{dt}\psi(x, y, z, t) = 0.$$

Recall, that $\frac{\partial\psi}{\partial t}$ is the partial derivative of ψ with respect to t , while the material derivative $\frac{d\psi}{dt}$ accounts for the rate of change of ψ experienced by a fluid particle moving with the fluid. Hence, the material derivative of ψ wrt. t is given by

$$\frac{d\psi}{dt} = \frac{\partial\psi}{\partial t} + \nabla\psi \cdot \mathbf{v}$$

Thus, the kinematic condition is given by

$$\frac{d}{dt}\psi = \psi_t + u\psi_x + v\psi_y + w\psi_z = 0. \quad (2.2.10)$$

Applying this to the free surface by substituting (2.2.8) into the kinematic condition (2.2.10) yields

$$(s_t + us_x + vs_y - w)|_{z=s} = 0. \quad (2.2.11)$$

Similarly, for the bottom, substituting (2.2.9) into the kinematic condition (2.2.10) gives

$$(ub_x + vb_y - w)|_{z=b} = 0. \quad (2.2.12)$$

The dynamical condition is related to the pressure distribution at the free surface. We assume that the pressure at the free surface is equal to the pressure in the air above the surface, that is, the atmospheric pressure. Since absolute pressure levels are irrelevant, as we are primarily concerned with pressure differences, we set the pressure at the free surface to zero. This leads to the following expression for the pressure at the free surface:

$$p(x, y, z, t)|_{z=s} = 0. \quad (2.2.13)$$

This condition, known as the dynamical condition, relates to the forces acting on the boundaries of the fluid.

Assumptions

To derive the SWE it is necessary to make some physical assumptions. The shallow water equations are an approximation to the full free-surface problem and result from the assumption that the vertical component of the acceleration is negligible. Therefore, we begin by assuming that the vertical acceleration, represented by the material derivative of the vertical velocity component w with respect to time, is negligible. This assumption leads to the condition

$$\frac{dw}{dt} = w_t + uw_x + vw_y + ww_z = 0.$$

Applying $\frac{dw}{dt} = 0$ in the z -momentum conservation equation (??) simplifies it to

$$p_z = -\rho g.$$

By using properties of an integral, together with (2.2.13) we get

$$\int_{b(x,y)}^z -\rho g \, dt + \int_z^{s(x,y,t)} -\rho g \, dt = \int_{b(x,y)}^{s(x,y,t)} -\rho g \, dt = p|_{z=s(x,y,t)} = 0.$$

This implies that the pressure distribution follows

$$p = \int_{b(x,y)}^z -\rho g \, dt = \int_z^{s(x,y,t)} \rho g \, dt = \rho g(s - z), \quad (2.2.14)$$

where s is the surface height. Differentiating (2.2.14) with respect to x and y yields

$$p_x = \rho g s_x, \quad p_y = \rho g s_y. \quad (2.2.15)$$

Substituting (2.2.15) into (2.2.7b) and (2.2.7c) gives

$$\left. \begin{aligned} u_t + uu_x + vu_y + wu_z &= -gs_x, \\ v_t + uv_x + vv_y + wv_z &= -gs_y, \end{aligned} \right\} \quad (2.2.16)$$

which can be further simplified. We realize that both p_x and p_y are independent of z , implying that $\frac{du}{dt}$ and $\frac{dv}{dt}$ are also independent of z . Hence $u_z = v_z = 0$, implying that (2.2.16) can be simplified to

$$\left. \begin{aligned} u_t + uu_x + vu_y &= -gs_x, \\ v_t + uv_x + vv_y &= -gs_y. \end{aligned} \right\} \quad (2.2.17)$$

These are the momentum equations for the shallow water equations in two spatial dimensions.

Integration over depth

The next step in deriving the SWE is to integrate the conservation equations over the vertical direction z . We integrate the mass conservation equation (2.2.4) and the momentum conservation equations (2.2.17), from the bottom, $z = b(x, y)$ to the free surface, $z = s(x, y, t)$. Starting with the mass conservation equation (2.2.4), we have

$$\int_b^s (u_x + v_y + w_z) \, dz = 0,$$

implying that, using linearity of the integral, we get

$$\int_b^s u_x \, dz + \int_b^s v_y \, dz + w|_{z=s} - w|_{z=b} = 0. \quad (2.2.18)$$

We will use Leibniz's integral rule [23], which is stated as follows:

$$\frac{d}{dx} \int_{a(x)}^{b(x)} f(x, t) dt = \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x, t) dt + f(x, b(x)) \frac{d}{dx} b(x) - f(x, a(x)) \frac{d}{dx} a(x), \quad (2.2.19)$$

to integrate the first two terms in (2.2.18), which yields

$$\left. \begin{aligned} \int_b^s u_x dz &= \frac{d}{dx} \int_b^s u dz - u|_{z=s} \frac{ds}{dx} + u|_{z=b} \frac{db}{dx}, \\ \int_b^s v_y dz &= \frac{d}{dy} \int_b^s v dz - v|_{z=s} \frac{ds}{dy} + v|_{z=b} \frac{db}{dy}. \end{aligned} \right\} \quad (2.2.20)$$

Note that since a change in x does not affect the y -component of the bottom or surface, we have that $\frac{ds}{dx} = s_x$ and $\frac{db}{dx} = b_x$, and correspondingly for s_y and b_y . Likewise we can substitute $\frac{d}{dx}$ with $\frac{\partial}{\partial x}$ in the integrals, since the integrals are with respect to z , and u and v are independent of z . Inserting these results in (2.2.20) gives

$$\left. \begin{aligned} \int_b^s u_x dz &= \frac{\partial}{\partial x} \int_b^s u dz - u|_{z=s} s_x + u|_{z=b} b_x, \\ \int_b^s v_y dz &= \frac{\partial}{\partial y} \int_b^s v dz - v|_{z=s} s_y + v|_{z=b} b_y. \end{aligned} \right\} \quad (2.2.21)$$

We can now insert the integrals (2.2.21) into the integrated mass conservation equation (2.2.18) to get

$$\frac{\partial}{\partial x} \int_b^s u dz - u|_{z=s} s_x + u|_{z=b} b_x + \frac{\partial}{\partial y} \int_b^s v dz - v|_{z=s} s_y + v|_{z=b} b_y + w|_{z=s} - w|_{z=b} = 0. \quad (2.2.22)$$

To simplify this equation further, we consider the boundary conditions. From (2.2.12) we have

$$w|_{z=b} = (ub_x + vb_y)|_{z=b}, \quad (2.2.23)$$

and from (2.2.11) we have

$$w|_{z=s} = (s_t + us_x + vs_y)|_{z=s}. \quad (2.2.24)$$

We use (2.2.12) and (2.2.11) to remove terms in (2.2.22). Hence we obtain

$$\frac{\partial}{\partial x} \int_b^s u dz + s_t + \frac{\partial}{\partial y} \int_b^s v dz = 0. \quad (2.2.25)$$

We note that $s = b + h$ and hence $s_t = h_t$, as the bottom is fixed. Recall that u and v are independent of z , and the water depth is $h = s - b$, meaning we have

$$\int_b^s u dz = u(s - b) = hu, \quad \int_b^s v dz = v(s - b) = hv.$$

Putting it all together the equation (2.2.25) simplifies to

$$h_t + (hu)_x + (hv)_y = 0, \quad (2.2.26)$$

which is also the first equation in the SWE stated in conservative form. When integrating the momentum equations (2.2.17) over the vertical direction, we see that since the equations are independent of z , the resulting equations are simply

$$\left. \begin{aligned} h(u_t + uu_x + vu_y + gs_x) &= 0, \\ h(v_t + uv_x + vv_y + gs_y) &= 0. \end{aligned} \right\} \quad (2.2.27)$$

We multiply (2.2.26) with u and v respectively, and add the resulting two equations to (2.2.27). By using the product rule for differentiation and collecting terms, we obtain the momentum equations stated in conservative form:

$$\left. \begin{aligned} (hu)_t + (hu^2 + \frac{1}{2}gh^2)_x + (huv)_y &= -ghb_x, \\ (hv)_t + (huv)_x + (hv^2 + \frac{1}{2}gh^2)_y &= -ghb_y. \end{aligned} \right\} \quad (2.2.28)$$

The three partial differential equations in (2.2.26) and (2.2.28) are the SWE stated in conservative form.

2.3 The SWE in vector form

We present the SWE in vector form, which is useful as it allows all three PDEs to be combined into a single unified equation. The SWE can be written in differential conservation law form as the vector equation

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{S}(\mathbf{U}), \quad (2.3.1)$$

where

$$\mathbf{U} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ hv \end{bmatrix}, \quad \mathbf{G}(\mathbf{U}) = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix} \quad \text{and} \quad \mathbf{S}(\mathbf{U}) = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -ghb_x \\ -ghb_y \end{bmatrix}.$$

We call \mathbf{U} the vector of conserved variables, $\mathbf{F}(\mathbf{U})$ and $\mathbf{G}(\mathbf{U})$ the flux vectors in the x - and y -direction, and $\mathbf{S}(\mathbf{U})$ the source term vector. Note that the source term vector $\mathbf{S}(\mathbf{U})$ often includes additional terms, such as friction and other external forces. However, since only gravity is considered in this case, the source term vector is simplified to the expression shown above. Moreover, if $\mathbf{S}(\mathbf{U}) = 0$, then we consider a homogeneous equation, and if $\mathbf{S}(\mathbf{U}) \neq 0$, we consider an inhomogeneous equation.

We also consider the one-dimensional case of the SWE in vector form, where the flow is only in the x -direction. The vector form of the SWE in 1D is given by

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}), \quad (2.3.2)$$

where

$$\mathbf{U} = \begin{bmatrix} h \\ hu \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \end{bmatrix} \quad \text{and} \quad \mathbf{S}(\mathbf{U}) = \begin{bmatrix} 0 \\ -ghb_x \end{bmatrix}. \quad (2.3.3)$$

When deriving numerical schemes in the finite volume method (FVM), the vector form of the 1D SWE can be used as a foundation. The 1D formulation is preferred over the 2D version for this purpose because the principles underlying the numerical schemes remain the same in both dimensions. Using the 1D SWE provides a simpler framework to introduce and analyze core aspects of the method, such as flux evaluation, conservation laws, and the treatment of source terms. Once these concepts are well-understood in 1D, the method can be extended to 2D by applying similar principles, with modifications to account for the additional spatial dimension and the directional components of fluxes and source terms.

2.4 The 1D SWE in integral form

It is often more convenient to work with the integral form of the SWE, since the integral form of equations of the form (2.3.1) and (2.3.2) allows discontinuous solutions. We derive the integral form of the 1D SWE in vector form (2.3.2). The integral form is obtained by integrating the vector form (2.3.2) over a control volume V in the x, t plane, given by

$$V = [x_L, x_R] \times [t_1, t_2].$$

The control volume is illustrated in Figure 2.2.

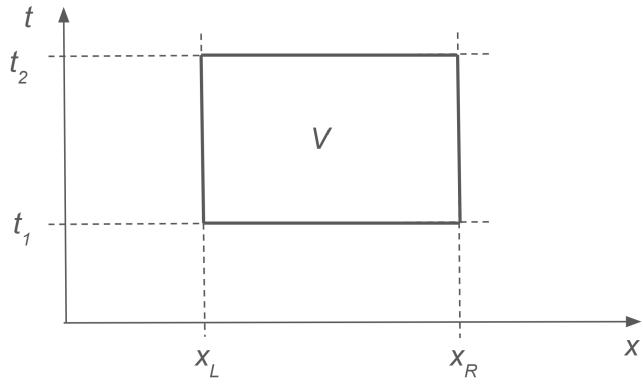


Figure 2.2: Illustration of a control volume V in the x, t plane. Illustration modified from [7].

First we integrate the vector form of the SWE (2.3.2) over x from x_L to x_R to obtain

$$\int_{x_L}^{x_R} \mathbf{U}_t \, dx + \int_{x_L}^{x_R} \mathbf{F}(\mathbf{U})_x \, dx = \int_{x_L}^{x_R} \mathbf{S}(\mathbf{U}) \, dx. \quad (2.4.1)$$

Using the fundamental theorem of calculus, we get that

$$\int_{x_L}^{x_R} \mathbf{F}(\mathbf{U}) \, dx = \mathbf{F}(\mathbf{U}(x_R, t)) - \mathbf{F}(\mathbf{U}(x_L, t)),$$

which we insert in (2.4.1):

$$\int_{x_L}^{x_R} \mathbf{U}_t \, dx = \mathbf{F}(\mathbf{U}(x_L, t)) - \mathbf{F}(\mathbf{U}(x_R, t)) + \int_{x_L}^{x_R} \mathbf{S}(\mathbf{U}) \, dx. \quad (2.4.2)$$

Then we integrate (2.4.2) over time from t_1 to t_2 to get

$$\int_{t_1}^{t_2} \int_{x_L}^{x_R} \mathbf{U}_t \, dx dt = \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_L, t)) \, dt - \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_R, t)) \, dt + \int_{t_1}^{t_2} \int_{x_L}^{x_R} \mathbf{S}(\mathbf{U}) \, dx dt.$$

Rewriting the left hand side using the fundamental theorem of calculus, we get

$$\int_{x_L}^{x_R} \mathbf{U}(x, t_2) \, dx = \int_{x_L}^{x_R} \mathbf{U}(x, t_1) \, dx + \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_L, t)) \, dt - \int_{t_1}^{t_2} \mathbf{F}(\mathbf{U}(x_R, t)) \, dt + \int_{t_1}^{t_2} \int_{x_1}^{x_2} \mathbf{S}(\mathbf{U}) \, dx dt, \quad (2.4.3)$$

which is the integral form of the conservation laws for the SWE in 1D. In chapter 3 we will use the integral form of the conserved quantities in \mathbf{U} over a domain (2.4.3) to derive the FVM for the SWE in 1D.

2.5 The Shallow Water Equations in Spherical Coordinates

Until now we have derived the shallow water equations in cartesian coordinates. In this section, we will derive the shallow water equations in spherical coordinate, which is necessary if we want to model water flow on Earth. We will follow the methods used in [9] [10], [11] and [24]. To illustrate the spherical coordinates, we will use the latitude and longitude system, visualized in Figure 2.3.

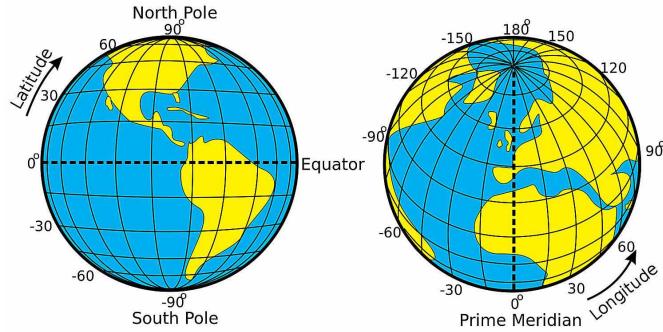


Figure 2.3: Illustration of latitude and longitude on the planet earth. Illustration from [25].

From Figure 2.3, we see that the latitude direction is the north-south component, whereas the longitude direction is the east-west component. The latitude angle, denoted by ϕ , goes from $-\frac{\pi}{2}$ radians at the south pole to $\frac{\pi}{2}$ radians at the north pole, and the longitude angle, denoted by θ , goes from 0 at the prime meridian through Europe and western Africa, increasing to the east, to 2π radians. The spherical coordinates we use are (r, θ, ϕ) , where r is the radius from the center of the sphere, θ is the longitude angle, and ϕ is the latitude angle. Both angles are measured in radians. This also means, that any point on the surface of the sphere can be represented by the coordinates (θ, ϕ) , since we assume the radius r is constant. We consider a small domain of the sphere, as illustrated in Figure 2.4.

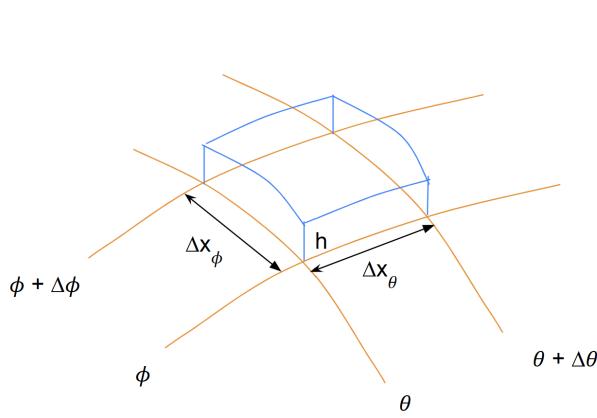


Figure 2.4: Illustrations of a small domain of the surface of the sphere.

The volume of the domain in Figure 2.4 is denoted by V . To find the volume, we need to find expressions for Δx_ϕ and Δx_θ , the distances in the ϕ and θ directions, respectively, as illustrated in Figure 2.4. We can find these distances by using the arc length formula. Recall that the circumference of a full circle is $2\pi r$, where r is the radius of the circle. The arc length is a fraction of the full circumference, and it is given by the formula $l = rv$, where l is the arc length, r is the radius, and v is the angle in radians. Assuming Earth's latitude side is a circle, we can find the distance Δx_ϕ by using the arc length formula, as:

$$\Delta x_\phi = r\Delta\phi,$$

where $\Delta\phi$ is the change in the latitude angle and r is the radius of the sphere. We assume that Earth is a perfect sphere, meaning that the radius is constant. Considering the distance in the longitude dimension Δx_θ , we need to

make some adjustments, as we can see that the circumreference at equator is larger than at the poles. That is, we need to consider the radius of the circle at the given latitude ϕ . To illustrate this, we consider Figure 2.5.

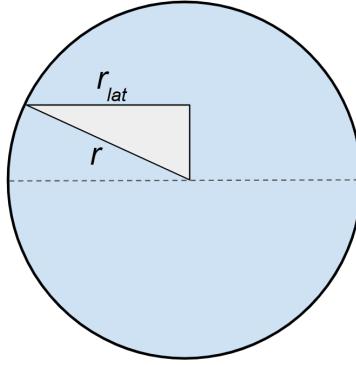


Figure 2.5: Illustration of the radius of the circle at the given latitude ϕ .

In Figure 2.5, we consider a right triangle where the length of the hypotenuse is equal to the radius of the sphere. The length of the adjacent side r_{lat} is the radius of the longitude circle at the given latitude ϕ . By using trigonometry, we can express the radius of the circle at the given latitude ϕ as

$$r_{lat} = r \cos(\phi).$$

Using that, together with the arc length formula, we can find the distance Δx_θ as:

$$\Delta x_\theta = r \cos(\phi) \Delta\theta.$$

We can then compute the volume of a small domain of the sphere, as shown in Figure 2.4. The volume V is given by

$$\begin{aligned} V &= \Delta x_\phi \Delta x_\theta h \\ &= r^2 h \cos(\phi) \Delta\phi \Delta\theta, \end{aligned}$$

assuming that the height of the domain is h , and that the domain is rectangular. This is a fair assumption for small values of Δx_ϕ and Δx_θ . We also assume that ϕ is not too close to the poles, as $\cos(\phi)$ will go to zero at the poles. We are interested in the rate of change of V with respect to time, and we can find this by taking the time derivative of the volume. That is, we consider the partial derivative with respect to time of the volume V :

$$\frac{\partial V}{\partial t} = r^2 \cos(\phi) \Delta\phi \Delta\theta \frac{\partial h}{\partial t}. \quad (2.5.1)$$

In (2.5.1) we have utilized that r is constant, and that $\cos(\phi)$, $\Delta\phi$ and $\Delta\theta$ are independent of the time t . We use u_θ and u_ϕ to denote the velocities in the θ - and ϕ -directions, not to be confused with partial derivatives. We are interested in the rate at which fluid volume enters the region from the sides. We can find this rate by considering the flux of fluid volume through the sides of the domain. That is, we consider how much fluid volume enters the domain from the θ -direction, and how much fluid volume enters the domain from the ϕ -direction. We also consider how much fluid volume leaves the domain in the θ - and ϕ -directions. The net flux of fluid volume into the domain is the difference between the influx and the outflux. The rate of change of the volume V with respect to time is equal to the net flux of fluid volume into the domain. That is, we have that

$$V_t = F_{net,\theta} + F_{net,\phi}, \quad (2.5.2)$$

which combined with (2.5.1) gives us the equation

$$r^2 \cos(\phi) \Delta\phi \Delta\theta \frac{\partial h}{\partial t} = F_{net,\theta} + F_{net,\phi}. \quad (2.5.3)$$

To find the net flux in the θ and ϕ directions, we consider the influx and outflux in these directions. We calculate the influx at the θ -line and the outflux at $(\theta + \Delta\theta)$ -line, see Figure 2.4, to find the net flux. The influx is the area of the θ line times the velocity in the θ direction at the θ line. The influx in the θ -direction is given by

$$F_{in,\theta} = u_\theta(\theta)h(\theta)r\Delta\phi,$$

where $h(\theta)$ is the height of the water at the θ -line, assumed to be constant along the line, and $u_\theta(\theta)$ is the velocity in the θ -direction at the θ -line. This way we can compute how much the volume changes due to the influx. For the outflux we do the same just for the $\theta + \Delta\theta$ line, introducing the notation $\theta' = \theta + \Delta\theta$. The outflux is

$$F_{out,\theta} = u_\theta(\theta + \Delta\theta)h(\theta + \Delta\theta)r\Delta\phi = u_\theta(\theta')h(\theta')r\Delta\phi.$$

The net flux in the θ direction is the difference between the influx and the outflux, and is given by

$$F_{net,\theta} = F_{in,\theta} - F_{out,\theta} = (u_\theta(\theta)h(\theta) - u_\theta(\theta')h(\theta'))r\Delta\phi. \quad (2.5.4)$$

We can do the same for the ϕ direction, also using the notation $\phi' = \phi + \Delta\phi$. Hence, we obtain the net flux in the ϕ direction as

$$\begin{aligned} F_{net,\phi} &= F_{in,\phi} - F_{out,\phi} \\ &= (u_\phi(\phi)h(\phi)\cos(\phi) - u_\phi(\phi')h(\phi')\cos(\phi'))r\Delta\theta. \end{aligned} \quad (2.5.5)$$

By inserting (2.5.4) and (2.5.5) into (2.5.3) we can write:

$$h_t r^2 \cos(\phi) \Delta\phi \Delta\theta = (u_\theta(\theta)h(\theta) - u_\theta(\theta')h(\theta'))r\Delta\phi + (u_\phi(\phi)h(\phi)\cos(\phi) - u_\phi(\phi')h(\phi')\cos(\phi'))r\Delta\theta. \quad (2.5.6)$$

Since we are interested in $\frac{\partial h}{\partial t} = h_t$, we divide (2.5.6) by the area of the element, given by $r^2 \cos(\phi) \Delta\phi \Delta\theta$. Hence we get

$$\frac{\partial h}{\partial t} = \frac{u_\theta(\theta)h(\theta) - u_\theta(\theta')h(\theta')}{r \cos(\phi) \Delta\theta} + \frac{u_\phi(\phi)h(\phi)\cos(\phi) - u_\phi(\phi')h(\phi')\cos(\phi')}{r \cos(\phi) \Delta\phi}, \quad (2.5.7)$$

where $\phi \neq \pm\frac{\phi}{2}$. By collecting terms to the left hand side and changing the order of the numerators, we can rewrite (2.5.7) as

$$\frac{\partial h}{\partial t} + \frac{u_\theta(\theta')h(\theta') - u_\theta(\theta)h(\theta)}{r \cos(\phi) \Delta\theta} + \frac{u_\phi(\phi')h(\phi')\cos(\phi') - u_\phi(\phi)h(\phi)\cos(\phi)}{r \cos(\phi) \Delta\phi} = 0. \quad (2.5.8)$$

Next step is to investigate the limit values, as $\Delta\theta$ and $\Delta\phi$ approaches zero. We can find the limit values by using the definition of the derivative. The derivative of a function $f(x)$ with respect to x is defined as

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

We can use this definition to find the limit values in (2.5.8) as

$$\lim_{\Delta\theta \rightarrow 0} \frac{u_\theta(\theta')h(\theta') - u_\theta(\theta)h(\theta)}{\Delta\theta} = \frac{\partial}{\partial\theta}(hu_\theta),$$

and

$$\lim_{\Delta\phi \rightarrow 0} \frac{u_\phi(\phi')h(\phi')\cos(\phi') - u_\phi(\phi)h(\phi)\cos(\phi)}{\Delta\phi} = \frac{\partial}{\partial\phi}(hu_\phi \cos(\phi)).$$

Inserting these results in (2.5.8) yields

$$h_t + \frac{1}{r \cos(\phi)} \left((hu_\theta)_\theta + (hu_\phi \cos(\phi))_\phi \right) = 0,$$

which is the mass conservation equation in spherical coordinates and is the first equation in the SWE in spherical coordinates. The next step is to derive the momentum equations in spherical coordinates. In this case, we focus on the horizontal velocity components, specifically the velocity tangential to the surface of the sphere, i.e., the θ and ϕ velocities. The vertical velocity is neglected, as the key assumption in the shallow water equations is that the vertical component of the acceleration is negligible. Additionally, when considering Earth, the water layer is thin compared to the radius of the Earth, referred to as a thin-layer approximation. We need to express the horizontal velocity u_h , which is dependent on the variables θ, ϕ and t . Since θ and ϕ are angles, we introduce the unit vectors \mathbf{e}_θ and \mathbf{e}_ϕ on the surface in the θ - and ϕ -directions. The unit vectors are illustrated in Figure 2.6.



Figure 2.6: Illustration of the unit vectors \mathbf{e}_θ and \mathbf{e}_ϕ , tangential to the surface.

We can express the horizontal velocity u_h in terms of the unit vectors as

$$u_h(\theta, \phi, t) = u_\theta \mathbf{e}_\theta + u_\phi \mathbf{e}_\phi. \quad (2.5.9)$$

We are then interested in the total derivative of the horizontal velocity u_h in (2.5.9) with respect to time. The total derivative is given by

$$\frac{du_h}{dt} = \frac{\partial u_h}{\partial t} + \frac{d\theta}{dt} \frac{\partial u_h}{\partial \theta} + \frac{d\phi}{dt} \frac{\partial u_h}{\partial \phi}. \quad (2.5.10)$$

If we differentiate the longitude angle θ with respect to time, we get the angular velocity ω_θ in the θ direction, i.e.,

$$\frac{d\theta}{dt} = \omega_\theta,$$

meaning that if $\omega_\theta > 0$, the point is moving eastwards, and if $\omega_\theta < 0$, the point is moving westwards. Similarly, if we differentiate the latitude angle ϕ with respect to time, we get the angular velocity ω_ϕ in the ϕ direction, i.e.,

$$\frac{d\phi}{dt} = \omega_\phi,$$

meaning that if $\omega_\phi > 0$, the point is moving northwards, and if $\omega_\phi < 0$, the point is moving southwards. By using the derivative of the arc length formula, we get that

$$\frac{d\theta}{dt} = \frac{u_\theta}{r \cos(\phi)}, \quad \frac{d\phi}{dt} = \frac{u_\phi}{r}. \quad (2.5.11)$$

We can now insert (2.5.11) into (2.5.10) to find the total derivative of the horizontal velocity split into the θ - and ϕ -directions:

$$\left. \begin{aligned} \frac{du_\theta}{dt} &= \frac{\partial u_\theta}{\partial t} + \frac{u_\theta}{r \cos(\phi)} \frac{\partial u_\theta}{\partial \theta} + \frac{u_\phi}{r} \frac{\partial u_\theta}{\partial \phi}, \\ \frac{du_\phi}{dt} &= \frac{\partial u_\phi}{\partial t} + \frac{u_\theta}{r \cos(\phi)} \frac{\partial u_\phi}{\partial \theta} + \frac{u_\phi}{r} \frac{\partial u_\phi}{\partial \phi}. \end{aligned} \right\} \quad (2.5.12)$$

We know that the right hand side of (2.5.12) are the given physical forces acting on the fluid. Earlier in this project, we focused on the shallow water equations in cartesian coordinates, accounting solely for gravitational forces. However, in spherical coordinates, additional physical forces must be considered. These include the Coriolis force, centripetal acceleration, and the effects of Earth's curvature. First we consider the gravitational force acting on the fluid, described as $-g\nabla h$, where g is the gravity acceleration, and ∇h is the gradient of the height h . We consider the gradient of h :

$$\begin{aligned}\nabla h &= \frac{\partial h}{\partial x_\theta} \mathbf{e}_\theta + \frac{\partial h}{\partial x_\phi} \mathbf{e}_\phi \\ &= \frac{1}{r \cos(\phi)} h_\theta \mathbf{e}_\theta + \frac{1}{r} h_\phi \mathbf{e}_\phi,\end{aligned}$$

meaning that the gravity force acting on the fluid in the θ and ϕ directions are given by:

$$\left. \begin{aligned}\theta - \text{direction: } & -\frac{g}{r \cos(\phi)} h_\theta, \\ \phi - \text{direction: } & -\frac{g}{r} h_\phi.\end{aligned}\right\}$$

Hence, we obtain the two momentum equations in spherical coordinates as

$$\left. \begin{aligned}(u_\theta)_t + \frac{u_\theta}{r \cos(\phi)} (u_\theta)_\theta + \frac{u_\phi}{r} (u_\theta)_\phi &= -\frac{g}{r \cos(\phi)} h_\theta + \text{other forces}, \\ (u_\phi)_t + \frac{u_\theta}{r \cos(\phi)} (u_\phi)_\theta + \frac{u_\phi}{r} (u_\phi)_\phi &= -\frac{g}{r} h_\phi + \text{other forces}.\end{aligned}\right\} \quad (2.5.13)$$

The next force we consider is the Coriolis force, which is a force that acts on moving objects on the surface of the earth [26]. The Coriolis force is given by $f = 2\Omega \sin(\phi)$, where Ω is the angular velocity of the earth. The Coriolis force in the θ and ϕ directions are then given by:

$$\left. \begin{aligned}\theta - \text{direction: } & fu_\phi, \\ \phi - \text{direction: } & -fu_\theta.\end{aligned}\right\}$$

The last thing we need to take into account when working in the spherical domain is the curvature of the earth. This adds the following terms:

$$\left. \begin{aligned}\theta - \text{direction: } & \frac{u_\theta u_\phi}{r} \tan(\phi), \\ \phi - \text{direction: } & -\frac{u_\theta^2}{r} \tan(\phi).\end{aligned}\right\}$$

There are several formulations of the SWE in spherical coordinates. Inserting these forces into the momentum equations (2.5.13), we get the following formulation of the shallow water equations in spherical coordinates:

$$\left. \begin{aligned}h_t + \frac{1}{r \cos(\phi)} \left((hu_\theta)_\theta + (hu_\phi \cos(\phi))_\phi \right) &= 0, \\ (u_\theta)_t + \frac{u_\theta}{r \cos(\phi)} (u_\theta)_\theta + \frac{u_\phi}{r} (u_\theta)_\phi - \frac{u_\theta u_\phi}{r} \tan(\phi) + \frac{g}{r \cos(\phi)} h_\theta - fu_\phi &= 0, \\ (u_\phi)_t + \frac{u_\theta}{r \cos(\phi)} (u_\phi)_\theta + \frac{u_\phi}{r} (u_\phi)_\phi + \frac{u_\theta^2}{r} \tan(\phi) + \frac{g}{r} h_\phi + fu_\theta &= 0,\end{aligned}\right\} \quad (2.5.14)$$

where r is the radius, (θ, ϕ) are the longitude and latitude angles, h is the height of the water, u_θ and u_ϕ are the velocities in the θ - and ϕ -directions and g is the gravitational constant. We can also write the spherical SWE (2.5.14) in vector form as

$$\mathbf{U}_t + \frac{1}{r \cos(\phi)} \mathbf{F}(\mathbf{U})_\theta + \frac{1}{r} \mathbf{G}(\mathbf{U})_\phi = \mathbf{S}, \quad (2.5.15)$$

where $\frac{1}{r \cos(\phi)} \mathbf{F}(\mathbf{U})_\theta$ and $\frac{1}{r} \mathbf{G}(\mathbf{U})_\phi$ are the flux terms in the θ - and ϕ -directions and \mathbf{S} is the source term.

2.6 The Linearized Shallow Water Equations in Spherical Coordinates

In this section, we derive the linearized shallow water equations (LSWE) in spherical coordinates, from the nonlinear SWE in spherical coordinates, given in (2.5.14). The derivation and methodology presented here are based on the sources [27] and [28]. The LSWE are employed for their simplicity, providing a foundation for understanding wave dynamics in a spherical geometry. Later in this project, we will apply the finite volume method to solve these equations numerically. We focus on one spatial dimension, namely the longitude θ . This approach assumes a constant latitude ϕ , which simplifies the equations to a one-dimensional framework. Hence we can neglect the third equation in (2.5.14) and the velocity component u_ϕ . Thus, we introduce $u = u_\theta$ to denote the fluid velocity in the θ direction. We also neglect all external forces, except for the gravitational force. Thus we have the following SWE in spherical coordinates for one spatial dimension:

$$\left. \begin{aligned} h_t + \frac{1}{r \cos(\phi)} (hu)_\theta &= 0, \\ u_t + \frac{u}{r \cos(\phi)} u_\theta + \frac{g}{r \cos(\phi)} h_\theta &= 0. \end{aligned} \right\} \quad (2.6.1)$$

To perform the linearization, we use perturbation theory, meaning we assume that the water height h and the velocity u are small perturbations from a state of rest. Let the perturbations in water height and velocity be represented by:

$$h = h_0 + h', \quad u = u_0 + u', \quad (2.6.2)$$

where h_0 and u_0 are the equilibrium height and velocity, and h' and u' are the perturbations. We assume that the equilibrium height h_0 is constant and that the equilibrium velocity u_0 is zero. By substituting the perturbed variables (2.6.2) into the 1D spherical SWE (2.6.1) and utilizing that $u_0 = 0$, we obtain

$$\left. \begin{aligned} (h_0 + h')_t + \frac{1}{r \cos(\phi)} ((h_0 + h')(u')_\theta + (h_0 + h')_\theta(u')) &= 0, \\ (u')_t + \frac{u'}{r \cos(\phi)} (u')_\theta + \frac{g}{r \cos(\phi)} h'_\theta &= 0. \end{aligned} \right\} \quad (2.6.3)$$

We neglect the terms involving products of perturbations arising in (2.6.3), as they are second-order terms. Hence, we obtain the LSWE in spherical coordinates with one spatial dimension θ and time t as:

$$\left. \begin{aligned} h'_t + \frac{h_0}{r \cos(\phi)} u'_\theta &= 0, \\ u'_t + \frac{g}{r \cos(\phi)} h'_\theta &= 0. \end{aligned} \right\} \quad (2.6.4)$$

These equations describe the evolution of small perturbations in water height and velocity on a spherical surface. We can also write the spherical LSWE in vector form as

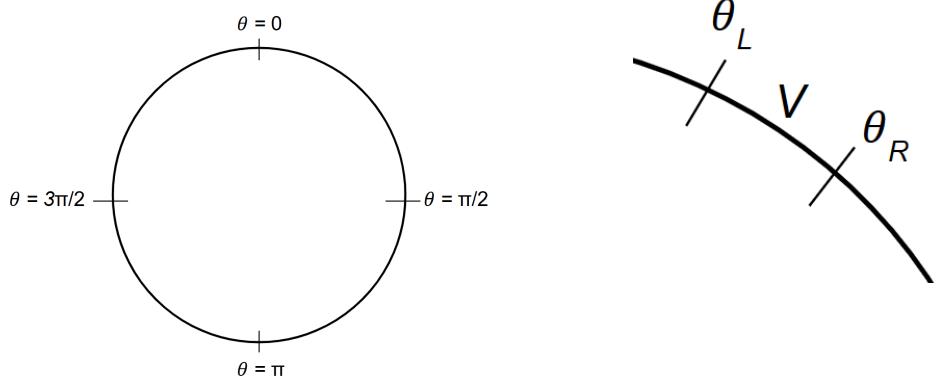
$$\mathbf{W}_t + \mathbf{A}\mathbf{W}_\theta = 0, \quad (2.6.5)$$

where $\mathbf{W} = \begin{bmatrix} h' \\ u' \end{bmatrix}$ and the coefficient matrix \mathbf{A} is constant and given as: $\mathbf{A} = \begin{bmatrix} 0 & \frac{h_0}{r \cos(\phi)} \\ \frac{g}{r \cos(\phi)} & 0 \end{bmatrix}$.

Integral form of the 1D spherical LSWE

As in the previous sections it is beneficial to state the integral form of the 1D spherical LSWE, as it will be used when deriving the finite volume scheme. We consider the longitude θ as the spatial dimension on a full circle, meaning

the domain is $[0, 2\pi]$ radians. The domain is divided into N cells or control volumes, that each has a length of $\Delta\theta = \frac{2\pi}{N}$ radians. We consider a control volume V with cell interfaces at θ_L and θ_R , as illustrated in Figure 2.7b.



(a) Illustration of the grid for the 1D SWE with small cells.

(b) Illustration of the grid for the 1D SWE with a small domain.

Figure 2.7: Grid illustrations for the 1D LSWE in spherical coordinates.

In Figure 2.7a, we see the circle, which is the full domain, and in Figure 2.7b we see a small subdomain of the circle.

The integral form of the LSWE is obtained by integrating the control volume given by $V = [\theta_L, \theta_R] \times [t_1, t_2]$, where t_2 denotes the time step after t_1 , over space and time. We begin by integrating the vector form of the 1D spherical LSWE in (2.6.5) over the spatial dimension θ from θ_L to θ_R to obtain

$$\int_{\theta_L}^{\theta_R} \mathbf{W}_t \, d\theta + \int_{\theta_L}^{\theta_R} \mathbf{A} \mathbf{W}_\theta \, d\theta = 0. \quad (2.6.6)$$

Since the matrix \mathbf{A} is constant, we can place it outside of the integral in (2.6.6) to get

$$\int_{\theta_L}^{\theta_R} \mathbf{W}_t \, d\theta + \mathbf{A} \int_{\theta_L}^{\theta_R} \mathbf{W}_\theta \, d\theta = 0. \quad (2.6.7)$$

We use the fundamental theorem of calculus to rewrite (2.6.7) as

$$\int_{\theta_L}^{\theta_R} \mathbf{W}_t \, d\theta = \mathbf{A} (\mathbf{W}(\theta_L, t) - \mathbf{W}(\theta_R, t)). \quad (2.6.8)$$

We then integrate (2.6.8) over time from t_1 to t_2 to get

$$\int_{t_1}^{t_2} \int_{\theta_L}^{\theta_R} \mathbf{W}_t \, d\theta \, dt = \mathbf{A} \left(\int_{t_1}^{t_2} \mathbf{W}(\theta_L, t) \, dt - \int_{t_1}^{t_2} \mathbf{W}(\theta_R, t) \, dt \right). \quad (2.6.9)$$

Rewriting (2.6.9) gives

$$\int_{\theta_L}^{\theta_R} \mathbf{W}(\theta, t_2) \, d\theta = \int_{\theta_L}^{\theta_R} \mathbf{W}(\theta, t_1) \, d\theta - \mathbf{A} \left(\int_{t_1}^{t_2} \mathbf{W}(\theta_R, t) \, dt - \int_{t_1}^{t_2} \mathbf{W}(\theta_L, t) \, dt \right), \quad (2.6.10)$$

which we refer to as the integral form of the LSWE in spherical coordinates with one spatial dimension. The integral form (2.6.10) is the foundation for the FVM, which can be used to solve the LSWE in spherical coordinates numerically.

Chapter 3

The Finite Volume Method

In this section, we present the finite volume method (FVM) for solving nonlinear systems of balance laws, specifically focusing on the shallow water equations (SWE). Nonlinear problems are more challenging than linear problems, as stability and convergence theory are more difficult to establish. In particular, when dealing with hyperbolic systems such as the SWE, the presence of discontinuities such as shock waves, hydraulic jumps, and rarefaction waves requires special treatment. Our focus is on discontinuous solutions, which can accurately capture these shocks and other discontinuities. The FVM provides a framework for dealing with such solutions in a robust and efficient manner. Unlike traditional methods, which may struggle to capture sharp gradients or discontinuities, FVM inherently incorporates a mechanism to resolve such features by solving the local Riemann problem at the interfaces between computational cells. This characteristic makes it particularly well-suited for problems involving abrupt changes in the flow, such as those encountered in the modeling of floods, tsunamis, or dam breaks.

The approach described here is based on the work of LeVeque [29], who developed a framework for using the FVM in the context of hyperbolic systems of PDEs. The key idea is to discretize the computational domain into small cells or control volumes and compute the fluxes across the boundaries of these cells. The solution within each cell is then updated using the information from neighboring cells, ensuring that conservation laws are satisfied. By solving the local Riemann problem at the cell interfaces, the FVM can accurately capture discontinuities and preserve the sharp transitions between different flow regimes. In addition to its ability to handle discontinuities, the FVM also has several other advantages, such as its ability to conserve mass, momentum, and energy in a discrete sense. This property is crucial when solving balance laws, as it ensures that the physical conservation laws are respected even at the discrete level. The method is also highly flexible, allowing for the use of unstructured grids, which can be beneficial for complex geometries and irregular domains. Despite its advantages, the FVM also has limitations. The accuracy of the method depends heavily on the choice of numerical flux function and the resolution of the computational grid. In the presence of strong discontinuities, additional techniques such as limiters or high-order schemes may be required to maintain stability and avoid spurious oscillations. Furthermore, while the FVM is well-suited for handling hyperbolic problems, it may not be as effective for problems with smooth solutions or for capturing fine-scale features in the solution without sufficient grid resolution.

In this chapter, we derive the finite volume scheme for the 1D SWE and extend it to 2D in cartesian coordinates. We introduce the MUSCL (monotonic upstream-centered scheme for conservation laws) scheme for higher-order accuracy and discuss the FVM for the 1D LSWE in spherical coordinates. Additionally, we examine the Riemann problem, including the dam break problem, and present numerical flux functions for solving the local Riemann problem at cell interfaces.

3.1 The Finite Volume Method for the 1D Shallow Water Equations

We begin by introducing the FVM for the SWE in one spatial dimension. Specifically, we work in the x, t -plane, discretizing the domain into finite control volumes or cells. In chapter 2, we derived the integral form for a single control volume $V = [x_L, x_R] \times [t_1, t_2]$. Here, we extend this approach to a global domain consisting of multiple cells. To accommodate this, we update the notation: $x_{i-1/2}$ and $x_{i+1/2}$ denote the cell interfaces, while t_n and t_{n+1} represent the time levels. The control volume for cell i at time level n is defined as:

$$V_i^n = [x_{i-1/2}, x_{i+1/2}] \times [t_n, t_{n+1}],$$

where $\Delta x = x_{i+1/2} - x_{i-1/2}$ is the length of the cell and $\Delta t = t_{n+1} - t_n$ is the time step size. The cell V_i^n is illustrated in Figure 3.1.

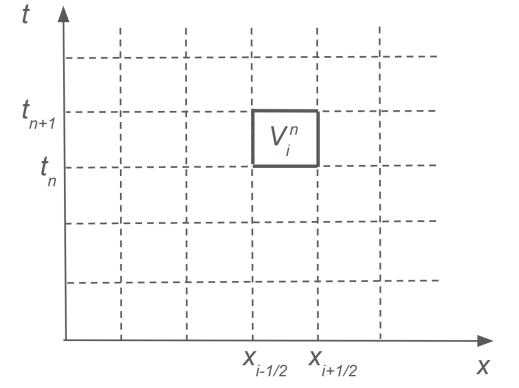


Figure 3.1: Illustration of the control volume V_i^n in the x, t plane.

The grid can be uniform or non-uniform, depending on the application. For now, we will assume a uniform grid for simplicity. The finite volume formula is derived from the integral form of the 1D SWE (2.4.3). The integral form stated in the new variables $(x_{i-1/2}, x_{i+1/2}, t_n, t_{n+1})$ over the cell V_i^n is given by

$$\begin{aligned} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t_{n+1}) dx &= \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t_n) dx + \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i-1/2}, t)) dt - \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i+1/2}, t)) dt \\ &\quad + \int_{t_n}^{t_{n+1}} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{S}(\mathbf{U})(x, t) dx dt. \end{aligned} \quad (3.1.1)$$

We divide the integral form (3.1.1) by the cell length Δx to obtain

$$\begin{aligned} \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t_{n+1}) dx &= \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t_n) dx \\ &\quad - \frac{\Delta t}{\Delta x} \left[\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i+1/2}, t)) dt - \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i-1/2}, t)) dt \right] \\ &\quad + \frac{\Delta t}{\Delta x \Delta t} \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{t_n}^{t_{n+1}} \mathbf{S}(\mathbf{U})(x, t) dx dt. \end{aligned}$$

To simplify the numerical implementation, we introduce cell-averaged values. These are defined as the averages of the conserved variables \mathbf{U} , the fluxes \mathbf{F} , and the source terms \mathbf{S} over the volume V_i^n . This leads to the explicit conservative form of the finite volume scheme:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} \left(\mathbf{F}_{i+1/2}^n - \mathbf{F}_{i-1/2}^n \right) + \Delta t \mathbf{S}_i. \quad (3.1.2)$$

The value \mathbf{U}_i^n is the average value over the i -th cell at time t_n :

$$\mathbf{U}_i^n = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t_n) dx,$$

also known as the cell average. The flux $\mathbf{F}_{i-1/2}^n$ is the average flux across the line $x = x_{i-1/2}$ from time t_n to t_{n+1} :

$$\mathbf{F}_{i-1/2}^n = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i-1/2}, t)) dt,$$

and correspondingly the flux $\mathbf{F}_{i+1/2}^n$ is the average flux across the line $x = x_{i+1/2}$ from time t_n to t_{n+1} :

$$\mathbf{F}_{i+1/2}^n = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{F}(\mathbf{U}(x_{i+1/2}, t)) dt.$$

The source term \mathbf{S}_i is the average source term over the i -th cell at time t_n :

$$\mathbf{S}_i = \frac{1}{\Delta t \Delta x} \int_{t_n}^{t_{n+1}} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{S}(x, t) dx dt.$$

The values are illustrated in Figure 3.2.

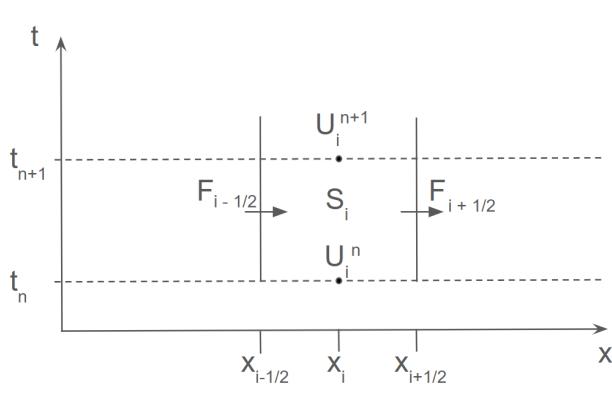


Figure 3.2: Illustration of the grid for the 1D SWE.

The central idea of the FVM is to define the numerical flux $\mathbf{F}_{i+1/2}^n$, at the cell interface, as a function of the cell averages \mathbf{U}_i^n and \mathbf{U}_{i+1}^n , since the solution is known only in terms of these cell averages. Consequently, the FVM does not provide pointwise values of the solution, i.e., $\mathbf{U}(x, t)$, but instead gives cell-averaged values, \mathbf{U}_i^n , over the control volume. One of the main challenges in the FVM is to determine appropriate numerical flux functions that, based on the available cell averages, can reasonably approximate the fluxes at the cell interfaces. Later in the thesis, we will consider several numerical flux functions that can be used to solve the local Riemann problem at the cell interfaces.

The FVM is closely related to the finite difference method (FDM), but they differ as the FVM is based on the integral form of the conservation laws. Where finite difference methods tend to break down near discontinuities in the solution, finite volume methods are more suited, since they are based on the integral form of the conservation laws. The key distinction between the FVM and the FDM lies in their formulation: while the FVM is based on the integral conservation over finite volumes, the FDM is based on satisfying the differential equations point-wise at mesh points. Similarly, the FVM and the finite element method (FEM) differ in some of their principles and applications. Like the FVM, the FEM divides the domain into smaller elements, but approximates the solution using interpolation functions. The FEM is especially effective for solving problems with complex geometries.

3.2 The Finite Volume Method for the 2D Shallow Water Equations

We now extend the FVM to two spatial dimensions, meaning we consider both the x - and y -direction. Consider the 2D SWE in vector form (2.3.1) with $\mathbf{S}(\mathbf{U}) = 0$:

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = 0. \quad (3.2.1)$$

Following the methods outlined in [6], an explicit finite volume scheme to solve (3.2.1) is given by

$$\mathbf{U}_{i,j}^{n+1} = \mathbf{U}_{i,j}^n - \frac{\Delta t}{\Delta x} (\mathbf{F}_{i+1/2,j} - \mathbf{F}_{i-1/2,j}) - \frac{\Delta t}{\Delta y} (\mathbf{G}_{i,j+1/2} - \mathbf{G}_{i,j-1/2}). \quad (3.2.2)$$

This is the unsplit finite volume method, meaning that, in a single step, the cell average $\mathbf{U}_{i,j}^n$ in cell $V_{i,j}$ is updated using the fluxes from all intercell boundaries. This means, that in each time step, we need to solve the Riemann problem at all cell interfaces, and then use the fluxes to update the cell averages. The fluxes $\mathbf{F}_{i-1/2,j}$ and $\mathbf{F}_{i+1/2,j}$ are the average fluxes across the lines $x = x_{i-1/2}$ and $x = x_{i+1/2}$. Correspondingly, the fluxes $\mathbf{G}_{i,j-1/2}$ and $\mathbf{G}_{i,j+1/2}$ are the average fluxes across the lines $y = y_{j-1/2}$ and $y = y_{j+1/2}$. We work on a domain discretized into rectangular cells of size $\Delta x \times \Delta y$, where $\Delta x = x_{j+1/2} - x_{j-1/2}$ and $\Delta y = y_{j+1/2} - y_{j-1/2}$. The fluxes are illustrated in Figure 3.3.

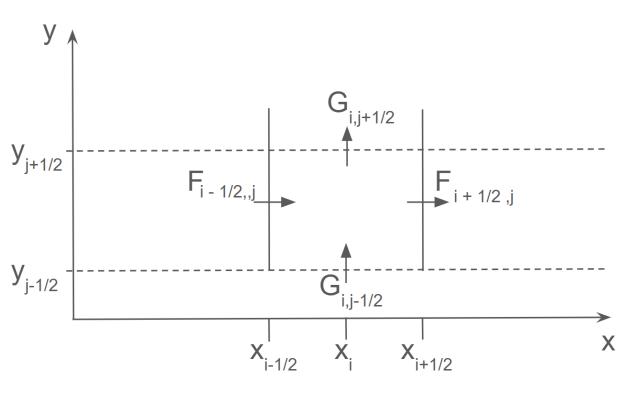


Figure 3.3: Illustration of the fluxes for the 2D SWE.

When using the simultaneous update scheme (3.2.2), we must consider potential stability issues. To address this, several methods can be employed.

The first method we consider is the weighted average flux (WAF) scheme. The basic WAF scheme, without any nonlinear total variation diminishing (TVD) modification, is prone to oscillations. To mitigate this, a TVD constraint can be enforced. A numerical method $U_{i,j}^{n+1}$ is called TVD if it satisfies the following condition [30]:

$$TV(U^{n+1}) \leq TV(U^n), \quad (3.2.3)$$

for all grid functions U^n . For a discrete function $U_{i,j}$ defined on a grid $i = 0, 1, \dots, N$ and $j = 0, 1, \dots, N$, the total variation $TV(U)$ is defined as

$$TV(U) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |U_{i+1,j} - U_{i,j}| + |U_{i,j+1} - U_{i,j}|. \quad (3.2.4)$$

The total variation measures the sum of the absolute differences between neighboring grid points. A TVD scheme ensures that the total variation of the solution does not increase in time, which is crucial for maintaining stability and preventing oscillations. The TVD property is used in shock-capturing schemes to prevent nonphysical oscillations near discontinuities, helping to accurately capture the shock location. To achieve the TVD condition,

additional constraints are imposed on the numerical flux or reconstruction process, often through the use of slope limiters. A limiter function is a technique used to reduce oscillations in the solution, especially near sharp edges or discontinuities, such as shock waves. For this project, we use the minmod limiter, defined as

$$\text{minmod}(a, b) = \begin{cases} \max(0, \min(a, b)) & \text{if } a > 0 \\ \min(0, \max(a, b)) & \text{if } a < 0. \end{cases}$$

Another method we consider is the MUSCL scheme. This method is second-order accurate in both time and space, while the WAF scheme is second-order accurate in space but relies on the time-stepping scheme for temporal accuracy. The MUSCL scheme, combined with slope limiters, ensures the solution satisfies the TVD constraint, effectively preventing oscillations that could occur with high-order spatial discretization near shocks, discontinuities, or sharp changes in the solution. The key idea in MUSCL is that, instead of using cell averages as in the 1D FVM, we reconstruct the solution in each cell using piecewise linear functions. This reconstruction provides a higher-order representation of the solution within each cell. In this project, we use the MUSCL scheme with the TVD criteria to ensure smooth and stable solutions.

3.3 The Finite Volume Method for the spherical 1D Linearized Shallow Water Equations

In this section, we derive the FVM for the 1D LSWE on a sphere, by following the same method as in section 3.1. We begin by stating the integral form of the 1D LSWE in spherical coordinates on a global domain, meaning we now consider a control volume given by

$$V_i^n = [\theta_{i-1/2}, \theta_{i+1/2}] \times [t_n, t_{n+1}],$$

where $\Delta\theta = \theta_{i+1/2} - \theta_{i-1/2}$ is the length of the cell and $\Delta t = t_{n+1} - t_n$ is the time step size. We rewrite the previous derived integral form (2.6.10) to be in global variables:

$$\int_{\theta_{i-1/2}}^{\theta_{i+1/2}} \mathbf{W}(\theta, t_{n+1}) d\theta = \int_{\theta_{i-1/2}}^{\theta_{i+1/2}} \mathbf{W}(\theta, t_n) d\theta - \mathbf{A} \left(\int_{t_n}^{t_{n+1}} \mathbf{W}(\theta_{i+1/2}, t) dt - \int_{t_n}^{t_{n+1}} \mathbf{W}(\theta_{i-1/2}, t) dt \right). \quad (3.3.1)$$

We divide the integral form (3.3.1) with the cell length $\Delta\theta$ to obtain

$$\begin{aligned} \frac{1}{\Delta\theta} \int_{\theta_{i-1/2}}^{\theta_{i+1/2}} \mathbf{W}(\theta, t_{n+1}) d\theta &= \frac{1}{\Delta\theta} \int_{\theta_{i-1/2}}^{\theta_{i+1/2}} \mathbf{W}(\theta, t_n) d\theta \\ &\quad - \frac{\Delta t}{\Delta\theta} \mathbf{A} \left(\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{W}(\theta_{i-1/2}, t) dt - \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} \mathbf{W}(\theta_{i+1/2}, t) dt \right). \end{aligned}$$

Averaging over the terms for a finite volume gives the first-order explicit time-stepping finite volume scheme:

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n - \frac{\Delta t}{\Delta\theta} (\mathbf{F}_{i+1/2}^n - \mathbf{F}_{i-1/2}^n). \quad (3.3.2)$$

The scheme uses the cell averages:

$$\mathbf{W}_i^n = \frac{1}{\Delta\theta} \int_{\theta_{i-1/2}}^{\theta_{i+1/2}} \mathbf{W}(\theta, t_n) d\theta$$

The flux $\mathbf{F}_{i-1/2}^n$ is the average flux across the line $\theta = \theta_{i-1/2}$ from time t_n to t_{n+1} :

$$\mathbf{F}_{i-1/2}^n = \frac{1}{\Delta t} \mathbf{A} \int_{t_n}^{t_{n+1}} (\mathbf{W}(\theta_{i-1/2}, t)) dt,$$

and correspondingly the flux $\mathbf{F}_{i+1/2}^n$ is the average flux across the line $\theta = \theta_{i+1/2}$ from time t_n to t_{n+1} :

$$\mathbf{F}_{i+1/2}^n = \frac{1}{\Delta t} \mathbf{A} \int_{t_n}^{t_{n+1}} (\mathbf{W}(\theta_{i+1/2}, t)) dt.$$

The Explicit Runge-Kutta 4th Order Method

We will now introduce how we can use the explicit Runge-Kutta 4th order (ERK4) method to solve the 1D spherical LSWE. The ERK4 method solves the general initial value problem (IVP) for ordinary differential equations (ODEs):

$$\frac{dy(t)}{dt} = f(y(t), t), \quad y(t_0) = y_0. \quad (3.3.3)$$

To solve the IVP (3.3.3) using ERK4 we choose a step size $\Delta t > 0$ and define the update formula:

$$\begin{aligned} y_{n+1} &= y_n + \frac{\Delta t}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ t_{n+1} &= t_n + \Delta t, \quad n = 0, 1, \dots, N - 1, \end{aligned}$$

where the stages k_1, k_2, k_3 and k_4 are given by [31]:

$$\begin{aligned} k_1 &= f(t_n, y_n), \\ k_2 &= f\left(t_n + \frac{\Delta t}{2}, y_n + \frac{\Delta t}{2}k_1\right), \\ k_3 &= f\left(t_n + \frac{\Delta t}{2}, y_n + \frac{\Delta t}{2}k_2\right), \\ k_4 &= f(t_n + \Delta t, y_n + \Delta t k_3). \end{aligned}$$

This means that to use ERK4 to solve the 1D spherical LSWE, we must bring the equations in the same form as the IVP (3.3.3). Recall, that the 1D LSWE in spherical coordinates are given by

$$\left. \begin{aligned} \frac{\partial h'}{\partial t} &= -\frac{h_0}{r \cos \phi} \frac{\partial u}{\partial \theta}, \\ \frac{\partial u}{\partial t} &= -\frac{g}{r \cos \phi} \frac{\partial h'}{\partial \theta}, \end{aligned} \right\} \quad (3.3.6)$$

where h' is the perturbation in water height, u is the velocity in the θ -direction, h_0 is the mean water depth, g is the gravitational acceleration, r is the Earth's radius, and ϕ is the fixed latitude. Since the ERK4 method solves ODEs, and the LSWE are PDEs, we must use the method of lines (MOL) to bring the LSWE in the form of the IVP (3.3.3). The MOL is a technique for solving PDEs by discretizing all variables but one (usually time), transforming the PDE into a system of ODEs. The system of ODEs can then be solved using standard ODE solvers, such as the ERK4 method. We discretize the spatial dimension θ into N cells, and approximate the spatial derivatives using central differences:

$$\left. \begin{aligned} \frac{\partial u}{\partial \theta} &\approx \frac{u_{i+1} - u_{i-1}}{2\Delta\theta}, \\ \frac{\partial h'}{\partial \theta} &\approx \frac{h'_{i+1} - h'_{i-1}}{2\Delta\theta}, \end{aligned} \right\} \quad (3.3.7)$$

for $i = 0, 1, \dots, N - 1$. To perform the time integration, we use the ERK4 method. This method is particularly effective for time-stepping and is an alternative approach to traditional time integration schemes. The ERK4 method provides higher accuracy in time while relying on the same spatial discretization, such as the FVM. Substituting the spatial derivatives (3.3.7) into the LSWE (3.3.6) gives the system of ODEs:

$$\begin{aligned} \frac{dh'}{dt} &= -\frac{h_0}{r \cos \phi} \frac{u_{i+1} - u_{i-1}}{2\Delta\theta}, \\ \frac{du}{dt} &= -\frac{g}{r \cos \phi} \frac{h'_{i+1} - h'_{i-1}}{2\Delta\theta}, \end{aligned} \quad (3.3.8)$$

for $i = 0, 1, \dots, N - 1$. Hence, we can write the 1D spherical LSWE in the form of the IVP (3.3.3) by defining

$$y(t) = \begin{bmatrix} h' \\ u \end{bmatrix}, \quad f(y(t), t) = \begin{bmatrix} -\frac{h_0}{r \cos \phi} \frac{u_{i+1} - u_{i-1}}{2\Delta\theta} \\ -\frac{g}{r \cos \phi} \frac{h'_{i+1} - h'_{i-1}}{2\Delta\theta} \end{bmatrix} \quad \text{for } i = 0, 1, \dots, N - 1. \quad (3.3.9)$$

As we consider a circle as the domain, we must impose periodic boundary conditions. We do this by setting $h'_{-1} = h'_{N-1}$, $h'_N = h'_0$, $u_{-1} = u_{N-1}$, and $u_N = u_0$. The initial conditions for y are a Gaussian function in water height and zero velocity:

$$y_0 = \begin{bmatrix} \exp(-\frac{(\theta_i - \theta_0)^2}{\sigma^2}) \\ 0 \end{bmatrix}, \quad i = 0, 1, \dots, N-1. \quad (3.3.10)$$

where θ_0 is the center of the Gaussian, and σ is the width of the Gaussian. This means, that we can employ the four-stage time-stepping scheme in the ERK4 method to update the solution for h' and u . The final values of h' and u are obtained by combining the contributions from all four stages, by using the general update formula:

$$\mathbf{W}_i^{n+1} = \mathbf{W}_i^n + \frac{\Delta t}{\Delta \theta} \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4),$$

where the state variable $\mathbf{W}_i = \begin{bmatrix} h'_i \\ u_i \end{bmatrix}$ represents the perturbation height and velocity at the i 'th cell. The intermediate stages k_1, k_2, k_3 and k_4 are computed at each time step. The scheme integrates the equations in time while relying on flux differences to update the solution. This approach allows the method to achieve high temporal accuracy while maintaining the spatial resolution provided by the finite volume discretization.

3.4 The Riemann problem

We will now define the Riemann problem, since it plays a crucial role in the finite volume method. The Riemann problem for the SWE with a zero source term, i.e. $\mathbf{S}(\mathbf{U}) = 0$, is defined as the initial-value problem (IVP) [7]:

$$\begin{aligned} \text{PDEs: } & \mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = 0, \\ \text{ICs: } & \mathbf{U}(x, 0) = \begin{cases} \mathbf{U}_L, & \text{if } x < x_0, \\ \mathbf{U}_R, & \text{if } x > x_0. \end{cases} \end{aligned} \quad (3.4.1)$$

The vectors \mathbf{U} and $\mathbf{F}(\mathbf{U})$ in (3.4.1) are given by

$$\mathbf{U} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ hvu \end{bmatrix}, \quad (3.4.2)$$

and the initial conditions \mathbf{U}_L and \mathbf{U}_R are

$$\mathbf{U}_L = \begin{bmatrix} h_L \\ h_L u_L \\ h_L v_L \end{bmatrix}, \quad \mathbf{U}_R = \begin{bmatrix} h_R \\ h_R u_R \\ h_R v_R \end{bmatrix},$$

which represents the conditions at time $t = 0$ s in the left and right states of $x = x_0$. The function \mathbf{U} is piecewise constant, with a discontinuity at $x = x_0$. The Riemann problem is solved at the interface between two cells, where the left and right states are known, and the solution is used to update the solution at each timestep in the FVM. The Riemann problem can be solved either exactly or approximately. When solving the Riemann problem exactly we distinguish between what we call a wet bed and a dry bed. A wet bed is the case where the water depth is positive everywhere, whereas a dry bed is the case where the water depth is zero in some cells. In this project, we focus on approximate Riemann solvers, which are able to solve the Riemann problem with high accuracy and efficiency. Various approximate Riemann solvers exist, based on finding an approximate solution to the Riemann problem, and some of these solvers will be considered in the next section. An interesting example of a Riemann problem is the so-called dam break problem, which is presented next.

3.4.1 The Dam-Break problem

We now introduce the dam-break problem, a scenario of significant physical interest. This problem models the sudden release of water following the collapse of a dam, making it highly relevant for studying natural disasters such as floods and tsunamis. The problem also serves as a classic test case for numerical methods, to test their ability to capture discontinuities in the solution. The dam-break problem is a special case of the Riemann problem (3.4.1). The difference is that in the dam-break problem, the initial velocity components, u_L, u_R, v_L and v_R , are zero, whereas in the Riemann problem they are allowed to be distinct from zero. The initial setup is visualized in Figure 3.4.

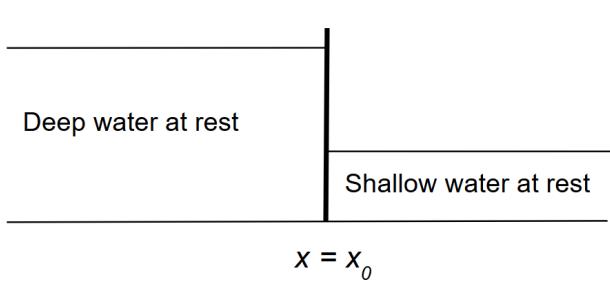


Figure 3.4: Initial conditions for the dam-break problem. An infinitely thin wall at $x = x_0$ divides two sections of water with different water levels. Illustration modified from [7].

We can use the SWE to model the flow of water in the dam-break problem, approximately, if we assume that the wall collapses instantaneously at $t = 0$ s. In this project we solve both the dam-break problem and cases of the Riemann problem, where the initial fluid velocity is nonzero. The results are presented in chapter 6.

3.4.2 Wave decomposition in the Riemann problem

To get a better understanding of the flow in shallow water, we provide some very short background information about the wave structures in the solution of the Riemann problem. In general, the wave structure in the solution of the Riemann problem (3.4.1) consists of three wave families separating four regions. The wave families are denoted W_1, W_2, W_3 and the regions are the spaces between the wave families, denoted by R_0, R_1, R_2 and R_3 . The wave structure is illustrated in Figure 3.5.

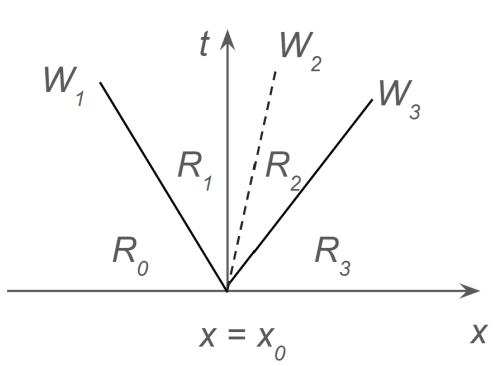


Figure 3.5: General wave structure in the solution of the Riemann problem. The wave families are denoted by W_1, W_2 and W_3 , and the regions between the wave families are denoted by R_1, R_2, R_3 and R_4 . The star region consists of the regions R_1 and R_2 .

From Figure 3.5 we see how the solution consists of three waves, W_1 , W_2 and W_3 , also referred to as the left wave, middle wave and right wave. The left and right waves are either shock waves or rarefaction waves, and correspond to the one-dimensional shallow water equations. The middle wave arises from the y -momentum equation in (3.4.1), and is always a shear wave. The region between the left and right wave is called the star region and is interesting, since we do not know the solution in this region. The star region is divided into two subregions R_1 and R_2 . The states in the regions are (from left to right) U_L, U_{*L}, U_{*R} and U_R , where U_L and U_R are known, as these are the initial conditions. The states U_{*L} and U_{*R} are in the regions R_1 and R_2 , i.e., the star region, and are unknown. In the star region, we use h_* to denote the water depth and u_* to denote the velocity. We can use h_* to determine whether the left and right waves are shock waves or rarefaction waves. Since we are considering the wet bed case, we can use the following characteristic:

$$\begin{cases} \text{The left wave is a shock wave} & \text{if } h_* > h_L, \\ \text{The left wave is a rarefaction wave} & \text{if } h_* \leq h_L, \end{cases} \quad (3.4.3)$$

and similarly for the right wave:

$$\begin{cases} \text{The right wave is a shock wave} & \text{if } h_* > h_R, \\ \text{The right wave is a rarefaction wave} & \text{if } h_* \leq h_R. \end{cases} \quad (3.4.4)$$

A shock wave is characterized by a discontinuity in the solution. On each side of the shock wave, the water properties, such as height and velocity, differ significantly. In contrast, a rarefaction wave represents a smooth transition in the solution. The water height and velocity change gradually across the wave, and the properties are more similar on each side of the wave, without the sharp edges seen in a shock wave. In general, the solution to the Riemann problem (3.4.1) can result in four possible wave patterns, which are combinations of shock waves and rarefaction waves. The four possible wave patterns are as follows:

- (a) Left rarefaction, right shock,
- (b) Left shock, right rarefaction,
- (c) Both left and right rarefaction,
- (d) Both left and right shock.

In the example of the dam-break problem, with initial conditions as in Figure 3.4, the solution for the water height h at time $t = t_*$ is shown in Figure 3.6.

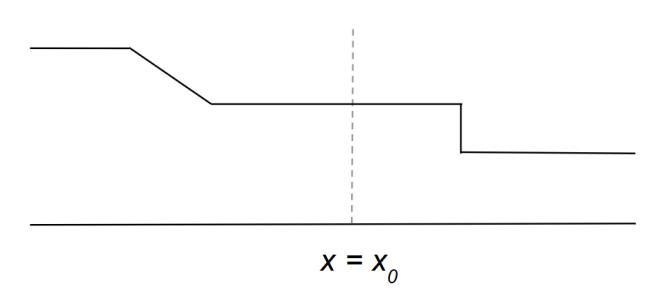


Figure 3.6: Water depth at time $t = t_*$ in the dam-break problem. The solution consists of a left rarefaction wave and a right shock wave.

From Figure 3.6 we see that the solution consists of a left rarefaction wave and a right shock wave. We can see that after the wall collapses, the water flows to the right, meaning the water height in the star/middle region,

h_* , will be between h_L and h_R . Based on the definitions (3.4.3) and (3.4.4), we can therefore conclude that the left wave is a rarefaction wave and the right wave is a shock wave. The right shock wave is characterized by a discontinuity in the solution and a high speed, whereas the left rarefaction wave is characterized by a more smooth transition in the solution and a lower speed.

In the finite volume framework, the Riemann solver is used to compute the flux across the interface between two cells, considering the left and right states. This way, we ensure that the numerical method handles the discontinuities in the solution correctly, and that the solution is stable and accurate.

3.5 Numerical fluxes

In this section we will study the numerical fluxes used to solve the SWE. At each cell interface, we need to solve the Riemann problem (3.4.1) to find the numerical flux. There are several numerical fluxes that can be used to solve the local Riemann problem, and we will consider some of them in this section. The fluxes we will consider are the Godunov method with an exact Riemann solver, the HLL, HLLC, Rusanov, Lax-Friedrichs, Lax-Wendroff and FORCE fluxes. All of them are later implemented and tested in the numerical experiments in chapter 6.

Godunov method with exact Riemann solver

We consider the Godunov upwind method, which is a first-order accurate method to solve non-linear systems of hyperbolic conservation laws [7]. In the method we solve the non-linear Riemann problem at each cell interface. The Godunov flux is given by

$$\mathbf{F}_{i+\frac{1}{2}} = \mathbf{F}(\mathbf{U}_{i+\frac{1}{2}}),$$

meaning that we solve the Riemann problem exactly to find h^* and u^* , and then use these values to compute the flux as

$$\mathbf{F}_{i+\frac{1}{2}} = \begin{bmatrix} h^* u^* \\ h^*(u^*)^2 + \frac{1}{2} g(h^*)^2 \end{bmatrix}.$$

HLL

The HLL (Harten, Lax and van Leer) approach assumes a two-wave structure of the Riemann problem. The solver is based on the data $\mathbf{U}_L := \mathbf{U}_i^n$, $\mathbf{U}_R := \mathbf{U}_{i+1}^n$ and fluxes $\mathbf{F}_L := \mathbf{F}(\mathbf{U}_L)$, $\mathbf{F}_R := \mathbf{F}(\mathbf{U}_R)$. The HLL flux is given by

$$\mathbf{F}_{i+\frac{1}{2}} = \begin{cases} \mathbf{F}_L & \text{if } S_L \geq 0, \\ \mathbf{F}^{HLL} \equiv \frac{S_R \mathbf{F}_L - S_L \mathbf{F}_R + S_L S_R (\mathbf{U}_R - \mathbf{U}_L)}{S_R - S_L} & \text{if } S_L \leq 0 \leq S_R, \\ \mathbf{F}_R & \text{if } S_R \leq 0. \end{cases} \quad (3.5.1)$$

The wave speeds S_L and S_R must be estimated in some way, and one possibility is to use

$$S_L = u_L - a_L q_L, \quad S_R = u_R + a_R q_R,$$

where $a_L = \sqrt{gh_L}$, $a_R = \sqrt{gh_R}$ and $q_K (K = L, R)$ is given by

$$q_K = \begin{cases} \sqrt{\frac{1}{2} \left(\frac{(\hat{h} + h_K) \hat{h}}{h_K^2} \right)} & \text{if } \hat{h} > h_K, \\ 1 & \text{if } \hat{h} \leq h_K. \end{cases}$$

Here \hat{h} is an estimate for the water depth in the star region, h_* . In the two-rarefaction Riemann solver, the water depth h in the star region is given by

$$h_* = \frac{1}{g} \left(\frac{1}{2}(a_L + a_R) + \frac{1}{4}(u_L - u_R) \right)^2, \quad (3.5.2)$$

which is what we use in this project for \hat{h} in the HLL solver. Since this is a two-wave model, it is complete for one dimensional problems, but for the augmented system of equations in two dimensions, the HLL solver is not complete, as it ignores the middle wave, the shear wave. This motivates the use of the HLLC solver, which is a modification of the HLL solver.

HLLC

The HLLC (Harten, Lax, van Leer, Contact) solver is an extension of the HLL solver, which includes the middle wave, i.e., it is a three-wave model. In addition to the wave speeds S_L and S_R , the HLLC solver also requires the speed of the middle wave S^* . We can write the HLLC numerical flux as

$$\mathbf{F}_{i+\frac{1}{2}}^{HLLC} = \begin{cases} \mathbf{F}_L & \text{if } 0 \leq S_L, \\ \mathbf{F}_{*L} & \text{if } S_L \leq 0 \leq S^*, \\ \mathbf{F}_{*R} & \text{if } S^* \leq 0 \leq S_R, \\ \mathbf{F}_R & \text{if } S_R \leq 0. \end{cases}$$

The fluxes \mathbf{F}_{*L} and \mathbf{F}_{*R} are given by

$$\begin{aligned} \mathbf{F}_{*L} &= \mathbf{F}_L + S_L(\mathbf{U}_L - \mathbf{U}_{*L}), \\ \mathbf{F}_{*R} &= \mathbf{F}_R + S_R(\mathbf{U}_R - \mathbf{U}_{*R}), \end{aligned}$$

and the middle states \mathbf{U}_{*L} and \mathbf{U}_{*R} are given by

$$\mathbf{U}_{*K} = h_K \left(\frac{S_K - u_K}{S_K - S_*} \right) \begin{bmatrix} 1 \\ S_* \\ \psi_K \end{bmatrix}.$$

The function ψ_K can represent either a passive scalar $\psi(x, t)$ or the velocity component $v(x, t)$ if we consider the two-dimensional shallow water equations. Mathematically $\psi(x, t)$ and $v(x, t)$ behave identically. An estimate for the middle wave speed S^* can be calculated as

$$S^* = \frac{S_L h_R (u_R - S_R) - S_R h_L (u_L - S_L)}{h_R (u_R - S_R) - h_L (u_L - S_L)},$$

where S_L and S_R are the wave speeds of the left and right waves, respectively.

Rusanov

The Rusanov flux uses the HLL framework, but with a different choice of wave speeds. To obtain the flux, we assume that an estimate S^+ for the positive wave speed is available. Then we set

$$S_L = -S^+, \quad S_R = S^+. \quad (3.5.3)$$

By substituting (3.5.3) into the \mathbf{F}^{HLL} in (3.5.1), we obtain the Rusanov flux as

$$\mathbf{F}_{i+\frac{1}{2}}^{Rus} = \frac{1}{2} (\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2} S^+ (\mathbf{U}_R - \mathbf{U}_L), \quad (3.5.4)$$

where a simple estimate for the wave speed S^+ is given by

$$S^+ = \max(|S_L|, |S_R|).$$

There are some requirement for S^+ in (3.5.4) to ensure stability. It must hold that

$$S^+ \leq \frac{\Delta x}{\Delta t},$$

where $\frac{\Delta x}{\Delta t}$ is called the mesh speed.

Lax-Friedrichs

In the Lax-Friedrichs method, we use the Rusanov flux, but with a different choice of wave speed. That is, we set the wave speed S^+ as the largest possible speed, while still ensuring stability, i.e.,

$$S^+ = \frac{\Delta x}{\Delta t}. \quad (3.5.5)$$

By inserting the wave speed (3.5.5) into the Rusanov flux (3.5.4), we obtain the Lax-Friedrichs flux as

$$\mathbf{F}_{i+\frac{1}{2}}^{LF} = \frac{1}{2} (\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2} \frac{\Delta x}{\Delta t} (\mathbf{U}_R - \mathbf{U}_L),$$

where $\mathbf{F}_L = \mathbf{F}(\mathbf{U}_L)$ and $\mathbf{F}_R = \mathbf{F}(\mathbf{U}_R)$. The Lax-Friedrichs method is a centred method, which is first-order accurate.

Lax-Wendroff

There are several versions of the Lax-Wendroff flux, but in this thesis we will use the following flux:

$$\begin{aligned} \mathbf{U}_{i+\frac{1}{2}}^{LW} &= \frac{1}{2} (\mathbf{U}_L + \mathbf{U}_R) - \frac{1}{2} \frac{\Delta t}{\Delta x} (\mathbf{F}_R - \mathbf{F}_L), \\ \mathbf{F}_{i+\frac{1}{2}}^{LW} &= \mathbf{F}(\mathbf{U})_{i+\frac{1}{2}}^{LW}. \end{aligned} \quad (3.5.6)$$

The Lax-Wendroff method is a centred method, which is second-order accurate in space and time.

FORCE

The FORCE scheme (First-Order Centred) is a combination of Lax-Friedrichs and Lax-Wendroff fluxes. The numerical flux is given by

$$\mathbf{F}_{i+\frac{1}{2}}^{FO} = \frac{1}{2} \left(\mathbf{F}_{i+\frac{1}{2}}^{LF} + \mathbf{F}_{i+\frac{1}{2}}^{LW} \right),$$

where $\mathbf{F}_{i+\frac{1}{2}}^{LF}$ is the Lax-Friedrichs flux and $\mathbf{F}_{i+\frac{1}{2}}^{LW}$ is the Lax-Wendroff flux. The FORCE scheme is first-order accurate. It is possible to extend the FORCE scheme to multiple dimensions on structured meshes by using dimensional splitting.

Chapter 4

Neural Networks and Fourier Neural Operators

So far, we have studied numerical methods for solving partial differential equations (PDEs). The finite volume method (FVM), along with other numerical solvers such as the finite difference method (FDM) and finite element method (FEM), solves PDEs by discretizing the domain into a grid. A finer grid improves the accuracy of the solution but also increases the computational cost, creating a trade-off between accuracy and efficiency. Complex PDEs often require a fine grid to accurately capture the solution, which can be computationally expensive. This motivates the use of data-driven methods, which have shown great potential in handling large data sets and learning complex patterns in the data.

In this chapter, we introduce the use of data-driven methods for solving PDEs. We will introduce the concepts of convolutional neural networks (CNNs) and Fourier neural operators (FNOs). A basic understanding of neural networks is assumed, for example through courses in machine learning or deep learning. Both methods rely solely on data, rather than the PDE itself, which is particularly useful when the governing PDE is unknown. The hope of data-driven methods is that they are able to reduce computational costs while maintaining high accuracy by learning the underlying dynamics of the solution. This could also help making predictions for the solution of the shallow water equations (SWE) outside the training data.

4.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a specialized class of artificial neural networks (ANNs) designed to process and analyze data with a grid-like topology, such as images or time-series data represented by 2D grids. CNNs excel at extracting spatial features from data through the use of convolutional layers, which apply learnable filters to detect patterns such as edges, shapes or textures. These layers are typically followed by pooling layers for dimensionality reduction and fully connected layers for classification or regression tasks. A key advantage of CNNs is their ability to reduce the number of parameters compared to fully connected neural networks by sharing weights across spatial regions, making them computationally efficient and less prone to overfitting when working with large inputs. Although CNNs are traditionally used for image recognition tasks, their architecture is adaptable for time-series analysis, especially when the data is structured spatially or sequentially [17].

In this project, CNNs are used to solve the SWE, by training on the solution data generated by the FVM solver. We aim to learn the underlying dynamics of the system by training the CNN on sequences of input-output pairs, where the input $\mathbf{x}(t)$ is the state of the system at time t and the output is the state at time $t + \Delta t$, denoted by

$\mathbf{x}(t + \Delta t)$. More specific, this means that the CNN model is trained on input-output pairs of sequences of data, with a sequence length of n , illustrated in the vectors below:

$$\begin{bmatrix} x_i \\ x_{i+1} \\ \vdots \\ x_{i+n-1} \end{bmatrix} \rightarrow \begin{bmatrix} x_{i+1} \\ x_{i+2} \\ \vdots \\ x_{i+n} \end{bmatrix}, \text{ for } i = 0, 1, \dots, N - n,$$

where x_i is the state of the system at time i . The input-output pairs are designed to train the CNN to predict the system's state at the next time step, based on the current state and the $n - 1$ previous states. Hence, the network is trained to construct a flow map, which is a mapping from the current state to the next state. The flow map $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined such that, for all $x \in X$, where X is the solution space, and all $t \in \mathbb{R}$:

$$\Phi(x_0) = x(t),$$

where $x(t)$ is the solution of the PDEs with initial condition $x(0) = x_0$. The flow map satisfies

$$\Phi(\Phi(x, t), s) = \Phi(x, s + t),$$

A CNN can approximate the flow map Φ by training on data that pairs a state $\mathbf{x}(t)$ with its state at later times $\mathbf{x}(t + \Delta t)$. The goal is to learn the mapping:

$$x_{i+1} = \Phi_{CNN}(x_i),$$

where Φ_{CNN} is the CNNS' approximation of the flow map. The network process a sequence of input data to predict the corresponding output data. The CNN's output is the solution at the next time step, effectively learning the dynamics of the SWE through the time-series data. This setup allows the model to capture both spatial and temporal dependencies in the data, leveraging the CNN's ability to learn localized features while processing sequential information. An advantage of CNNs is their efficiency. By processing data in parallel using convolutional layers, the CNN efficiently handles large datasets without requiring excessive computational resources. Another strength is their adaptability. The model's ability to learn from sequential data makes it adaptable for time-series predictions in dynamic systems like the shallow water equations. However, a drawback is the challenge of data representation. Representing time-series data as sequences may require preprocessing, which can introduce complexity or a potential loss of information. The results of the CNN models will be presented in chapter 7.

4.2 Fourier Neural Operators

In this section, we introduce the concept of Fourier Neural Operators (FNOs), based on the method and theory presented in [18]. FNOs are distinctive since, unlike classical neural networks that primarily learn mappings between finite-dimensional spaces, they approximate mappings between infinite-dimensional function spaces. The goal of an FNO is to learn a mapping between two infinite-dimensional spaces from a finite collection of input-output pairs. Consider the operator $G : A \rightarrow U$, which maps functions from an infinite-dimensional function space A to another infinite-dimensional function space U . Our objective is to approximate the exact operator G by constructing the map

$$G_\theta : A \mapsto U, \quad \theta \in \Theta, \tag{4.2.1}$$

where Θ is a finite-dimensional parameter space. Let $a \in A$ and $u \in U$ represent the input and output functions, respectively. We assume access to data in the form of pointwise evaluations of these functions, i.e., we have access to the observations $\{a_j, u_j\}_{j=0}^{N-1}$, in a domain $D \subset \mathbb{R}$, which is a bounded open set. The observations are time shifted, meaning $u_j = G(a_j)$, where G is the true operator, and $u_j = a_{j+1}$ for all $j = 0, 1, \dots, N - 1$. Thus, the goal is to approximate the mapping $a_j \mapsto u_j$ for all $j = 0, 1, \dots, N - 1$.

The process begins with the input layer P , typically a shallow fully connected neural network, which generates $v_0(x) = P(a(x))$. The FNO is an iterative neural operator, applying several updates to compute intermediate representations v_1, v_2, \dots, v_T . In the context of this project, we have that $a(x) = v_t(x)$ and $u(x) = v_{t+1}(x)$. An update $v_t \mapsto v_{t+1}$ is defined as

$$v_{t+1}(x) := \sigma(Wv_t(x) + (\mathcal{K}(a; \phi)v_t)(x)), \quad \forall x \in D, \quad (4.2.2)$$

where $W : \mathbb{R} \rightarrow \mathbb{R}$ is a linear transformation, $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ is a non-linear activation function, and $\mathcal{K}(a; \phi)$ is a kernel function parameterized by ϕ . The final update $v_T(x)$ is transformed by the output layer Q to produce the output $u(x) = Q(v_T(x))$, ensuring the correct output dimensions. Similar to the CNN approach, state pairs $\{v_t, v_{t+1}\}_{j=0}^N$ are collected for training the flow map, represented by G_θ . Feeding this data into the FNO model enables it to learn the flow map for the system.

There are various types of operators, but the core of the Fourier neural operator lies in its kernel function \mathcal{K} . We define the Fourier integral operator \mathcal{K} as

$$(\mathcal{K}(\phi)v_t)(x) := \mathcal{F}^{-1}(R_\phi \cdot (\mathcal{F}v_t))(x), \quad \forall x \in D, \quad (4.2.3)$$

where \mathcal{F} is the Fourier transform, \mathcal{F}^{-1} is the inverse Fourier transform, and R is the linear transformation applied on the lower Fourier modes. Recall that the Fourier transform of a function $f(t)$ is defined as [32]:

$$\mathcal{F}(f(t)) = \hat{f}(s) = \int_{-\infty}^{\infty} e^{-2\pi i s t} f(t) dt, \quad (4.2.4)$$

and correspondingly the inverse Fourier transform of a function is

$$\mathcal{F}^{-1}(\hat{f}(s)) = f(t) = \int_{-\infty}^{\infty} e^{2\pi i s t} \hat{f}(s) ds.$$

By transforming the data into the Fourier domain, FNOs can take advantage of the periodicity and smoothness properties of the Fourier basis, which simplifies the learning process for functions defined over continuous domains. The network architecture for the FNO model is illustrated in Figure 4.1.

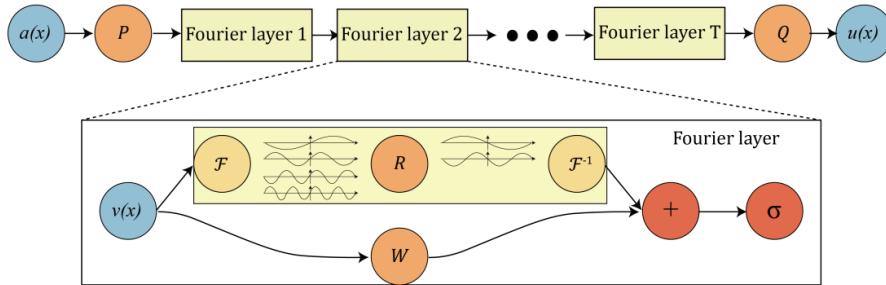


Figure 4.1: An overview of the network architecture with several Fourier layers. Illustration from [18].
Top: Overall structure with input function a , input layer P , several Fourier layers, output layer Q and output function u .

Bottom: A Fourier layer consists of two parallel paths. Top is a Fourier transformation \mathcal{F} , a linear transformation R and an inverse Fourier transformation \mathcal{F}^{-1} . The bottom path consists of a linear transformation W . The parts meet and undertake an activation function σ .

From the top in Figure 4.1 we see that the network consists of an input function $a(x)$, an input layer P , several Fourier layers, an output layer Q and some output function $u(x)$. In the bottom of Figure 4.1 we see the structure of a Fourier layer, which consists of two parallel paths. In the top path, the data undergoes a Fourier transform \mathcal{F} , decomposing it into a sum of Fourier basis functions (sines and cosines) with varying frequencies, amplitudes

and phases. A linear transformation R is then applied to filter out the higher Fourier modes, as illustrated in the figure, where high-frequency components are removed. When implementing the model, we choose the number of Fourier modes to retain, depending on how much information we want to preserve. Retaining more modes keeps more information, but may also introduce more noise and oscillations. After filtering, the inverse Fourier transform \mathcal{F}^{-1} is applied to reconstruct the data in its original form. The bottom path involves a linear transformation W , and the two paths merge before applying a non-linear activation function σ .

A key advantage of FNOs is their ability to learn mappings between function spaces, making them independent of any specific grid or mesh. This allows them to transfer solutions across different grids, a capability known as zero-shot super-resolution. For instance, the learned operator can generalize from a coarse grid to a fine grid without retraining. This feature significantly reduces the computational cost associated with fine-grid simulations. In chapter 7, we will evaluate the performance of the implemented FNO model in this context, testing its ability to maintain high accuracy when making predictions on finer grids.

Multistep prediction

Current literature highlights the strong performance of FNOs in long-term predictions. One potential application of the FNO model is predicting the solution of the SWE at multiple future time steps. We aim to develop a multi-step prediction model capable of forecasting a specified number of time steps ahead. In its original form, the FNO is designed to predict one time step ahead. However, for many practical applications, multi-step predictions are more useful. The first, naive approach to achieve this is to use the model iteratively: predicting the solution at the next time step, then using that prediction as input to predict the solution at the next time step, and so on. However, as one can imagine this approach quite fast lead to inaccurate predictions, as errors accumulate over time.

To improve the accuracy of multi-step predictions, we train the model on sequences of input data, where the input data consists of several previous time steps. The sequential data is formatted similarly to that used in the CNN, with input sequences of length n . The model is trained to predict the output state based on this sequence of previous time steps. We will conduct experiments to determine the optimal sequence length for our model. That is, we train the model on input-output pairs of sequences, each with a sequence length of n , as illustrated in the vectors below:

$$\begin{bmatrix} v_{t-n} \\ v_{t-n+1} \\ \vdots \\ v_{t-1} \\ v_t \end{bmatrix} \rightarrow \begin{bmatrix} v_{t-n+1} \\ v_{t-n+2} \\ \vdots \\ v_t \\ v_{t+1} \end{bmatrix}, \text{ for } t = 0, 1, \dots, N - n.$$

Training the model on sequences of input data allows it to learn the dynamics of the system over time, capturing both short-term and long-term dependencies in the data. It also makes the model more accurate by providing more context for making predictions, ensuring that recent predictions do not disproportionately influence future predictions. When making multi-step predictions, the model's output is added to the input as the newest data point, replacing the oldest data. This process is repeated until predictions for the desired number of future time steps are made. The process is illustrated in Figure 4.2.

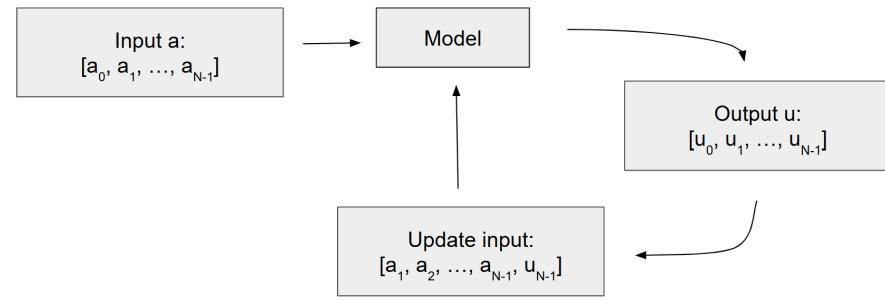


Figure 4.2: Flowchart of the multi-step prediction process.

In Figure 4.2, we present the flowchart of the multi-step prediction process. Starting with the input data a , the model makes predictions u , which are then used to update the input data for the next predictions. This process is applied to both the CNN and FNO models. The results of the multi-step prediction model will be presented in chapter 7.

Chapter 5

Data generation

In this chapter, we outline the process of generating the data required for the data-driven methods, which is a crucial part of the project. We generate all the data ourselves using our numerical solver, and the quality and relevance of the data directly impact the training and performance of the models. We begin by detailing the data generation process for the 1D SWE using the FVM with fluxes determined from an exact Riemann solver. Next, we describe the approach for generating data for the 1D LSWE in spherical coordinates. We also present the methodology for data generation in the context of the 2D SWE. Finally, we provide a brief introduction of how we plan to generate data for training a spherical Fourier neural operator (SFNO) for applications on a planetary scale.

5.1 Data generation using the Finite Volume Method

In this section we clarify the process of generating data using our numerical solver, that is, the FVM. The FVM is used to solve the 1D SWE, 1D LSWE on a sphere and the 2D SWE. We specify the necessary information, such as the initial conditions, the domain, and the parameters used in the data generation process.

1D SWE with exact Riemann solver

In this section, we present how the so-called true solution is found in the code by solving the Riemann problem exactly. The true solution is found by solving the Riemann problem exact using a high-resolution grid, and distinguishing between the wetbed or drybed case, and also identifying the shock and rarefaction waves. First we calculate the wave speeds for the left and right states, respectively, as

$$c_L = \sqrt{gh_L}, \quad c_R = \sqrt{gh_R},$$

which are used to determine the critical water height h_{crit} as

$$h_{\text{crit}} = (u_R - u_L) - 2(c_L + c_R).$$

If either $h_L \leq 0$ or $h_R \leq 0$, we are in a drybed case. If $h_{\text{crit}} \geq 0$, it indicates that the water depth is critical, and we are also in a drybed case. If none of the above conditions are met, we are in a wetbed case. Summarized:

$$\begin{cases} \text{Dry-bed case} & \text{if } h_L \leq 0, \quad h_R \leq 0 \text{ or } h_{\text{crit}} \geq 0, \\ \text{Wet-bed case} & \text{otherwise.} \end{cases}$$

In a dry bed case, we first identify the location of the dry region, whether it is on the left, right, or in the middle, and then calculate the wave speeds accordingly. In a wet bed case, we compute the characteristics h_* and u_* for the star region. We then identify the shock and rarefaction waves, and calculate the wave speeds for the left and right states, respectively. The process is illustrated in Figure 5.1.

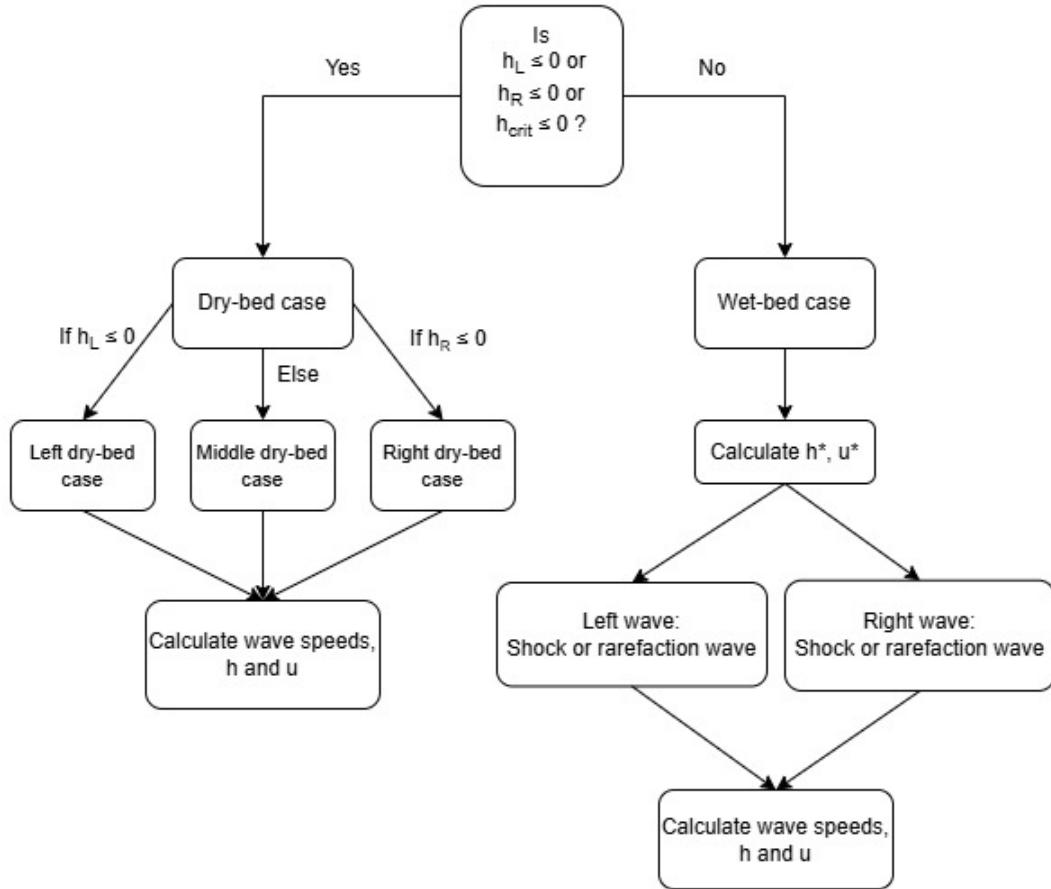


Figure 5.1: Flowchart for generating the solution.

When generating the data, we need an initial condition for the water height h and the velocity u . In this study, we use the Gaussian function as the initial condition for the water height h . That is, we define the initial condition for the water height as

$$h(x, 0) = a \exp\left(\frac{-(x - \mu)^2}{2\sigma^2}\right), \quad (5.1.1)$$

where a is the amplitude of the Gaussian, μ is the center of the distribution, and σ is the standard deviation. We are working on the domain $x \in [0, 1]$ m and the parameters are set to $a = 1$, $g = 9.81 \text{m/s}^2$ and $\sigma = 0.1$ m. The value of μ is varied to generate different initial conditions, we generate data for $\mu = 0.3$ m and $\mu = 0.5$ m. The initial conditions for the water height can be seen in Figure 5.2.

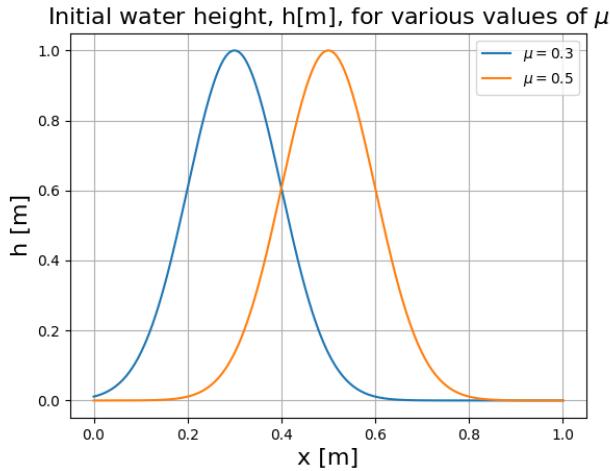


Figure 5.2: The generated data has a Gaussian distribution for the initial water height, with $\mu = 0.3$ m and $\mu = 0.5$ m.

For the initial velocity u , we set it to zero, i.e., $u(x, 0) = 0$ m/s, meaning that the water is initially at rest. The solver is validated by comparing the results with known test cases, such as the dam break problem. We use a CFL number of 0.9 and variable time steps. The data is generate over the time interval $t = 0.0$ to $t_{\text{end}} = 1.0$ s.

Truncation error

When generating data, it is essential to be aware of the truncation error. Truncation errors arises from approximating the solution of the PDEs using a numerical method, in this case the FVM. It specifically refers to the difference between the exact solution and the numerical approximation. The critical question is: after a certain number of time steps, how significant is this error? If the error becomes too large, it raises concerns about the reliability of the generated data. Excessive truncation error could compromise the accuracy of the model trained on this data. Therefore, we must carefully evaluate and mitigate these risks to ensure the data's quality. To assess the truncation error, we generate a more accurate solution using a finer grid. This high-resolution solution serves as a reference for evaluating the numerical approximation. By comparing the high-resolution solution with the numerical solution, we gain insights into the error introduced by the approximation.

We generate a reference solution for solving the 1D SWE using $N = 1000$ grid points and compare it with the solution for $N = 100$ grid points at the final time step, $t = 1.0$ s. The results are shown in Figure 5.3.

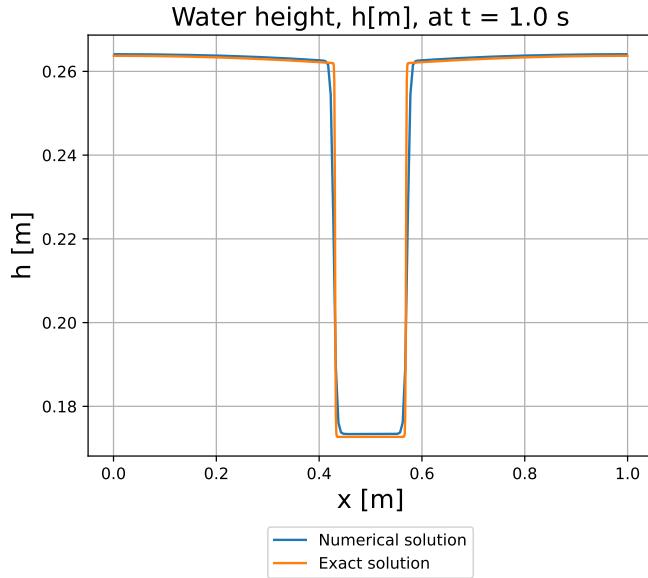


Figure 5.3: Truncation error for the 1D SWE.

The high-resolution solution (labelled as exact solution) and the low-resolution solution (labelled as numerical solution) are plotted in Figure 5.3. We observe that there is a small difference between the high-resolution solution and the numerical solution, as the high-resolution solution has a more discontinuous behavior. However, overall the two solutions are almost identical, indicating that the truncation error is negligible. This suggests that the data generated using the FVM is of high quality and can be used for training the data-driven models.

1D LSWE on a sphere

We also consider the 1D linearized shallow water equations in spherical coordinates, focusing on a circular domain. The length of the domain corresponds to a full circle, $L = 2\pi$ radians and is discretized into $N = 500$ points. The initial condition for the water height h is specified as a Gaussian function wrapped around the circle, expressed as:

$$h(\theta, 0) = a \exp\left(\frac{-(\theta - \mu)^2}{2\sigma^2}\right), \quad (5.1.2)$$

where h is the water height in meters, a is the amplitude of the Gaussian, μ is the mean value, and σ is the standard deviation. The parameters are $a = 1$ and $\mu = \frac{\pi}{4}$ radians. We generate data for varying values of σ to investigate the effect of the standard deviation on the model performance. The data is generated for $\sigma = \frac{\pi}{8}$, $\sigma = \frac{\pi}{16}$ and $\sigma = \frac{\pi}{32}$. The initial velocity u is set to zero, i.e., $u(\theta, 0) = 0$ m/s. The time step size is fixed and set to $\Delta t = 0.0025$ s. The initial conditions for the three different σ values can be seen in Figure 5.4.

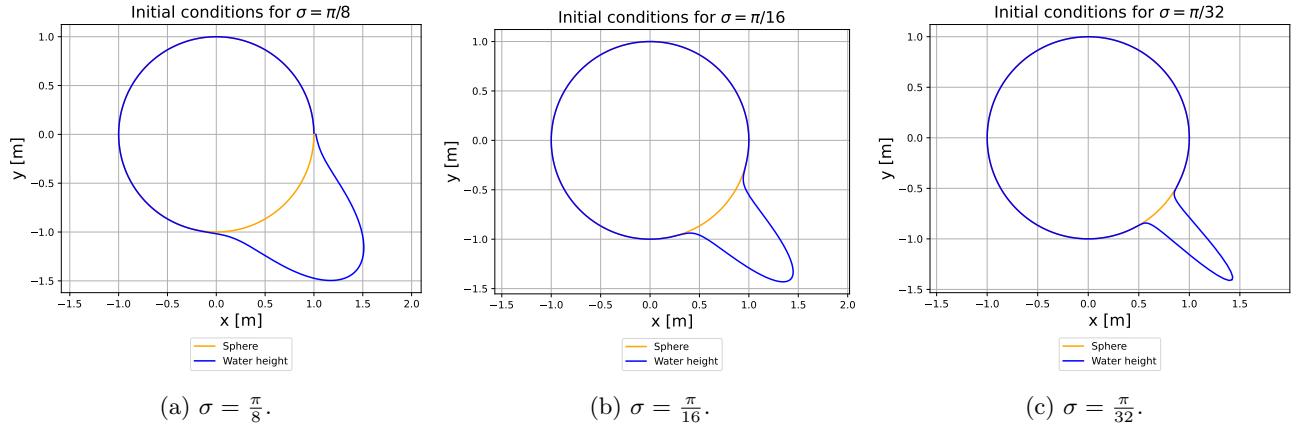


Figure 5.4: Initial conditions for the 1D LSWE in spherical coordinates for different σ values.

From Figure 5.4, we observe that the standard deviation σ affects the width of the Gaussian function. The smaller the σ , the narrower the Gaussian function, meaning the curves are steeper. This is to test the different models abilities to handle steep gradients. The data is generated from $t = 0.0$ to $t_{\text{end}} = 1.0$ s.

2D SWE

For the 2D SWE, we also use the Gaussian function as initial condition for the water height h , but now in two dimensions:

$$h(x, y, 0) = h_0 + a \cdot \exp\left(-\frac{(x - x_c)^2 + (y - y_c)^2}{2\sigma^2}\right), \quad (5.1.3)$$

where h_0 is the initial water height outside of the Gaussian, a is the amplitude of the Gaussian, (x_c, y_c) is the center of the Gaussian in meters, and σ is the standard deviation. The domain is $x, y \in [0, 40]$ m and is discretized into N points in each direction. We use the parameters $h_0 = 1$ m, $a = 2$, $(x_c, y_c) = (20\text{m}, 20\text{m})$, and $\sigma = 2$ m. The initial velocity u is set to zero, i.e., $u(x, y, 0) = 0$. The initial conditions for the water height can be seen in Figure 7.17.

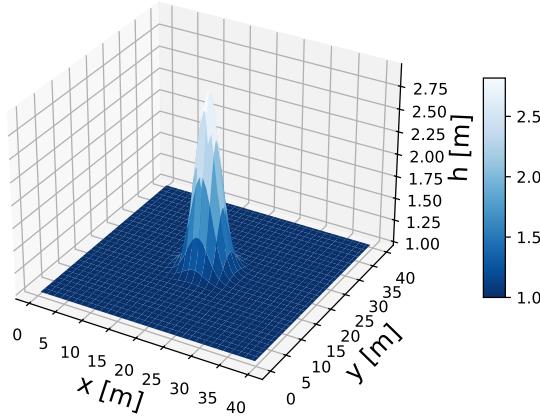
Initial water height, $h[m]$ 

Figure 5.5: Initial condition for the 2D problem.

To generate the data, we use our FVM solver to solve the 2D SWE with the initial conditions just specified. The solver is validated by comparing its results with known test cases. We generate two distinct data sets containing solutions of the 2D SWE. The first data set is created with a variable time step size Δt , determined dynamically using the CFL condition, where the CFL number is set to 0.9. Data is generated from $t = 0.0$ to $t_{\text{end}} = 5.0$ s. To investigate the impact of grid resolution on the model performance, we generate data for different grid resolutions, with values of $N = 64$, $N = 128$, and $N = 256$. This approach also allows us to test the models ability to transfer solutions to different grid resolutions, a critical capability for generalization.

The second data set is designed for long-term predictions, where the time step size is fixed. Since predictions are made beyond the data, the time step size must be known. For this data set, we use a constant time step size of $\Delta t = 0.025$ s. This value was determined by analyzing the time step sizes used in the variable step data generation. To ensure stability, the time step size must be sufficiently small. By halving the smallest observed time step size, we obtained $\Delta t = 0.025$ s. The used grid resolution is $N = 64$. The data is generated from $t = 0.0$ to $t_{\text{end}} = 15.0$ s, providing a robust data set for training models for long-term predictions.

5.2 Data generation for the sphere

A significant challenge in this project has been obtaining high-quality data for training data-driven methods on a sphere. To tackle this, we are exploring various approaches for data generation, including the potential to create our own data and utilize external data sources.

Projecting 2D data to the sphere

Our initial approach involved projecting the solution data from the 2D SWE onto the sphere. To ensure stability and generate high-quality data, we set the CFL number to 0.8. The coordinates θ (longitude) and ϕ (latitude) are treated similarly to x and y in the 2D case. The grid is configured with $N_\theta = 200$ grid points in the θ -direction and $N_\phi = 100$ grid points in the ϕ -direction. The higher resolution in the θ -direction accounts for its double distance

compared to the ϕ -direction. We use a Gaussian function as initial condition

$$h(\theta, \phi, 0) = h_0 + a \cdot \exp\left(-\frac{(\theta - \theta_c)^2 + (\phi - \phi_c)^2}{(2\sigma)^2}\right),$$

where $h_0 = 1$ m, $a = 3$, $\theta_c = \frac{3\pi}{2}$, $\phi_c = \frac{\pi}{3}$, and $\sigma = \frac{\pi}{16}$. The SWE are solved from $t = 0$ s to $t_{\text{end}} = 0.5$ s, using a variable step size Δt . The simulation takes 199 time steps, resulting in an average time step size of $\Delta t \approx 0.0025$ s. This value is significantly smaller than in the 2D case, likely due to the higher number of grid points, another domain and lower CFL number. The results after some given time steps are presented in Figure 5.6.

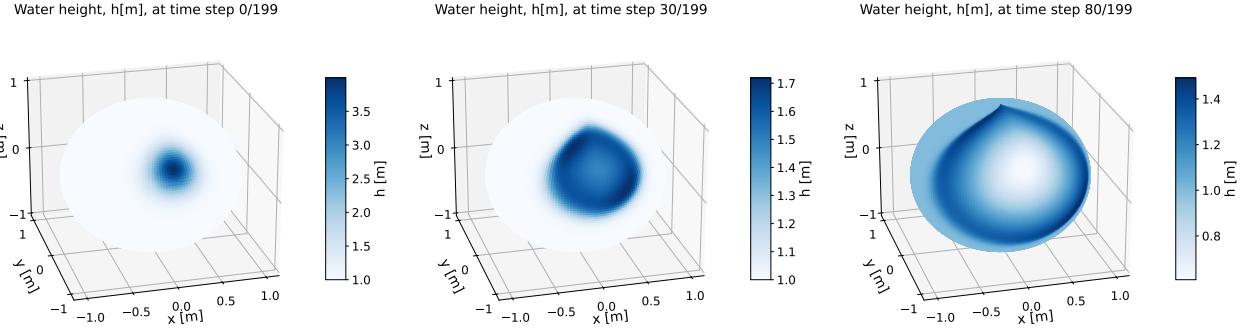


Figure 5.6: Water height on the sphere for different timesteps.

In Figure 5.6 we observe the evolution of the water height over time on the sphere. The initial Gaussian bump propagates across the sphere, but singularities are present, especially at the poles. These issues arise, among other things, from neglecting the curvature of the sphere. While this approach might be acceptable for some applications, for instance when focusing on a small region near the equator, where the projection is more accurate, it is not suitable for our project. Since our aim is to model the entire sphere, accounting for its curvature is essential. Consequently, we cannot use this data.

Mesh generation for the sphere

To solve the SWE on the sphere, we must use a different grid than the regular grid used in the 2D case. One approach is to use an icosahedral grid, which approximates the sphere with triangles. The grid structure allows for various levels of refinement, depending on the desired level of accuracy. At each refinement level, the number of triangles increases by a factor of four, as each triangle is divided into four smaller triangles, meaning that the number of triangles increases drastically with each level of refinement. The grid is generated using Matlab code from the Github repository [33] and then rewritten to Python. The first four levels of refinement are shown in Figure 5.7.

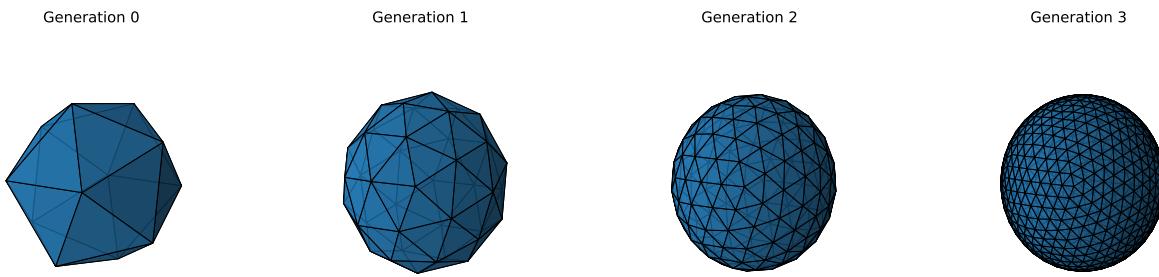


Figure 5.7: Icosahedral grid for the first 4 levels of refinement.

In Figure 5.7 we see how the grid is refined at each level, resulting in a progressively more accurate representation of the sphere. We will briefly outline the process of solving the SWE on this grid structure, inspired by the finite element method (FEM). The main idea is to keep track of which neighboring triangles each triangle has, as well as which interfaces they share. Each triangle is numbered, as well as each vertex. This information is organized into three key tables:

- Element-to-Vertex (EToV) table. This table stores the vertices of each triangle. Each row corresponds to a triangle, and the three columns correspond to the three vertices of the triangle.
- Element-to-Element (EToE) table. Derived from the EToV table, this table stores the neighboring triangles of each triangle. Each row corresponds to a triangle, and the three columns correspond to the three neighboring triangles.
- Element-to-Face (EToF) table. Keeps track of which faces (edges) are shared between two triangles. Each row corresponds to a triangle, and the columns correspond to the neighboring triangles and the face they share.

The method involves computing fluxes across interfaces between triangles. For each interface, the flux leaving one triangle and entering its neighbor is calculated to ensure the total flux across the system is zero, conserving the total volume of water. However, since the triangles are not aligned in straight lines, as in the 2D case, we must consider the contributions in both the θ - and ϕ -directions for each interface flux. Previously, we could view the problem as two 1D problems, but now they are intertwined, highlighting the complexity of the problem.

External data sources

Another approach for obtaining data is to use external data sources. One possible source is the Copernicus climate data store, specifically the ORAS5 global ocean reanalysis dataset, which provides monthly data from 1958 to the present [34]. For this study, we focus on the sea surface height, defined as the vertical distance between the actual sea surface and a reference surface, representing a mean sea level if the ocean were at rest. The data is a 2D field, depending on the longitude and latitude. To get an overview of the data, we downloaded the data from January to December 2014 and plotted the sea surface height as the difference from the reference sea surface height. The plots for January to April are shown in Figure 5.8.

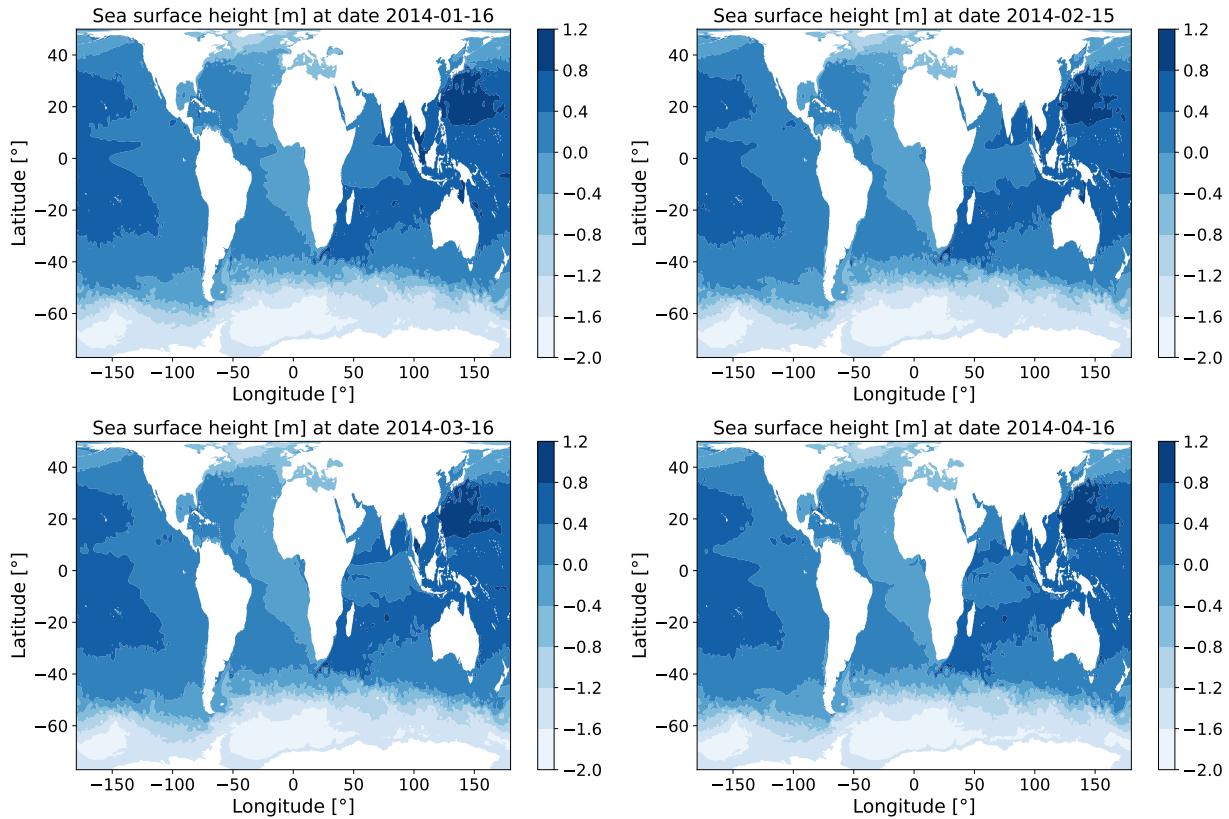


Figure 5.8: Sea surface height as the difference from reference sea surface height for the months Jan-Apr 2014.

Note that the data was recently updated (2025-01-15). At the time of downloading, we obtained the newest data available. While training an SFNO on real-world data could be an exciting prospect, we have decided not to proceed with this data set due to several uncertainties. In particular, since we cannot fully account for the quality and reliability of the data, we have chosen not to pursue this approach.

Chapter 6

Numerical results

In this chapter we present the results of the numerical experiments, where we have implemented the FVM for solving the SWE and tested it on several problems. A key focus is to validate the implementation, as it will generate data for the data-driven methods, including the convolutional neural networks and Fourier neural operators. In section 6.1, we solve the 1D dam break problem and compare the numerical solution against the true solution. In section 6.2, we present the results from the five test cases from Toro (2001) [5]. These problems are all discontinuous in either the water height h or the fluid velocity u . The idea is, that if the numerical solution can capture the discontinuities, it should be well-suited to handle smoother solutions as well. We also compare the results from the different fluxes used in the FVM. In section 6.3 we present the results from the 1D LSWE on a sphere.

We extent to consider 2D problems in cartesian coordinates, and present the results from the 2D idealised circular dam break problem in section 6.4. The results from the 2D problem are compared to the results from Toro (2024) [7] to validate the implementation of the 2D FVM. In section 6.5, we test the scalability of the FVM to solve the 2D SWE, by running the 2D problem with a Gaussian initial condition for different values of N , i.e., the number of cells in each direction. Finally, in section 6.6, we present animations for the 2D idealised circular dam break problem and the solution of the SWE with a Gaussian initial condition projected on a sphere.

The used computer for running simulations is a Windows 11 computer with an Intel Core i7 CPU and 16 GB of RAM. All code and data can be found at the Github repository [35].

6.1 The 1D Dam Break Problem

We consider the 1D dam break problem, a special case of the Riemann problem (3.4.1), with the following initial conditions:

$$h(x, 0) = \begin{cases} h_L, & \text{if } x < x_0, \\ h_R, & \text{if } x > x_0, \end{cases}$$

where $x \in [0 \text{ m}, 50 \text{ m}]$, $h_L = 3.5 \text{ m}$, $h_R = 1.25 \text{ m}$ and $x_0 = 20 \text{ m}$. Since it is a dam break problem the initial fluid velocity is zero, i.e., $u(x, 0) = 0 \text{ m/s}$. We solve the problem starting at $t = 0 \text{ s}$ and ending at $t = 2.5 \text{ s}$. The numerical solution to the 1D dam break problem using the FVM, together with the true solution, provided from the Ph.D. course *An Introduction to Discontinuous Galerkin Methods for solving Partial Differential Equations* [36], can be seen in Figure 6.1.

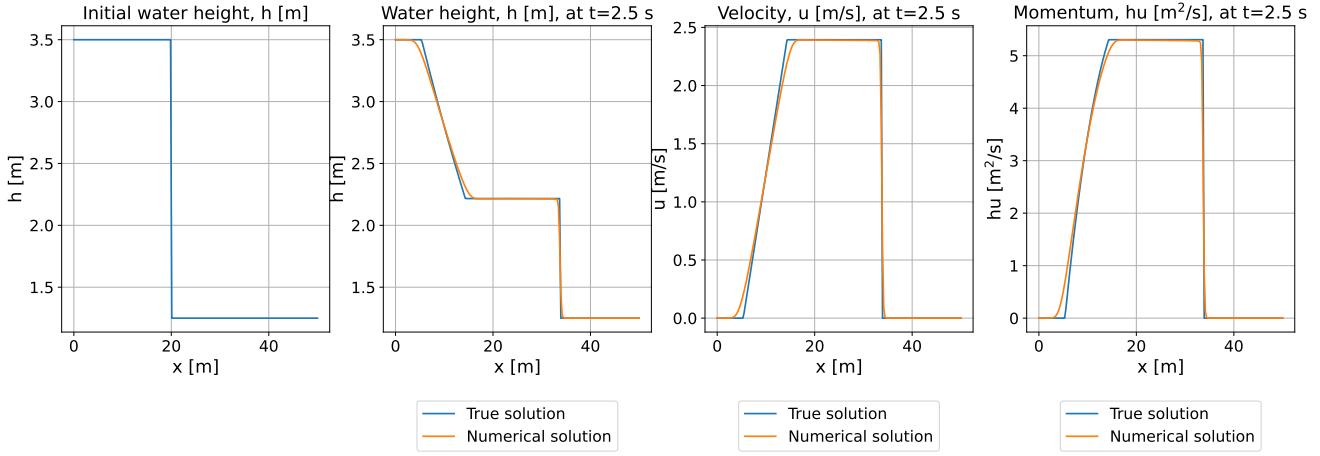


Figure 6.1: The initial water height h at $t = 0$ s, together with the water height, the fluid velocity u and the momentum hu at $t = 2.5$ s.

From Figure 6.1 we see that the numerical solution aligns well with the true solution, and overall successfully captures the discontinuity. We also see that the solution consists of a right shock wave and a left rarefaction wave, as expected from the initial conditions. This is seen as, the shock wave to the right has a high speed and represent a discontinuity in the solution for the water height. The rarefaction wave to the left is a smooth transition from the high water height to a lower water height.

6.2 Toro test cases

We have tested the FVM on the five test cases for Riemann problems from Toro (2001) [5]. The initial conditions for the five test cases are given in Table 6.1.

Test case	h_L [m]	u_L [m/s]	h_R [m]	u_R [m/s]	x_0 [m]	t_{end} [s]
1	1.0	2.5	0.1	0.0	10.0	7.0
2	1.0	-5.0	1.0	5.0	25.0	2.5
3	1.0	0.0	0.0	0.0	20.0	4.0
4	0.0	0.0	1.0	0.0	30.0	4.0
5	0.1	-3.0	0.1	3.0	25.0	5.0

Table 6.1: Initial conditions for the five test cases.

The domain is $x \in [0 \text{ m}, 50 \text{ m}]$ for all test cases. The Riemann problems are chosen to test the numerical method on different combinations of shock waves and rarefaction waves. To solve the test cases we have used the following fluxes:

1. Godunov method with exact Riemann solver,
2. Lax-Friedrich flux,
3. Lax-Wendroff flux,
4. FORCE flux,

5. HLL flux,

Test case 1

The initial conditions for test case 1 are visualised in Figure 6.2 and the solution after $t = 7.0$ s is given in Figure 6.3.

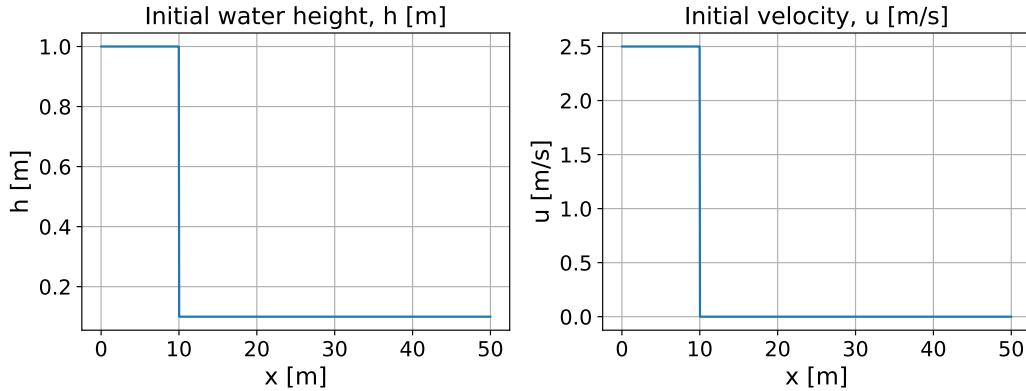


Figure 6.2: Initial conditions for the test case.

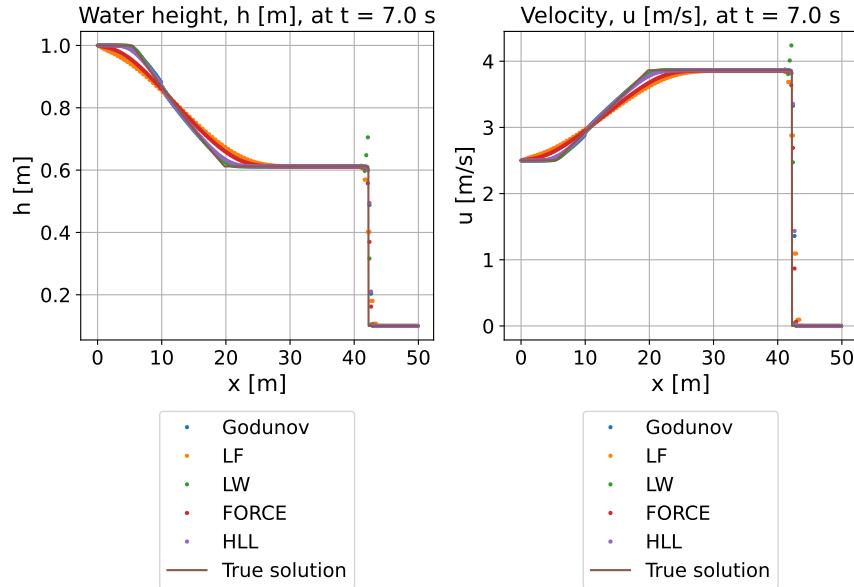


Figure 6.3: Final solution for the test case after $t = 7.0$ s.

For this test case all the fluxes work well, but there are minor differences in the solution, which can be seen in Figure 6.3. For instance, we see that the Lax-Wendroff (LW) flux has some oscillations in the solution at the discontinuity, which is not present in the other fluxes. Similarly to the case of the 1D dam break problem in section 6.1, we see that the solution consists of a right shock wave and a left rarefaction wave.

Test case 2

The initial conditions for test case 2 are illustrated in Figure 6.4.

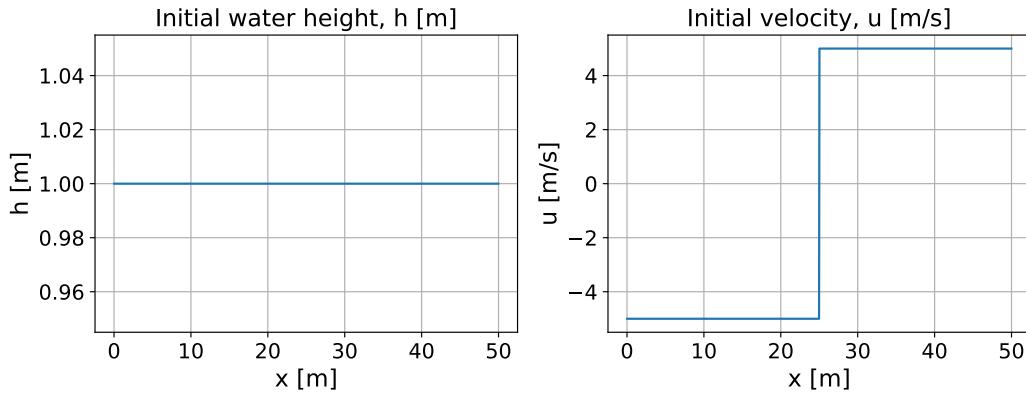


Figure 6.4: Initial conditions for the test case.

In test case 2 we have two rarefaction waves, one on the left side and one on the right side of x_0 . As they are travelling in opposite directions (away from each other), due to the initial velocity, there will be created a nearly dry bed in the middle of the domain. Many methods have difficulties with this test case as they may compute a negative water height. For these experiments we were able to get close to the true solution, using Lax-Friedrich flux, FORCE flux and HLL flux. The solution after $t = 2.5$ s is illustrated in Figure 6.5.

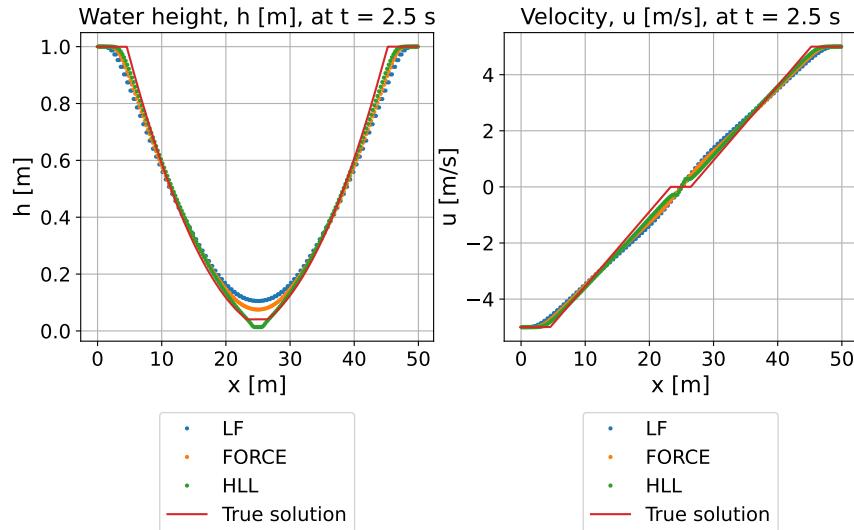


Figure 6.5: Final solution for the test case.

For the fluxes, Godunov method with exact Riemann solver, and Lax-Wendroff, it was not possible to get an acceptable solution. Among the LF, FORCE and HLL fluxes we also see differences in how close the numerical solution comes to $h = 0$ m at $x = 25$ m.

Test case 3

The initial conditions for test case 3 are given in Figure 6.6.

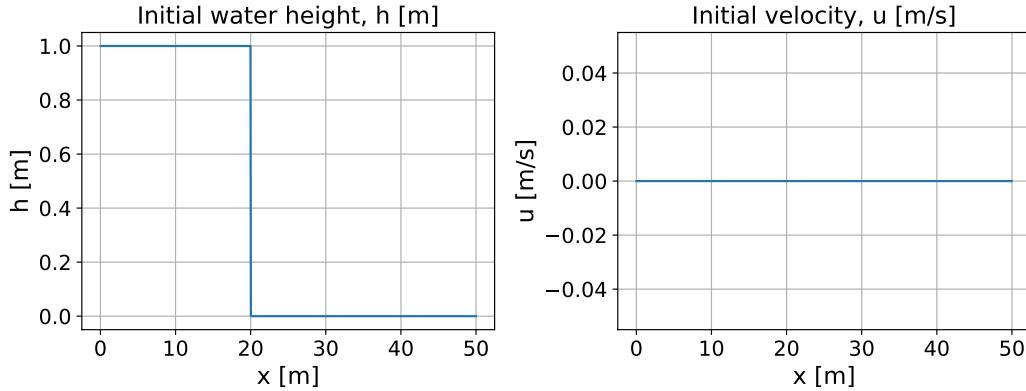


Figure 6.6: Initial conditions for the test case.

In Figure 6.6 we see that this is a dam break problem, as the initial velocity is zero everywhere. The solution after $t = 4.0$ seconds is given in Figure 6.7.

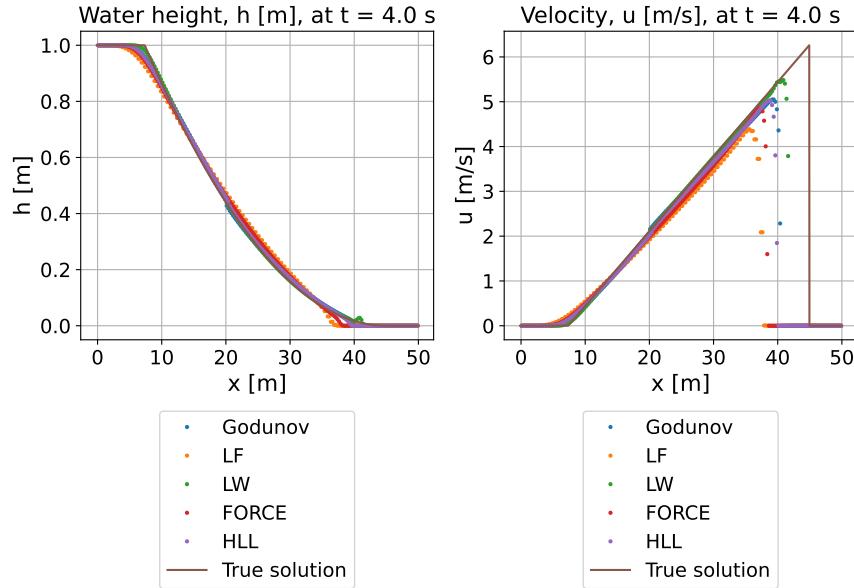


Figure 6.7: Final solution for the test case.

In this case, we are challenged with a dry-bed region in the right part of the domain. To solve case 3, with the FVM we must add a small amount to h_R , since the methpd does not handle $h_R = 0$ well. We set $h_R = 0.00005$ to solve it numerically, but the true solution is for $h_R = 0$. By running experiments with different values of h_R , we see that the solution converges to the true solution as h_R approaches 0. From Figure 6.7 we see that when it comes to predicting the velocity, there are some differences in how the different fluxes perform. Overall, using the Lax-Wendroff (LW) flux gives the closest solution to the true solution for the velocity.

Test case 4

The initial conditions for test case 4 are given in Figure 6.8, and the solution after $t = 4.0$ s is given in Figure 6.9.

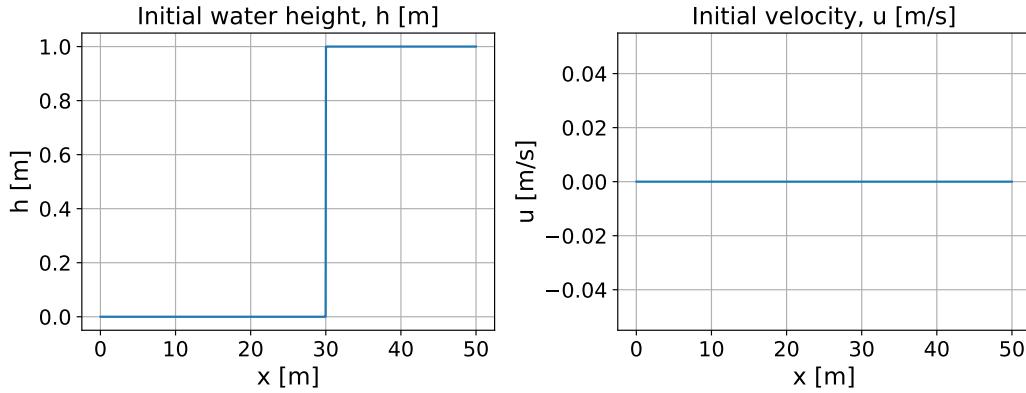


Figure 6.8: Initial conditions for the test case.

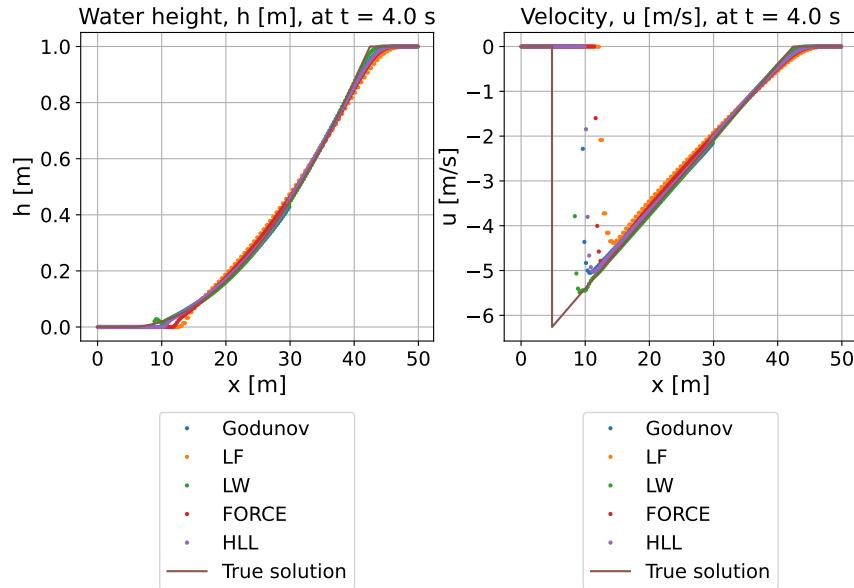


Figure 6.9: Final solution for the test case.

This test problem is symmetric to test case 3, meaning we face the same challenges of a dry-bed region, now located in the left part of the domain. We set $h_L = 0.00005$, and the solution converges to the true solution as h_L approaches 0. This case is included to test if the results are as expected. As in test case 4, we observe differences in the fluxes performance, and again we observe that the Lax-Wendroff (LW) flux comes closest to the true solution for the velocity.

Test case 5

The initial conditions for test case 5 are given in Figure 6.10, and the final solution after $t = 5.0$ s is given in Figure 6.11.

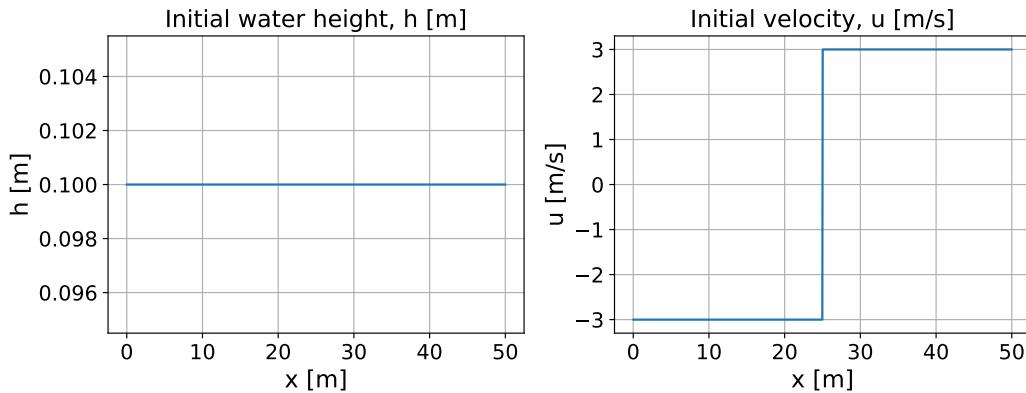


Figure 6.10: Initial conditions for the test case.

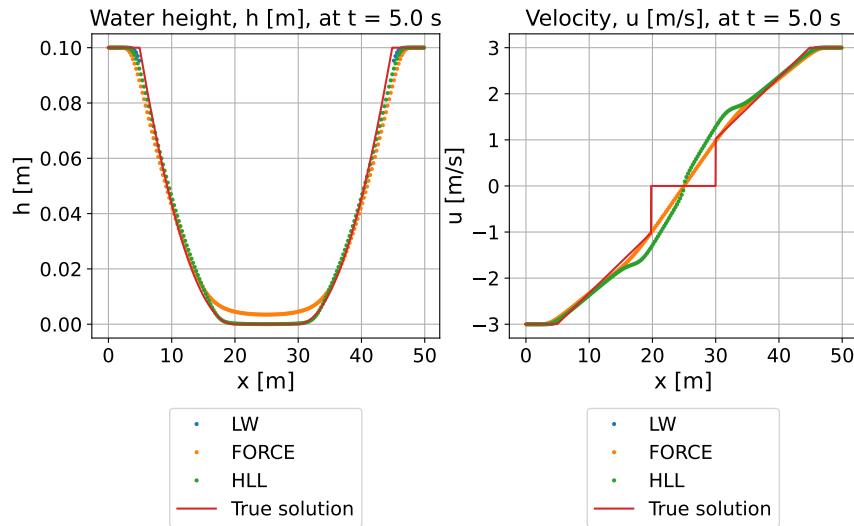


Figure 6.11: Final solution for the test case.

From Figure 6.11 we see that the numerical solution for the velocity v at $t = 5.0$ is smooth, where the true solution is discontinuous. In this test case there are also challenges with some of the fluxes due to the generation of a dry-bed region in the middle of the domain. The fluxes that are not able to solve this case are: the Godunov method with exact Riemann solver and the Lax-Wendroff flux, the same as in test case 2. The solution consists of two rarefaction waves, one on the left side and one on the right side, and a dry-bed region in the middle. To get an overview of which fluxes that were able to produce solutions for the test cases, consider Table 6.2.

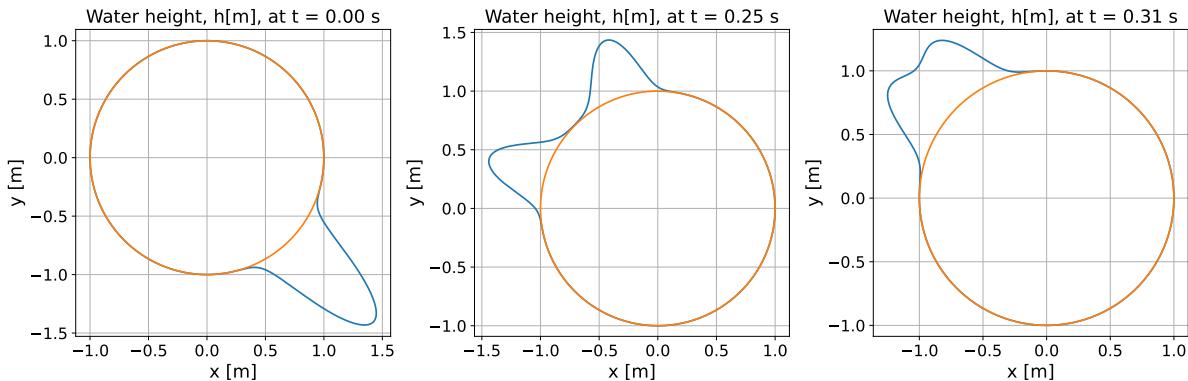
Test case	Godunov	LF	LW	FORCE	HLL
1	✓	✓	✓	✓	✓
2	✗	✓	✗	✓	✓
3	✓	✓	✓	✓	✓
4	✓	✓	✓	✓	✓
5	✗	✓	✗	✓	✓

Table 6.2: Overview of which fluxes that were able to produce solutions for the test cases.

Note, that as we see in the results, there are still differences in the solutions accuracy between the fluxes that were able to solve the test cases.

6.3 1D linearized Shallow Water Equations on a sphere

In this section we consider the 1D LSWE on a sphere. We consider the spatial dimension θ , which is the longitude angle and keep the latitude ϕ constant. The LSWE on a sphere are given by (2.6.4) with the initial conditions given by (5.1.2). The LSWE on a sphere are solved using the FVM with the ERK4 time-stepping method. The initial conditions, and the solution after $t = 0.25$ s and $t = 0.5$ s are visualised in Figure 6.12.

Figure 6.12: The water height, $h[m]$, at $t = 0$ s, $t = 0.25$ s and $t = 0.31$ s.

From Figure 6.12 we see how the water height evolves over time. We observe that the initial wave split into two waves, travelling in opposite directions. When the two waves meet, we see how the periodic boundary conditions operate, as the waves are melting together, and then split again into two waves, as in the beginning. This is more clear visualised in the animation for this case, a link and qr-code can be found in Figure 6.15. Since we do not account for friction and have neglecting external forces except gravity, the waves will continue to travel around the sphere. In this case we do not have a true solution to compare with, but the numerical solution seems reasonable.

6.4 2D idealised Circular Dam Break Problem

We now proceed to the 2D case, considering an idealised circular dam break problem over a horizontal bottom, a problem from Toro (2024) [7]. We assume there is an infinitely thin circular wall at radius $R = 2.5$ m in a square domain of size 40 m \times 40 m with centre at $(x_c, y_c) = (20$ m, 20 m). The initial conditions are a Gaussian function

for the water height h and zero initial velocity for the fluid.

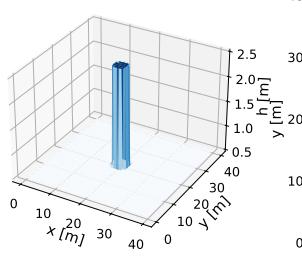
$$h(x, y, 0) = \begin{cases} 2.5 \text{ m}, & \text{if } \sqrt{(x - x_c)^2 + (y - y_c)^2} \leq R, \\ 0.5 \text{ m}, & \text{otherwise,} \end{cases}$$

$$u(x, y, 0) = 0 \text{ m/s},$$

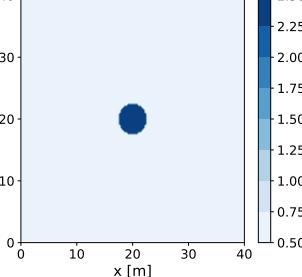
$$v(x, y, 0) = 0 \text{ m/s}.$$

We use a mesh of size 64×64 . The problem is solved using the FVM with the Rusanov flux, and the results after $t = 0.0 \text{ s}$, 0.4 s , 0.7 s and 1.4 s are visualised in Figure 6.13.

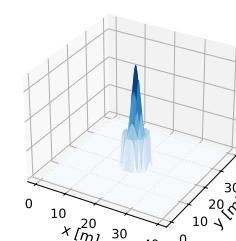
Water level, h [m], at $t = 0.00 \text{ s}$



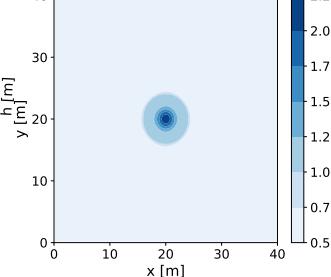
Water level, h [m], at $t = 0.00 \text{ s}$



Water level, h [m], at $t = 0.40 \text{ s}$



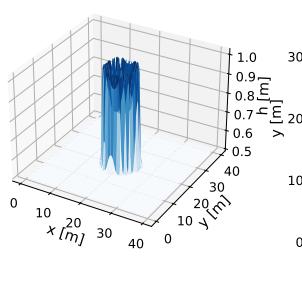
Water level, h [m], at $t = 0.40 \text{ s}$



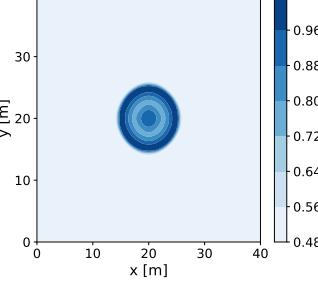
(a) 2D dam break problem after $t = 0 \text{ s}$.

(b) 2D dam break problem after $t = 0.4 \text{ s}$.

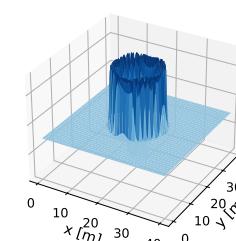
Water level, h [m], at $t = 0.70 \text{ s}$



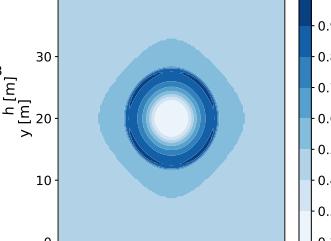
Water level, h [m], at $t = 0.70 \text{ s}$



Water level, h [m], at $t = 1.40 \text{ s}$



Water level, h [m], at $t = 1.40 \text{ s}$



(c) 2D dam break problem after $t = 0.7 \text{ s}$.

(d) 2D dam break problem after $t = 1.4 \text{ s}$.

Figure 6.13: Snapshots of the 2D dam break problem at different times.

By comparing Figure 6.13 with the results from Toro (2024) [7], we see that the numerical solution aligns well with the true solution.

6.5 Scalability

Numerical methods are good as they can be more or less as accurate as we want them to be, but the computational cost increases with the number of cells, i.e., the more high-resolution grid we use. Previously in the thesis, we have indicated that a disadvantage of the FVM is that the computational cost increases with the number of cells. To test the scalability of the FVM to solve the 2D SWE, we have solved the 2D SWE with a Gaussian initial condition, as described in chapter 5, for different values of N , i.e., the number of cells in each direction, and measured the run time. The results are presented in Table 6.3.

N	16	32	64	128	256
Time (s)	0.37	2.73	19.90	170.08	4205.45

Table 6.3: Running time for the FVM to solve 2D SWE for different values of N .

The run time [s] dependent on the number of cells N is illustrated in Figure 6.14.

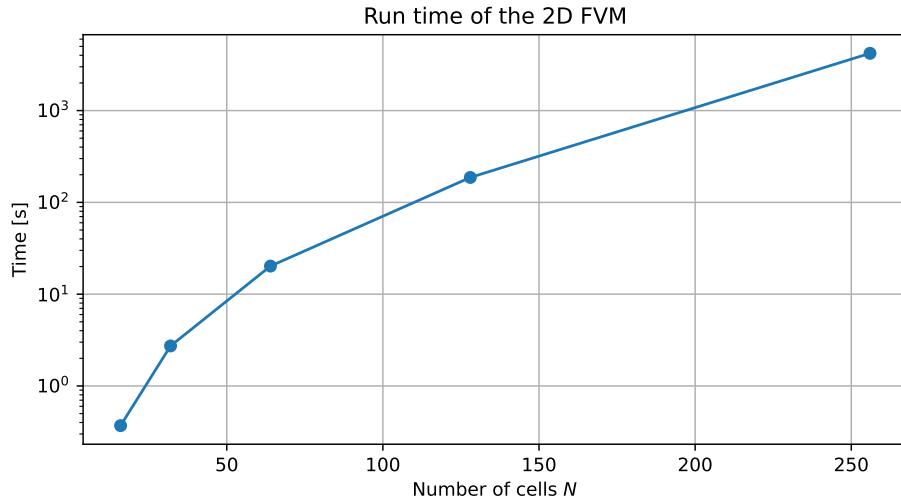


Figure 6.14: Run time [s] of the 2D FVM to solve the SWE, depending on number of grid cells N . Note that the axis with time is in log-scale.

From Table 6.3 and Figure 6.14 we see that the run time increases drastically with the number of cells N . This behavior is expected since an increase in the number of cells leads to a higher number of computations. Consequently, the computational time rises, making the method less scalable for real-life simulations. To effectively model phenomena such as floods or tsunamis, we need a scalable method. This challenge motivates the exploration of data-driven methods, as we aim to investigate whether they offer a more scalable solution.

6.6 Animations

To visualise the results of the numerical experiments, we have created some animations. These include an animation of the 1D LSWE on a sphere, demonstrating the periodic boundary conditions. There are also animations for the 2D idealised circular dam break problem and the SWE with a Gaussian initial condition projected on a sphere, as described in section 5.2. The following QR-codes provide access to the animations. A link to the animations is also provided in the caption of the QR-codes, directing to a dedicated GitHub repository.



(a) QR-code for the 1D LSWE on a sphere.



(b) QR-code for the 2D idealized circular dam break problem.



(c) QR-code for the SWE with initial Gaussian conditions, projected on a sphere.

Figure 6.15: QR-codes for the animations. All animations can be found at: <https://github.com/MelissaJessen/Shallow-Water-Equations-Animations/blob/main/README.md>.

In the animation for the 2D idealised circular dam break problem, we observe what happens after the time steps in Figure 6.13. We note that when the waves hit the boundaries, they are reflected back into the domain, demonstrating the behaviour of the boundary conditions. In the animation for the SWE with a Gaussian initial condition on a sphere, we note that since the SWE are solved in a 2D cartesian domain and wrapped to a sphere, we observe non-physical behaviour, especially close to the poles.

Chapter 7

Data-driven results

In this chapter we present the results of the data-driven models, that is, the convolutional neural network (CNN) and the Fourier neural operator (FNO), for solving the shallow water equations. We analyse three main cases:

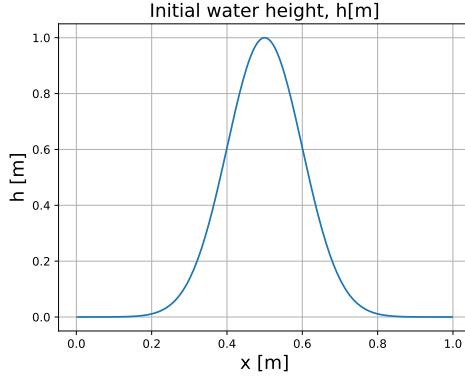
1. 1D SWE with Gaussian initial conditions.
2. 1D LSWE in spherical coordinates.
3. 2D SWE with Gaussian initial conditions.

In each case, we evaluate the performance of the CNN and the FNO in terms of mean squared error (MSE), mean absolute error (MAE), training time and prediction time. In addition, we analyze the models' predictions for various initial conditions and their ability to generalize to unseen data. We also explore their capability to transfer solutions across different grid resolutions, such as transitioning from a coarse grid to a finer grid. Finally, we examine the models' long-term predictive performance, assessing how well they maintain accuracy over extended time steps. The convolutional neural networks are implemented in python version 3.12.6, using PyTorch version 2.4.1. The Fourier neural operators are implemented using the Github repository *neuraloperator*, which is a library for learning neural operators in PyTorch [37].

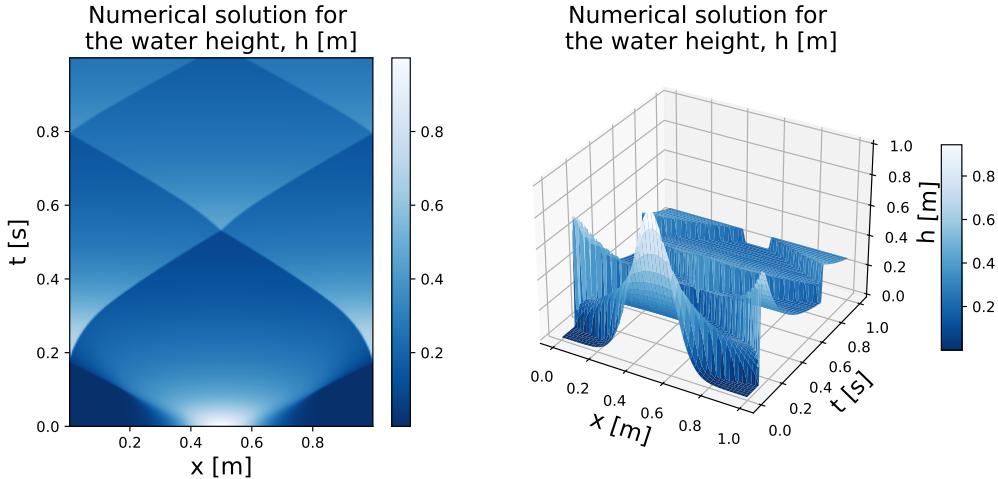
In chapter 6, we presented the results of discontinuous test cases, operating under the premise that a solver capable of handling discontinuous solutions should also perform well with smooth solutions. In this chapter, we focus on evaluating the performance of data-driven models under smooth initial conditions, as they provide a more suitable starting point.

7.1 1D Shallow Water Equations with Gaussian initial conditions

In this section, we consider the 1D SWE with Gaussian initial conditions, where the data is generated as described in section 5.1. The initial condition for the water height h is given in (5.1.1) and illustrated in Figure 7.1.

Figure 7.1: The initial conditions for the water level h in the 1D SWE.

The domain is $x \in [0, 1]$ m and the data covers the time interval $t \in [0, 1]$ s. In Figure 7.1, we see that the initial condition is a Gaussian function with center in the middle of the domain. To get an overview of how the solution evolves over time, we have plotted the numerical solution in the x, t -plane, shown in Figure 7.2, in both a contour plot and a 3D plot.

Figure 7.2: Numerical solution of the 1D SWE from $t = 0$ s to $t = 1$ s.

In Figure 7.2, we see how the water height in the domain evolves over time. The 3D plot provides an overview of the solution, whereas the contour plot offers a more detailed view of the water level. From the contour plot, we observe that even though the initial condition is smooth, the solution develops sharp edges, close to discontinuities, over time, represented by the lines in the plot. This behavior is typical for the nonlinear shallow water equations, where the solution can become discontinuous due to the formation of shock waves. Some details of the data can be seen in the following table.

Case	n_train	n_val	n_test	N	Δx	Δt
1D SWE	369	123	123	200	0.005 m	[0.0008 s, 0.00225 s]

Table 7.1: Details of the used data for the case with the 1D SWE with a Gaussian initial condition.

In Table 7.1, there are listed some details of the CNN model. The time step size Δt is not constant, but varies between 0.0008 s and 0.00225 s. This is due to the CFL condition, which requires the time step to be small enough to ensure stability.

CNN Model

In the convolutional neural network, we train the model using the generated data. The input and output data are the same, shifted by one time step, allowing the model to predict the solution at the next time step. This approach enables the model to learn the flowmap, as described in section 4.1. The model takes input with 10 channels, corresponding to the sequence length, and processes it through a series of three 1D convolutional layers with ReLU activation functions. The final convolutional layer reduces the output to a single channel, which is then mapped to the prediction using a fully connected layer. Training is performed using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. The loss function minimizes the mean squared error (MSE). The data set is split into 60% training, 20% validation, and 20% testing data, with the exact number of training points provided in Table 7.1. The model is trained for 500 epochs, where the model's parameters are continuously saved if the validation loss improves upon the previous best. This is done to prevent overfitting. The training and validation losses for the CNN model are shown in Figure 7.3.

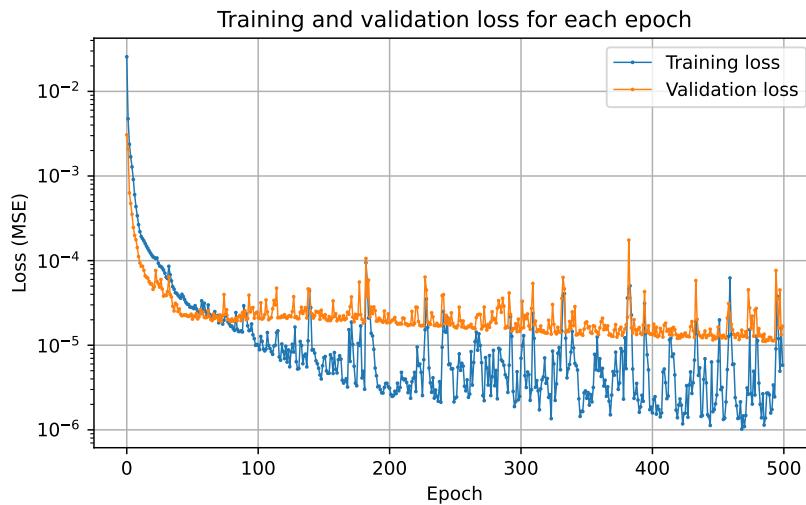


Figure 7.3: Training and validation loss (MSE) for the CNN model.

In Figure 7.3, we see that the training and validation loss decrease over the epochs, demonstrating that the model is learning the dynamics of the solution. However, while the training loss continues to decrease, the validation loss has largely stabilized. This indicates that further training is unlikely to improve the model's performance and may lead to overfitting. Additionally, we assess the accuracy of the model's predictions by examining the error, as shown in Figure 7.4.

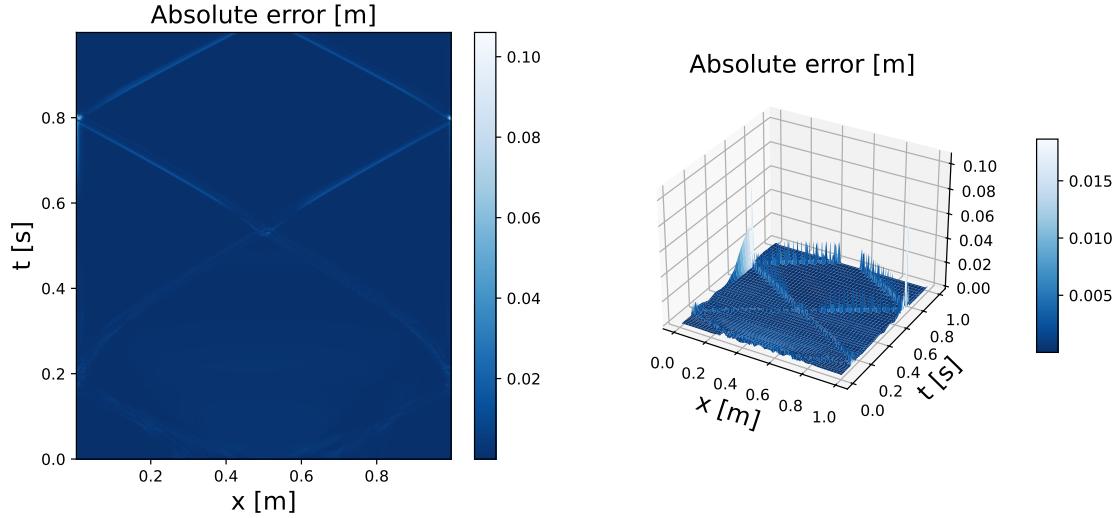


Figure 7.4: Error plot for the predictions for the CNN model.

From Figure 7.4, we see that the largest errors are observed in regions where the solution exhibits discontinuities, which is expected as the model struggles to make accurate predictions in these areas. To gain deeper insight into the model's performance, we examine its predictions at specific time steps, as shown in Figure 7.5.

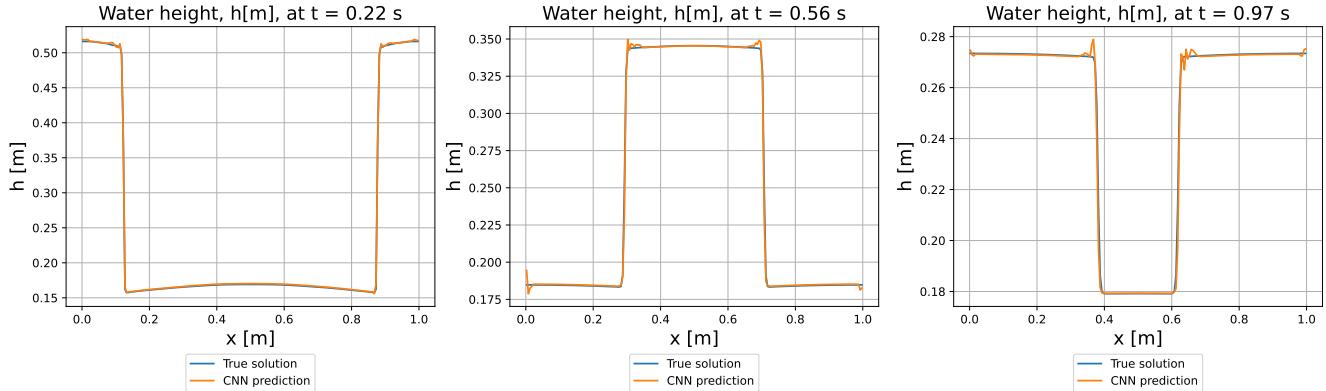


Figure 7.5: Predictions for the CNN model for some given time steps.

From Figure 7.5, we observe that the CNN model overall captures the dynamics of the solution, but struggles to predict the sharp edges. This is especially illustrated in the prediction at $t = 0.97$ s, where we observe oscillations in the solution that are not present in the true solution.

FNO Model

We define a FNO model, which consists of an input channel, 64 hidden channels and an output channel. We use a Fourier basis with 16 modes and a batch size of 32. The model is trained using the Adam optimizer with a learning rate of 0.001 and the criteria is to minimize the mean squared error (MSE). We use the same train/validation/test split as for the CNN model. The model is trained for 500 epochs, but the current best model is saved if the

validation loss is lower than the previous best validation loss. The training and validation loss for the FNO model is shown in Figure 7.6.

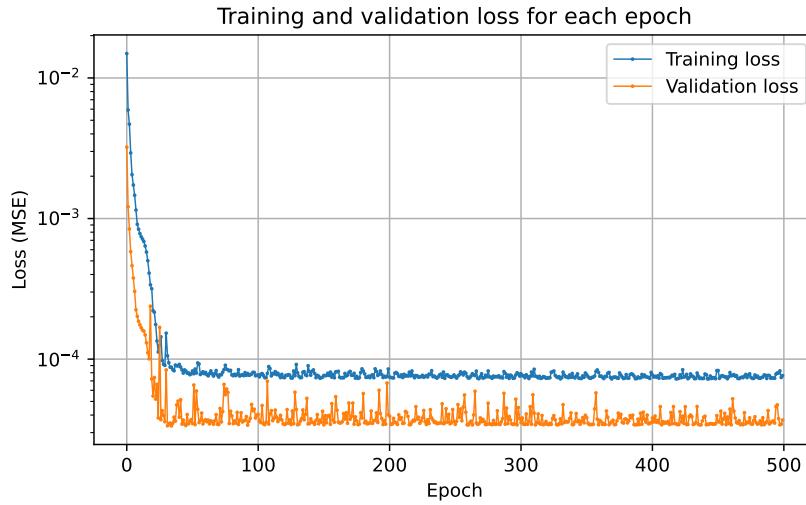


Figure 7.6: Training and validation loss for the FNO model.

From Figure 7.6, we see that the training and validation loss decrease over the epochs, indicating that the model is learning the dynamics of the solution. The losses drop quickly and are then more or less stable, suggesting that further training is unlikely to improve the model's performance. To see how the errors are distributed in the solution, we plot the error in Figure 7.7.

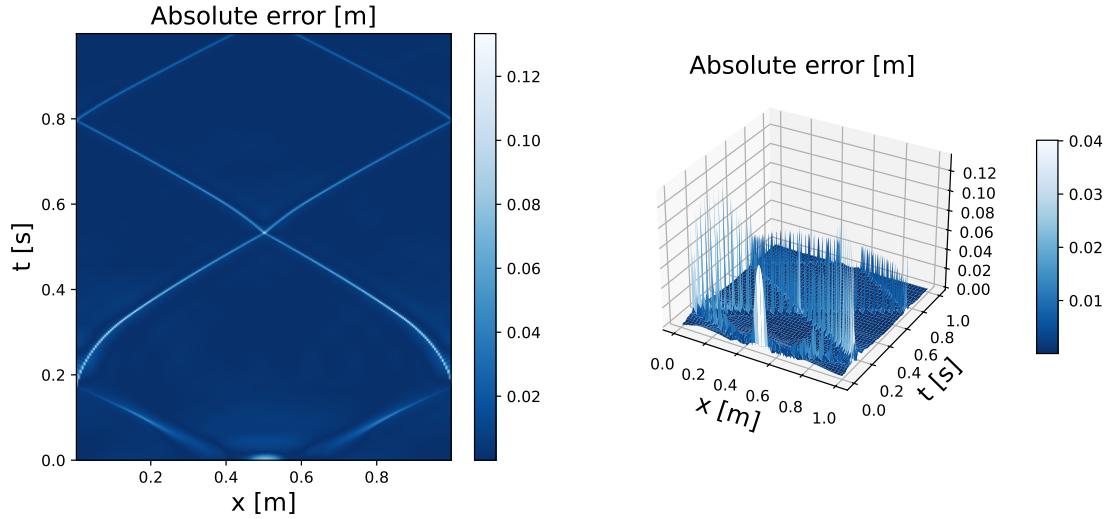


Figure 7.7: Error plot for the predictions for the FNO model.

In Figure 7.7, we see more or less the same error distribution as for the CNN model, with the largest errors at the discontinuities. We also consider the predictions for some given time steps, shown in Figure 7.8.

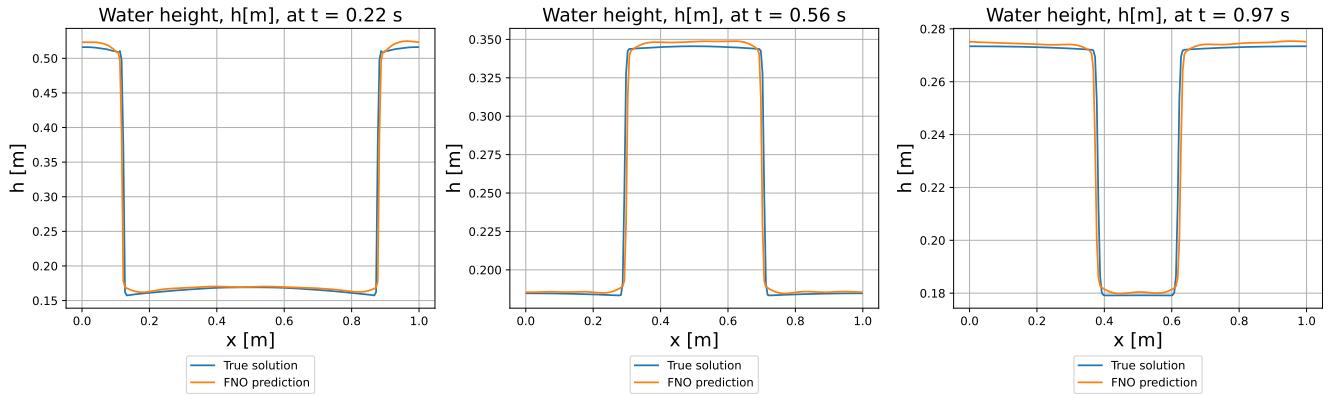


Figure 7.8: Predictions for the FNO model for some given time steps.

From Figure 7.8, we see that the FNO model produces smooth predictions, avoiding the oscillations seen in the CNN model. However the FNO model shows general inaccuracies in the solution.

New initial condition

To evaluate the models' ability to generalize to unseen data, we introduce a new initial condition for the water height h . This new condition retains the Gaussian form described in (5.1.1), but with a different mean parameter μ . Specifically, μ is set to $\mu = 0.3$ m, shifting the initial condition to the left. The new initial condition is illustrated in Figure 7.9.

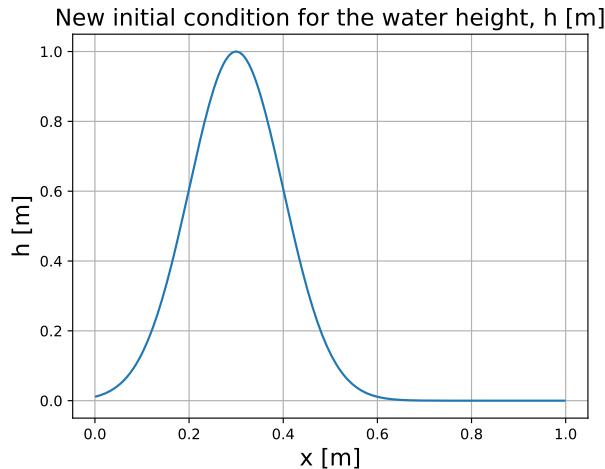


Figure 7.9: New initial condition for the 1D SWE.

The models are then tasked with making predictions for the new initial condition without retraining. These predictions cover a single time step, and the results are summarized in Table 7.2.

Comparison

To compare the performance of the CNN and FNO models, we evaluate the MSE and MAE for the test data predictions of the 1D SWE, with Gaussian initial conditions of $\mu = 0.5$ m, as well as for the new initial condition with $\mu = 0.3$ m. The reason we consider both MSE and MAE is that the MSE is more sensitive to outliers, while the MAE provides a more general overview of the error. The results are summarized in Table 7.2.

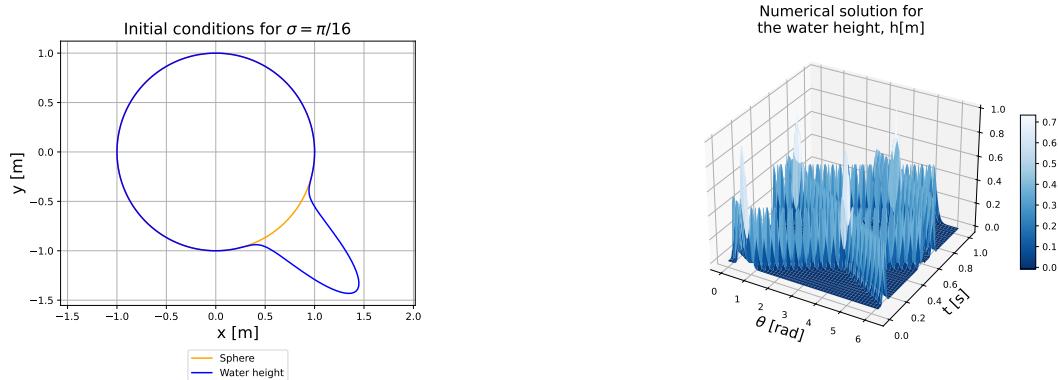
Model	Gauss initial condition			New initial condition			
	Epochs	MSE	MAE	Training time (s)	MSE	MAE	Prediction time (s)
CNN	500	5.16e-06	7.86e-04	116.47	2.60e-05	1.29e-03	0.08
FNO	500	1.75e-05	1.35e-03	970.65	1.38e-05	1.40e-03	0.39

Table 7.2: Test loss in terms of MSE and MAE, and time for training the models for the 2D SWE.

From Table 7.2, we observe that both models achieve a low MSE and MAE for the Gaussian initial conditions, indicating strong performance. However, the training time for the FNO model is significantly higher than that of the CNN model. For the new initial condition, the errors are nearly identical, suggesting that the FNO model has caught up.

7.2 1D linearized Shallow Water Equations in Spherical Coordinates

In this section, we consider the 1D LSWE in spherical coordinates on a circular domain. The length of the domain corresponds to the circumference of the circle, $L = 2\pi$ m, and is discretized into $N = 500$ points. The initial conditions is specified as a Gaussian function wrapped around the circle as given in (5.1.2). We use the middle value of σ , i.e., $\sigma = \frac{\pi}{16}$, to generate the initial conditions. The initial conditions can be seen in Figure 7.10a. To get an overview of how the solution evolves, we have plotted the numerical solution in the θ, t -plane from $t = 0$ s to $t = 1.0$ s, shown in Figure 7.10b.



(a) Initial conditions for the 1D LSWE on a sphere.
(b) Numerical solution of the 1D spherical LSWE in the θ, t -space.

Figure 7.10: Visualization of the 1D LSWE on a sphere: (a) shows the initial conditions, and (b) presents the numerical solution.

In Figure 7.10b, we see that the solution has some steep descents, and it is interesting to see how the data-driven models handle these sharp edges. Some details of the data such as the number of training, validation and test points, and the grid spacing can be seen in Table 7.3.

Case	n_train	n_val	n_test	N	$\Delta\theta$	Δt
1D spherical LSWE	240	80	81	500	$\frac{2\pi}{500} = 0.0126 \text{ rad}$	0.0025 s

Table 7.3: Details of the used data for the case with the 1D LSWE in spherical coordinates.

In Table 7.3, we see that in contrast to the 1D SWE, the time step size Δt is constant. From the number of training, validation and test points, we see that the data is split into 60% training, 20% validation and 20% test data, like for the 1D SWE with Gaussian initial conditions.

CNN Model

We define and train a CNN model to solve the spherical 1D LSWE. The model takes input with 10 channels, corresponding to the sequence length, and processes it through a series of three 1D convolutional layers with ReLU activation functions. The criteria is to minimize the MSE, and the model is trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. Training is conducted over 500 epochs, continuously saving the current best model. The training and validation loss is shown in Figure 7.11.

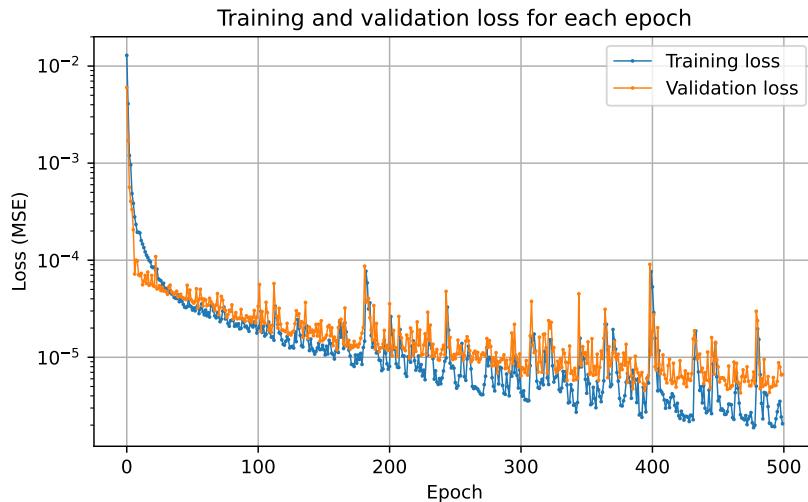


Figure 7.11: Training and validation loss for the CNN model for the 1D spherical LSWE.

In Figure 7.11, we see that the training and validation loss decrease over the epochs, indicating that the model is learning the dynamics of the solution. We note that the cross between the training and validation loss is quite early, which we also observed for the CNN model in the case with the 1D SWE with Gaussian initial conditions. To see how the model performs, we consider the error plots in Figure 7.12.

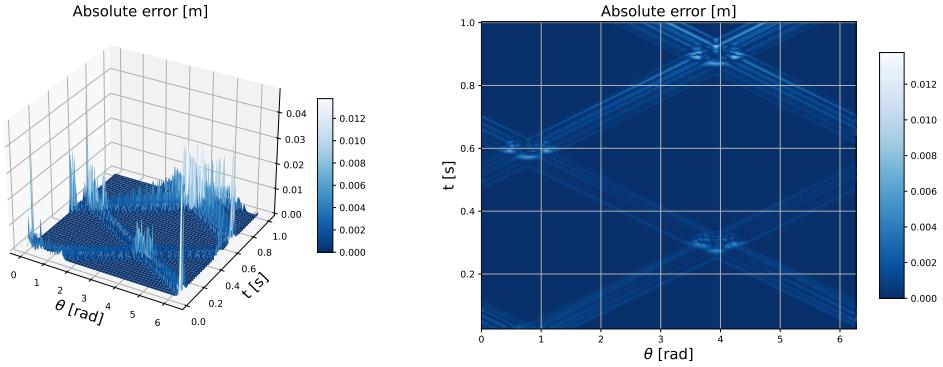


Figure 7.12: Error plots for the predictions of the CNN model for solving the 1D linearized spherical SWE.

In Figure 7.12, we see that errors are largest at the sharp edges of the solution, which is expected as the solution tends to be discontinuous. The predictions for some given time steps are shown in Figure 7.13.

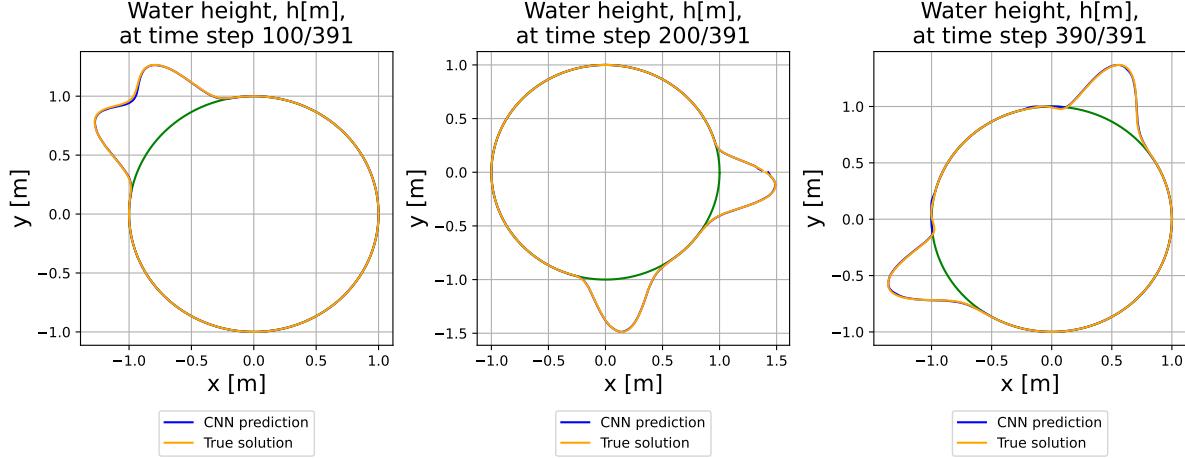


Figure 7.13: Predictions for the spherical 1D LSWE using the CNN model for some given time steps.

From Figure 7.13, we see that the predictions capture the waves.

FNO model

The FNO model consists of an input channel, 64 hidden channels and an output channel. We use a Fourier basis with 16 modes and a batch size of 32. The model is trained using the Adam optimizer with a learning rate of 0.001 and the criteria is to minimize the mean squared error (MSE). The model is trained on the data from $t = 0$ s to $t = 0.6$ s, validated on the data from $t = 0.6$ s to $t = 0.8$ s, and tested on the data from $t = 0.8$ s to $t = 1.0$ s. The model is trained for 200 epochs, where the current best model is saved throughout the training. The training and validation loss is shown in Figure 7.14.

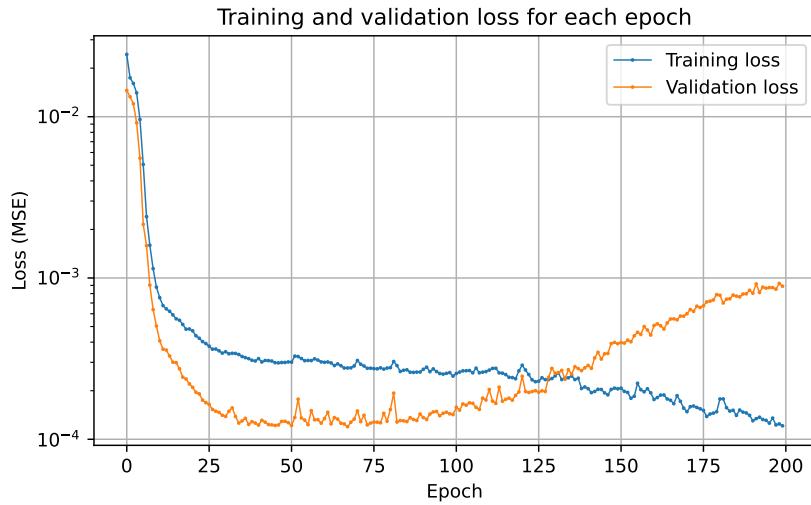


Figure 7.14: Training and validation loss for the FNO model for the spherical shallow water equations in 1D.

Figure 7.14 shows that the training and validation loss decrease over the epochs, indicating that the model is learning the dynamics of the solution. We see that after some time the validation loss increases, indicating that the model is overfitting the training data. This also shows why it can be beneficial to use early stopping. The error plots are shown in Figure 7.15.

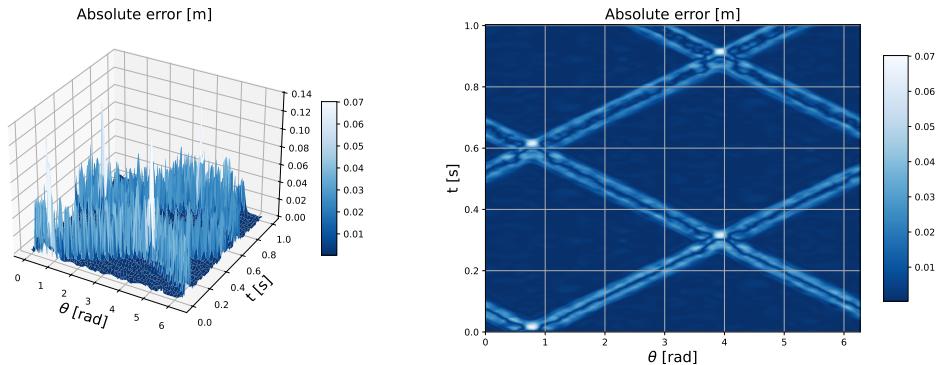


Figure 7.15: Error plots for the predictions of the 1D linearized spherical SWE.

Similar to the CNN model, we observe from Figure 7.15 that the error is largest at the sharp edges of the solution. The predictions for some given time steps are shown in Figure 7.16.

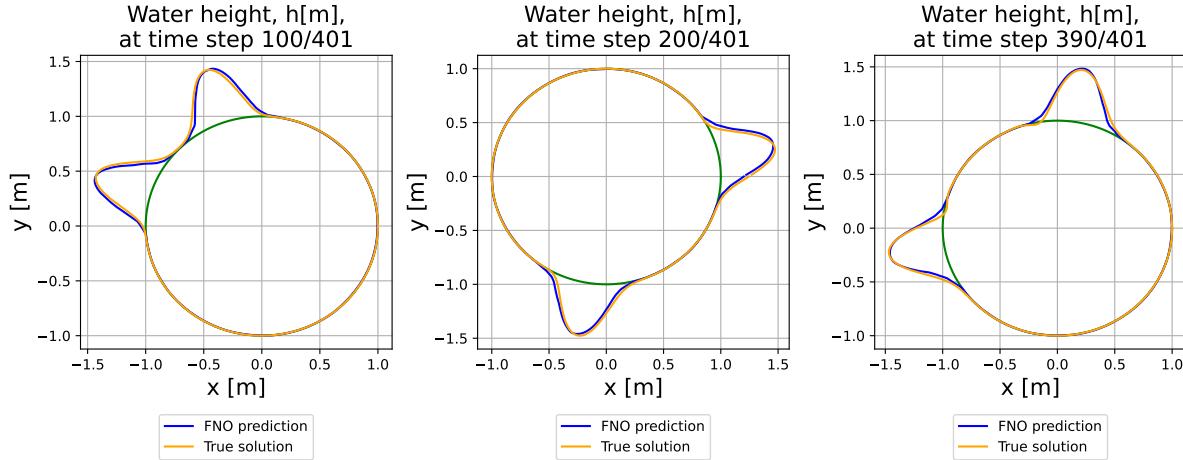


Figure 7.16: Predictions for the spherical shallow water equations in 1D using the FNO model for some given time steps.

Figure 7.16 demonstrates that the FNO model overall captures the waves, but has some minor delays.

Comparison

To compare the performance of the CNN and FNO models, we consider the MSE and MAE for the predictions of the 1D spherical LSWE, as well as the training time for the models. We are also interested in measuring how the models handle sharp edges in the solution, tending to discontinuities. Therefore, we introduce two new initial conditions for the water height h , as described in (5.1.2), with $\sigma = \frac{\pi}{8}$ and $\sigma = \frac{\pi}{32}$. This is done to see how the models perform for different types of initial conditions, depending on how smooth or discontinuous the solution is. The results are summarized in Table 7.4.

Model	$\sigma = \pi/8$			$\sigma = \pi/16$			$\sigma = \pi/32$		
	MSE	MAE	Time [s]	MSE	MAE	Time [s]	MSE	MAE	Time [s]
CNN	2.66e-06	8.27e-04	134.43	1.34e-05	1.60e-03	159.14	8.26e-04	1.11e-02	142.33
FNO	1.63e-04	7.51e-03	229.75	2.49e-04	7.80e-03	186.91	8.90e-04	1.46e-02	201.29

Table 7.4: Test loss in terms of MSE and MAE, and time for training the models for the 1D spherical SWE.

From Table 7.4 we observe that the CNN model is slightly faster and achieves a lower MSE and MAE than the FNO model across all three initial conditions. However, the difference for $\sigma = \pi/32$ is minimal, suggesting that the FNO model performs better at capturing discontinuities in the solution. Additionally, the FNO model exhibits more stable performance with new initial conditions, while the CNN model shows more variation.

It is also notable that the MAE is generally higher than the MSE and increases as σ decreases. This difference highlights that the MAE places greater emphasis on small errors compared to the MSE.

7.3 2D Shallow Water Equations with Gaussian initial conditions

In this section we present the results for the 2D SWE using data-driven models. The initial condition, a Gaussian function, is the same as in the 1D case but extended to two dimensions. We solve the 2D SWE using both a CNN

and and FNO model, comparing their performance in terms of run time and accuracy. Additionally, we compare their run time to the FVM to evaluate whether the data-driven models can serve as a faster alternative. We also assess the models' ability to transfer solutions across grids, by training the models on a coarse grid and making predictions on a finer grid. Finally, we evaluate the models' capability to generalize further in time, testing their long-term predictive performance.

The initial condition for the 2D problem is a Gaussian function as given in (5.1.3). The initial condition is illustrated in Figure 7.17.

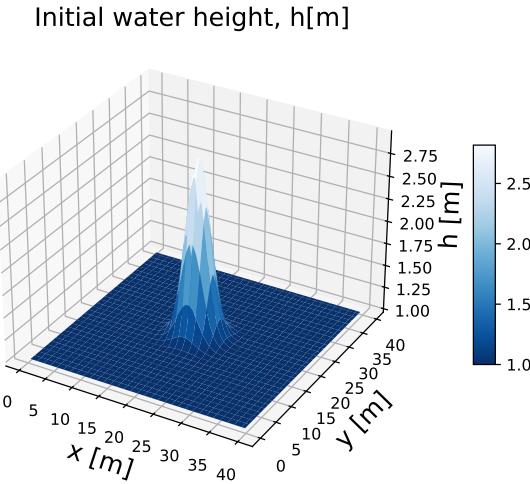


Figure 7.17: Initial condition for the 2D SWE.

In Figure 7.17 we see the initial condition for the 2D SWE problem, which is close to the 2D dam break problem, only here with a Gaussian function. The solution is generated from $t = 0$ to $t = 5$ s. Some information of the data used for the 2D SWE generated by the FVM can be found in Table 7.5.

Case	n_train	n_val	n_test	Δx	Δt
2D SWE, $N = 64$	48	16	17	0.625 m	[0.050 s, 0.073 s]
2D SWE, $N = 128$	99	33	33	0.3125 m	[0.025 s, 0.035 s]

Table 7.5: Details of the used data for the case with the 2D SWE for both a number of grid points $N = 64$ and $N = 128$.

In Table 7.5 we see that the train/validation/test data split is 60%/20%/20% for both $N = 64$ and $N = 128$. In the FVM we use a variable time step size Δt to ensure stability, which is why the time step size varies. We see that for $N = 64$ the time step size varies between 0.050 s and 0.073 s, while for $N = 128$ the time step size varies between 0.025 s and 0.035 s. The time step size for the higher N is in general smaller than for the lower N to ensure stability. This is also what gives more data points for $N = 128$ compared to $N = 64$.

CNN Model

We train a CNN model with several convolutional layers and ReLU activation functions. We use the Adam optimizer with a learning rate of 0.001 and a batch size of 32. We make predictions from $t = 0$ to $t = 5$. The model is trained

for 500 epochs. The training and validation loss for the 2D CNN model can be seen in Figure 7.18.

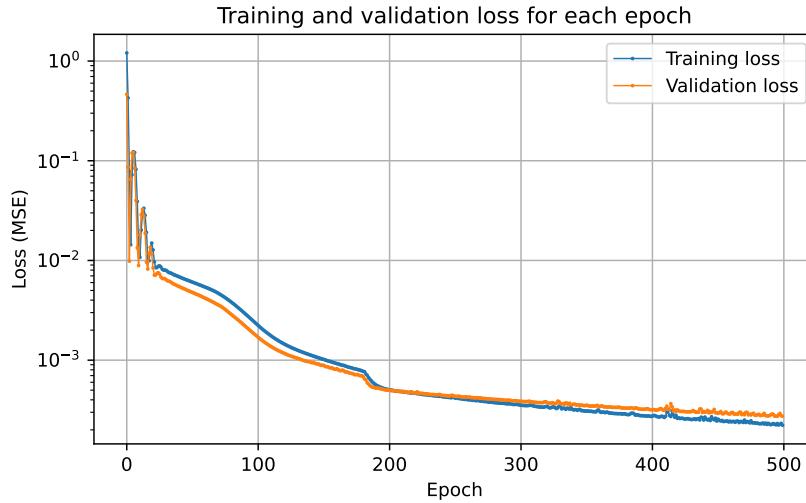


Figure 7.18: Training and validation loss for the 2D CNN model.

Figure 7.18 shows that the training and validation loss are decreasing, as a function of the number of epochs. The error plot for the last prediction for the 2D CNN can be seen in Figure 7.19.

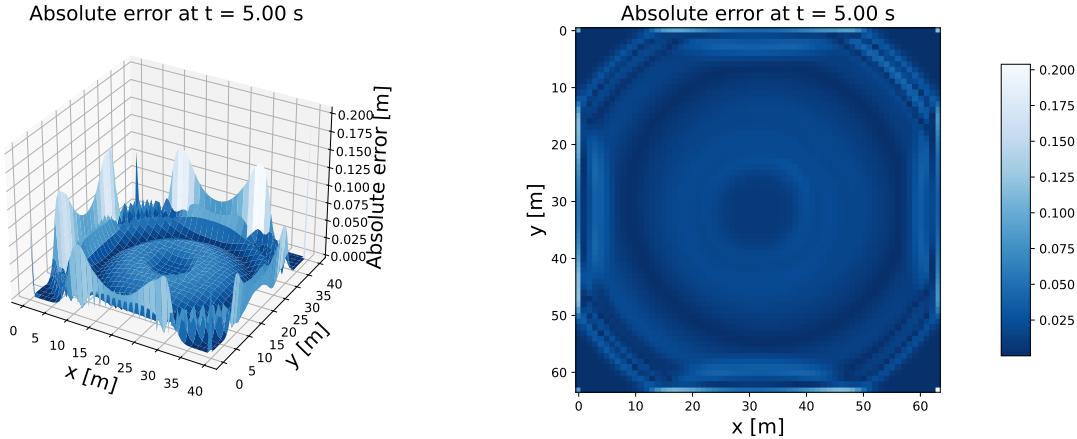


Figure 7.19: Error plot for the last prediction for the 2D CNN.

From Figure 7.19, we see that overall the absolute error is small. The error is largest at the boundaries of the domain. This may be due to the fact that in this case we are working with boundary conditions simulating a wall. Hence, when the wave hits the wall, the wave is reflected, and the model has to learn the reflection of the wave. This may also lead to discontinuities in the solution, which can be difficult for the model to handle.

FNO Model

We train a FNO model with the same training/validation/testing split as for the CNN model. The FNO model uses 8 Fourier modes, the Adam optimizer with a learning rate of 0.001, and a batch size of 10. The model is trained

for 500 epochs. The training and validation loss for the 2D FNO model can be seen in Figure 7.20.

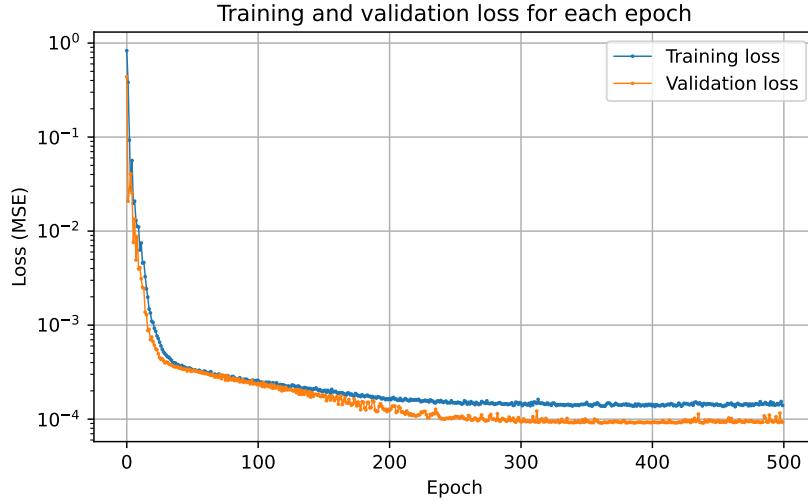


Figure 7.20: Training and validation loss for the 2D FNO model.

From Figure 7.20 we see that the training and validation loss are decreasing, and more or less stable with some fluctuations. The plot suggests that the model has converged. The error plot for the last prediction for the 2D FNO can be seen in Figure 7.21.

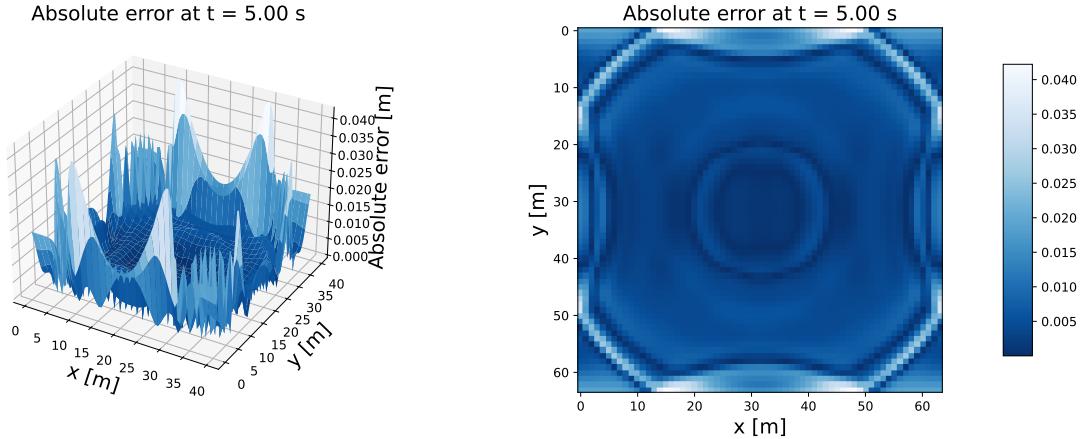


Figure 7.21: Error plot for the last prediction for the 2D FNO.

As for the CNN model, the error is largest at the boundaries. However, we notice that the absolute error is smaller for the FNO model compared to the CNN model.

Comparison

We compare the performance of the CNN and FNO models in terms of the MSE and MAE for the predictions. We test for both $N = 64$ and $N = 128$ grid points in each direction. The errors are calculated for the test data, i.e.,

the last 20% of the time steps. The results are summarized in Table 7.6, together with the time for training the models.

Model	$N = 64$				$N = 128$			
	Epochs	MSE	MAE	Time (s)	Epochs	MSE	MAE	Time (s)
CNN	500	4.96e-04	1.53e-02	65.18	100	1.34e-03	2.16e-02	114.00
FNO	500	1.56e-04	8.34e-03	549.55	100	4.63e-04	2.63e-02	1411.42

Table 7.6: Test loss in terms of MSE and MAE, and time for training the models for the 2D SWE.

The errors are low and similar for both models. However, we also observe that the FNO model generally requires more time to train. Based on theory, FNOs have demonstrated potential in transferring solutions across grids. To test this, we train the models on a coarse grid and then make predictions on a finer grid. The table below presents the results when the models are trained on a grid with $N = 64$ and subsequently used to make predictions on finer grids with $N = 128$ and $N = 256$.

Model	$N = 128$			$N = 256$		
	MSE	MAE	Prediction time [s]	MSE	MAE	Prediction time [s]
CNN	3.97e-04	1.47e-02	0.10	3.80e-04	1.48e-02	1.43
FNO	9.32e-05	6.20e-03	1.34	7.30e-05	5.15e-03	20.43

Table 7.7: Test loss in terms of MSE and MAE, and time for training the FNO model on a grid with $N = 64$ and then making predictions on a grid with $N = 128$ and $N = 256$.

From the results in Table 7.7 we see that FNO model achieves small errors both for $N = 128$ and $N = 256$, indicating its ability to generalize to a finer grid. The CNN model shows higher errors, but might still be accurate enough depending on the application. This ability to train the FNO model (and partially, also the CNN model) on a coarse grid and make predictions on a finer grid is a significant advantage. In this study, it is particularly useful for solving the SWE numerically, as solving the SWE on a fine grid using the FVM is computationally expensive. By comparing to the run time of the FVM, shown in Table 6.3, we see that the data-driven models are much faster, and thus, might provide an alternative to the FVM, depending on the specific application. This approach could help address the scalability challenges faced when solving the SWE numerically.

7.3.1 Long-term predictions for the 2D Shallow Water Equations

In this section, we evaluate the models' ability to generalize to longer time horizons. The previous models were trained on data generated using the FVM with variable time step sizes. To control the specific time points for long-term predictions, a fixed time step size is required. Consequently, we have generated data for the 2D SWE using $N = 64$ points in each spatial direction and a fixed time step size of $\Delta t = 0.025$ s. To facilitate testing long-term predictions, the data set spans a longer time interval, covering $t = 0$ s to $t = 15$ s. The models are trained on the data from $t = 0$ s to $t = 8$ s, validated on the data from $t = 8$ s to $t = 10$ s, and tested for long-term prediction capability using data from $t = 10$ s to $t = 15$ s. Information of the data used for the 2D SWE with a fixed time step size can be found in Table 7.8.

Case	n_train	n_val	n_predictions	Δx	Δt
2D SWE	480	120	200	0.625 m	0.025

Table 7.8: Details of the used data for the case with the 2D SWE with a fixed time step size.

We generate predictions for up to $n = 200$ time steps into the future, corresponding to the interval from $t = 10$

s to $t = 15$ s. The validation data points are crucial for selecting the best model for long-term predictions and prevent overfitting. Initially, we adopted an iterative approach, where the model made a prediction for the next time step and used that prediction as input for the next time step. While this approach produces acceptable results for the initial time steps, the error increases as we make predictions further in time, probably due to accumulating errors. To address this issue, we implemented an alternative approach by creating sequences of data and training the model to predict the next time step based on the entire sequence. In this method, the most recent prediction becomes part of a sequence, distributing the influence across subsequent predictions, rather than heavily impacting the next prediction alone. This approach aims to stabilize long-term predictions.

CNN Model

We train the CNN model on the described data set for 100 epochs. The model consists of several convolutional layers with ReLU activation functions. The sequence length is set to 50 and the model is trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 16. We make predictions for up to $n = 200$ time steps into the future, and the error plot for the prediction after $n = 20$ time steps, corresponding to 0.5 s, can be seen in Figure 7.22.

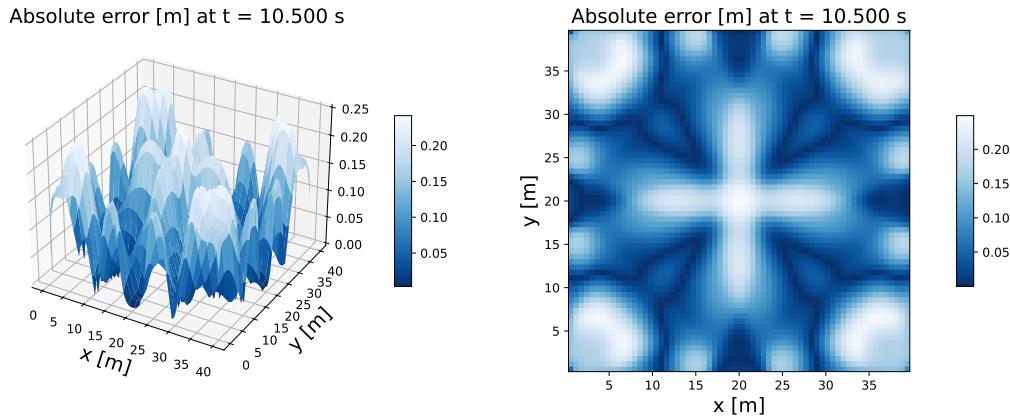


Figure 7.22: Error plot for the long-term prediction for the 2D CNN model.

In Figure 7.22 we see that the error is more equally distributed over the domain compared to the short-term predictions in Figure 7.19. The prediction, together with the numerical solution can be seen in Figure A.9 in Appendix section A.2.

FNO Model

We run the same test for the FNO model, training it on the same data set as the CNN model, and also for 100 epochs. The model has 12 Fourier modes and is trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 16. The sequence length is set to 50. The error plots of the long-term predictions from the FNO model after 20 time steps, equivalent to 0.5 s, can be seen in Figure 7.23.

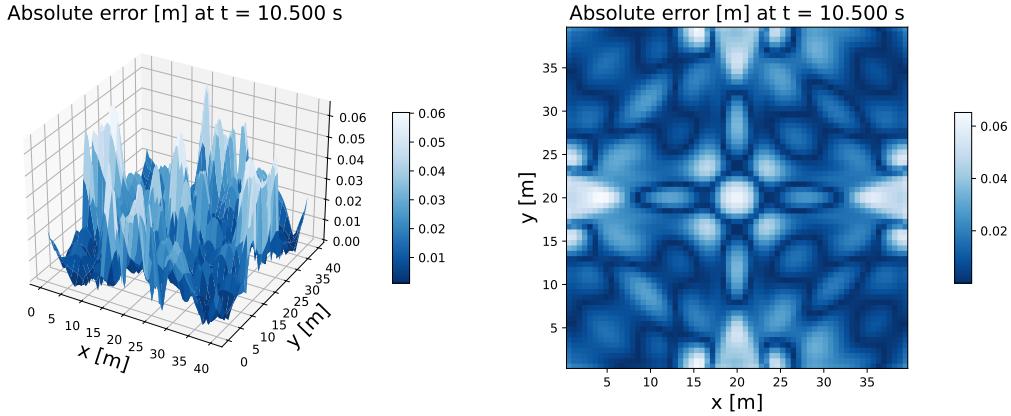


Figure 7.23: Error plot for the long-term prediction for the 2D FNO model.

From the error plot in Figure 7.23, we see that the error is bigger than for the short-term predictions in Figure 7.21, but for many applications still acceptable. We also see that error for the FNO model is smaller than for the CNN model, indicating that the FNO model is better able to generalize further in time and make long-term predictions. A plot of the predictions and the ground truth can be found in Figure A.10 in Appendix section A.2, where we see that the FNO model is able to capture the dynamics of the solution well.

Comparison

To assess the models performance over time, we plot the error at each time step for long-term predictions up to $n = 200$ time steps into the future. Then CNN and FNO models are compared in terms of MSE, MAE and the absolute max error for long-term predictions, as shown in Figure 7.24.

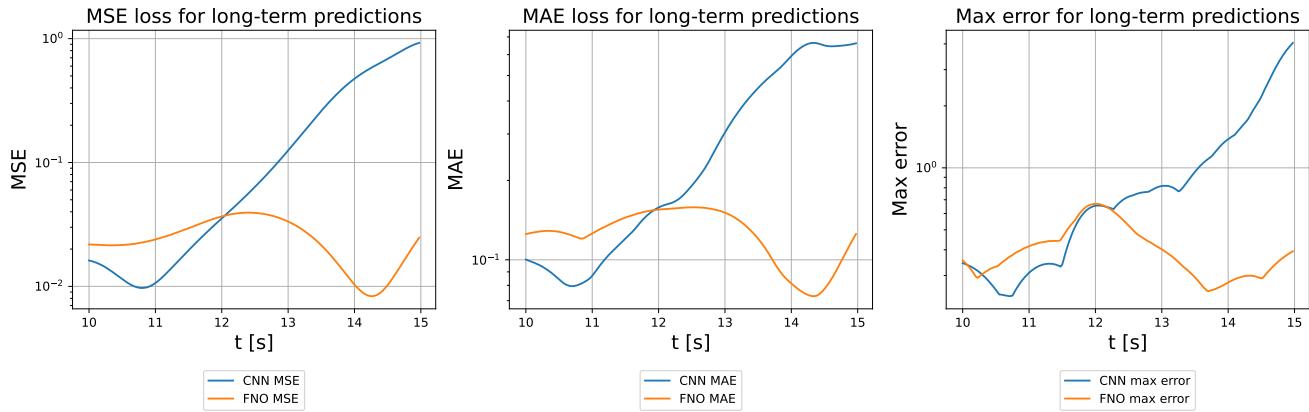


Figure 7.24: MSE, MAE and max error for the long-term predictions for the 2D CNN and FNO models.

When making predictions, it is essential to consider the physical properties of the system, rather than relying solely on the loss functions. The choice of error metric should align with the goals of the predictions. For instance, it may be more important to minimize the MSE or the maximum absolute error, depending on the physical properties of the system. This raises a key question: is it better to have one large error or many small errors?

As shown in Figure 7.24, the error increases as we make predictions further in time. Overall, the FNO model demonstrates lower errors compared to the CNN model, indicating it may offer more accurate long-term predictions.

Chapter 8

Discussion

- How good are the data-driven methods compared to the classic numerical methods, in this case the FVM?
- Shorter computation time = sustainability
- Advantages/disadvantages numerical / data-driven methods - compare with the literature For the numerical methods we can often estimate the error. IS that possible using data-driven methods?
- The data-driven methods could possibly perform better if trained on more data and for more epochs.
- Has somehow shown potential for what the data-driven methods can do? Supplement to numerical methods, depending on the application.
- The data-driven methods are not as transparent as the numerical methods. Or accurate.
- The data-driven methods are not as general as the numerical methods.
- Truncation errors. Generating data from the FVM, how much accuracy do we lose?
- Perspectives to the literature.

Chapter 9

Conclusion

- By training the FNO on a coarse grid and evaluate on a fine grid, the run time is XX times faster, while maintaining the same accuracy.
- Long-term prediction: the FNO can predict the solution for a longer time than the training time. It was demonstrated that training on a coarse grid was not suitable to facilitate a long-term prediction for a fine grid.
- Long-term prediction: the FNO model outperforms the CNN model for long-term prediction.
- Grid independence: when training on a coarse grid and evaluating on a fine grid, the predictions by the FNO is much better than the predictions by the CNN. Aligning with the theory that FNOs are grid independent.
- Efficiency: CNN is faster than FNO, but FNO is more accurate. Which we prefer depends on the application.
- CNN has shown potential for many of the same abilities as the FNO - especially in the 1D case.
- Some of the theoretical advantages of the FNO are better realized in the 2D case than in the 1D case.
- Which model we choose is depending on the application.

9.1 Further work

Bibliography

- [1] Niloofar Najibi and Naresh Devineni. Recent trends in the frequency and duration of global floods. <https://esd.copernicus.org/articles/9/757/2018/>, 2018.
- [2] European Space Agency (ESA). Valencia flood disaster, 2024. Accessed: 2024-12-27.
- [3] Fathom. The german flash floods 2021, 2021. Accessed: 2025-01-13.
- [4] The Guardian. Climate scientists shocked by scale of floods in germany. 2021. Accessed: 2025-01-13.
- [5] E.F. Toro. *Shock-Capturing Methods for Free-Surface Shallow Flows*. Oxford University Press, 2001.
- [6] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, 2009.
- [7] E.F. Toro. *Computational Algorithms for Shallow Water Equations*. Springer, 2024.
- [8] Ilya Peshkov. Advanced numerical methods for environmental modeling. <https://www.unitn.it/dricam/1116/advanced-numerical-methods-environmental-modeling>, 2023. Lecture notes.
- [9] Manuel J. Castro, Sergio Ortega, and Carlos Parés. Well-balanced methods for the shallow water equations in spherical coordinates. <https://www.sciencedirect.com/science/article/pii/S0045793017303237>, 2017.
- [10] Alex Bihlo and Roman O. Popovych. Physics-informed neural networks for the shallow-water equations on the sphere. *Journal of Computational Physics*, 456:111024, 2022.
- [11] David J. Raymond. Shallow water on a sphere. <http://kestrel.nmt.edu/~raymond/classes/ph332/notes/sphere/sphere.pdf>, New Mexico Tech. Accessed 2024.
- [12] Geophysical Fluid Dynamics Laboratory (GFDL). Shallow water equations, n.d. Accessed: 2025-01-12.
- [13] L. Gavete, B. Alonso, M.L. Gavete, F. Ureña, and J.J. Benito. An adaptive solver for the spherical shallow water equations. *Mathematics and Computers in Simulation*, 79(12):3466–3477, 2009. The International Conference on Approximation Methods and numerical Modeling in Environment and Natural Resources.
- [14] Joseph Galewsky, Richard K. Scott, and Lorenzo M. Polvani. An initial-value problem for testing numerical models of the global shallow-water equations. *Tellus A: Dynamic Meteorology and Oceanography*, Jan 2004.
- [15] Jan S. Hesthaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer, 2008.
- [16] Keiron O’Shea and Ryan Nash. An introduction to convolutional neural networks, 2015.
- [17] François Fleuret. A comprehensive guide to convolutional neural networks — the eli5 way. *Towards Data Science*, 2017. Accessed: 2024-12-27.

BIBLIOGRAPHY

- [18] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. [https://arxiv.org/pdf/2010.08895](https://arxiv.org/pdf/2010.08895.pdf), 2021.
- [19] Nathan Baker, Frank Alexander, Timo Bremer, Aric Hagberg, Yannis Kevrekidis, Habib Najm, Manish Parashar, Abani Patra, James Sethian, Stefan Wild, et al. Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence. Technical report, USDOE Office of Science (SC), Washington, D.C. (United States), 02 2019.
- [20] Boris Bonev, Christian Hundt, Thorsten Kurth, Jaideep Pathak, et al. Modeling earth's atmosphere with spherical fourier neural operators. <https://developer.nvidia.com/blog/modeling-earths-atmosphere-with-spherical-fourier-neural-operators/>, 2023.
- [21] Boris Bonev, Thorsten Kurth, Christian Hundt, Jaideep Pathak, Maximilian Baust, Karthik Kashinath, and Anima Anandkumar. Spherical fourier neural operators: Learning stable dynamics on the sphere. <https://arxiv.org/abs/2306.03838>, 2023.
- [22] C.B. Vreugdenhil. *Numerical Methods for Shallow-Water Flow*. Kluwer Academic Publishers, 1994.
- [23] Wikipedia. Leibniz integral rule. https://en.wikipedia.org/wiki/Leibniz_integral_rule, 7/10-2024.
- [24] Adrian E. Gill. *Atmosphere-Ocean Dynamics*. Academic Press, 1982.
- [25] Longitude and latitude map with degrees. <https://animalia-life.club/qa/pictures/longitude-and-latitude-map-with-degrees>, Accessed: 2024.
- [26] Wikipedia. Coriolis force. https://en.wikipedia.org/wiki/Coriolis_force, 14/11-2024.
- [27] Boris Bonev, Jan S. Hesthaven, Francis X. Giraldo, and Michal A. Kopera. Discontinuous galerkin scheme for the spherical shallow water equations with applications to tsunami modeling and prediction. *Journal of Computational Physics*, 362:425–448, 2018.
- [28] Claes Eskilsson and Spencer J. Sherwin. Discontinuous galerkin spectral/hp element modelling of dispersive shallow water systems. *Journal of Scientific Computing*, 22:269–288, 2005.
- [29] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002.
- [30] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser, Basel, 1992.
- [31] Fourth order runge-kutta. <https://lpsa.swarthmore.edu/NumInt/NumIntFourth.html>, Accessed: 15-01-2025.
- [32] Fourier transform. https://en.wikipedia.org/wiki/Fourier_transform, Accessed: 15-01-2025.
- [33] Siddarth Maddali. Spheremesh. <https://github.com/siddharth-maddali/SphereMesh/tree/master>.
- [34] Copernicus Climate Data Store. Climate data store (cds). <https://cds.climate.copernicus.eu/datasets/reanalysis-oras5?tab=overview>, Accessed: 2024-11-18.
- [35] Melissa Jessen. Shallow-water-equations. <https://github.com/MelissaJessen/Shallow-Water-Equations>, 2024-2025.
- [36] A. P Engsig-Karup and J. S. Hesthaven. An introduction to discontinuous galerkin methods for solving partial differential equations. <https://www2.compute.dtu.dk/~apek/DGFEMCourse2009/>, 2009. Lecture notes.
- [37] David Pitt. Neural operator: Learning operators for pdes. <https://github.com/neuraloperator/neuraloperator>, 2024. GitHub repository. Accessed: January 14, 2025.

Appendix A

Appendix

The code can be found at: <https://github.com/MelissaJessen/Shallow-Water-Equations>.

A.1 Additional figures from 1D SWE spherical CNN and FNO

Results for the other sigma values.

CNN, $\sigma = \frac{\pi}{8}$.

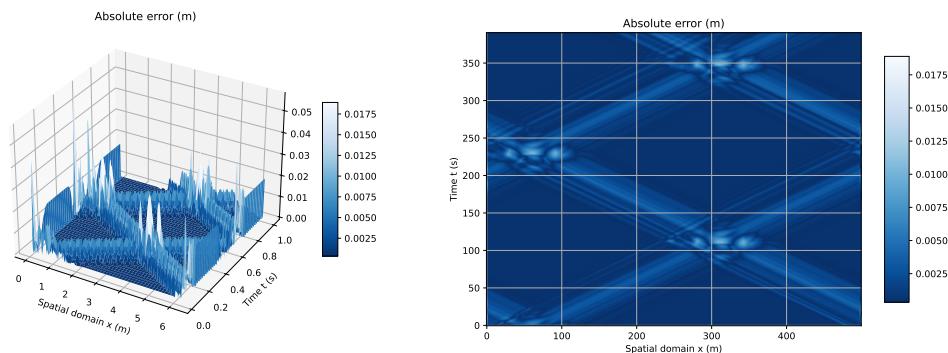


Figure A.1: Error for the 1D SWE spherical CNN with $\sigma = \frac{\pi}{8}$.

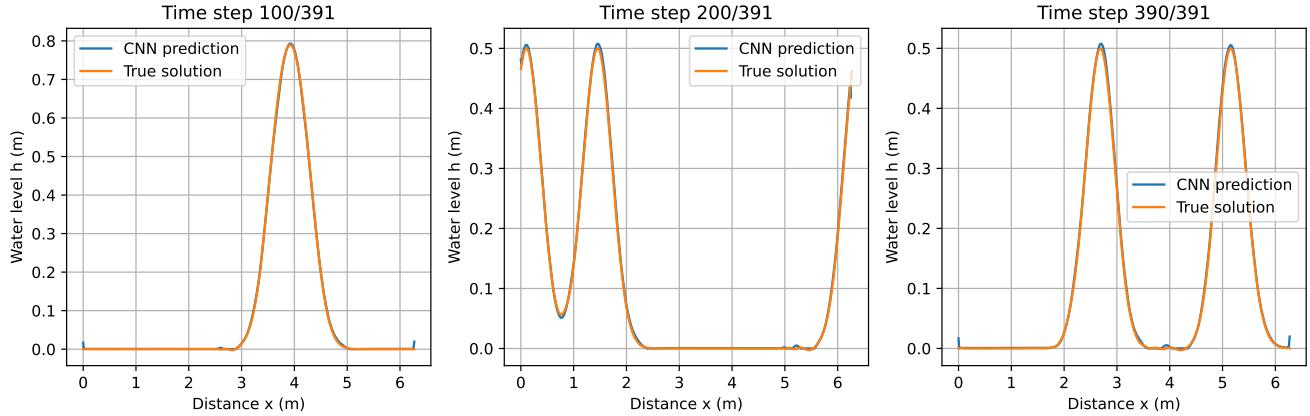


Figure A.2: Results for the 1D SWE spherical CNN with $\sigma = \frac{\pi}{8}$.

FNO, $\sigma = \frac{\pi}{8}$.

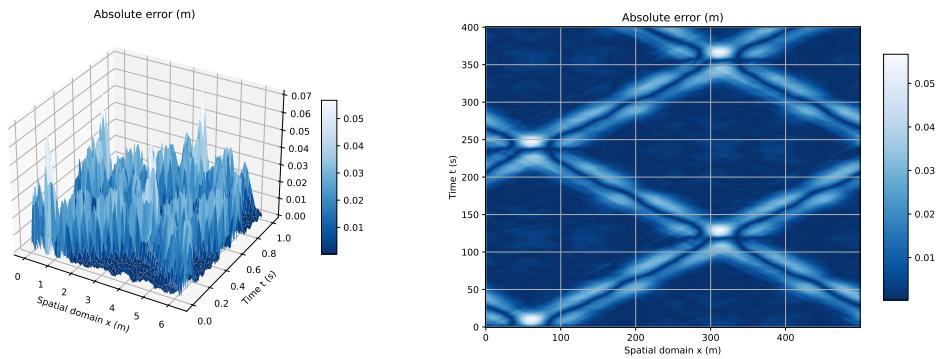


Figure A.3: Error for the 1D SWE spherical FNO with $\sigma = \frac{\pi}{8}$.

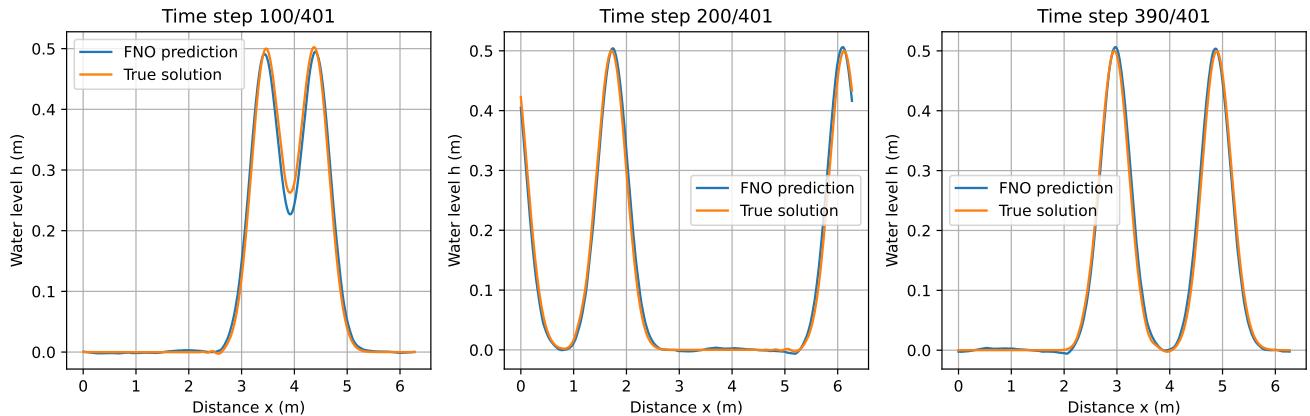


Figure A.4: Results for the 1D SWE spherical FNO with $\sigma = \frac{\pi}{8}$.

CNN $\sigma = \frac{\pi}{32}$.

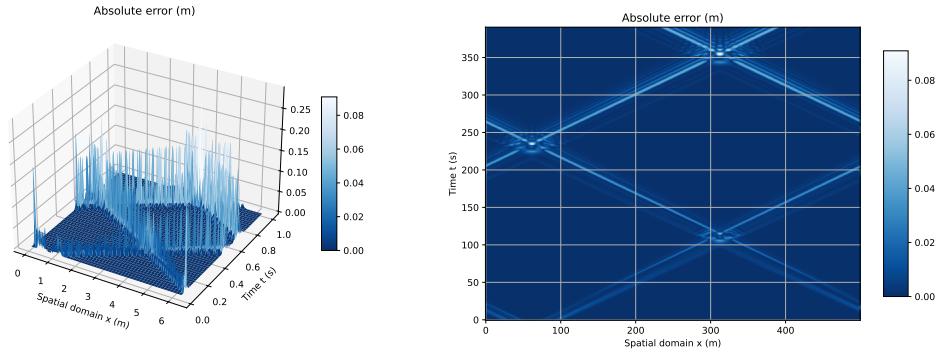


Figure A.5: Error for the 1D SWE spherical CNN with $\sigma = \frac{\pi}{32}$.

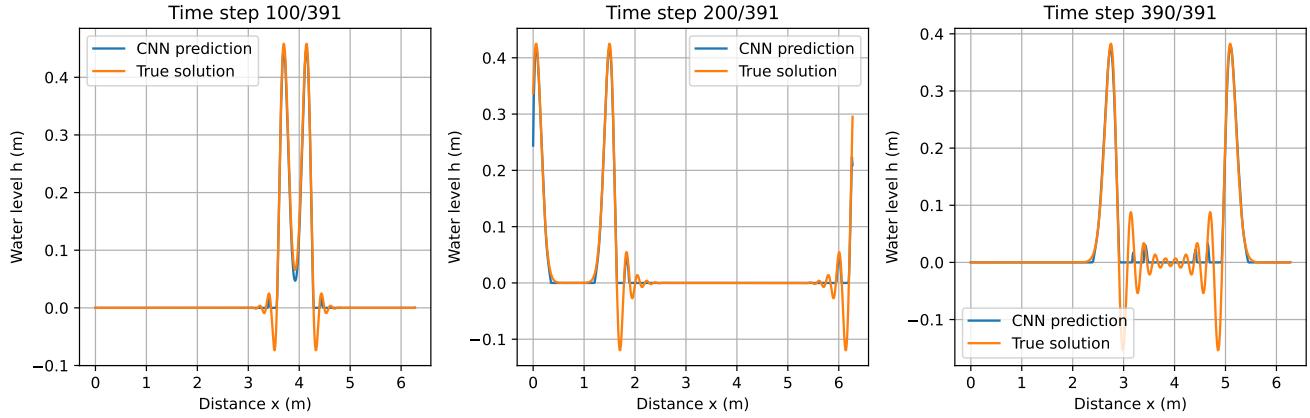


Figure A.6: Results for the 1D SWE spherical CNN with $\sigma = \frac{\pi}{32}$.

FNO $\sigma = \frac{\pi}{32}$.

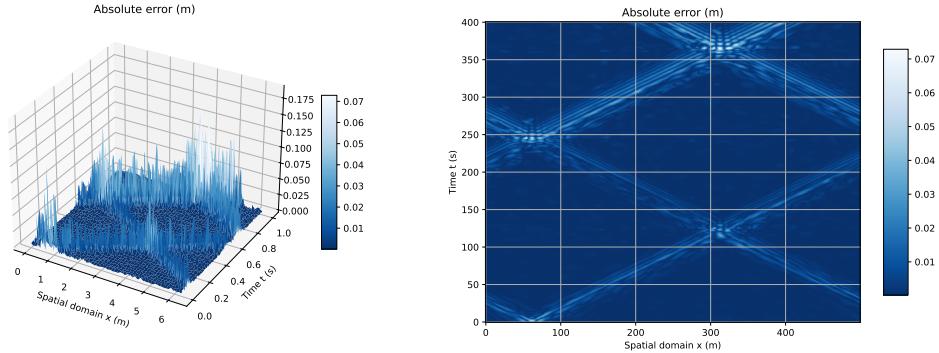


Figure A.7: Error for the 1D SWE spherical FNO with $\sigma = \frac{\pi}{32}$.

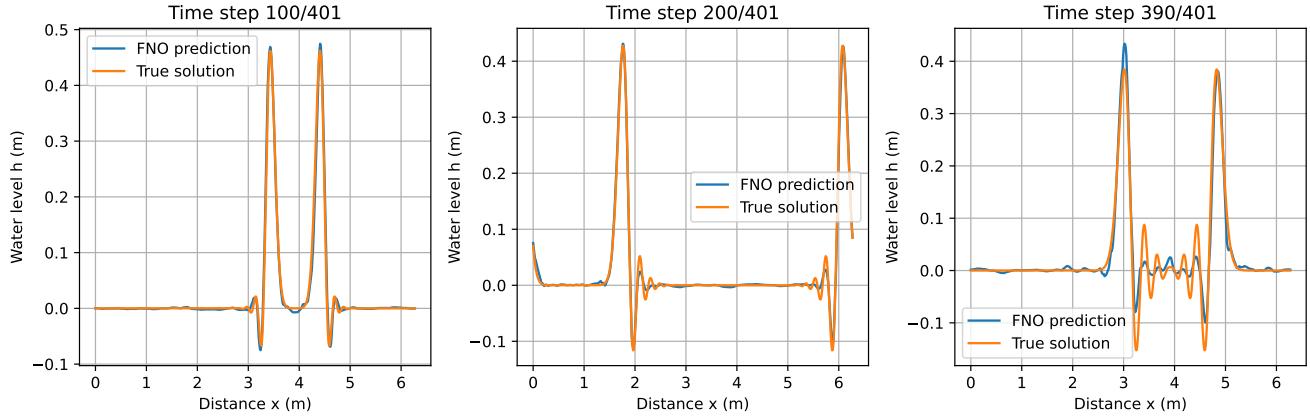


Figure A.8: Results for the 1D SWE spherical FNO with $\sigma = \frac{\pi}{32}$.

A.2 2D SWE long term prediction

The results for the long-term predictions for the CNN model can be seen in Figure A.9.

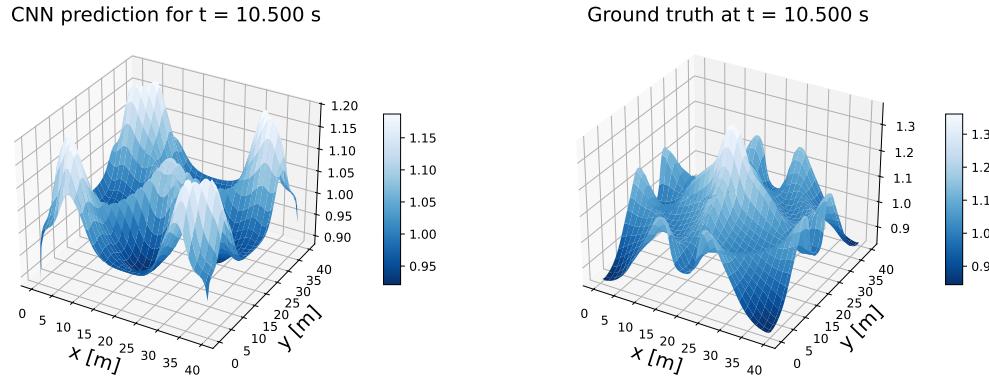


Figure A.9: Long-term prediction for the 2D CNN model.

The results for the long-term predictions for the FNO model can be seen in Figure A.10.

FNO prediction for $t = 10.500$ s Ground truth at $t = 10.500$ s

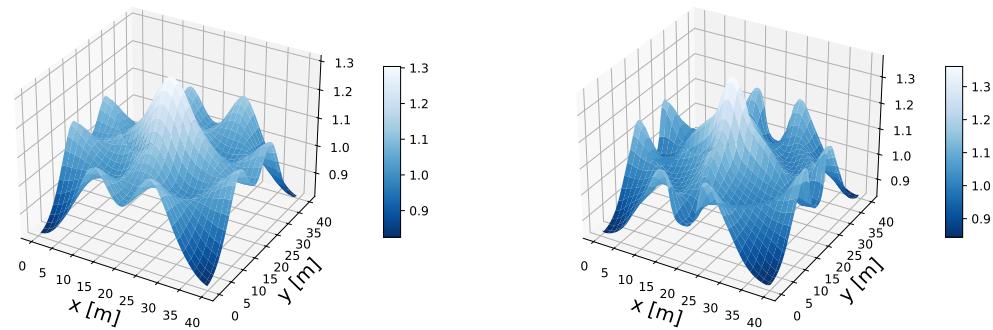


Figure A.10: Long-term prediction for the 2D FNO model.